الجمهورية الجزائرية الديمقراطية الشعبية

People's Democratic Republic of Algeria

وزارة التعليم العالي و البحث العلمي

Ministry of Higher Education and Scientific Research

Mohamed Khider University - Biskra
Faculty of Science and Technology
Department of Electrical Engineering
Ref: ....

جامعة محمد خضر – بسكرة
كلية العلوم والتكنولوجيا
قسم الهندسة الكهربائية
المرجع: ....

The final thesis submitted to the department of Electrical Engineering in candidacy for the

PhD Degree in Electrical Engineering

Major: Automation

Option: Automation

Initial title:

Intelligent control of agriculture production in greenhouses

Presented by:

Mounir Guesbaya

Discussed publicly on December 19th, 2022

The jury consists of:

Pr. Megherbi Ahmed Chaouki	**Prof**	**President**
Pr. Megherbi Hassina	**Prof**	**Supervisor**
Pr. Francisco Rodríguez Díaz	**Full Prof**	**Co-supervisor**
Dr. Touba Mohamed Mostafa	**MCA**	**Examiner**
Dr. Cherroun Lakhmissi	**MCA**	**Examiner**

الجمهورية الجزائرية الديمقراطية الشعبية
People's Democratic Republic of Algeria
وزارة التعليم العالي و البحث العلمي
Ministry of Higher Education and Scientific Research

Mohamed Khider University - Biskra
Faculty of Science and Technology
Department of Electrical Engineering
Ref: ....

جامعة محمد خضر – بسكرة
كلية العلوم والتكنولوجيا
قسم الهندسة الكهربائية
المرجع: ....

The final thesis submitted to the department of Electrical Engineering in candidacy for the

PhD Degree in Electrical Engineering

Major: Automation

Option: Automation

Final title:

Adaptive modelling of greenhouse climate using metaheuristic algorithms and machine learning techniques

Presented by:

Mounir Guesbaya

Discussed publicly on December 19th, 2022

The jury consists of:

Pr. Megherbi Ahmed Chaouki	**Prof**	**President**
Pr. Megherbi Hassina	**Prof**	**Supervisor**
Pr. Francisco Rodríguez Díaz	**Full Prof**	**Co-supervisor**
Dr. Touba Mohamed Mostafa	**MCA**	**Examiner**
Dr. Cherroun Lakhmissi	**MCA**	**Examiner**

To my parents

To my little sisters and brothers

DEDICATION

I dedicate this modest work to:

My beloved parents

My little sisters and brothers

All my relatives

All my colleagues and friends

All my professors throughout my academic journey

ACKNOWLEDGEMENTS

First of all, praise is to Allah with all my heart for giving me the courage and patience to accomplish this modest work. I wish to express my profound thanks to my supervisors, **Pr. Hassina Megherbi** from the University of Biskra and **Pr. Francisco Rodríguez Díaz** from the University of Almería (Spain) for believing in me and for all kinds of support, assistance and guidance.

I thank the jury members for having agreed to review and evaluate this work:

Pr. Megherbi Ahmed Chaouki, professor at the University of Biskra and jury president.

Dr. Touba Mohamed Mostafa, professor at the University of Biskra.

Dr. Cherroun Lakhmissi, professor at the University of Djelfa.

I tremendously and deeply thank my family for their encouragement, advice, sacrifice and emotional and financial support.

Finally, I thank my colleagues and friends for all kinds of support.

ABSTRACT

The agricultural greenhouse system has undergone significant developments in recent years. Greenhouse microclimate is the phenomenon under study in this work. Its modelling and control processes are complex tasks to be performed mainly due to the strong nonlinearity of the phenomenon. In this thesis, a set of contributions in greenhouse microclimate modelling and control, including implementing computational intelligence algorithms, have been accomplished. The second chapter briefly describes the experimental greenhouses used in this thesis. Initially, due to the lack of an experimental greenhouse, a wooden-structured polyethene-covered greenhouse prototype was constructed and used as a small-scale nursery under arid climate conditions (moderate desert climate) in Meziraa, Biskra, Algeria. A low-cost microcontroller-based data acquisition system with a wireless connection was designed (hardware and software) and installed in the greenhouse with several low-cost sensors. It was used to gather instant information on the essential inside and outside climate variables. A dataset of five days was successfully acquired for modelling, estimation and experimental validation purposes. Secondly, a metal-structured polyethene-covered commercial-sized experimental greenhouse under Mediterranean climate conditions was exploited. It is located at "Las Palmerillas" Experimental Station, a property of the Cajamar Foundation in Almería, Spain. It is equipped with all the necessary professional sensors, actuators and data acquisition systems. A set of sufficient reliable datasets of fifteen days were obtained in different agri-seasons and used for different purposes such as microclimate modelling and control, online parameter estimation and real-time experimental validation.

In the third chapter, two contributions were achieved. Firstly, a grey-box model for greenhouse temperature prediction under moderate desert climate conditions has been proposed. This contribution stands on reformulating a white-box model to make it independent of the availability of accurate values of the static parameters of its elements. The model has become less complicated by alleviating the coupling between its parameters, which makes it easier for the identification algorithm to find the optimal parameter values. A variant of the Particle swarm optimisation algorithm (PSO) called Random Inertia Weight PSO (RIWPSO) was used to identify the parameters of the proposed model by calibrating it against the experimental data. The constructed greenhouse prototype has been used to validate the proposed temperature model. The simulation results show that particle swarm optimisation has successfully achieved the desired optimality. The experimental validation process has confirmed the suitability of this model to be implemented to study and predict the greenhouse temperature, and it has emphasised the successful prediction with satisfactory accuracy. Secondly, an enhanced variant of the bio-inspired metaheuristic Bat Algorithm (BA) has been proposed and called the Random Scaling-based Bat Algorithm (RSBA). The proposition includes modifying the exploitation of the standard BA by randomly making the scaling parameter changes over the iterations. It has been dedicated to the same task of calibrating the proposed thermal grey-box model. It has been assessed as the same as PSO, primarily on the same simulated greenhouse temperature model with the assumed parameters. The simulation results have shown the superiority of the proposed RSBA compared to the standard BA in terms

of convergence and performance accuracy. To experimentally investigate the proposed RSBA algorithm, the same experimental dataset from the greenhouse prototype has been used. The obtained prediction results are found to be in good agreement with the measured ones, which show the effectiveness of the proposed RSBA in identifying the real greenhouse thermal model. Finally, a comparative study was conducted between the RSBA and the RIWPSO. The BA has shown a faster convergence than PSO at the start of optimisation, but its convergence speed was reduced at the end. BA and PSO have shown superb performance in accurately finding the optimal solutions. However, PSO has shown a superior performance than BA in terms of time consumption regarding the problem of interest.

Greenhouse microclimate modelling is a difficult task mainly due to the strong nonlinearity of the phenomenon and the uncertainty of the involved physical and non-physical parameters. The uncertainty stems from the fact that most of these parameters are unmeasurable or difficult to measure, and some are time-varying, signifying the necessity to estimate them. As the first contribution in the fourth chapter of the thesis, a methodology for online parameter estimation is proposed to estimate the time-varying parameters of a simplified greenhouse temperature model for real-time model adaptation purposes. An online estimator is developed based on an enhanced variant of the Bat Algorithm called the Random Scaling-based Bat Algorithm. It allows the continuous adaptation of the internal air temperature model and the internal solar radiation sub-model by estimating their parameters simultaneously by minimising a cost function, intending to achieve global optimality. Constraints on the search ranges are imposed to respect the physical sense. The adaptation of the models was tested with recorded datasets of different agri-seasons and on a real greenhouse in real time. The evolutions of the time-varying parameters were graphically presented and thoroughly discussed. The experimental results illustrate the successful model adaptation, presenting an average error of less than $0.28\ ^{\circ}\mathrm{C}$ for air temperature prediction and $20\ \mathrm{W\ m^{-2}}$ for solar radiation simulation. It proves the usefulness of the proposed methodology under changing environmental conditions.

Natural ventilation flux is an important variable to measure or estimate for its significant effect on greenhouse microclimate modelling and control. It is commonly known that it can be mathematically estimated depending on the type and dimension of the greenhouse and its vents and, most importantly, on the vents opening percentage. However, most commercial greenhouses are not equipped with an automatic vent opening system which obligates the grower to perform manual control, in addition to the lack of vent position sensors, due to economic and management reasons. It leads to the absence of the control signal variable representing the vents opening percentage necessary for ventilation flux estimation. This issue has been encountered in this work after attempting to implement the developed adaptive microclimate model based on the online parameter estimator through an IoF2020 platform (internet of food and farm) in a set of commercial greenhouses with manually controlled vents located in Almeria province, Spain. To cope with this issue, the estimation of ventilation flux without using the vent opening percentage was investigated. As a second contribution in the fourth chapter, a virtual sensor for greenhouse ventilation flux estimation is proposed. It has been developed using a nonlinear autoregressive

neural network with exogenous inputs based on principal component analysis using the available measured data and the evolutions of the heat fluxes representing the greenhouse energy balance. Preliminary results show an encouraging performance of the virtual sensor in estimating the ventilation flux with a mean absolute error of 0.41 $m^3 s^{-1}$.

Keywords: Protected agriculture, greenhouse system, evolutionary algorithms, online estimation, model adaptation, machine learning, principal component analysis, artificial neural networks, virtual sensors.

الملخص

شهد نظام الدفيئة الزراعية تطورات كبيرة في السنوات الأخيرة. المناخ المحلي للدفيئه الزراعيه هو الظاهرة قيد الدراسة في هذا العمل. السيطرة على هذه الظاهرة يعني بشكل مباشر التحكم في نمو المحصول. تعد نمذجة المناخ المحلي للدفيئه الزراعيه والتحكم فيه من المهام الصعبة التي يجب القيام بها بشكل أساسي بسبب عدم الخطية القوية للظاهرة. في هذه الأطروحة، تم إنجاز مجموعة من المساهمات في نمذجة المناخ المحلي للدفيئه الزراعيه والتحكم فيه بما في ذلك تنفيذ خوارزميات الذكاء الحسابي. في الفصل الثاني تم وصف البيوت البلاستيكية التجريبية المستخدمة في هذه الرسالة بإيجاز. في البداية، نظرًا لعدم توفر دفيئه زراعية تجريبية، تم إنشاء نموذج أولي لبيت زجاجي مصنوع من الخشب ومغطى بالبولي إيثيلين واستخدامه كمشتل صغير الحجم في ظل ظروف مناخية قاحلة (مناخ صحراوي معتدل) يقع في مدينة مزيرعة، بسكرة، الجزائر. تم تصميم نظام اكتساب بيانات منخفض التكلفة قائم على وحدة تحكم دقيقة مع اتصال لاسلكي (أجهزة وبرامج) وتم تثبيته في الدفيئة مع العديد من أجهزة الاستشعار منخفضة التكلفة. تم استخدامه لجمع معلومات فورية عن المتغيرات المناخية الأساسية داخل وخارج الدفيئه. تم الحصول على مجموعة بيانات مدتها خمسة أيام بنجاح لأجل استخدامها في أغراض النمذجة والمحاكاة والتحقق التجريبي. ثانيًا، تم استغلال دفيئه زراعية تجريبية ذات حجم تجاري ومغطاة بالبولي إيثيلين ذات هيكل معدني تحت ظروف مناخ البحر الأبيض المتوسط. تقع الدفيئه المستخدمه في محطة "Las Palmerillas" التجريبية وهي ملك لمؤسسة Cajamar في ألميريا، إسبانيا. هي مجهزة بجميع المستشعرات والمشغلات وأنظمة الحصول على البيانات اللازمة. تم الحصول على عدة مجموعات من البيانات الموثوقة لمدة خمسة عشر يومًا في مواسم زراعية مختلفة واستخدمت لأغراض مختلفة مثل نمذجة المناخ المحلي، وتقدير المعلمات عبر الاتصال المباشر والتحقق التجريبي منها في الوقت الفعلي.

في الفصل الثالث، تم تحقيق مساهمتين. أولاً، تم اقتراح نموذج المربع الرمادي للتنبؤ بدرجة حرارة الدفيئة في ظل ظروف مناخية صحراوية معتدلة. تعتمد هذه المساهمة على إعادة صياغة نموذج الصندوق الأبيض لجعله مستقلاً عن توافر القيم الدقيقة للمعلمات الثابتة لعناصره. أصبح النموذج أقل تعقيدًا من خلال التخفيف من الاقتران بين معلماته، مما يسهل العثور على قيم المعلمات المثلى بواسطة خوارزمية تحديد المعلمات. تم استخدام متغير من خوارزمية Particle Swarm Optimization (PSO) يسمى Random Inertia Weight PSO (RIWPSO) لتحديد معلمات النموذج المقترح من خلال معايرته باستخدام البيانات التجريبية. تم استخدام النموذج الأولي للدفيئه الزراعيه المصمم للتحقق من صحة نموذج درجة الحرارة المقترح. تظهر نتائج المحاكاة أن النتيجه الأمثل المطلوبه قد تم تحقيقها بنجاح باستخدام خوارزمية تحسين سرب الجسيمات. أكدت عملية التحقق التجريبية ملاءمة هذا النموذج ليتم تنفيذه لدراسة درجة حرارة الدفيئة والتنبؤ بها، كما أكدت على التنبؤ الناجح بدقة مرضية. ثانيًا، تم اقتراح متغير محسّن من Bat Algorithm (BA) وتسمى Random Scaling-based Bat (RSBA) Algorithm. يتضمن الاقتراح تعديل خاصية الاستغلال في النسخه الاصليه للخوارزمية عن طريق إجراء تغييرات في معلمة القياس بشكل عشوائي على طول التكرارات. وقد تم تخصيصه لنفس المهمة وهي معايرة نموذج الصندوق الرمادي الحراري المقترح سابقا. تم تقييمه ينفس طريقة تقييم PSO، بشكل أساسي على نفس نموذج محاكاة درجة حرارة الدفيئة مع المعلمات المفترضة. أظهرت نتائج المحاكاة تفوق RSBA المقترح مقارنة بالنسخه الاصليه BA من حيث التقارب ودقة الأداء. للتحقيق التجريبي في خوارزمية RSBA المقترحة، تم استخدام نفس مجموعة البيانات التجريبية من النموذج الأولي للدفيئه الزراعيه. تم العثور على نتائج التنبؤ التي تم الحصول عليها في اتفاق جيد مع تلك التي تم قياسها والتي تظهر فعالية RSBA المقترحة في معايرة النموذج الحراري الحقيقي للدفيئه الزراعيه. أخيرًا، تم إجراء دراسة مقارنة بين RSBA و

RIWPSO حيث أظهر RSBA تقاربًا أسرع من RIWPSO في بداية التحسين ولكن انخفضت بشكل ملحوظ في النهاية. أظهر كل من BA و RIWPSO أداءً رائعًا في إيجاد الحلول المثلى. ومع ذلك، فقد أظهرت RIWPSO اداءً أفضل من RSBA من حيث استهلاك الوقت ودقه النتائج فيما يتعلق بالمشكله المدروسه.

تعتبر نمذجة المناخ المحلي للدفيئه الزراعيه مهمة صعبة بسبب عدم الخطية القوية للظاهرة وعدم اليقين من المعلمات الفيزيائية وغير الفيزيائية المعنية. ينبع عدم اليقين من حقيقة أن غالبية هذه المعلمات غير قابلة للقياس أو يصعب قياسها وبعضها متغير بمرور الوقت، مما يدل على ضرورة تقديرها. كأول مساهمة في الفصل الرابع من الأطروحة، تم اقتراح منهجية لتقدير المعلمات عبر الاتصال المباشر للتعامل مع تقدير المعلمات المتغيرة بمرور الوقت لنموذج مبسط لحرارة الدفيئة لأغراض التكيف في الوقت الحقيقي. تم تطوير مقدر عبر الاتصال المباشر بناءً على الخوارزمية المحسّنه والمقترحه سابقا RSBA. المقدر المطور يسمح بالتكيف المستمر لنموذج حرارة الهواء الداخلي والنموذج الفرعي للإشعاع الشمسي الداخلي، من خلال تقدير معلماتهما في نفس الخطوة الزمنية عن طريق تقليل دالة التكلفة، بهدف تحقيق النتائج المثلى. يتم فرض قيود على نطاقات البحث لاحترام المعنى الفزيائي. تم اختبار تكيف النموذج مع مجموعات البيانات المسجلة للمواسم الزراعية المختلفة وعلى دفيئة حقيقية في الوقت الحقيقي. تم عرض تطورات المعلمات المتغيرة بمرور الوقت بشكل بياني ومناقشتها بدقة. توضح النتائج التجريبية التكيف الناجح للنموذج، حيث تعرض متوسط خطأ أقل من $0.28\ °\mathrm{C}$ للتنبؤ بدرجة حرارة الهواء و $20\ \mathrm{W\ m^{-2}}$ لمحاكاة الإشعاع الشمسي. هذا يثبت فائدة المنهجية المقترحة في ظل الظروف البيئية المتغيرة.

يعتبر تدفق التهوية الطبيعي متغيرًا مهمًا يجب قياسه أو تقديره لمراعاة تأثيره الكبير على نمذجة المناخ المحلي للدفيئه الزراعيه والتحكم فيه. من المعروف أنه يمكن تقديرها رياضيًا اعتمادًا على نوع وأبعاد الدفيئه الزراعيه وفتحاتها، والأهم من ذلك على نسبة فتح الفتحات. ومع ذلك، فإن معظم الصوبات التجارية غير مجهزة بنظام فتح أوتوماتيكي للتهوية والذي يلزم المزارع بأداء التحكم اليدوي، بالإضافة إلى عدم وجود حساسات موضع التنفيس لأسباب اقتصادية. يؤدي هذا إلى عدم وجود متغير إشارة التحكم الذي يمثل النسبة المئوية لفتح الفتحات اللازمة لتقدير تدفق التهوية. تمت مواجهة هذه المشكلة في هذا العمل بعد محاولة تنفيذ نموذج المناخ المحلي التكيفي المطور استنادًا إلى مقدر المعلمة عبر الاتصال المباشر من خلال منصة IoF2020 (إنترنت للأغذية والمزرعة) في مجموعة من البيوت الزجاجية التجارية مع فتحات يتم التحكم فيها يدويًا وتقع في مقاطعة ألميريا، إسبانيا. للتغلب على هذه المشكلة، تم فحص تقدير تدفق التهوية بدون استخدام نسبة فتحة التهوية. كمساهمة ثانية في الفصل الرابع، تم اقتراح جهاز استشعار افتراضي لتقدير تدفق تهوية الدفيئة. تم تطويره باستخدام الشبكة العصبية ذاتية الانحدار غير الخطية مع مدخلات خارجية (NARX) بناءً على تحليل المكون الرئيسي (PCA) باستخدام البيانات المقاسة المتاحة وتطورات التدفقات الحرارية التي تمثل توازن طاقة الدفيئة. تظهر النتائج الأولية أداءً مشجعًا للحساس الافتراضي في تقدير تدفق التهوية بمتوسط خطأ مطلق يبلغ $0.41\ \mathrm{m^3\ s^{-1}}$.

الكلمات المفتاحيه: الزراعة المحمية ، نظام الدفيئة ، الخوارزميات التطورية ، التقدير عبر الإنترنت ، تكييف النموذج ، التعلم الآلي ، تحليل المكونات الرئيسية ، الشبكات العصبية الاصطناعية ، أجهزة الاستشعار الافتراضية.

RÉSUMÉ

Le système de serre agricole a connu des développements majeurs ces dernières années. Le microclimat en serre est le phénomène étudié dans ce travail. Son contrôle signifie essentiellement et directement le contrôle de la croissance des cultures. La modélisation et le contrôle du microclimat en serre sont des tâches difficiles à réaliser principalement en raison de la forte non-linéarité du phénomène. Dans cette thèse, un ensemble de contributions à la modélisation et au contrôle du microclimat à effet de serre, y compris la mise en œuvre d'algorithmes d'intelligence informatique, a été réalisée. Dans le deuxième chapitre, les serres expérimentales utilisées dans cette thèse sont brièvement décrites. Initialement, en raison de l'absence d'une serre expérimentale, un prototype de serre à structure en bois recouvert de polyéthylène a été construit et utilisé comme pépinière à petite échelle dans des conditions climatiques arides (climat désertique modéré) situé à Meziraa, Biskra, Algérie. Un système d'acquisition de données à faible coût basé sur un microcontrôleur avec connexion sans fil a été conçu (matériel et logiciel) et installé dans la serre avec plusieurs capteurs à faible coût. Il a été utilisé pour recueillir des informations instantanées sur les variables climatiques intérieures et extérieures essentielles. Un ensemble de données de cinq jours a été acquis avec succès pour être utilisé à des fins de modélisation, de simulation et de validation expérimentale. Deuxièmement, une serre expérimentale de taille commerciale recouverte de polyéthylène à structure métallique dans des conditions climatiques méditerranéennes a été exploitée. Il est situé à la station expérimentale « Las Palmerillas » qui est une propriété de la Fondation Cajamar à Almería, en Espagne. Il est équipé de tous les capteurs, actionneurs et systèmes d'acquisition de données professionnels nécessaires. Un ensemble d'ensembles de données fiables et suffisants de quinze jours a été obtenu au cours de différentes saisons agricoles et utilisé à différentes fins telles que la modélisation et le contrôle du microclimat, l'estimation des paramètres en ligne et la validation expérimentale en temps réel.

Dans le troisième chapitre, deux contributions ont été obtenues. Premièrement, un modèle de boîte grise pour la prévision de la température des serres dans des conditions climatiques désertiques modérées a été proposé. Cette contribution repose sur la reformulation d'un modèle boîte blanche pour le rendre indépendant de la disponibilité de valeurs précises des paramètres statiques de ses éléments. Le modèle est devenu moins compliqué en atténuant le couplage entre ses paramètres, ce qui permet à l'algorithme d'identification de trouver plus facilement les valeurs optimales des paramètres. Une variante de l'algorithme Particle Swarm Optimization (PSO) appelé Random Inertia Weight PSO (RIWPSO) a été utilisée pour identifier les paramètres du modèle proposé en le calibrant par rapport aux données expérimentales. Le prototype de serre construit a été utilisé pour valider le modèle de température proposé. Les résultats de la simulation montrent que l'optimalité souhaitée a été atteinte avec succès en utilisant l'optimisation de l'essaim de particules. Le processus de validation expérimentale a confirmé la pertinence de ce modèle à mettre en œuvre pour étudier et prédire la température de la serre, et il a souligné la réussite de la prédiction avec une précision satisfaisante. Deuxièmement, une variante améliorée de l'algorithme métaheuristique de chauve-souris bio-inspiré (BA) a été proposée et appelée algorithme de chauve-

souris basé sur l'échelle aléatoire (RSBA). La proposition comprend la modification de l'exploitation du BA standard en faisant changer le paramètre d'échelle de manière aléatoire au cours des itérations. Il a été consacré à la même tâche d'étalonnage du modèle de boîte grise thermique proposé. Il a été évalué de la même manière que le PSO, principalement sur le même modèle de température de serre simulée avec les paramètres supposés. Les résultats de la simulation ont montré la supériorité du RSBA proposé par rapport au BA standard en termes de convergence et de précision des performances. Pour étudier expérimentalement l'algorithme RSBA proposé, le même ensemble de données expérimentales du prototype de serre a été utilisé. Les résultats de prédiction obtenus sont en bon accord avec ceux mesurés qui montrent l'efficacité du RSBA proposé pour identifier le modèle thermique réel de la serre. Enfin, une étude comparative a été menée entre le RSBA et le RIWPSO. Le BA a montré une convergence plus rapide que le PSO au début de l'optimisation mais sa vitesse de convergence a été réduite à la fin. BA et PSO ont tous deux montré de superbes performances pour trouver les solutions optimales avec précision. Cependant, PSO a montré une performance supérieure à BA en termes de consommation de temps concernant le type de problème d'intérêt.

La modélisation du microclimat en serre est une tâche difficile principalement en raison de la forte non-linéarité du phénomène et de l'incertitude des paramètres physiques et non physiques impliqués. L'incertitude vient du fait que la majorité de ces paramètres sont non mesurables ou difficiles à mesurer et certains d'entre eux sont variables dans le temps, signifiant la nécessité de les estimer. Comme première contribution dans le quatrième chapitre de la thèse, une méthodologie pour l'estimation de paramètres en ligne est proposée pour traiter l'estimation des paramètres variant dans le temps d'un modèle simplifié de température de serre à des fins d'adaptation de modèle en temps réel. Un estimateur en ligne est développé sur la base d'une variante améliorée de l'algorithme de chauve-souris appelée algorithme de chauve-souris basé sur une échelle aléatoire. Il permet l'adaptation continue du modèle de température de l'air interne et du sous-modèle de rayonnement solaire interne, en estimant leurs paramètres au même pas de temps en minimisant une fonction de coût, dans le but d'atteindre l'optimalité globale. Des contraintes sur les plages de recherche sont imposées pour respecter le sens physique. L'adaptation des modèles a été testée avec des ensembles de données enregistrés de différentes saisons agricoles et sur une vraie serre en temps réel. Les évolutions des paramètres variant dans le temps ont été présentées graphiquement et discutées en détail. Les résultats expérimentaux illustrent l'adaptation réussie du modèle, présentant une erreur moyenne de moins de $0.28\,°\mathrm{C}$ pour la prévision de la température de l'air et de $20\,\mathrm{W\,m^{-2}}$ pour la simulation du rayonnement solaire. Cela prouve l'utilité de la méthodologie proposée dans des conditions environnementales changeantes. Le flux de ventilation naturelle est une variable importante à mesurer ou à estimer pour considérer son effet significatif sur la modélisation et le contrôle du microclimat des serres. Il est communément connu qu'il peut être estimé mathématiquement en fonction du type et de la dimension de la serre et de ses évents, et surtout du pourcentage d'ouverture des évents. Cependant, la plupart des serres commerciales ne sont pas équipées d'un système d'ouverture d'évent automatique qui oblige le producteur à effectuer un contrôle manuel, en plus de l'absence de capteurs de position d'évent, pour des raisons économiques et de gestion. Ceci conduit à

l'absence de la variable du signal de commande qui représente le pourcentage d'ouverture des évents nécessaire à l'estimation du flux de ventilation. Ce problème a été rencontré dans ce travail après avoir tenté de mettre en œuvre le modèle de microclimat adaptatif développé basé sur l'estimateur de paramètres en ligne via une plate-forme IoF2020 (internet de l'alimentation et de la ferme) dans un ensemble de serres commerciales avec des évents contrôlés manuellement situés dans la province d'Almeria, Espagne. Pour faire face à ce problème, l'estimation du flux de ventilation sans utiliser le pourcentage d'ouverture d'évent a été étudiée. Comme deuxième contribution dans le quatrième chapitre, un capteur virtuel pour l'estimation du flux de ventilation des serres est proposé. Il a été développé à l'aide d'un réseau autorégressif non linéaire à apports exogènes basé sur une analyse en composantes principales utilisant les données mesurées disponibles et les évolutions des flux de chaleur représentant le bilan énergétique de l'effet de serre. Les résultats préliminaires montrent une performance encourageante du capteur virtuel dans l'estimation du flux de ventilation avec une erreur absolue moyenne de $0.41\ m^3\ s^{-1}$.

Mots clé: Agriculture protégée, système de serre agricol, algorithmes évolutifs, estimation en ligne, adaptation de modèles, apprentissage automatique, analyse en composantes principales, réseaux de neurones artificiels, capteurs virtuels.

RESUMEN

La agricultura bajo invernadero ha experimentado importantes desarrollos en los últimos años. Este trabajo se ha centrado en el modelado y control del clima interior del invernadero ya que afecta directamente en el crecimiento del cultivo, por lo que si se controlan las variables climáticas se puede controlar el crecimiento de las plantas. El problema del modelado de estas variables no es sencillo ya que presentan una fuerte componente no lineal con todo lo que esto supone. En esta tesis se han realizado un serie de contribuciones en esta línea utilizando técnicas de inteligencia artificial.

La memoria de esta tesis se ha organizado de forma que en el primer capítulo se presenta el problema, la motivación y el contexto de esta investigación, así como un resumen de las principales contribuciones. A continuación, ya en el segundo capítulo, se describen brevemente los invernaderos experimentales en los que se han realizado los ensayos presentados en esta tesis. Inicialmente, debido a la falta de un invernadero experimental, se construyó un prototipo de invernadero con estructura de madera y cubierta de polietileno y, se utilizó como un semillero a pequeña escala en condiciones de clima árido (clima desértico moderado) ubicado en Meziraa, Biskra, Argelia. Se diseñó un sistema de adquisición de datos basado en microcontroladores de bajo coste con conexión inalámbrica (hardware y software) al que se conectaron sensores también de bajo coste. Se utilizó para recopilar información en tiempo real de las variables climáticas internas y externas del sistema invernadero con fines de modelado, simulación y validación experimental. Ya en condiciones reales, se realizaron ensayos en un invernadero experimental de tamaño comercial cubierto de polietileno, con estructura metálica, en condiciones climáticas mediterráneas. Está ubicado en la Estación Experimental "Las Palmerillas" de la Fundación Cajamar en Almería (España). En este caso, está equipado con todos los sensores, actuadores y sistemas de adquisición de datos profesionales y de investigación necesarios. Se obtuvo un conjunto de datos suficientes y confiables en diferentes campañas agrícolas. Se utilizó para diferentes propósitos, como el modelado y control del microclima interior así como la estimación de parámetros en línea y la validación experimental en tiempo real.

En el tercer capítulo se presentan dos contribuciones importantes de este trabajo. Por una parte, se ha desarrollado y validado un modelo de caja gris para la predicción de la temperatura del invernadero en condiciones climáticas desérticas moderadas. Esta contribución se basa en reformular un modelo de caja blanca para hacerlo independiente de la disponibilidad de valores precisos de los parámetros estáticos de sus elementos. Este modelo es menos complicado al evitar parte del acoplamiento entre sus parámetros, lo que facilita que el algoritmo de identificación obtenga los valores óptimos de los mismos. Se ha utilizado una variante del algoritmo de optimización de enjambre de partículas (PSO) denominado Random Inercia Weight PSO (RIWPSO) para identificar los parámetros del modelo propuesto calibrándolo con los datos experimentales. El prototipo de invernadero construido se ha utilizado para validar el modelo de

temperatura propuesto. Los resultados de la simulación muestran que la optimización deseada se ha logrado con éxito mediante el uso de este tipo de técnicas. El proceso de validación experimental ha confirmado la idoneidad de este modelo para ser implementado tanto para analizar como predecir, con una precisión aceptable, la temperatura del invernadero. Por otra parte, se ha propuesto una mejora al algoritmo metaheurístico bioinspirado en el comportamiento de los murciélagos (BA), que se ha denominado algoritmo de murciélago basado en escala aleatoria (RSBA). La propuesta incluye modificar la explotación del BA estándar haciendo que los parámetros de escala cambien aleatoriamente a lo largo de las iteraciones. Se ha usado en la misma tarea de calibrar el modelo de caja gris de temperatura propuesto. Los resultados de la simulación han demostrado un mejor comportamiento y desempeño de las técnicas RSBA frente al algoritmo BA estándar en términos de convergencia y precisión de rendimiento. Para corroborar estos resultados experimentalmente, el algoritmo RSBA propuesto, se ha utilizado con el mismo conjunto de datos del prototipo de invernadero, obteniéndose resultados de predicción que se encuentran en buena concordancia con los medidos lo que demuestra la bondad de las técnicas RSBA desarrolladas para la identificación del modelo de temperatura interior en el invernadero real. Finalmente, se realizó un estudio comparativo entre la RSBA y la RIWPSO. El BA ha mostrado una convergencia más rápida que el PSO al inicio de la optimización, pero su velocidad de convergencia se redujo al final. Tanto BA como PSO han demostrado un excelente desempeño para encontrar las soluciones óptimas con precisión. Sin embargo, PSO ha mostrado un desempeño superior al BA en términos de coste computacional en cuanto al tipo de problema de interés.

Otra contribución interesante y de mucha aplicación práctica, se muestra en el capítulo cuarto. Como se ha comentado anteriormente, el modelado del microclima de invernadero es una tarea complicada debido principalmente a la fuerte no linealidad del fenómeno y, a la incertidumbre de los parámetros involucrados en el proceso de modelado, ya que no se pueden medir directamente y que son variables en el tiempo, por lo que es imprescindible estimarlos. Para ello, se propuso una metodología de estimación de los parámetros en línea para hacer frente a la obtención de los parámetros variables en el tiempo de un modelo simplificado de temperatura de invernadero con fines de adaptación del propio modelo en tiempo real. Se desarrolló un estimador en línea basado en una variante RSBA presentada anteriormente, que ha permitido la adaptación continua de los modelos de temperatura y radiación interior del invernadero, mediante la estimación de sus parámetros en el mismo paso de tiempo minimizando una función de costo, con la intención de lograr una optimización global. Evidentemente, se impusieron unas restricciones en el espacio de búsqueda para respetar el sentido físico de estos parámetros. La adaptación de los modelos se validó en tiempo real con conjuntos de datos adquiridos en diferentes campañas agrícolas y en el invernadero comercial. La evolución de cada uno de los parámetros variables en el tiempo se presentan gráficamente y se discuten a fondo en este capítulo. Los resultados experimentales muestran la exitosa adaptación del modelo, presentando un error medio menor de 0.28 °C para la predicción de la temperatura del aire y 20 W m^{-2} para la simulación de radiación solar, lo que demuestra la utilidad y bondad de la metodología propuesta en condiciones ambientales reales y cambiantes.

Como última contribución en este campo, el capítulo cuarto también muestra el desarrollo de un sensor virtual del flujo de ventilación natural. Ya que es el principal actuador de refrigeración que se usa en los invernaderos ubicados en las latitudes que se han tratado en esta tesis. Evidentemente se puede estimar matemáticamente según el tipo y la dimensión del invernadero y sus ventilaciones, y lo más importante del porcentaje de apertura de las mismas. Sin embargo, la mayoría de los invernaderos comerciales no están equipados con un sistema automático de apertura de ventilación, lo que obliga al productor a realizar un control manual; además de la falta de sensores de posición de ventilación, por razones económicas y de gestión. Estas situaciones provocan la ausencia de la variable de señal de control que representa el porcentaje de apertura de las ventanas necesario para la estimación del flujo de ventilación. Este problema se ha encontrado en el desarrollo de esta tesis después de intentar implementar el modelo de microclima adaptativo desarrollado basado en el estimador de parámetros en línea a través de una plataforma IoF2020 (Internet for Farm and Foods) en un conjunto de invernaderos comerciales con ventilaciones controladas manualmente ubicados en la provincia de Almería, España. Para hacer frente a este problema, se analizó la estimación del flujo de ventilación sin utilizar el porcentaje de apertura de ventilación, y se propuso un sensor virtual para la estimación esta variable. Se ha desarrollado utilizando una red autorregresiva no lineal con entradas exógenas basadas en el análisis de componentes principales (PCA) utilizando los datos medidos disponibles y las evoluciones de los flujos de calor que representan el balance energético del invernadero. Los resultados preliminares muestran un rendimiento alentador del sensor virtual en la estimación del flujo de ventilación con un error absoluto medio de 0.41 $m^3\ s^{-1}$.

Palabras clave: Agricultura protegida, sistema de invernadero, algoritmos evolutivos, estimación en línea, adaptación de modelos, aprendizaje automático, análisis de componentes principales, redes neuronales artificiales, sensores virtuales.

List of Publications

Publications in journals

- **Guesbaya, M.**, García-Mañas, F., Megherbi, H. and Rodríguez, F., 2021. Real-time adaptation of a greenhouse microclimate model using an online parameter estimator based on a bat algorithm variant. Computers and electronics in agriculture, 192, p.106627. [Journal Impact Factor: 6.757, Quartile: Q1, cited by 13949 in 2021 in Scopus, CiteScore of 11.8 in 2018-2021 in Elsevier]. DOI: https://doi.org/10.1016/j.compag.2021.106627

Publications in international conferences

- **Guesbaya, M.**, García-Mañas, F., Rodríguez, F., Megherbi, H. and Ouamane, M., 2021. Virtual sensor for ventilation flux estimation in greenhouses. The XI Iberian Congress of Argo-engineering, 2021, Valladolid, Spain. Link: https://www.researchgate.net/publication/356216686_Virtual_sensor_for_ventilation_flux_estimation_in_greenhouses
- **Guesbaya, M.**, Megherbi, H. and Megherbi, A.C., 2019, November. Random Scaling-Based Bat Algorithm for Greenhouse Thermal Model Identification and Experimental Validation. In *the International Conference on Electrical Engineering and Control Applications (ICEECA)* (pp. 47-62). Springer, Singapore. DOI: https://doi.org/10.1007/978-981-15-6403-1_4
- **Guesbaya, M.** and Megherbi, H., 2019, November. Thermal Modeling and Prediction of Soilless Greenhouse in Arid Region Based on Particle Swarm Optimization. Experimentally Validated. In *the International Conference on Advanced Electrical Engineering (ICAEE)* (pp. 1-6). IEEE. DOI: https://doi.org/10.1109/ICAEE47123.2019.9015190

Table of Contents

Liste of Acronyms

MIMO	Multi-input multi-output;
MISO	Multi-input single-output;
PSO	Particle swarm optimisation algorithm;
RIWPSO	Random inertia weight PSO;
DE	Differential evolution algorithm;
ARX	Auto-regressive model with exogenous variables;
ARMAX	Autoregressive–moving-average model;
EKF	Extended Kalman flter;
BA	Bat algorithm;
RSBA	Random scaling-based bat algorithm;
GA	Genetic algorithm;
HS	Harmony search;
SA	Simulated annealing;
LAI	Leaf area index;
SCADA	Supervisory and control data acquisition;
MSE	Mean Square Error;
RMSE	Root Mean Square Error;
MAE	Mean Absolute Error;
MaxAE	Max Absolute Error;
R^2	Coeffcient of Determination;
RE	Residual Error;
ANN	Artifical neural networks;
UDP	User datagram protocol;
ARM	Automatic Control, Robotics, and Mechatronic Research Group
UAL	University of Almería;
CHROMAE	Control and optimal management of heterogeneous resources in agroindustrial production districts integrating renewable energies;
IEEE	Institute of Electrical and Electronics Engineers;
AR	Auto-Regressive;
ARMAX	Autoregressive–moving-average model with exogenous inputs;
AgEng	Agricultural Engineering;
CeiA3	The Agrifood Campus of International Excellence - Campus de Excelencia Internacional Agroalimentario de Andalucía;

Liste of Figures

Pages

Liste of Tables

Pages

CHAPTER I INTRODUCTION

I.1 Research motivation

Nowadays, the world is witnessing serious issues with the change of climate, the increase in population and the lack of resources. All the countries are investigating sustainable solutions in all domains, especially in agriculture. Providing food for the growing population is one of the main issues for sustainable development in the United Nations' 2030 agenda (Desa., 2016). Greenhouse systems have become a prominent means in the agricultural field to fulfil the targeted sustainability. The agricultural greenhouse is an enclosure generally based on a metal structure covered by a transparent plastic or glass cover that allows solar radiation to pass through. Inside the enclosed structure, an isolated environment is affected by a set of physical phenomenons representing the heat and mass balance inside the greenhouse. It is created to suit the cultivated plant favourably and it is commonly called the greenhouse microclimate. Modern agriculture is outstandingly affected by the greenhouse system because it plays an essential role in enhancing the management, qualities, and quantities of agricultural production. The optimisation of the greenhouse production system is required to suit the increasing population and the strict standards of the local and international markets. To tackle these agricultural necessities and achieve sustainability, the continuous development of greenhouse systems is important. The fundamental level in the management hierarchical structure of the greenhouse production system is the analysis, modelling, estimation, prediction and control of greenhouse microclimate variables that strongly affect crop growth and yield. Continuous adaptivity of the greenhouse system to the change of climate conditions is an essential aspect to be achieved for the best continuous optimization of the studied outputs. Efficiently fulfilling the continuous optimization based on adaptive models and controllers for these challenging processes will tremendously help in leading the outcome to hit the largest quantities, finest qualities with the lowest costs, as the ultimate target of the field.

I.2 Aim and objectives

The greenhouse microclimate consists of a set of physical phenomenons representing the heat and mass balance inside the greenhouse. The elements of this balance are a set of commonly known fluxes generated by a set of physical processes: Solar and thermal radiation, convection, conduction, condensation, evaporation, transpiration, ventilation and infiltration (see Fig. 1). These processes can be considered as the driving force of climate change inside the greenhouse (microclimate change). Their change in turn urges the change of the essential microclimate variables measured by sensors: Air temperature, relative humidity, global radiation, soil temperature, soil moisture, CO_2 concentration, crop transpiration …etc.

This work aims to develop an online parameter estimation technique as an adaptation tool for real-time implementation as it appears in Fig. 2. The vision is to take advantage of the advancing metaheuristic bio-inspired population-based algorithms by implementing them in an online parameter estimation scheme. Continuous real-time adaptation is an essential aspect of microclimate models and controllers for prediction and control optimization purposes. It permits handling the multiple objectives of the greenhouse system. Since the management of the greenhouse systems is two parts which are modelling and control, this work mainly aims to efficiently model

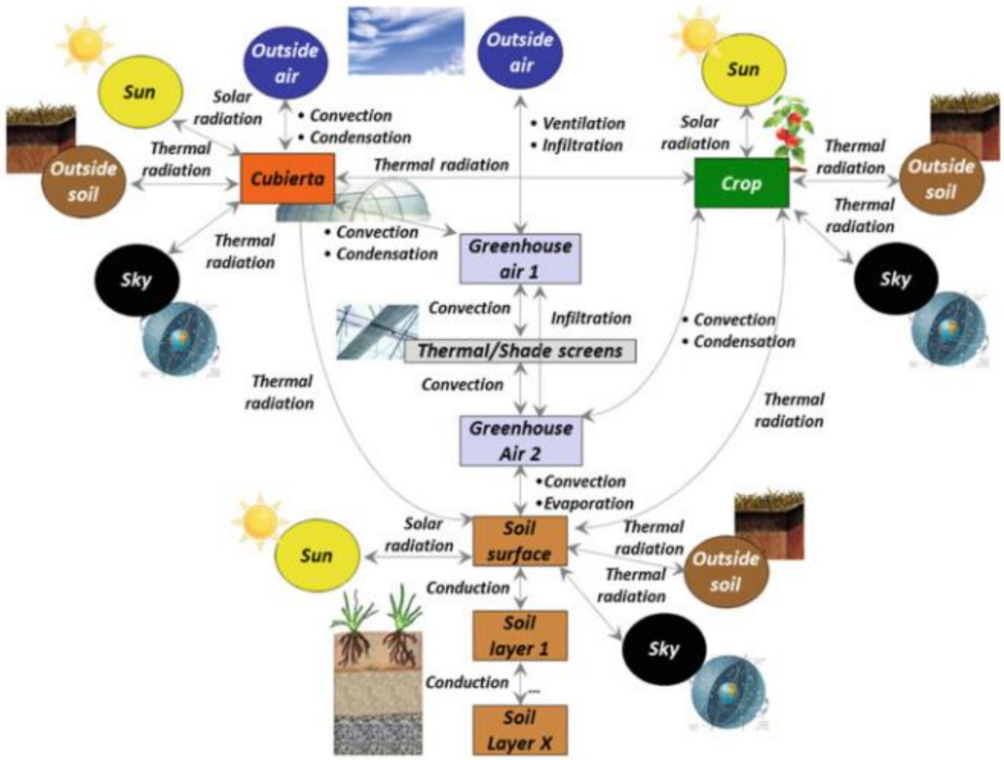

Figure 1. Relationship between greenhouse elements (Rodríguez et al., A, 2015).

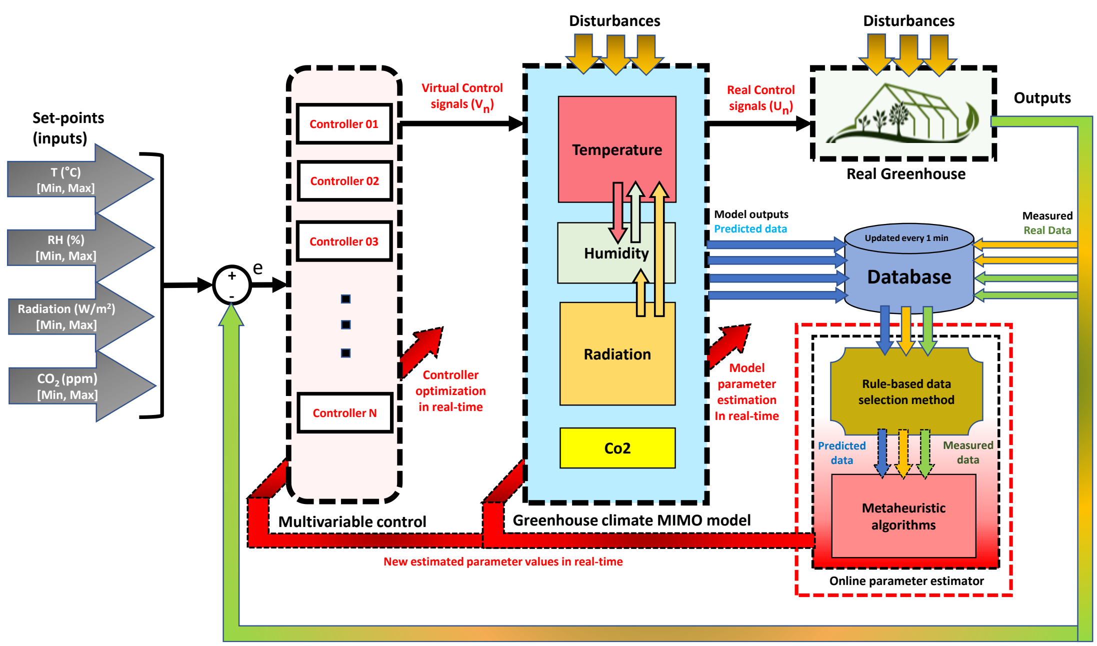

Figure 2. General greenhouse climate prediction and control scheme using online estimation

the essential climate variables inside the greenhouse and then apply some control techniques to the obtained greenhouse model. The microclimate modelling stage includes implicitly the estimation of a set of heat and mass fluxes for the prediction of the targeted climate variables in this work: inside air temperature, inside solar radiation and CO_2 concentration. Advanced and intelligent methods such as metaheuristic bio-inspired optimization algorithms and artificial neural networks have been used for modelling and control purposes and principal component analysis for data analysis purposes. Based on the fact that the greenhouse systems are complicated and nonlinear, system adaptivity is an important aspect that should be addressed to provide a robust tool against the change of climate conditions inside and outside the greenhouse for the long-term successful implementation of greenhouse models and controllers. The aim is to develop an online parameter estimation method as an adaptation tool based on a metaheuristic optimization algorithm to be dedicated for modelling and control purposes.

Finally, the objectives of this thesis can be summarized as follows:

1. Air temperature modelling and prediction using the acquired data using the prototype
2. Application of metaheuristic optimization algorithms for offline parameter estimation in the context of model calibration.
3. Development of an online parameter estimator based on a metaheuristic bio-inspired optimization algorithm to be used for continuous model adaptation.
4. Implementation of the developed online parameter estimator using experimental datasets of different agri-seasons acquired from a commercial greenhouse.
5. Real-time implementation of the developed adaptive climate model on a commercial greenhouse for the prediction of air temperature and the simulation of solar radiation
6. Development of a virtual sensor for ventilation flux estimation for greenhouses with manual control of natural ventilation.

I.3 Research context

Greenhouse microclimate modelling and control are the fundamental aspects of the greenhouse system. It has included extensive development and many outstanding results. In general, three kinds of greenhouse models are recognized (Rodríguez et al., 2015): the physical-based model (white-box model), the input-output based model (black-box model), and the grey-box model which is a model where all the available knowledge of the process mechanisms is used to build a white-box part, while the missing information is approximated by black-boxes fitted to the process data. The advantage of the grey-box model over others is its advantage in addressing some of the most restrictive factors of the white-box and black-box approaches, as it also seeks to combine the advantages of both model types (Sjöberg et al., 1995). Greenhouse models are proposed for two purposes: on one hand, the analysis of a certain physical phenomenon (Temperature, radiation, humidity, transpiration, Co_2, …etc) which is the most complicated type of model. It has to be a white-box model including all the needed sub-equations describing the majority of the physical effects on the simulated target to avoid neglecting any of them. It can also be a grey-box model but still has to be dependent more on the physical-based equations and it obliges having constraints on the empirical parts of the model to ensure respecting the physical sense of the calculated information leading to the simulated target. On the other hand, there are the simplified and pseudo-physical models which are also considered grey-box models but they are commonly more useful for control purposes. Some advanced and computational intelligence techniques were proposed in the literature (Sjöberg et al., 1995; Kennedy and Eberhart., 1995; Ali et al., 2015; Ali et al., 2018) and one of them is proposed in this thesis in the context of making more use of those simplified models for both purposes of greenhouse microclimate modelling, analysis and prediction (Guesbaya and Megherbi., 2019).

The modelling of the microclimate variables inside the greenhouse has been studied in the literature using different mathematical and data-driven models, simplified and complicated ones (Fourati., 2014; Rodríguez et al., 2015; Ben Ali et al., 2018; Choab et al., 2019; Hoyo et al., 2019; Atia and El-madany., 2017; Li et al., 2020; Laktionov et al., 2020). A simplified nonlinear grey-box model is proposed in this thesis (Lamrani et al. 2001). It predicts the inside air temperature based on a set of heat exchange processes generated by the differences in energy content between the inside and outside air. This model was derived by reformulating a physical-based dynamic model. The reformulation includes having new static parameters linearly dependent that have to be identified with offline parameter estimation. This fact has led to obtaining a less complicated model that requires less computational costs and, hence, a better chance to achieve successful air temperature prediction. The model was experimentally validated in a constructed gable-shaped greenhouse prototype that has been built to be used as a nursery, equipped with a designed microcontroller-based low-cost data acquisition system. The lack of information on the system parameters is considered a kind of uncertainty in greenhouse models. Thus, here we highlight the need for an identification technique to solve this problem. Recently, considerable attention has been paid to nature-inspired metaheuristic algorithms to solve optimization issues in many domains thanks to their characteristics and efficient search methods. *The particle swarm optimization* (PSO) algorithm is among those very common population-based stochastic algorithms (Yang, X. S, 2014). A variant of PSO called the *Random Inertia Weight particle swarm optimization algorithm* (RIWPSO) has been implemented in this study as a parameter identification technique for the static parameters of the proposed grey-box model under arid desert climate conditions (Imran et al., 2013). In addition, another metaheuristic bio-inspired algorithm that has been used for the same purpose is the *Bat algorithm* (BA) (Yang, X. S, 2014). In another study, a variant of the BA called the *random scaling-based bat algorithm* (RSBA) has been proposed and validated using the same prototype and datasets (Guesbaya et al 2019). The idea behind the RSBA proposition is to adjust the scaling parameter to make the step size have relatively large and small values to maintain an effective step size control relative to the closeness of the optimal solution. A comparative study has been carried out at the end on the performance of PSO and RSBA in identifying the parameters of the grey-box model in the context of offline parameter estimation and greenhouse air temperature prediction.

To reach an accurate performance despite the high nonlinearity of the phenomena and the presence of uncertainties, multi-input multi-output (MIMO) and multi-input single-output (MISO) greenhouse models are usually calibrated using various offline optimisation or estimation approaches based on either numerical or artificial intelligence algorithms (Hasni et al., 2011; Yu et al., 2016; Sanchez-Molina et al., 2017) as in our proposed works (Guesbaya and Megherbi., 2019; Guesbaya et al., 2019). In all these cited works, datasets of only a few days were used in the calibration phase and the validation was carried out in a short period of a specific agri-season. Thus, for the long-term successful implementation of models during different agri-seasons, larger datasets have to be provided but even though, the results carried out could probably still be unreliable because of the indispensable presence of time-varying parameters that depend on weather conditions and the state of the crop (Cunha et al., 1997; Vanthoor et al., 2011; Pérez-González et al., 2018). The microclimate model has to consider the effect of those time-varying parameters that are usually unmeasurable, or their measuring instrumentation or procedures are unaffordable (Choabet al., 2019; Guesbaya et al., 2019; Ma et al., 2019). The periodical offline calibration of the time-varying parameters is a candidate solution but is commonly considered a laborious procedure that consumes time and computational resources and requires at least a one-year dataset or large datasets from every season (Rodríguez et al., 2015; Speetjens et al., 2009). Therefore, it is convenient to develop an online parameter estimation technique to avoid the laborious periodical offline model calibration and continuously adjust the time-varying parameters depending on the evolution of climate conditions in real time. In this thesis, the presented problem has been addressed in the context of achieving the continuous adaptivity of a greenhouse microclimate model in different agri-seasons aiming to analyse the effect of the time-varying parameters and the optimal performance of the adaptive microclimate model.

To sustain the best agricultural outcome, greenhouse ventilation flux is a very important phenomenon that should be studied individually and controlled according to its vital role in influencing the crop through most of the microclimate variables. To measure this variable, it is necessary to use special anemometers such as ultrasound or thermal effect-based ones. Since the installation of such sensors is unusual due to their high costs, ventilation flux can be estimated using some proposed methods in the literature (Boulard et al., 1995; Kittas et al., 1997). However, the common methods of estimating or predicting the greenhouse ventilation flux are all dependent on the total surface of the openings of vents. This could be considered as a problem of high complexity in greenhouses where the exact vents opening is unknown because it is performed manually by growers. Manual control means the absence of the control signal which in turn means the lack of the exact and continuous recording of vent opening percentage making it impossible to estimate the ventilation flux based on the common explicit methods in hands which makes it a serious issue. As far as we know, this problem was not investigated in the literature and for this reason, the present work is focused on studying new possibilities to estimate the ventilation flux in greenhouses using alternative techniques like artificial neural networks. This problem has been encountered on the way to proving the robustness of the proposed RSBA-based online parameter estimator with different commercial Mediterranean greenhouses connected by IoT at the IoF2020 platform (Guesbaya et al, 2021; Guesbaya et al., 2021) and located in the surroundings of Almeria, Spain. These commercial greenhouses lack the automatic control of vents as same as the majority of commercial greenhouses worldwide. In this context, an attempt to develop a virtual sensor for greenhouse ventilation flux estimation without the dependence on the vent position information (Guesbaya et al., 2021) based on principle components analysis (PCA) and artificial neural networks (ANN) is performed and some preliminary results have been presented in this thesis.

I.4 Research scope

This PhD thesis has been accomplished within the framework of a set of projects and a long-period scholarship which are presented as follows:

- As a collaborator in CHROMAE Project (DPI2017-85007-R) entitled: Control and Optimal Management of Heterogeneous Resources in Agro-industrial Production Districts Integrating Renewable Energy. Funded by the Spanish Ministry of Science, Innovation and Universities. Through direct collaboration with the research group of Automatic Control, Robotics and Mechatronics (ARM) research group (TEP-197) at the University of Almeria. URL: http://www2.ual.es/chromae/researchers/mounir-guesbaya/
- As a beneficiary of the Algerian scholarship of "*Exceptional National Program 2019/2020*" (PNE internship) funded by the Algerian Ministry of High Education and Scientific Research.
- As a member in "*Projet de Recherche-Formation Universitaire*" (PRFU project) entitled: Application of intelligent techniques for modelling and control of mobile robots, renewable energy systems and agricultural greenhouses. Under the reference: A01L08UN070120180001. This project is funded by the Algerian Ministry of High Education and Scientific Research.

I.5 Main contributions

Published works and achievements during the accomplishment of this PhD thesis are presented as follows:

In the second and third chapters, three contributions are presented as follows:

- Firstly, a gable-shaped small-scale greenhouse prototype was constructed to be used as a nursery in an arid region Mziraa, Biskra, Algeria. It included designing a low-cost microcontroller-based data acquisition system for the wireless monitoring of the prototype. They have been used to acquire a modest dataset including climate variables from inside and outside environments of the greenhouse prototype under moderate desert climate conditions. This project was chosen as the best project in the exhibition of the International Symposium of Technology and Sustainable Industry Development 2019 (ISTSID), El-Oued, Biskra, Algeria. The successfully acquired dataset was used in the investigations of two international conferences.

  Related certificates and publications:

  - Guesbaya, M. and Megherbi, H., 2019. "Un prototype de serre agricole à faible coût surveillé et contrôlé sans fil". The International Symposium of Technology and Sustainable Industry Development 2019 (ISTSID), El-Oued, Biskra, Algeria.
  - (Guesbaya et al., 2019) Guesbaya, M. and Megherbi, H., 2019. Thermal modeling and prediction of soilless greenhouse in arid region based on particle swarm optimization:

Experimentally validated. *International Conference on Advanced Electrical Engineering* (ICAEE), Algiers, IEEE.

- (Guesbaya et al., 2019) Guesbaya, M., Megherbi, H. and Megherbi, A.C., 2019. Random scaling-based bat algorithm for greenhouse thermal model identification and experimental validation. *The 4th International Conference on Electrical Engineering and Control Applications* (ICEECA), Constantine, Springer, pp. 47-62

- Secondly, a nonlinear grey-box model of greenhouse air temperature is proposed. It describes the inside air temperature as a set of heat exchange processes generated by the differences in energy content between the inside and outside air. This model was derived by reformulating a physical-based model. The reformulation includes having new static parameters linearly dependent that have to be identified based on offline parameter estimation using the acquired dataset from the greenhouse prototype. A less complicated model has been derived to be used for greenhouse air temperature prediction. The lack of information on the system parameters is considered a kind of uncertainty in greenhouse models. For this issue, the commonly known metaheuristic bio-inspired algorithm called the *Random inertia weight particle swarm optimization algorithm* (RIWPSO) was chosen to be implemented to the proposed model for model calibration as an offline parameter estimation method.

  Related publications:

  - (Guesbaya et al., 2019) Guesbaya, M. and Megherbi, H., 2019. Thermal modeling and prediction of soilless greenhouse in arid region based on particle swarm optimization: Experimentally validated. *International Conference on Advanced Electrical Engineering* (ICAEE), Algiers, IEEE.

- Thirdly, another well-known metaheuristic bio-inspired and population-based optimation algorithm was also implemented in the proposed grey-box temperature model, called the *Bat Algorithm* (BA). An enhanced variant of this algorithm is proposed in this thesis and called the *Random scaling-based bat algorithm* (RSBA). It was used for the same purpose of identifying the unknown values of the proposed static parameters of the model through an offline estimation process. The RSBA was proven to have superior performance over the standard BA in terms of accuracy and speed of convergence. Finally, a comparative study has been carried out between the performance of RIWPSO and RSBA in identifying the parameters of the grey-box model and the prediction accuracy of the different obtained greenhouse air temperature models. The results have shown the superiority of the RIWPSO over the RSBA in solving the problem at hand. However, the RSBA still could be more useful against other different problems such as the online parameter estimation in real-time where the advantage of the early convergence to optimality can be necessary due to time constraints.

  Related publications:

  - (Guesbaya et al., 2019) Guesbaya, M., Megherbi, H. and Megherbi, A.C., 2019. Random scaling-based bat algorithm for greenhouse thermal model identification and experimental validation. *The 4th International Conference on Electrical Engineering and Control Applications* (ICEECA), Constantine, Springer, pp. 47-62.

In the fourth chapter, two contributions are presented as follows:

- In the first part of the fourth chapter of this thesis, a methodology for online parameter estimation is proposed for greenhouse microclimate model adaptation purposes as one of those possible greenhouse system optimizations. It is proposed as an alternative to the laborious periodical offline calibration of the time-varying parameters which is commonly considered a laborious procedure that consumes time and computational resources. Specifically, an online parameter estimator is developed to achieve the real-time adaptation of a greenhouse microclimate model and intends to thoroughly study the time-varying parameters, aiming for optimal prediction performance. The online estimator works based on the RSBA as an enhanced variant of the nature-inspired BA algorithm. The performance of the developed online parameter estimator in adapting the greenhouse microclimate model has been evaluated from both physical and statistical points of view. The evolution of the estimated time-varying parameters has proven that transpiration parameters could be considered constants for simplicity. However, the parameters of convection and conduction processes should be time-varying. The resulting prediction error was very low less than $0.28\,°C$ for air temperature prediction and $20\,Wm^{-2}$ for solar radiation simulation. Research works that include a graphical illustration and a detailed discussion of the evolution of the time-varying parameters have not been encountered in literature to be compared to the results presented in this paper. The more accurate the model performance, the more accurate the yield control and the better the economic profits quantitatively and qualitatively.

  Related publications:

  - Guesbaya, M., García-Mañas, F., Megherbi, H. and Rodríguez, F., 2022. Real-time adaptation of a greenhouse microclimate model using an online parameter estimator based on a bat algorithm variant. Computers and electronics in agriculture, 192, p.106627.

- In the second part of the fourth chapter of this thesis, a virtual sensor for greenhouse ventilation flux has been developed based on PCA-NARX modelling following the methodology explained in detail in the folds of this chapter. A dataset has been generated from a Mediterranean multi-span greenhouse located at "Las Palmerillas" Experimental Station which is a property of the Cajamar Foundation (36.79316 latitude, -2.72014 longitude). The dataset includes a combination of measured microclimate variables and the evolutions of greenhouse heat fluxes. The heat fluxes were estimated using an adaptive air temperature model due to its capability of providing their optimal estimations with an error of <5% between a set of the same tests and its reliability in estimating the greenhouse ventilation flux without installing expensive sensors. All the obtained variables were firstly processed by: signal filtring, centralization, reduction and standardisation. Secondly, the treated dataset was used to generate the PCs for data reduction using PCA. These PCs are then considered the new inputs of the neural network. Thus, the network was trained based on the PCs to fit the target which is the estimated heat loss using the opening percentage of the roof and side vents based on the previously mentioned explicit approach. Finally, the estimated heat loss flux was used to inversely calculate the ventilation flux representing the ultimate objective of the proposed virtual sensing method. The validation of this developed virtual sensor has shown very promising preliminary results.

Related publications:

- (Guesbaya et al., 2021) Guesbaya, M., García-Mañas, F., Rodríguez, F., Megherbi, H., Ouamane, M, R. 2021. Virtual sensor for ventilation flux estimation in greenhouses. *The XI Iberian Congress of Agroengineering*, Valladolid, Spain.

I.6 Outline of the thesis

This PhD thesis is organized into five chapters. The description of each of the following chapters is briefly presented. An abstract is included at the beginning of each chapter to provide more information about its content. In Chapter 1, factors of motivation, research objectives, research context and main contributions of this dissertation are all presented. A list of the published main contributions is also included. Chapter 2 overviews the greenhouse facilities used to acquire the experimental datasets and perform the experiments. It illustrates the main characteristics of the used greenhouse systems, as well as the collected experimental datasets. Chapter 3, describes the proposed simplified grey-box thermal model, the used pseudo-physical model of greenhouse air temperature and solar radiation. It also presents two metaheuristic bio-inspired algorithms and their application for the offline estimation of model parameters for model calibration purposes. They are the *Random inertia weight particle swarm optimization algorithm* (RIWPSO) and a variant of the *Bat algorithm* (BA) proposed in this thesis and called the *Random scaling-based bat algorithm* (RSBA). This was finalized by a comparative study of the performances of the algorithms in estimating the model parameters and the models in predicting the greenhouse air temperature after a set of repetitive experimental validation processes. In chapter 4, an online parameter estimator is developed as a model adaptation technique based on the proposed RSBA. The online estimator was validated in simulation with different datasets from different agri-seasons, and then, it was implemented in real time with the commercial greenhouse system. Furthermore, a virtual sensor for greenhouse ventilation flux estimation was developed as a supplement for the robustness of the RSBA-based online parameter for commercial greenhouses with manual control of vents. Finally, the main results, discussions, major conclusions and future perspectives are all drawn in Chapter 5.

CHAPTER II DESCRIPTION OF EXPERIMENTAL FACILITIES

II.1 Introduction

This chapter briefly describes the experimental greenhouses used in this thesis. Initially, due to the lack of an experimental greenhouse, a wooden-structured polyethene-covered greenhouse prototype was constructed and used as a small-scale nursery under arid climate conditions (moderate desert climate) located in Meziraa, Biskra, Algeria. A low-cost data acquisition system was designed (hardware and software) and installed in the greenhouse with several low-cost sensors to gather instant information on the essential inside and outside climate variables. A dataset of five days was successfully acquired for modelling, simulation and experimental validation purposes. Secondly, a metal-structured polyethene-covered commercial-sized experimental greenhouse under Mediterranean climate conditions was exploited. It is located at “Las Palmerillas” Experimental Station, a property of the Cajamar Foundation (36.79316 latitude, -2.72014 longitude) in Almería, Spain. It is equipped with all the necessary professional sensors, actuators and data acquisition systems. A set of sufficient and reliable datasets of 15 days were obtained in different agri-seasons and used for different purposes such as microclimate modelling, online parameter estimation, real-time experimental validation and soft sensor development, as described in detail in the following chapters.

II.2 Greenhouse prototype

II.2.1 Description of the greenhouse prototype

The experimental set-up has included the construction of a small-size gable-shaped (single-span), wooden-structured greenhouse prototype as a nursery (Guesbaya and Megherbi., 2019). It is covered by polyethene with 0.2mm of thickness as it appears in Fig. 3. A nursery is a kind of greenhouse specialized for seedling nursing and management, designed to produce favourable conditions for seedling until it becomes a ready healthy transplantable plant. It was implemented in the municipality of M’ziraa, affiliated to Biskra province in Algeria (34°43′19.7″ N 6°17′39.2″ E), which is characterized by its moderate desert climate in the winter season. The experimental nursery has a wooden soilless floor that embraces 3 seedling bunches (45x20 cm) full of treated soil.

Figure 3. External and internal view of greenhouse prototype

II.2.2 Low-cost data acquisition system

The data acquisition system was designed in this work (see Fig. 4) based on the programmable microcontroller-based board called Arduino Mega 2560. It gathers measurements from all the installed sensors. The measurements are then sent to a personal computer via Wi-Fi with a communication protocol called *user datagram protocol* (UDP) using another programable microcontroller-based board called NodeMCU V0.1. Finally, the data is received and treated by MATLAB to generate the initial dataset used in this study.

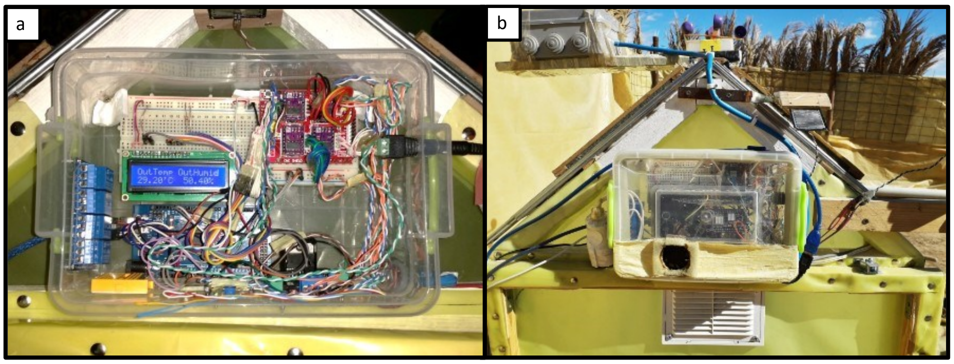

Figure 4. The developed low-cost data acquisition system. (a) Internal view; (b) External view

The required sensors have been installed and connected to the I/O pins of Arduino Mega. These sensors are dedicated to obtain sufficient information about the instant state of the greenhouse internal and external climate. Two DHT22 are installed indoor the greenhouse at a high of 0.3m and outdoor the greenhouse at a high of 1.25m to measure air temperature, where the accuracy of humidity is ±5%RH and temperature accuracy is <±0.5°C. Both DHT22 sensors are equipped with a low-cost design of a plate-shaped radiation shield (Holden et al., 2013). A low-cost pyranometer based on BPW34 silicon photodiode with an accuracy of 2.3% error (Čekon et al., 2016) is installed outdoor the greenhouse at a height of 1.4 m to measure the solar irradiation. Another low-cost element is the anemometer based on a DC motor installed outside at a height of 1.55 m (Horsey., 2016). MH-RD Rain module is also used to be warned from heavy rain. All the sensors are presented in Fig. 5. The sensitiveness of each of the sensors are:

- The temperature measurement sensitiveness is about 0.1 °C
- The humidity measurement sensitiveness is about 0.1 %
- The pyranometer measurement sensitiveness is about 9.6 w/m^2
- The anemometer measurement sensitiveness is about 0.47 m/s

II.3 Commercial Mediterranean greenhouse

II.3.1 Structure and actuators

The greenhouse utilised in this work is presented in Fig. 6. It is a traditional Mediterranean greenhouse, commonly named "Almería-type" greenhouse. It is located at "Las Palmerillas" Experimental Station which is a property of the Cajamar Foundation (36.79316 latitude, -2.72014 longitude), in Almería, Spain, at an altitude of 151 m. The total surface of the greenhouse is 877 m^2

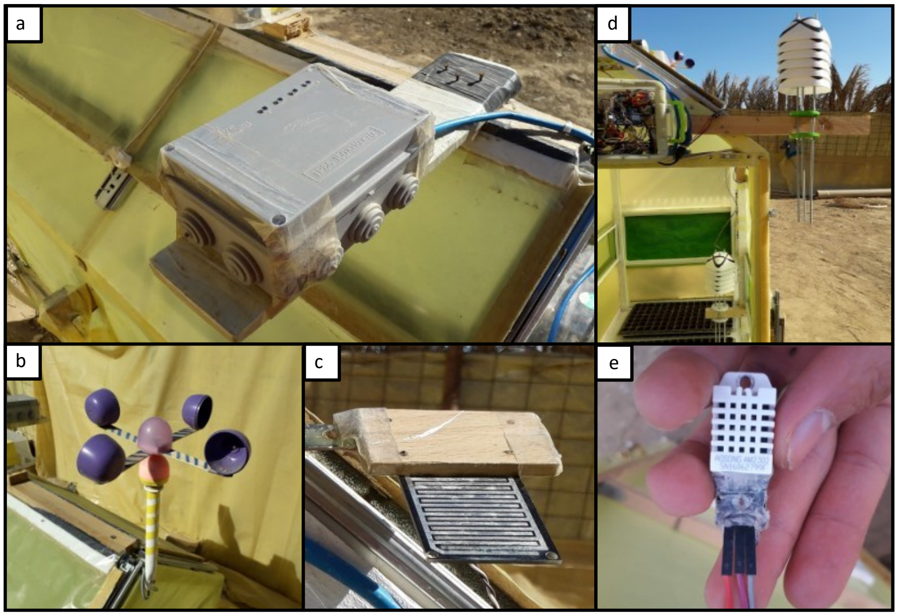

Figure 5. Sensors installed inside and outside the greenhouse prototype. (a) A designed low-cost BPW34 photodiode-based pyranomter; (b) A designed DC motor-based wind-velocity sensor; (c) Rain sensor for Alarm system; (d) A hand-made plat-shaped radiation shield for DHT22 inside and outside the greenhouse; (e) DHT22 temperature and humidity sensor.

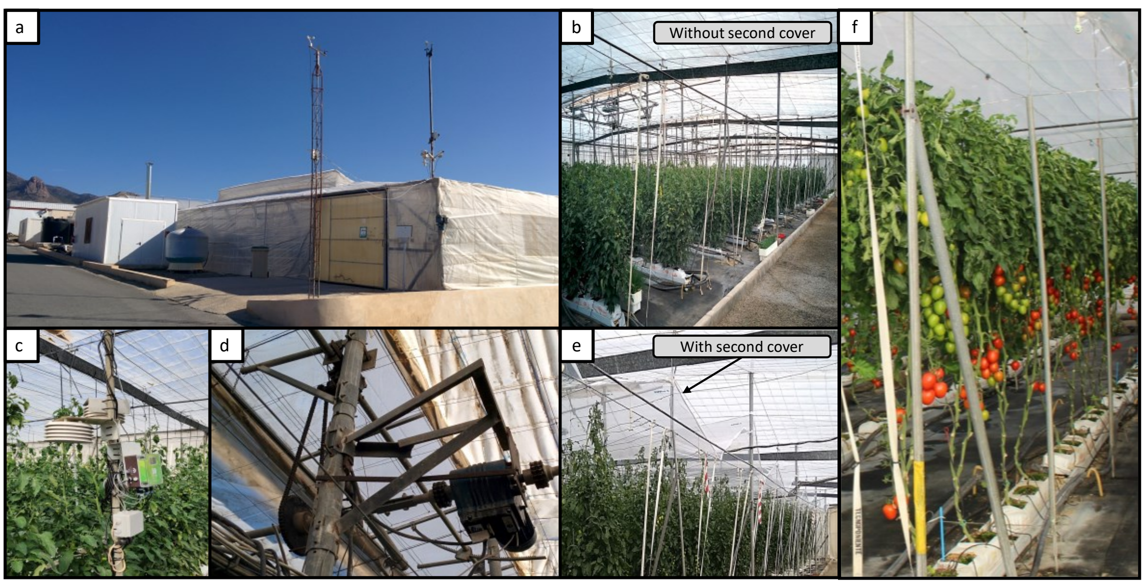

Figure 6. Greenhouse facilities used for the experimental tests with grown tomato crop. (a) greenhouse exterior view with roof and sidewall vents; (b) interior view without second cover; (c) example of a commercial data acquisition device, and inside temperature and humidity sensors; (d) interior view of roof vent; (e) interior view with the second cover installed; (f) Tomato crop.

(37.80 m × 23.20 m) and it is protected by a polyethene cover. Under the cover, an approximate area of 600 m^2 is reserved for tomato crops. The plants are cultivated in coconut coir bags aligned in rows orientated from north to south with a slope of 1%.

The greenhouse is equipped with several actuators to control the microclimate under the cover, providing adequate conditions required by the plants for optimal crop growth. Thus, the greenhouse facilities are complemented with a humidification and dehumidification system, a carbon dioxide enrichment system, a pipe heating system based on a biomass boiler, and a natural ventilation system, among others. For the natural ventilation system, five zenithal windows (8.36 m × 0.73 m) are installed on the roof of the structure and two lateral windows (32.75 m × 1.90 m) are situated along the north and south sidewalls of the cover.

II.3.2 Data acquisition system

A wide set of sensors is deployed inside and outside the greenhouse to measure the climatic variables affecting the crop every 30 seconds. An external weather station measures air temperature and humidity, solar radiation, air CO_2 concentration, and wind velocity. Inside the greenhouse, a protected probe is employed to measure the inside air temperature and relative humidity (see Fig. 5c). Several sensors are installed in different rows of the crop to measure the air CO_2 concentration, the solar radiation under the cover and the temperature of the soil surface.

All the distributed sensors shown in Table 1 are connected to a series of data acquisition devices (Compact FieldPoints, National Instruments, Austin, TX, USA), which transmit the measurements through an Industrial Ethernet network to a supervisory and control data acquisition system (SCADA) based on LabVIEW (National Instruments) (see Fig. 7).

Table 1. Greenhouse sensors specifications

Sensors	Brand	Model	Precision	Physical operational range
Outside and inside temperature and humidity	Campbell Scientific	HC2S3	± 0.1 ºC	-40 to 60 ºC
Outside and inside global solar radiation	Hukseflux	LP02	< ± 1 %	0 to 2000 Wm^{-2}
Soil surface temperature	Campbell Scientific	108	±0.7°C	-5 to 95 °C
Outside and inside relative humidity	Campbell Scientific	HC2S3	± 0.1 %	0 to 100 %
Wind velocity	Vector Instruments	A100L2/PC3	< 2 %	0 to 75 m/s
Inside and outside CO_2 concentration	E+E Elektronik	EE820-C2	< ± 50 ppm	0 to 2000 ppm

The computational unit used for the real-time application is a computer located in the experimental station near the greenhouse. The computer specifications are Intel Core i7-7700, quad-core and 8 threads with 3.60 GHz (up to 4.20 GHz), 16 GB RAM DDR4 2133 MHz, and equipped with Windows™ 10 64-bit, MATLAB R2017b and LabVIEW™ 2015.

II.3.3 Maintenance and cultural tasks

During a crop season, different maintenance and cultural tasks are usually practised to the plants to ensure a healthy evolution toward the desired growth yield. For this work, two types of maintenance tasks were registered since they affect the state of the crop and/or they have an impact on the greenhouse microclimate. An example of the type of maintenance task is the periodical pruning of the plants' leaves to reduce the crop leaf area index (LAI). Another type of cultural task is related for example to the whitening of the cover and the necessity of regulating the solar radiation transmission through the cover of the greenhouse. The whitening of the cover is usually performed

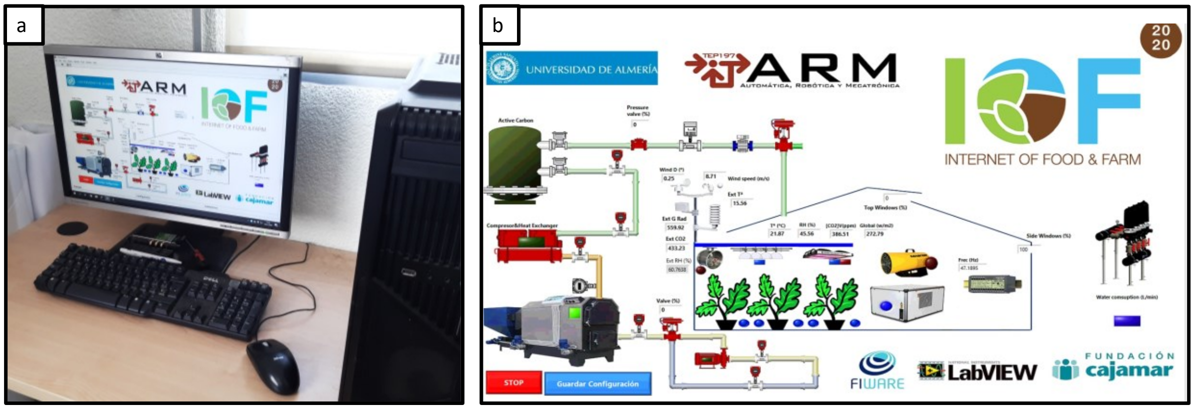

Figure 5. (a) A central unit for greenhouse monitoring and control; (b) Supervisory and control data acquisition system (SCADA) based on LabVIEW.

in the months with the highest values of solar radiation (spring and summer) to reduce the net radiation reaching the crop. For the autumn and winter period, the whitening is removed. Also, during the coldest periods, a floating plastic cover can be installed inside close to the plants to increase the crop isolation from external weather. All these tasks are considered in this work to explain the recorded evolution of the greenhouse microclimate.

II.4 Conclusions

Due to the lack of an experimental dataset for greenhouse climate modelling, a greenhouse prototype was constructed and used under arid climate conditions in Meziraa, Biskra, Algeria. A low-cost data acquisition system was designed and installed in the greenhouse prototype with several low-cost sensors to gather instant information on the essential inside and outside climate variables. A dataset of five days was successfully acquired and used to perform predictive modelling.

After starting my research collaboration with the University of Almería, a commercial-sized experimental greenhouse under Mediterranean climate conditions was exploited. It is located at “Las Palmerillas” Experimental Station in Almería, Spain. It has all the required professional sensors, actuators and commercial data acquisition systems. A set of sufficient and reliable datasets of 15 days were obtained in different agri-seasons and used for different purposes, such as microclimate modelling, online parameter estimation, real-time experimental validation and soft sensor development.

The use of the greenhouses and the acquired datasets in investigations about adaptive and predictive climate modelling using metaheuristic algorithms and machine learning techniques is described in detail in the following chapters.

CHAPTER III GREENHOUSE CLIMATE MODELLING AND OFFLINE CALIBRATION

III.1 Introduction

In this chapter, a set of contributions were achieved. Firstly, a grey-box model for greenhouse temperature prediction under moderate desert climate conditions has been proposed. This contribution stands on reformulating a white-box model to make it independent of the availability of accurate values of the static parameters of its elements. The model has become less complicated by alleviating the coupling between its parameters, which makes it easier for the identification algorithm to find the optimal parameter values. A variant of the Particle swarm optimisation algorithm (PSO) called Random Inertia Weight PSO (RIWPSO) was used to identify the parameters of the proposed model by calibrating it against the experimental data. The greenhouse prototype is located in M'ziraa, Biskra, Algeria, and the designed low-cost data acquisition system has been used to validate the proposed thermal prediction method. The simulation results show that particle swarm optimisation has successfully achieved the desired optimality. The experimental validation process has confirmed the suitability of this model to be implemented to study and predict the greenhouse temperature, and it has emphasised the successful prediction with satisfactory accuracy. Secondly, an enhanced variant of the bio-inspired metaheuristic Bat Algorithm (BA) has been proposed and called the Random Scaling-based Bat Algorithm (RSBA). The proposition includes modifying the exploitation of the standard BA by making the scaling parameter randomly changes over the iterations. It has been dedicated to the same task of calibrating the proposed thermal grey-box model. It has been assessed as same as PSO, primarily on an assumed greenhouse thermal model with known parameters. The simulation results have shown the superiority of the proposed RSBA compared to the standard BA in terms of convergence and performance accuracy. The same dataset from the greenhouse prototype has been used to experimentally investigate the proposed identification method. The obtained prediction results are found to be in good agreement with the measured ones, which show the effectiveness of the proposed RSBA in identifying the real greenhouse thermal model. Finally, a comparative study was conducted between the RSBA and the RIWPSO. The BA has shown a faster convergence than PSO at the beginning of optimisation, but its convergence speed was reduced at the end. BA and PSO have shown superb performance in accurately finding the optimal solutions. However, PSO has shown a superior performance than BA in terms of time consumption regarding the problem of interest.

III.2 Greenhouse air Temperature model

In general, the greenhouse microclimate system is divided into four homogeneous subsystems: the cover, the internal air, the canopy and the soil. In this study, only two components are studied: the cover and the internal air. As matter of fact, the greenhouse is used as a nursery, so the effect of the canopy which is constituted of seedlings could be neglected and since these seedlings are planted in bunches, the soil could also be neglected. The greenhouse temperature model is a mean that quantitatively describes the energy exchanges (Ali et al., 2018). The dynamic temperature behaviour is a combination of physical interactions, including conduction, convection, solar and thermal radiation and infiltration as depicted in Fig. 8. These processes are mainly affected by the outside environmental conditions and the structure of the greenhouse (Rodríguez et al., 2015). The following assumptions have been taken into account:

- The heat exchange of plants is neglected because the main idea was to model an empty greenhouse. In a further step, the effect of the crop through latent heat of transpiration will be included.
- The heat exchange between the soil and the inside air is neglected, due to the role of the wooden floor as a separator.
- There is no stratification in greenhouse air temperature.
- The convective heat transfer coefficient and the absorbed solar radiation are uniform throughout the cover.
- The temperature of the treated soil in seedling trays is not involved, because it has a small effect compared with the other elements due to its small global surface
- The heat storage of plants, internal air and cover can be neglected, due to the quite small heat capacities of these elements compared to existing heat fluxes.
- The greenhouse is East-West oriented.

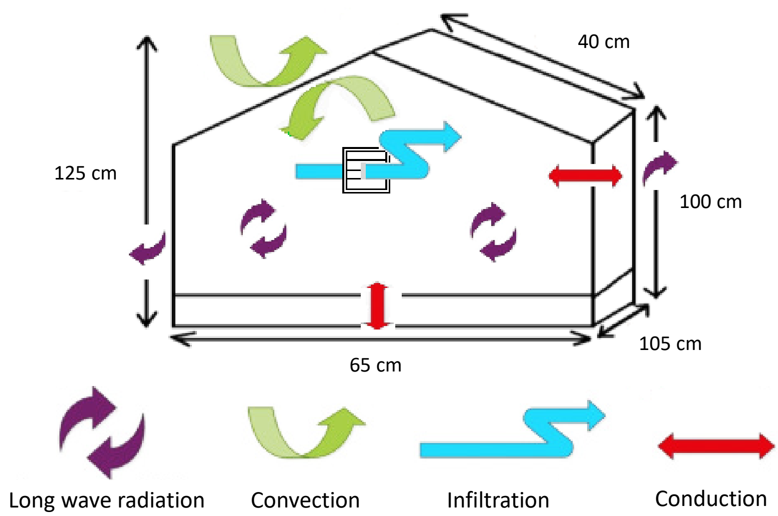

Figure 8. Heat fluxes interactions with greenhouse

III.2.1 Physical-based model

A common simplified internal heat balance equation (Ali et al., 2018) is given by:

$$\rho_a C_a \frac{V}{S_f}\frac{\mathrm{d}T_{in}}{\mathrm{d}t} = Q^{solar} - Q^{cnv,cnd} - Q^{\mathrm{loss}} - Q^{\mathrm{thermal}} \qquad \text{(III.1)}$$

where t is the time in seconds, T_{in} is the internal air temperature (K), ρ_a is the internal air density (Kg m^{-3}), C_a is the specific heat of the air (J Kg^{-1} K^{-1}), V is the volume of the greenhouse (m^3) and S_f is the floor surface (m^2).

The heat fluxes Q (W m^{-2}) are as follows:

- Q^{solar}, is the shortwave radiation absorbed by the greenhouse
- $Q^{cnv,cnd}$, is the convection and conduction heat exchange rate
- Q^{loss}, is the leakage rate of air through the greenhouse

- Q^{thermal}, is the longwave radiation absorbed by the greenhouse

The internal air density ρ_a is considered to be variant with the change of the inside absolute humidity. It has been approved by Iga et al. (2008) that it has a better effect on temperature prediction than being a constant. It is given by:

$$\rho_a = \gamma_0 + H_a \qquad \text{(III.2)}$$

where, γ_0 is the dry air density (Kg m^{-3}) and H_a is the absolute humidity (Kg m^{-3}), which has been obtained by converting the relative humidity (Snyder., 2005), as follows:

$$H_a = \frac{(0.62198\, P_w)}{P_{atm} - P_w} \qquad \text{(III.3)}$$

where, P_{atm} (kPa) is the atmospheric pressure, P_w is the instant vapour pressure, calculated as follows:

$$P_w = P_{ws} \cdot \frac{H_r}{100} \qquad \text{(III.4)}$$

where, H_r is the relative humidity of the greenhouse (%) and P_{ws} is the saturation vapour pressure (kPa °C^{-1}); it is defined by:

$$P_{ws} = 0.61078\, e^{\frac{(17.2694(Tin-273.15))}{Tin}} \qquad \text{(III.5)}$$

The energy heat fluxes are defined in the sequel:

$$Q^{solar} = \alpha_c\, \tau_c\, I \qquad \text{(III.6)}$$

where, α_c is the solar radiation absorptivity of the cover, τ_c is the cover transmissivity, and I is the external global solar radiation (W m^{-2}).

$$Q^{cnv,cnd} = U(T_{in} - Ta) \qquad \text{(III.7)}$$

where, T_a is the ambient temperature (K) and U is the overall heat transfer coefficient through the greenhouse cover proposed by (Ghosal et al., 2005) and calculated as follows:

$$U = [\frac{1}{h_o} + \frac{1}{h_i}]^{-1} \qquad \text{(III.8)}$$

where, h_o and h_i are respectively, the convective heat transfer coefficients of the outside and inside greenhouse cover (W m^{-2} K^{-1}) (Lamrani et al., 2001, Guesbaya et al., 2019). They are computed by:

$$h_o = 2.8 + 1.2\, W_v \qquad \text{(III.9)}$$

$$h_i = 5.2 \cdot |T_{in} - T_a|^{0.33} \qquad \text{(III.10)}$$

where, W_v presents external wind velocity (m s^{-1}).

where ACH is the number of air changes per hour (h^{-1}).

where, T_{sky} is the sky temperature (K) (Guesbaya and Megherbi., 2019).

$$Q^{thermal} = h_o(1 - \tau_c)(T_{in} - T_{sky}) \tag{III.12}$$

$$T_{sky} = 0.0552(T_a)^{1.5} \tag{III.13}$$

III.2.2 Grey-box model

The above physical-based model includes static parameters (V, S_f, γ_0, P_{atm}, C_a, α_c, τ_c, ACH) and dynamic ones (h_o, h_i) that have to be accurately known. In our case, the static parameters C_a, α_c, τ_c and ACH are not known accurately, which highlights the need for their identification. Nevertheless, these parameters are non-linearly dependent on the physical model which will complicate the identification process. To overcome this problem, we suggest in this work combining the dependent static parameters into one parameter in every heat flux equation (Guesbaya and Megherbi., 2019). This consideration has led us to reformulate the model as follows:

$$\rho_a \frac{V}{S_f}\frac{\mathrm{d}T_{in}}{\mathrm{d}t} = Q^{solar} - Q^{cnv,cnd} - Q^{\mathrm{loss}} - Q^{\mathrm{thermal}} \tag{III.14}$$

Where,

$$Q^{solar} = C_{solar}\, I \tag{III.15}$$

$$C_{solar} = \frac{\alpha_c\, \tau_c}{C_a} \tag{III.16}$$

$$Q^{cnv,cnd} = C_{cnd,cnv}\, U(T_{in} - T_a) \tag{III.17}$$

$$C_{cnd,cnv} = \frac{1}{C_a} \tag{III.18}$$

$$Q^{loss} = C_{loss}\, \rho_a\, (T_{in} - T_a) \tag{III.19}$$

$$C_{loss} = \frac{ACH}{3600} \tag{III.20}$$

$$Q^{thermal} = C_{thermal}\, h_o(T_{in} - T_{sky}) \tag{III.21}$$

$$C_{thermal} = \frac{(1 - \tau_c)}{C_a} \tag{III.22}$$

After the above arrangement, we get four new static parameters that have to be identified, which are: C_{solar}, $C_{cnv,cnd}$, C_{loss} and $C_{thermal}$and the other static parameters are fixed and given in Table 2.

Table 2. Input parameters used for computation

Symbol	Value	Unit	Description
S_f	0.68	m^2	Greenhouse floor surface
V	0.76	m^3	Greenhouse volume
γ_0	1.205	Kg m^{-3}	The inside dry air density
P_{atm}	101	kPa	The atmospheric pressure

III.2.3 Pseudo-physical model

The used greenhouse air temperature model is considered as a nonlinear simplified grey-box model (Rodríguezet al., 2015). It includes empirical terms that physically represent heat flux balances inside the greenhouse. The model is described based on the following differential equation:

$$C_{ter}\frac{dXT_{in}}{dt} = Q_{sol,a} + Q_{cnv,ss,a} - Q_{cnv,cnd,ae} - Q_{trp,cr} - Q_{vent,a} \tag{III.23}$$

$$C_{ter} = C_{sph}\, C_{den}\, \frac{C_{vol}}{C_{area}} \tag{III.24}$$

where, XT_{in} represents the predicted inside air temperature of the greenhouse, calculated in $K°$. C_{ter} is the product of specific heat of air, air density and effective height of the greenhouse. Q refers to the heat fluxes occurring inside the greenhouse in $W\,m^{-2}$. Where, $Q_{sol,a}$ is the solar radiation flux absorbed by the air although it is inert to radiation, however, most of the simplified models consider this assumption. $Q_{cnv,ss,a}$ is the convective flux between the soil surface and inside air. $Q_{cnv,cnd,ae}$ represents the convective and conduction fluxes together between inside and outside air (at the cover level). $Q_{trp,cr}$ describes the latent heat effect of crop transpiration. $Q_{vent,a}$ is the heat lost by natural ventilation. All the model parameters are described in Table 3 with their units.

To simplify and reduce the complexity of the model without neglecting the main interactions, it is necessary to consider some empiric approximations based on the thermic components of the greenhouse system. Thus, modelling the main fluxes can be achieved based on different empiric combinations of equations. In this work, we use the following terms:

$$Q_{sol,a} = C_{asw}\, V_{sr,cr} \tag{III.25}$$

where $V_{asw,a}$ is the greenhouse air absorption coefficient of the short-wave radiation. $V_{sr,cr}$ is the internal solar radiation reaching the crop.

$$Q_{cnv,ss,a} = C_{cnv,ss,a}\,(DT_{ss} - XT_{in}) \tag{III.26}$$

Table 3. Nomenclature

Symbol	Description	Unit
Constant arameters		
C_{vol}	Greenhouse volume	m^3
C_{area}	Greenhouse surface	m^2
C_{ven_h}	The vertical distance between the midpoints of the lateral and roof vents	m
C_{ven_l}	Vent length	m
C_{ven_w}	Vent width	m
C_{sph}	Specific heat of air	$J\ kg^{-1}\ K^{-1}$
C_{den}	Air density	$kg\ m^{-3}$
$C_{extsw,cr}$	Short wave crop extinction coefficient (Tomato crop in our case)	(-)
C_g	Gravity constant	$m\ s^{-2}$
C_{nven_lat}	Number of lateral vents	(-)
C_{nven_roof}	Number of roof vents	(-)
Disturbances		
DT_{ss}	The soil surface temperature	$K°$
DT_{out}	The external air temperature	$K°$
$DLAI$	The leaf area index	$m^2\ m^{-2}$
DT_{cr}	The crop temperature	$C°$
D_{ws}	The wind velocity	$m\ s^{-1}$
$D_{sr,e}$	The external solar radiation	$m\ s^{-2}$
Time-varying parameters that have to be calculated		
$V_{lt,vap}$	The latent heat of evaporation	$J\ kg^{-1}$
V_{vpd}	vapor pressure deficit	Hpa
$V_{vent,flux}$	The ventilation flux inside the greenhouse	$m^3\ s^{-1}$
$V_{ven\_area\_roof}$	The opening area of the roof ventilation	m^2
$V_{ven\_area\_lat}$	The opening area of the sidewall ventilation	m^2
Time-varying parameters that have to be estimated in real-time		
$C_{asw,a}$	The greenhouse air absorption coefficient of the short-wave radiation	(-)
$C_{cnv,ss,a}$	Coefficient of convection between the soil surface and internal air	(-)
$C_{cnv,cnd,ae}$	Coefficient of convection and conduction between internal and external air	(-)
C_A	Transpiration coefficient dependent on the crop state and internal radiation	(-)
$C_{B_{d/n}}$	Transpiration coefficient dependent on the crop state and vapor pressure deficit for both periods diurnal and nocturnal	$kg^{-2}\ h^{-1}\ kPa^{-1}$
$C_{ven,d}$	The discharge coefficient	(-)
$C_{ven,w}$	The wind effect coefficient	(-)
C_{loss}	The ventilation loss through greenhouse air leakage	$m^3\ s^{-1}$
$C_{tsw,cv}$	The cover solar transmission coefficient	(-)

where $V_{\text{cnv,ss,a}}$ is the convection coefficient which is considered constant in most of the other investigations. However, it is considered to be dynamic in this work and it has to be online estimated.

$$Q_{cnv,cnd,ae} = C_{cnv,cnd,ae}\,(XT_{in} - DT_{out}) \qquad \text{(III.27)}$$

where $V_{cnv,cnd,ae}$ is the coefficient of the thermal loss considering convection and conduction processes between internal and external air. It is considered to be dynamic in this work and it will be estimated in real-time.

Crop transpiration is one of the main influential effects on inside air temperature. It is represented by the following empirical equation (Rodríguez et al., 2015):

$$Q_{trp,cr} = M_{trp,cr}\,V_{lt,vap} \qquad \text{(III.28)}$$

$$M_{trp,cr} = C_A\ V_{sr,cr}\left(1 - e^{(-C_{ext,cr}\,DLAI)}\right) + C_{B_{d/n}}\,V_{vpd}\,DLAI \qquad \text{(III.29)}$$

$$V_{lt,vap} = 4185.5 * (597 - 0.56 * DT_{cr}) \qquad \text{(III.30)}$$

where $M_{trp,cr}$ is the crop evapotranspiration equation (Sánchez et al., 2012). In our case, it only represents the transpiration process since the soil surface is mulched. V_A (unitless) and $V_{B_{d/n}}$ ($kg^{-2}\,h^{-1}\,kPa^{-1}$) are commonly considered as empirical constant parameters based on the simplified version of the Penman-Monteith evapotranspiration equation. In our case, these two parameters are considered as time-varying parameters and they are estimated in real-time. The equation used to calculate V_{vpd} can be found in the appendix in addition to other needed equations. The crop temperature DT_{cr} is assumed equal to the inside air temperature XT_{in}.

The ventilation flux is calculated by Eq. III.31 and the heat loss by ventilation is obtained based on Eq. III.32 as follows:

$$V_{vent,flux} = C_{ven,d}\left[\left(\frac{V_{ven,area,lat}\,V_{ven,area,roof}}{\sqrt{\left({V_{ven,area,lat}}^2 + {V_{ven,area,roof}}^2\right)^2}}\right)\left(2\,C_g\,C_{ven,h}\,\frac{XT_{in} - DT_{out}}{DT_{out}}\right) + \left(\frac{V_{ven,area,lat}\ \ V_{ven,area,roof}}{2}\right)^2 C_{ven,w}\,{D_{ws}}^2\right]^{0.5} + V_{loss} \qquad \text{(III.31)}$$

$$Q_{vent,a} = C_{den}\,\frac{C_{sph}}{C_{area}}\,V_{vent,flux}\,(XT_{in} - DT_{out}) \qquad \text{(III.32)}$$

where $\text{V}_{\text{ven\_area\_lat}}$ and $\text{V}_{\text{ven\_area\_roof}}$ are the opening areas of the roof and sidewall ventilation, calculated based on the control signals U_{vent} expressed in (%) by the Eqs. III.33 and III.34.

$$V_{ven\_area\_lat} = C_{n,ven,lat}\,C_{ven,l,lat}\,C_{ven,w,lat}\left(\frac{U_{ven\_lat}}{100}\right) \qquad \text{(III.33)}$$

$$V_{ven,area,roof} = C_{n,ven,roof}\,2C_{ven,l,roof}\ C_{ven,w,roof}\left(sin\left(\left(\frac{U_{ven,roof}}{100} * \frac{U_{ven,max}}{2}\right) * \frac{\pi}{180}\right)\right) \qquad \text{(III.34)}$$

where in our work, $V_{ven,d}$ and $V_{ven,w}$ which represent the discharge coefficient and the wind effect coefficient respectively are considered as highly time-varying parameters (dur to wind effect) that have to be online estimated. Moreover, $\mathrm{V_{loss}}$ represents ventilation loss through greenhouse air leakage is also estimated in real-time based on the system needs.

III.3 Solar Radiation model

This simple model simulates the solar radiation passing through the cover and reaching the crop (Rodríguezet al., 2015). It is executed in connection with the air temperature model as one of its sub-equations based on a very simple empirical term described as follows:

$$V_{sr,cr} = C_{tsw,cv}\, D_{sr,e} \qquad \text{(III.35)}$$

where $V_{tsw,cv}$ is the cover solar transmission coefficient which is used to be considered constant in literature. In this work, it is considered as a time-varying parameter and will be online estimated for adaptation purposes to any changes in greenhouse materials (cover, whitening, shading, dirt…etc). Inhere, the estimation technique aims to minimise the error representing the difference between the measured and simulated internal solar radiation variables. Furthermore, the estimation of the radiation parameter at the same time step when also the estimation of air temperature model parameters is performed will prove the capability of the proposed algorithm in handling multi-objective problems of the greenhouse as a MIMO system.

III.4 Model calibration

Model calibration is a prominent procedure that has to be performed by estimating the model parameter values to fit the model to the characteristics of the greenhouse and the climate conditions aiming for the optimal model output whether it is a simulation, estimation, prediction or control. In our case, it is performed to the used grey-box models including empirical equations and parameters in the context of offline parameter estimation based on metaheuristic optimisation algorithms as it is described in the following stages.

III.4.1 Metaheuristic optimisation algorithms:

This section presents the optimisation algorithms used for the calibration of the greenhouse temperature models through a set of offline parameter estimation processes.

III.4.1.1 Random inertia weight particle swarm optimisation algorithm

The PSO algorithm was first chosen to realise an off-line parametric identification of the proposed grey-box model. It in described in Eq. III.14-III.22. PSO algorithm is a prominent metaheuristic, bio-inspired and swarm-intelligence-based algorithm. It is based on imitating the swarm behaviour in nature, such as the behaviour of the bird's flock when searching for food. PSO algorithm is widely used thanks to its ease of implementation and efficiency (Yang., 2014).

The RIWPSO is the specific algorithm that has been used in this work (Guesbaya., 2019). It adjusts the trajectories of individual agents, called *particles*. Each particle, have a precise position in the search space and move at a certain velocity. The position of every particle presents the potential values of the static parameters of the model that have to be identified (C_{solar}, $C_{cnv,cnd}$, C_{loss} and $C_{thermal}$) along with the iterations. The position (x^i), the velocity (v^i) of the i^{th} particle and the inertia parameter (w) are updated as follows (Imran., 2013):

$$x^i(t+1) = x^i(t) + v^i(t+1) \tag{III.36}$$

$$v^i(t+1) = w \cdot v^i(t) + C_1 \cdot r_1 \left(p^i(t) - x^i(t)\right) + C_2 \cdot r_2 (p^G - x^i(t)) \tag{III.37}$$

$$w = w_{min} + (w_{max} - w_{min}) \cdot r_3 \tag{III.38}$$

where, C_1 and C_2 are respectively the constants that affect the cognitive and the social behaviour of particles. p^i and p^G are the best solution achieved by the i^{th}particle, and the whole swarm, respectively

An algorithm has been implemented, to replace the parameter values out of the search range [x^i_{min}, x^i_{max}] with random ones. It is given as follows:

$$If\ (x^i > x^i_{max})\ or\ \left(x^i < x^i_{min}\right)\ then$$

$$x^i = x^i_{min} + \left(x^i_{max} - x^i_{min}\right) \cdot rand$$

The objective function to be minimised by the *random inertia weight PSO algorithm* is the commonly used least-squares criterion, defined as:

$$J = \sum_{i=1}^{N} (x_{real}(i) - x_{sim}(i))^2 \tag{III.39}$$

where, x_{real} and x_{sim} are respectively the real and simulated samples.

III.4.1.2 Random scaling-based bat algorithm

This sub-section concerns the application of the proposed RSBA in an off-line parametric identification process (Guesbaya., 2019). The problem of parameter identification in this paper consists of finding the optimal static parameters (C_{solar}, $C_{cnv,cnd}$, $C_{loss}$and $C_{thermal}$) of the thermal model described in Eq. III.14- III.22, that fit the data samples of the inside air temperature of the greenhouse.

BA is a bio-inspired population-based metaheuristic algorithm. It has been developed based on bats behaviour of how they search for their targets using the echolocation capability (Yang., 2014). Virtual bats are used in simulation and a set of bases are defined for clarification as follows:

- All bats sense distance and differentiate between prey and background barriers by using echolocation.
- Bats fly randomly (Random walks technique) with velocity v_i at position x_i. The frequency (or wavelength) of their emitted pulses is automatically adjusted.

- The rate of pulse emission r can also be tuned automatically according to the closeness of the target, or considered as a constant.
- The loudness can be considered variant, starting from a large positive value A_0 to a minimum value A_{min} or it can also be a constant.

Bats position (solution) x_i^t and velocity v_i^t in a predefined d-dimensional search space are updated according to the following equations:

$$f_i = f_{min} + (f_{max} - f_{min})\beta \tag{III.40}$$

$$v_i^{t+1} = v_i^t + (x_i^t - x_*)f_i \tag{III.41}$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \tag{III.42}$$

where, $f_i \in [f_{min}, f_{max}]$ is the randomly assigned frequency to each bat. $\beta \in [0, 1]$ is a random variable drawn from a uniform distribution, x_* is the current global best solution which is selected after comparing all the solutions of all the n bats.

When a solution is chosen among the current best solutions, a new solution for some bats (according to pulse emission rate) is generated locally during the exploitation stage using random walk based on the following equation:

$$x_{new} = x_{old} + \sigma\epsilon^t\overline{A^t} \tag{III.43}$$

where, x_{new} and x_{old} are respectively the new and the old best local solutions, $\epsilon^t \in$ [-1, 1] is a random number, σ is a scaling parameter to control the step size, and $\overline{A^t}$ is the average of loudness.

The scaling parameter is declared constant for the standard BA. It should be linked to the scalings of the design variables of the problem under consideration (Yang., 2014). However, this is not enough to reach the optimality, because when the optimal solution is near to be reached, the constant step size of search (local random walk) remains relatively large even with the effect of loudness on local search. It results in a reduction in convergence speed and low-accurate solution as a sub-stagnation state at the end. Based on this fact, the scaling parameter σ has been proposed to be (Guesbaya., 2019):

1. Fully responsible about step size control by eliminating the effect of loudness on local search equation given by:

$$x_{new} = x_{old} + \sigma^t\epsilon^t \tag{III.44}$$

2. Dynamically updated over iterations based on random scaling parameter mechanism, given by:

$$\sigma^{t+1} = \sigma_{min} + (\sigma_{max} - \sigma_{min})\beta \tag{III.45}$$

Where the scaling parameter $\sigma \in [\sigma_{min}, \sigma_{max}]$.

As in standard BA, the loudness A_i and the pulse emission rate r_i have to be updated as the iterations proceed. The loudness generally decreases once a bat finds the target, whereas the rate of pulse emission increases. They are updated based on the following equations:

$$A_i^{t+1} = \alpha A_i^t \qquad \text{(III.46)}$$
$$r_i^{t+1} = r_i^0[1 - exp\,(-\gamma\, t)]$$

where α and γ are constants. For any $0 < \alpha < 1$ and $\gamma > 0$, the change in loudness and pulse rate is directed as follows:

$$A_i^t \to 0, \qquad r_i^t \to r_i^0, \qquad \text{as } t \to \infty$$

In order to replace the parameter values x_{ij}^t out of the search range [max x_j, min x_j] with random ones, an algorithm has been implemented, it is given as follows:

$$If\ \left(x_{ij}^t > \max x_j\right)\ or\ \left(x_{ij}^t < \min x_j\right)\ then$$
$$x_{ij}^t = \min x_j + \left(\max x_j - \min x_j\right) \cdot rand(0{,}1)$$

The objective function to be minimised by BAs is the commonly used least-squares criterion, defined as:

$$J = \sum_{i=1}^{N} (y_{real}^i - y_{pre}^i)^2 \qquad \text{(III.47)}$$

where, y_{real}^i and y_{pre}^i are respectively the real and the predicted data samples and N is the number of data samples.

Based on the aforementioned assumptions and rules the essential stages of the proposed RSBA are summarised as the schematic pseudo-code presented in Fig. 9.

```
Random Scaling-based Bat Algorithm

Initialize the bat population x_i and v_i (i = 1, 2, ..., n)
Initialize frequencies f_i, pulse rates r_i and the loudness A_i
while (Current iteration < Maximum number of iterations)
    Generate new solutions (positions) by frequency tuning
    Update velocities and solutions Eqs. (III.40-III.42)
    if (rand > r_i)
        Select a solution among the best solutions
        Generate a local solution around the selected best solution Eq. (III.44)
        Update scaling parameter value Eq. (III.45)
        end if
        Generate a new solution by flying randomly
        if (rand < A_i and f(x_i) < f(x_*) )
        Accept the new solutions
        Increase  r_i and reduce A_i Eq. (III.46)
    end if
    Find the current best x_*
end while
```

Figure 9. Pseudo code of the proposed RSBA

III.4.2 Results and discussion

The aim of this section is three folds as illustrated in Fig. 10 and described in brief as follows:

- Firstly, the investigation of the efficiency of the proposed RSBA against the standard BA and the RIWPSO in a set of offline parameter estimation processes. The parameters to be offline estimated are a set of values that have been logically assumed to create a simulated greenhouse temperature model that was used in turn to obtain a simulated output that represents the target in simulation. The purpose of this fold is to obtain the best search control parameters of the RSBA and RIWPSO and make a comparative study between the performances of all the used metaheuristic algorithms in optimally identifying the assumed parameters.
- Secondly, the implementation of the metaheuristic algorithm with the best performance in the offline parameter estimation to calibrate the real temperature model using an experimental dataset obtained from the constructed greenhouse prototype.
- Finally, the calibrated grey-box temperature model will be experimentally validated using a different experimental dataset to accurately predict the inside air temperature.

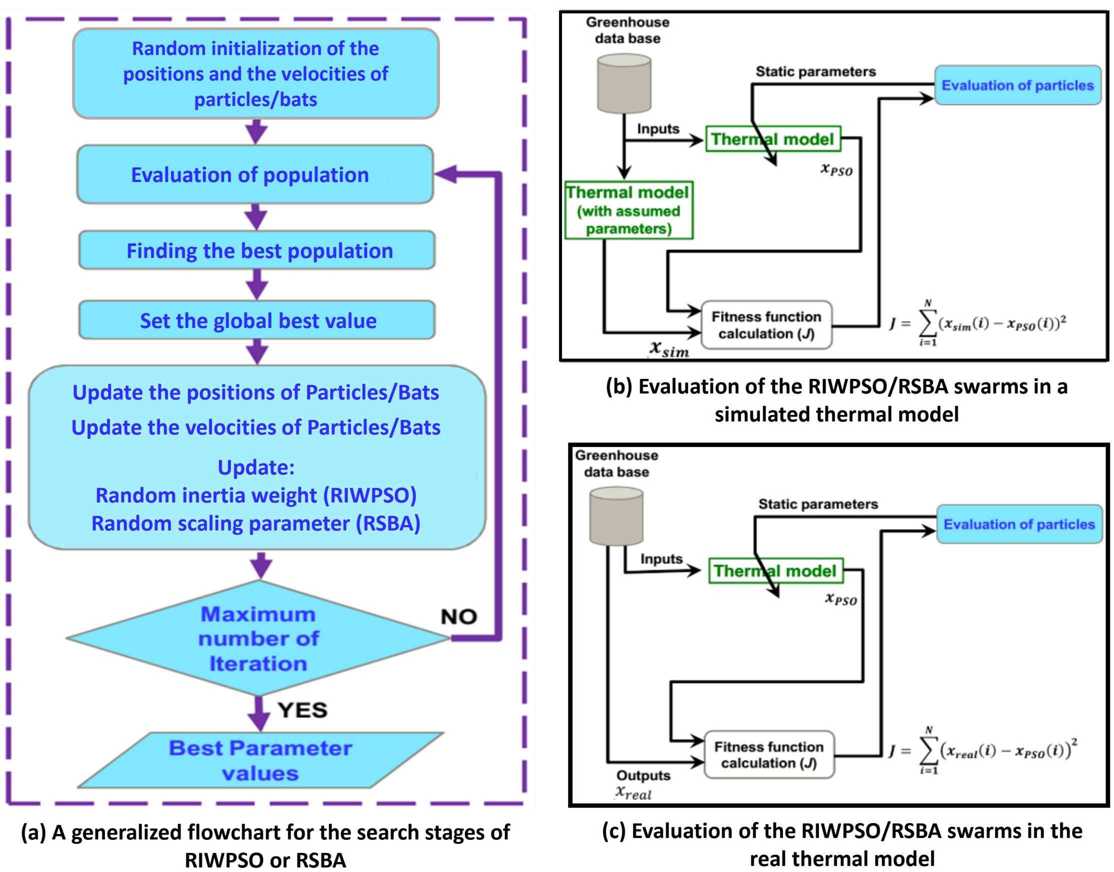

Figure 10. Offline parameter estimation methadology using the RIWPSO and RSBA for the calibration of the proposed temperature model

III.4.2.1 Implementation of the grey-box temperature model

III.4.2.1.1 Greenhouse experimental dataset

The database has been successfully acquired using the constructed greenhouse prototype wirelessly monitored by the designed low-cost data acquisition system. It consists of measurements of five successive days in the winter season ranging from 26th to 30th of January 2019, as it appears in Fig. 11-12. The measuring step size is 1 sample/min, meaning that every day includes 1440 samples. The data of the five days have been divided into two parts; the data of the second and third days have been used in the identification due to its climate diversification. This will ensure an effective selection of model parameter values and flexibility of the prediction process for various climate states (The second day has a calm climate, whereas the third day has a turbulent climate due to wind fluctuations and clouds), as for the data of the remaining three days are kept for the experimental validation (first day as a calm day and the fourth and fifth as turbulent days).

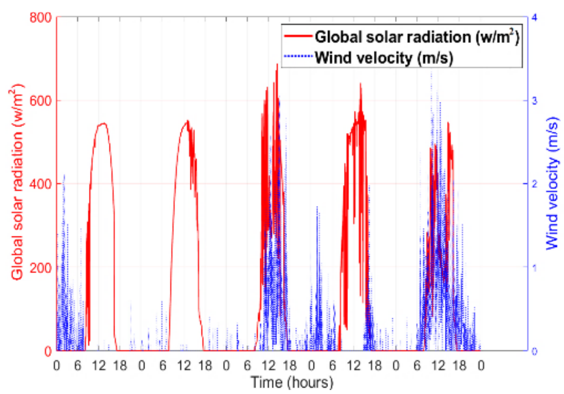

Figure 12. External solar radiation and wind velocity

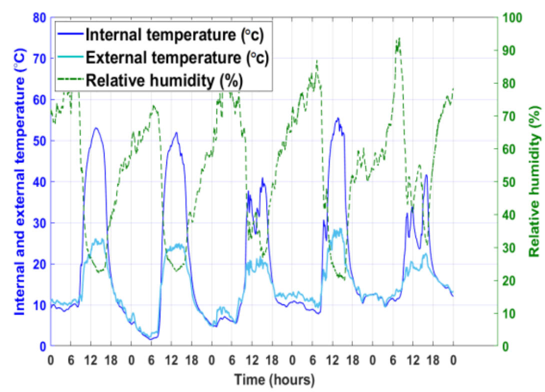

Figure 11. Internal and external temperature and internal relative humidity

III.4.2.1.2 Simulation process

A simulated temperature grey-box model was obtained. It is a model with the same equations as the original model but with assumed values of the static parameter and a simulated output based on that. This aims to assess the RIWPSO and RSBA in calibrating the assumed temperature model and finding the same assumed values of the simulated model. The success of this process with a certain set of algorithm search settings means that the use of the same settings will be efficient in a real process with the original model and using an experimental dataset.

A. Algorithm search specifications

The searching ranges of the model parameters are defined the same for all the used metaheuristic algorithms as follows:

$C_{solar} \in [0, 2.10^{-4}]$, $C_{cnv,cnd} \in [0, 2.10^{-3}]$, $C_{loss} \in [0, 2.10^{-3}]$ and $C_{thermal} \in [0, 2.10^{-4}]$.

The control parameters of the RIWPSO algorithm are: the number of particles of the swarm is equal to 100 particles, the range of inertia is $\in [0,1]$, and the cognitive and the social behaviour coefficients (C_1, C_2) are respectively 2 and 1.5, and the tolerance has been specified to be equal to 10^{-10}.

The common control parameters between standard BA and RSBA are: the number of population is n = 100 bats, the minimum and maximum frequency respectively are f_{min} = 0 and f_{max} = 1.5, the loudness of the initial bats is A_i^0= 1, the rate of pulse emission of the initial bats is generated randomly $r_i \in [0, 1]$ and $r_i^0 = 0.1$, and the constants $\alpha = \gamma = 0.9$. The scaling parameter of the standard BA is constant $\sigma = 10^{-2}$. Whereas the range of its variations regarding RSBA is $[10^{-7}, 10^{-2}]$.

The relative error criterion is considered to compare the parameter values estimated by the metaheuristic algorithms and the optimal assumed ones. It is given by:

$$RE = \frac{|p_{sim} - p_{real}|}{p_{real}} \cdot 100 \quad \text{(III.48)}$$

where, p_{real} is the assumed parameter value and p_{sim} is the parameter value optimised by the used PSO algorithm.

B. Analysis of search dynamics

Five runs of the used RIWPSO algorithm have been achieved using the same control parameters with different initial swarms. the evolution of the best fitness values along the iterations is illustrated in Fig. 13.

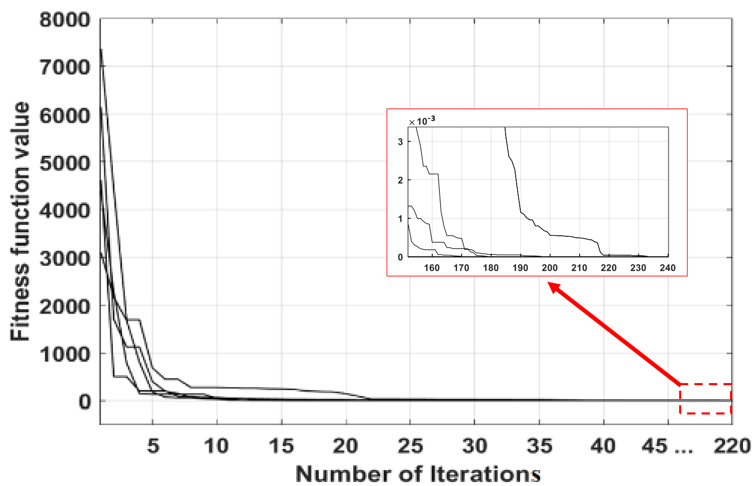

Figure 13. Evolution of the best fitness function values with different initial swarms using RIWPSO

It can be noticed that the convergence is achieved for all the identification attempts before 220 iterations. The fastest run of RIWPSO has attained tolerance after 174 iterations. A successful repetitive convergence appears in this series of RIWPSO-based identification processes.

The analysis is done based on ten runs for the standard BA and the RSBA. The results of identification are analysed according to a limited number of iterations (500 iterations) and achieved using different random initial populations. The evolution of the best fitness function values for all the runs of BAs along the iterations are illustrated in Fig. 14.

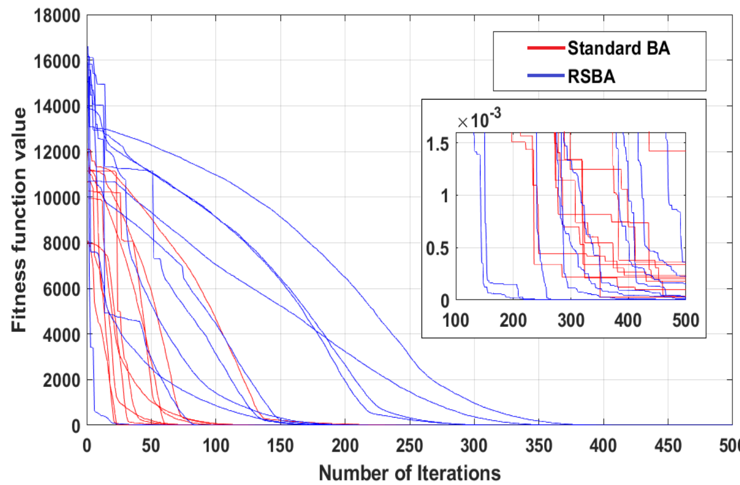

Figure 14. Evolution of the best fitness function values using the RSBA and BA

It can be noticed that standard BA seems to have a weak performance for all the runs in terms of solutions precision. In contrast, RSBA runs have achieved better performance in terms of convergence speed at the late stage and solution accuracy. Table 5 includes all the fitness function values and illustrates that RSBA scored the best fitness value, as a result of the influence of random scaling parameters on the local search step size. Moreover, it shows a good balance between convergence speed and result quality. As a result, RSBA proves to be more efficient than the standard BA in finding the parameters of the thermal model.

C. The evaluation of parameter values

The best-identified parameter values that produced the best fitness function values using both the RIWPSO and RSBA are listed in Table 4 in contrast to their corresponding values and the relative errors (the precent error). The evolution of the relative error of every optimised parameter r by the RIWPSO is shown in Fig.15. The RIWPSO algorithm succeeded at the end in optimally finding the assumed parameter values. It is commonly admitted that the acceptable identification result corresponds to relative errors of less than 0.5%. Thus, the results show that the parameters have been high-accurately found. The evolution of the relative error of every identified parameter value from the best RSBA execution is shown in Fig. 16. It is clear that the RSBA also succeeded at the end in optimally identifying the parameter values and it also illustrates the fast convergence of the RSBA to the optimal values at a very early stage of the execution.

Table 4. Comparison between the assumed and the identified model parameter values

Parameter	Assumed value	RIWPSO		RSBA	
		Identified value	Relative error (%)	Identified value	Relative error (%)
C_{solar}	5.8e-05	5.8e-05	4.23e-06	5.8e-05	6e-05
$C_{cnv,cnd}$	5.3e-04	5.3e-04	2.7e-05	5.3e-04	1e-03
C_{loss}	1.7e-04	1.69e-04	1.33e-04	1.69e-04	6.25e-03
$C_{thermal}$	2.2e-05	2.19e-05	7.67e-06	2.19e-05	5.37e-05

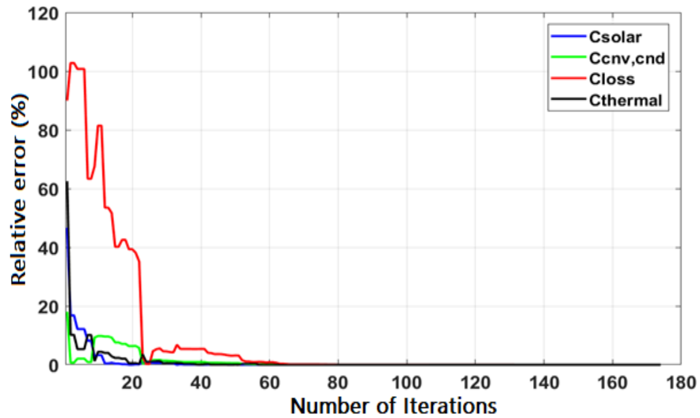

Figure 15. Evolution of the relative error of every parameter using the RIWPSO

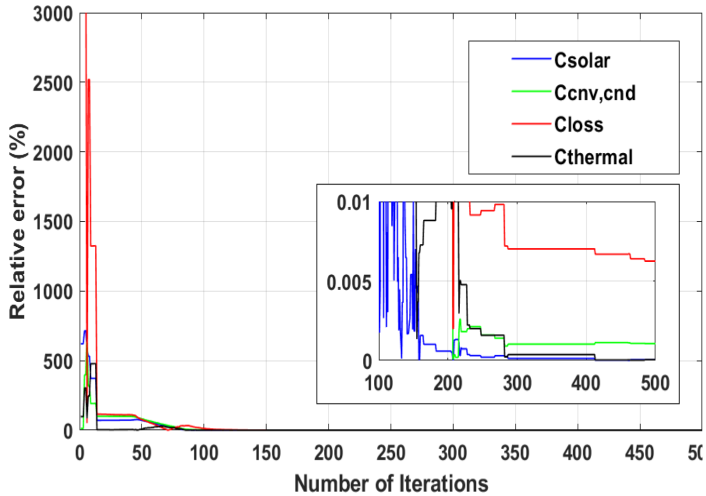

Figure 16. Evolution of the relative error of every parameter from the best identification run of the RSBA

It is shown in all cases that the value of C_{loss} is the identified parameter with the lowest accuracy due to its small effect on the output but still considered very satisfactory, in contrast to $C_{thermal}$ and C_{solar} which are the most accurate identified parameters because of their strong effect on the inside air temperature. All the identified values show errors of less than 0.5%. Hence, the results show that the parameters have been successfully identified with satisfactory accuracy.

A comparison study has been conducted between the used metaheuristic algorithms. It can be noticed that the enhanced RSBA has outperformed the standard BA. Furthermore, it can also be noticed that the RIWPSO has outperformed the RSBA in terms of:

- Accuracy of the identified (offline estimated) parameters.
- Speed of the process where RIWPSO needed less than 180 iterations to reach the optimal values.

The RSBA has a prominent advantage which is:

- The fast convergence to the optimal values at an early stage of the execution.

Table 5. Best fitness function values of identification processes

Identification Process	Bat Algorithms	
	Standard BA	*RSBA*
1st run	3.4003e-04	9.7949e-08
2nd run	1.4187e-03	2.6286e-05
3rd run	3.6587e-04	1.0008e-06
4th run	2.1736e-04	1.4771e-04
5th run	2.0633e-04	1.3245e-05
6th run	2.2142e-04	2.9300e-05
7th run	2.8384e-05	2.2254e-04
8th run	3.7457e-05	7.1692e-07
9th run	1.7585e-04	3.5900e-04
10th run	1.0056e-04	1.5055e-06

This proves the superiority of the RIWPSO over the RSBA in solving this studied problem. The RSBA could be more useful than the RIWPSO against other different problems such as the online parameter estimation in simulation or in real time, where the advantage of the early convergence can be nessccary due to the time limits and constraints.

III.4.2.1.3 Experimental validation process

This section aims to validate the simulation outcomes by implementing the proposed temperature model, the RIWPSO and the proposed RSBA using the experimental dataset obtained from the greenhouse prototype. The calibration of the real model has been achieved using the same control parameters as for the identification of the simulated greenhouse model but this time, the fitness function included the real inside air temperature of the real greenhouse prototype, not the simulated one as in the previous section

A. Calibration of the real greenhouse model using RIWPSO and RSBA

The RIWPSO and RSBA-based model calibration processes of the temperature model has been achieved using the same search control parameters as for the calibration of the simulated temperature model. Whereas, in this section, the target is the offline model calibration and the prediction of the real inside air temperature. The identified parameter values by both metaheuristic algorithms using the data of the second and third days of the experimental database are listed in Table 6. The results of offline model calibration and thermal prediction using the real identified parameter are shown in Fig 17-18.

Table 6. Identified parameter values of the real greenhouse model

Parameter	Experimentally identified values by the RIWPSO	Experimentally identified value using RSBA
C_{solar}	6.83449810^{-5}	$6{,}83455\ 10^{-5}$
$C_{cnv,cnd}$	$4.142013\ 10^{-4}$	$4{,}14191\ 10^{-4}$
C_{loss}	$2.556947\ 10^{-4}$	$2{,}55726\ 10^{-4}$
$C_{thermal}$	$2.556266\ 10^{-5}$	$2{,}55627\ 10^{-5}$

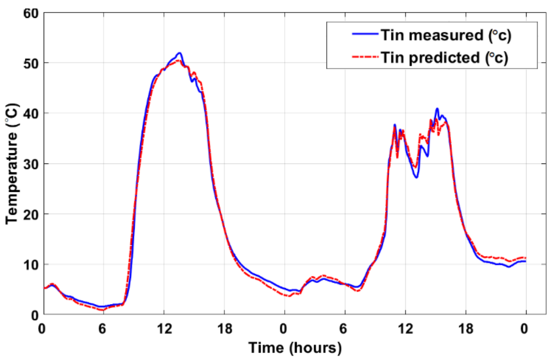

Figure 17. RIWPSO-based model calibration: the measured and predicted inside temperature of the 2nd and 3rd days

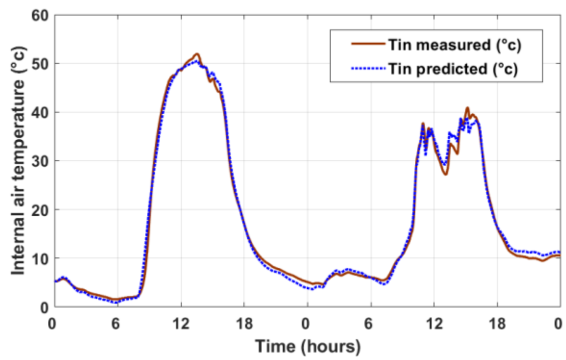

Figure 18. RSBA-based model calibration: the measured and predicted inside temperature of the 2nd and 3rd days

The grey-box model has been experimentally validated using the real offline identified parameter values found by both RIWPSO and RSBA, to predict the inside air temperature of the first, fourth and fifth days of the available dataset. Fig 19-20 show the variation of the predicted and measured temperature inside the greenhouse prototype. Five criteria are used to evaluate the performance of the calibrated model: *Mean Absolute Error* (MAE), *Max Absolute Error* (MaxAE), *coefficient of determination* (R^2) and *Model Efficiency* (EF). They are calculated based on acquired data, by comparing the predicted and the measured temperature using Eqs. III.49-III.52.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|T_m(i) - T_p(i)\right| \tag{III.49}$$

$$\text{MaxAE} = \max\left|T_m(i) - T_p(i)\right| \tag{III.50}$$

$$R^2 = \left[\frac{\sum_{i=1}^{n}(T_m(i) - \overline{T_m})(T_p(i) - \overline{T_p})}{\sum_{i=1}^{n}(T_m(i) - \overline{T_m}) \cdot \sum_{i=1}^{n}(T_p(i) - \overline{T_p})}\right]^2 \tag{III.51}$$

$$EF = \frac{\sum_{i=1}^{n}(T_m(i) - \overline{T_m})^2 - \sum_{i=1}^{n}(T_p(i) - T_m(i))^2}{\sum_{i=1}^{n}(T_m(i) - \overline{T_m})^2} \tag{III.52}$$

where, T_m is the measured temperature, $\overline{T_m}$is the mean value of the measured temperature, T_p is the predicted temperature, $\overline{T_p}$is the mean value of the predicted temperature and n is the number of data samples.

Table 7 presents the results of this statistical analysis that aims at evaluating the performance of the calibrated models by both used metaheuristic algorithms. The surprising thing is that the qualitative and quantitative results are the same which proves the efficiency and usefulness of both RIWPSO and RSBA. The MAE shows a satisfactory evaluation represented by a very small difference between the measured and predicted inside air temperature. We can also see that R^2 gives a very high value for regression analyses meaning that 99.5% of the variance in the measured inside air

Table 7. Statistical evaluation of the performance of the calibrated models in predicting the inside air temperature

Criterion	MAE	R^2	EF	$MaxAE$
RIWPSO	0.7937	0.9956	0.9951	3.3910
RSBA	0.7937	0.9956	0.9951	3.3910

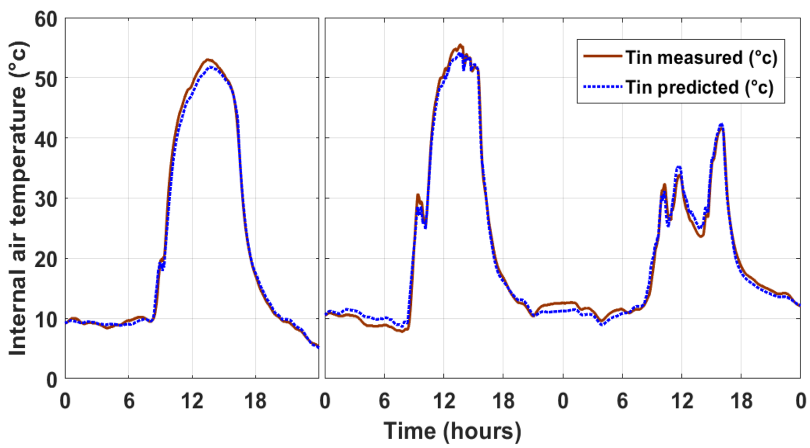

Figure 19. Experimental validation of the calibrated model using RSBA using a dataset of the 1st, 4th and 5th days: measured and predicted inside air temperature

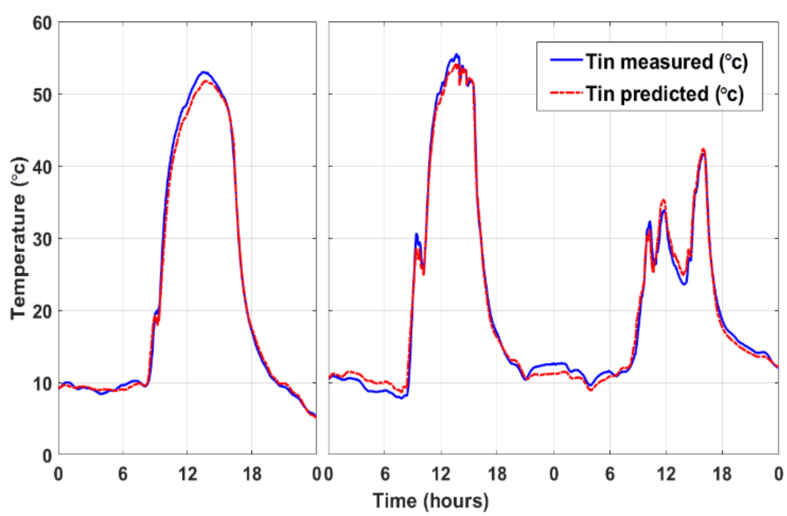

Figure 20. Experimental validation of the calibrated model using RIWPSO using a dataset of the 1st, 4th and 5th days

temperature of the greenhouse prototype has been predicted by the real thermal model. Moreover, EF shows a very good evaluation where the closer the value to 1, the more accurate the model. The $MaxAE$ shows that the biggest error value among all the data samples equals 3.39°C which is logically small. The aforementioned evaluation criteria illustrate that the model is a promising tool to study this phenomenon in terms of usefulness, efficiency and applicability.

III.5 Conclusions

In this chapter, a grey-box model is proposed to simulate the temperature inside the greenhouse in an arid region. The unknown parameters of this model have been identified using the Random Inertia Weight PSO algorithm. The efficiency of this optimisation algorithm was assessed in calibrating an assumed and a real temperature model. The simulation results using the assumed greenhouse temperature have shown the usefulness of the used PSO algorithm in identifying the model parameters with high accuracy. The calibrated model using the real greenhouse temperature has been validated, and the results have shown a satisfactory fit between the measured and predicted inside temperature. The proposed modelling methodology can be used for energy balance studies of greenhouses. It could also be adopted as a tool to study different climate conditions or for practical applications of control systems. This study paves the way for future investigations on applying a cooling system to overcome the harsh summer climate of arid regions.

As a second contribution in this chapter, a variant of BA called RSBA has been proposed and used to identify the parameters of the same temperature model. The proposed RSBA differ from the standard BA in using a dynamic scaling parameter with a random mechanism. A comparative study has been conducted between the proposed RSBA and the standard BA using an assumed output of the temperature model with known parameters. By examining the search evolution and the optimisation results, it has been found that the RSBA outperforms the standard BA, and its exploitation capability has been enhanced. Specifically, the solution accuracy and the convergence speed when the optimal solution is near to be found have been effectively increased. The RSBA was then applied to identify a real greenhouse thermal model. The results of the identified model validation using the experimental database exhibited a satisfactory fit between the measured and predicted inside air temperature. This study paves the way for future investigation on developing a novel mechanism that adapts the scaling parameter changes with the closeness to finding the optimal solution.

Another comparative study has been conducted between the proposed RSBA and the used RIWPSO. The results have shown the superiority of the RIWPSO over the RSBA in solving the problem at hand. However, the RSBA could still be more useful than the RIWPSO against other different problems, such as the online parameter estimation in simulation or in real-time, where the advantage of the early convergence to optimality can be necessary due to the time limits and constraints.

CHAPTER IV

CLIMATE MODEL UTILITY AND RELA-TIME MODEL ADAPTATION

IV.1 Introduction

Greenhouse microclimate modelling is a difficult task mainly due to the strong nonlinearity of the phenomenon and the uncertainty of the involved physical and non-physical parameters. The uncertainty stems from the fact that most of these parameters are unmeasurable or challenging to measure, and some are time-varying, signifying the necessity to estimate them. As the first contribution in this chapter of the thesis, a methodology for online parameter estimation is proposed to deal with the estimation of the time-varying parameters of a simplified greenhouse temperature model for real-time model adaptation purposes. An online estimator is developed based on an enhanced variant of the Bat Algorithm called the Random Scaling-based Bat Algorithm. It allows the continuous adaptation of the internal air temperature model and the internal solar radiation sub-model by estimating their parameters simultaneously by minimising a cost function, intending to achieve global optimality. Constraints on the search ranges are imposed to respect the physical sense. The adaptation of the models was tested with recorded datasets of different agri-seasons and on a real greenhouse in real time. The evolutions of the time-varying parameters were graphically presented and thoroughly discussed. The experimental results illustrate the successful model adaptation, presenting an average error of less than 0.28 °C for air temperature prediction and 20 W m^{-2} for solar radiation simulation. It proves the usefulness of the proposed methodology under changing environmental conditions.

Natural ventilation flux is an important variable to measure or estimate for its significant effect on greenhouse microclimate modelling and control. It is commonly known that it can be mathematically estimated depending on the type and dimension of the greenhouse and its vents and, most importantly, on the vents opening percentage. However, most commercial greenhouses are not equipped with an automatic vent opening system which obligates the grower to perform manual control, in addition to the lack of vent position sensors, due to economic and management reasons. It leads to the absence of the control signal variable representing the vent opening percentage necessary for ventilation flux estimation. This issue has been encountered in this work after attempting to implement the developed adaptive microclimate model based on the online parameter estimator through an IoF2020 platform (internet of food and farm) in a set of commercial greenhouses with manually controlled vents located in Almeria province, Spain. To cope with this issue, the estimation of ventilation flux without using the vent opening percentage was investigated. A virtual sensor for greenhouse ventilation flux estimation is proposed as a second contribution in this chapter. It has been developed using a nonlinear autoregressive network with exogenous inputs based on principal component analysis using the available measured data and the evolutions of the heat fluxes representing the greenhouse energy balance. Preliminary results show an encouraging performance of the virtual sensor in estimating the ventilation flux with a mean absolute error of 0.41 m^3 s^{-1}.

IV.2 Real-time climate model adaptation

Online parameter estimation consists of estimating the values of the parameters of a model in parallel with the operation of a real system, using the available data from the real system to achieve the model adaptation. Very few proposals on online parameter estimation and model adaptation techniques for greenhouse microclimate modelling have been reported in the literature (Pérez-Gonzalez et al., 2018).

Metaheuristic optimisation algorithms seem to offer promising results. In recent years, the popularity of nature-inspired optimisation algorithms is expanding, and these algorithms are being developed at an increasing rate (Yang., 2014). One of these well-known algorithms is the bat algorithm (BA), which has been enhanced in different ways (Yang and He, 2013). In this chapter, an enhanced variant of BA, proposed in previous work (Guesbaya et al., 2019), called the random scaling-based bat algorithm (RSBA), is used and it constitutes the core of the developed online parameter estimator. The BA algorithm has been chosen for its simple application and robustness thanks to its interesting features and for outperforming some well-recognised optimisation algorithms such as genetic algorithms (GA), particle swarm optimisation (PSO), harmony search (HS) and simulated annealing (SA) in handling complex optimisation problems (Khan and Sahai., 2012). The proposal of this work is summarised in Fig. 21.

The main objective in this section of this chapter of the thesis is to deal with the need to periodically adjust the values of the time-varying parameters of a simplified greenhouse temperature model and its solar radiation sub-model by using online parameter estimation to reach their optimal performance. Both microclimate variables were chosen to be studied due to their significant and direct effect on crop growth. The online estimation methodology followed in this work consists of two main stages:

(I) Pre-processing stage includes offline model calibration and model sensitivity analysis.

(II) Online parameter estimation stage includes simulations with datasets of different agri-seasons and real-time implementation of a real greenhouse system.

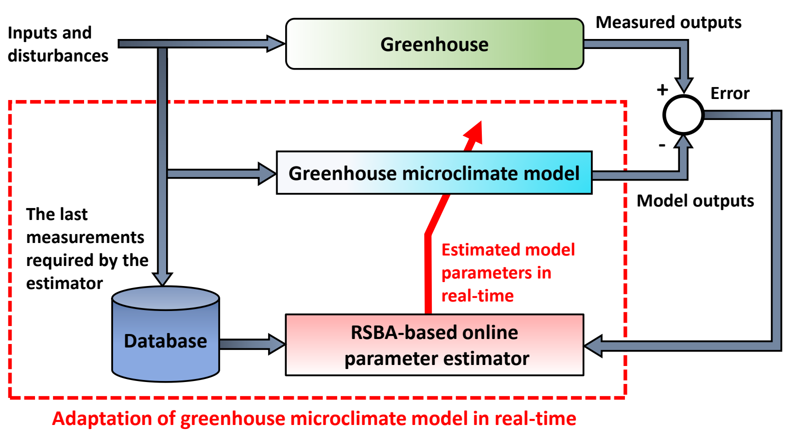

Figure 21. Principle of the adaptation of a greenhouse microclimate model in real time using the developed online parameter estimator

IV.2.1 Materials and methods

IV.2.1.1 Greenhouse experimental datasets

In this work, three different datasets containing the greenhouse climate variables are used. Two datasets were already acquired in different periods of the year. The third dataset was acquired during the testing of the developed online estimator in real time. The period contained in all the datasets was chosen to be 15 days because it is a sufficient period to catch the different dynamics of the climate in the location of the greenhouse. The first and second datasets were acquired in the transitional periods between seasons when the weather is more diversified. Although the sampling time of the data acquisition system is 30 seconds, the datasets were recorded by measuring the variables every one minute, which is sufficient for the model adaptation problem. The wind velocity variable was filtrated using a low-pass filter. The first dataset was acquired during the transitional period between the winter and spring seasons, starting from 27 March 2020 to 11 April 2020 (15 days, 21500 samples), as presented in Fig. 22. The second dataset was acquired during the transitional period between the summer and autumn seasons, starting from 01 September 2020 to 15 September 2020 (15 days, 21500 samples) as presented in Fig. 23. The third dataset was acquired during the winter season starting from 07 January 2021 to 22 January 2021 (15 days, 22000 samples) as presented in Fig. 24.

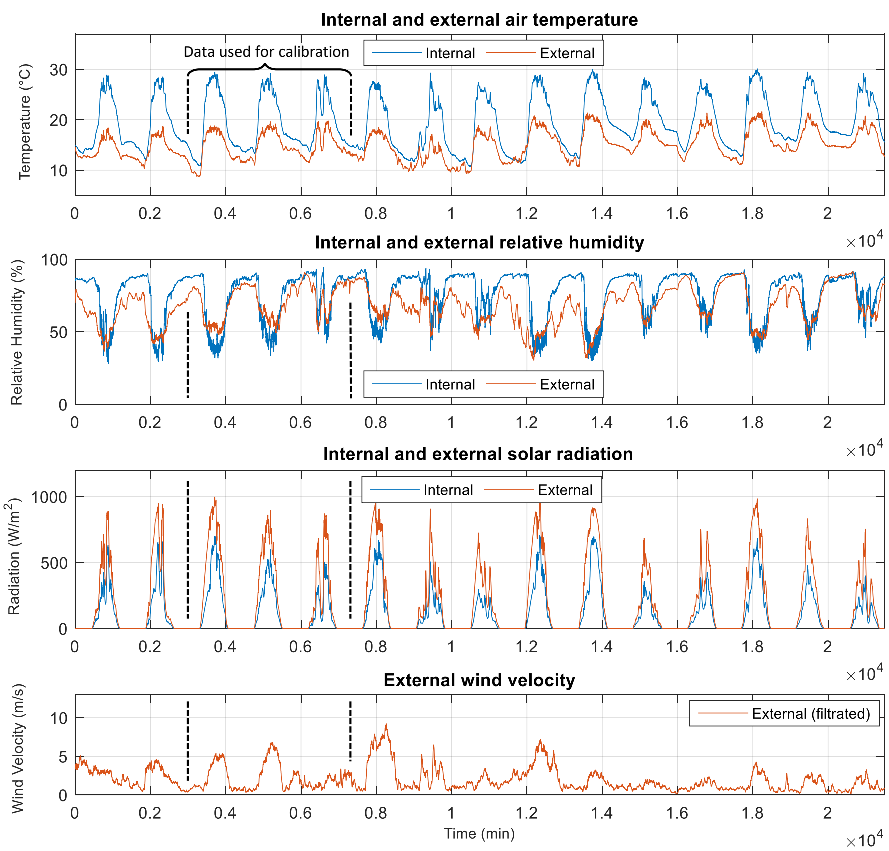

Figure 22. Dataset of the transitional period between winter and spring seasons starting from 27 March 2020 to 11 April 2020

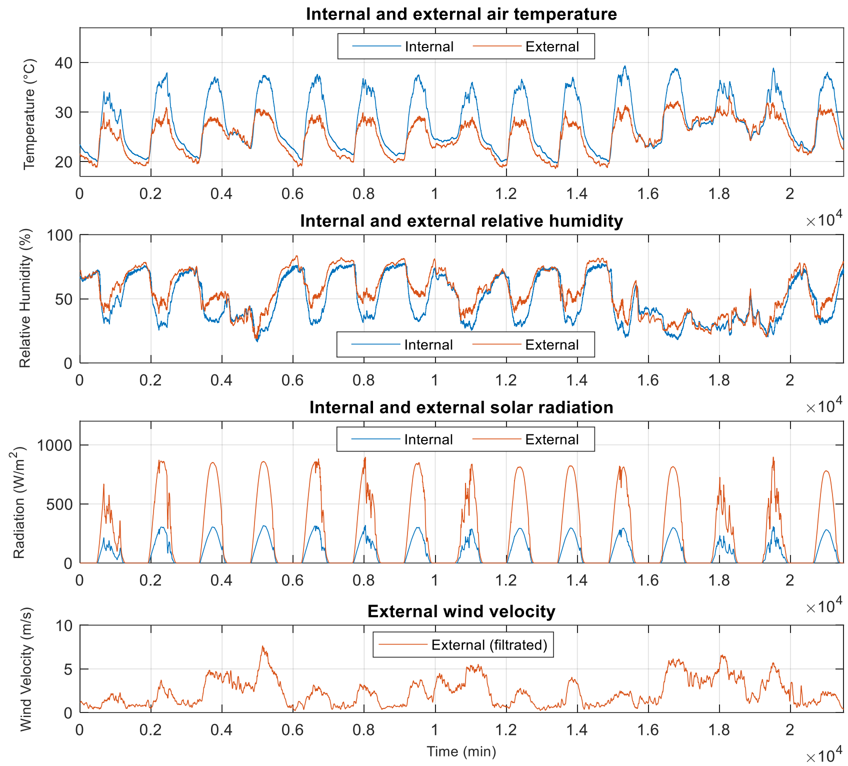

Figure 23. Dataset of the transitional period between summer and autumn seasons starting from 01 September 2020 to 15 September 2020

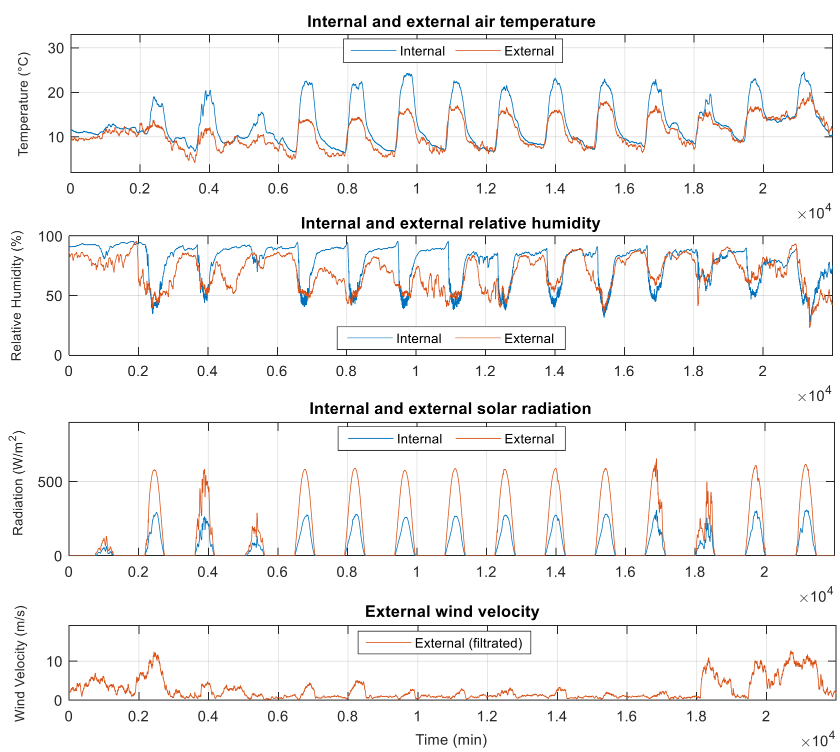

Figure 24. Dataset acquired during the real-time application of the online estimator in winter season starting from 07 January 2021 to 22 January 2021

IV.2.1.2 Methodology for online parameter estimation

This section presents the proposed model adaptation methodology and the developed online parameter estimator based on RSBA for greenhouse microclimate adaptive modelling purposes. Implementation efforts have been focused on demonstrating the potential of the developed estimator to achieve an optimal adaptation of the used microclimate model in real time according to Fig. 21. The online parameter estimation enhances the accuracy of the internal air temperature prediction and the internal solar radiation simulation at the same time step. The main stages of the methodology and their purposes are illustrated in Fig. 25. All steps are explained in the following points:

1. Offline model calibration: This stage consists of the application of the RSBA to calibrate the greenhouse microclimate model with an offline parameter identification process using recorded experimental data from the greenhouse. For this offline calibration, all the parameters of the model are considered constant. The identified parameter values are calculated so that they can be used in the next stages for the model sensitivity analysis and as initial parameter values for the next online estimation processes using datasets of different seasons. The cost function used in this stage to evaluate the calibration of the model is the Mean Square Error (MSE).

2. Model sensitivity analysis: In this stage, the sensitivity of the model is studied to understand the influence of its parameters. Two different sensitivity analyses are performed to compare results when assuming constant versus time-varying parameters. On the one hand, it aims to investigate how much each parameter affects the model outputs. On the other hand, using different sets of parameter values with the same greenhouse climate dataset of one day helps to examine the change in the model sensitivity affected by the time-varying parameter values. These tests are also performed to facilitate the selection of the variation ratios for the parameters, depending on how sensitive the model is toward each parameter.

3. Simulation of the online parameter estimation: In this stage, the model adaptation is performed in simulation and the final structure of the developed online parameter estimator is accomplished as illustrated in Fig. 26. It involves a combination of mechanisms that are designed based on the results of the previous stages and some trial-and-error procedures. For the estimation process, ten parameters of the greenhouse microclimate model are considered

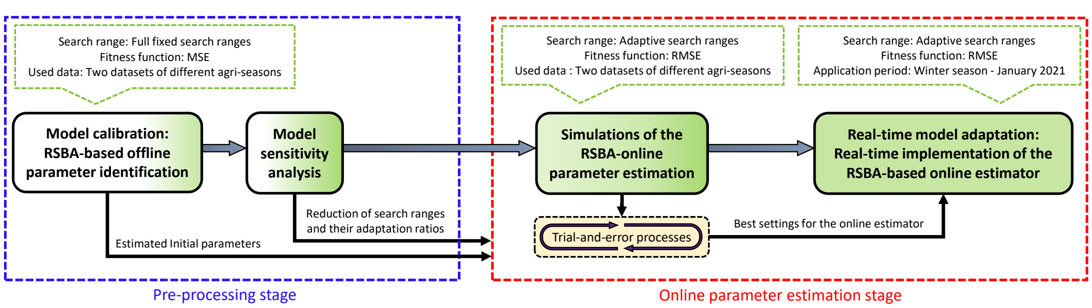

Figure 25. Proposed methodology for the adaptation of the greenhouse microclimate model using the developed online estimator

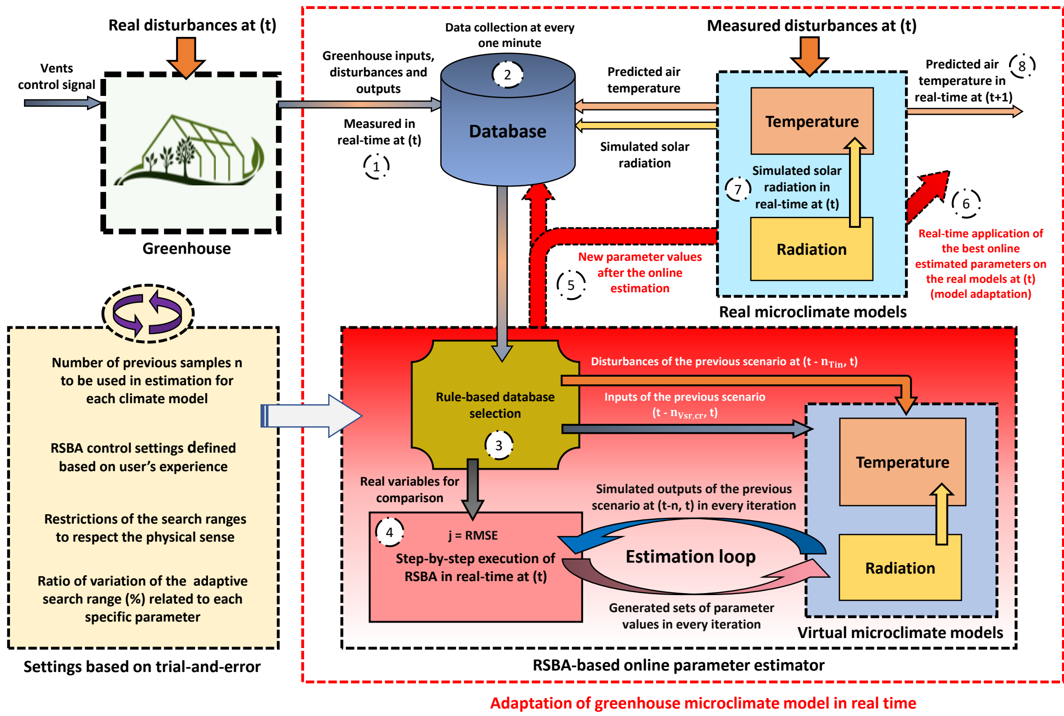

Figure 26. Online estimator mechanism and its application scheme for microclimate model adaptation in real time

time-varying. The microclimate model is effectively adapted by online estimating the values of the time-varying parameters to minimise the cost function for this stage which is the Root Mean Square Error (RMSE) representing the error between the real measured data and the model output. The RMSE penalises errors greater than 1, which helps in avoiding undesirable large error oscillations. The online estimation is performed with real datasets of different seasons to assess the adaptation capability of the estimator against different climate conditions. The proposed estimation mechanisms, settings and constraints are described as follows:

a. The air temperature model and solar radiation sub-model are adapted together as two targets but with different execution times for their respective online parameter estimation processes.

b. The online estimation processes for both models are performed based on two other "virtual models" identical to the original ones. The "virtual models" are used to simulate the previous scenario consisting of the n last time instants at $[t-n,\ t]$ based on last previous inputs, outputs and disturbances at $[t-n,\ t]$. This means that the selected n data samples are used for the estimation process in a sub-algorithm with an identical greenhouse microclimate model. This sub-algorithm is used as a testbed where all the potential solutions (sets of parameter values) generated by the RSBA are evaluated to find optimal values for the model parameters by minimising a cost function for the error between the real n data samples and the model simulated variables. This aims to optimise the performance of the "virtual models" in simulating the previous scenarios at $[t-n,\ t]$ according to a specific number of iterations; then adapting the original models by applying the best-estimated values of the time-varying parameters in real time at t before predicting the next sample at $(t + 1)$.

c. The number n of previous time instants representing the last scenario at $[t - n,\ t]$ can be adjusted to suit the characteristics of each phenomenon to be simulated thanks to a rule-based data selection algorithm. The rule-based data selection algorithm is programmed in a nested way with the "virtual model" sub-algorithm to provide it with the previously measured data that represents the selected past time instants.

d. The RSBA is used as it is proposed in (Guesbaya., 2019) except for the search ranges which are originally constant but, in this work, most of them are programmed to be dynamic and adaptive based on the physical nature of each parameter which determines how it varies in time. As described in Fig. 27, the adaptive search range of each parameter j varies between the boundaries of a larger range that represents the constraints for the adaptive search range. Each adaptive search range at t is determined based on the previous best parameter value according to a specific variation ratio $\pm R_j\%$ of the best parameter value itself (neighbourhood of variation) as presented in the following terms:

$$LB_j^{t+1} = C_j^t(1 - R_j\%) \tag{IV.53}$$

$$UB_j^{t+1} = C_j^t(1 + R_j\%) \tag{IV.54}$$

where LB_j^{t+1} and UB_j^{t+1} are respectively the new lower and upper boundaries of the adaptive search range and C_j^t is the current value of the specific j^{th} estimated parameter at t.

e. A set of constraints are defined to restrict each adaptive search range (see Fig. 27). They are defined based on the common ranges mentioned in the literature for greenhouse microclimate modelling (Rodríguez et al., 2015; Choab et al., 2019), the physical nature of each parameter, and some trial-and-error procedures performed with the microclimate model during the development of this tool. It is important to highlight that, at each time instant, the parameters being estimated are only the ones related to an active physical

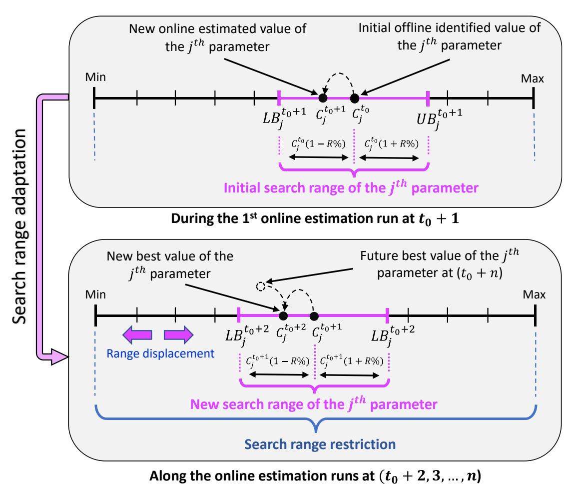

Figure 27. The adaptation mechanism of the parameter search range with respect to its restriction

process of the greenhouse microclimate at that moment t. For instance, the parameter related to radiation is online estimated only when radiation is greater than 5 $\mathrm{W\,m^{-2}}$, the parameters related to ventilation are online estimated only when the vents of the greenhouse are open according to a control signal greater than 0%, and the parameters related to transpiration are online estimated only when the crop exists in the greenhouse with an LAI greater than 0.1 $\mathrm{m^2_{leaf}\,m^{-2}_{ground}}$. Otherwise, the values of those parameters are constant, equal to the last estimated value when the corresponding physical process was active.

4. Experimental implementation: The last stage is dedicated to the implementation of the developed online parameter estimator on the greenhouse system in real time. It is considered a crucial stage to validate the real-time adaptation of the greenhouse microclimate model under real crop growth conditions.

IV.2.2 Results and discussion

This section presents the quantitative and qualitative results obtained for each development stage of the explained online parameter estimator. The statistical criteria used to evaluate all the results in the following sub-sections are: Mean Absolute Error (MAE), Max Absolute Error (MaxAE), Coefficient of Determination (R^2), Residual Error (RE), MSE and RMSE. For the simulation tests, the computational unit utilised was a computer consisting of an Intel Core i7-4810MQ with an octa-core processor, 2.8 GHz, 16 GB RAM DDR3 1600 MHz, running a Windows™ 10 64-bit with MATLAB R2017b. The online parameter estimator has been coded, tested and executed in MATLAB.

IV.2.2.1 Pre-processing stage

In this stage, analysis of the model performance with constant parameters was conducted, the most influential parameters on the output were observed, and the model sensitivity to time-varying parameters was analysed.

A. Offline model calibration

The offline model calibration procedure to identify the values of all the parameters of the greenhouse microclimate model can be found in (Rodríguez et al., 2015). Offline model calibration based on the RSBA algorithm is performed intending to obtain the best possible prediction of the internal air temperature and simulation of the solar radiation, assuming constant parameters. Furthermore, the analysis aims to determine adequate search ranges for the parameters. The prediction results with offline calibrated constant parameters will be compared to the prediction results using the online estimated parameters to demonstrate the capability of the developed estimator.

To offline calibrate the greenhouse model, two different datasets were used in this stage, one from the winter-spring period (see Fig. 22) and another one from the summer-autumn period (see Fig. 23). Three days of the winter-spring dataset (3rd, 4th and 5th days) were selected for the calibration process as a climate-diversified target, and both complete datasets (15 days per each) were used for validation. The simulation step time and prediction horizon were fixed as one minute. The settings of the RSBA were chosen based on the personal experience with the algorithm and some trial-and-error processes, resulting as follows: the number of bats is 20, the maximum number of iterations is 500, the minimum and maximum frequency respectively are $f_{min} = 0$ and $f_{max} = 2$, the loudness of the initial bats is $A_i^0 = 1$, the rate of pulse emission of the initial bats is $r_i = 0$ and $r_i^0 = 0.2$ and the constants α and γ are equal to 0.8. The scaling parameter randomly variates in a range of $\sigma \in [1, 10^{-3}]$.

The search ranges and the calibrated parameter values after the offline calibration are presented in Table 8. The results of the offline calibration process of air temperature prediction and radiation simulation are highlighted in Fig. 28 and Fig. 29, contained among the validation results of the calibrated model using the complete dataset. It is observed that the model calibration process was successful according to the acceptable fit between the measured variables and the outputs. Table 9 contains the statistical results with a $MAE = 0.75$ °C, $MSE = 1.12$ °C^2 and $R^2 = 0.96$.

Fig. 28 presents the graphical results of validating the offline calibrated air temperature model with the complete winter-spring dataset. It can be noticed that the prediction accuracy for the air temperature model in other days of the same dataset has decreased, showing larger errors between the model outputs and measured variables. As shown in Table 9, the error increases even more in the validation using the summer-autumn dataset (the graphical result is not presented), evidencing an unsuccessful prediction for the offline calibrated model using a different dataset. Similar conclusions concerning the validation of the calibrated radiation sub-model can be obtained according to Fig. 29 and Table 10. The poorly simulated radiation thresholds in most of the days can negatively affect the temperature prediction. Therefore, the long-term prediction of temperature and simulation of radiation should be improved for the efficient applicability and usefulness of the model, which highlights the necessity for a model adaptation method.

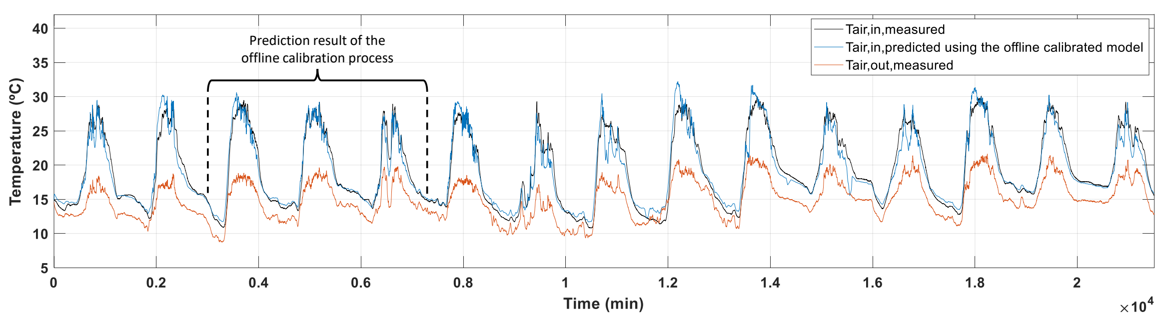

Figure 28. Validation of the offline calibrated model of internal air temperature using the complete winter-spring dataset

Table 8. Search ranges and offline clibrated parameter values

Parameters	$C_{asw,a}$	$C_{cnv,ss-a}$	$C_{cnd-cnv,a-e}$	C_A	C_{B_d}	C_{B_n}	$C_{ven,d}$	$C_{ven,wd}$	C_{loss}	$C_{tsw,cv}$
Range	[0.1, 0.9]	[1, 35]	[1, 30]	[0.2, 0.7]	[4, 26]	[4, 26]	[15, 35].10^{-4}	[0.1, 1]	[0.1, 1]	[0.1, 1]
Calibrated value	0.42	13.43	10.32	0.26	8.27	10.28	0.0016	0.11	0.2	0.56

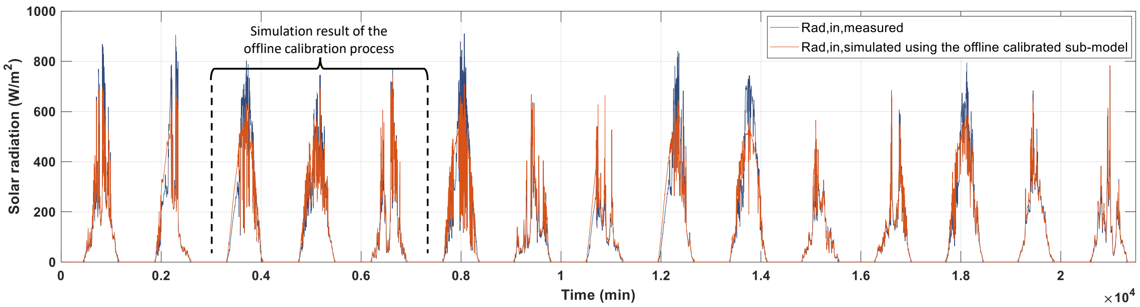

Figure 29. Validation of the offline calibrated sub-model of internal solar radiation using the winter-spring dataset

Table 9. Statistical evaluation of internal temperature prediction: calibration process result and validation of the offline calibrated model

	MAE (°C)	MSE (°C^2)	RMSE (°C)	MaxAE (°C)	Interval (°C)
Calibration in winter-spring	0.75	1.12	1.06	3.88	[10.9, 29.5]
Validation in winter-spring	0.98	1.78	1.33	6.39	[10.8, 30]
Validation in summer-autumn	1.65	6.22	2.49	8.05	[19.7, 39.3]

Table 10. Statistical evaluation of internal solar radiation simulation: calibration process result and validation of the offline calibrated sub-model

	MAE (W m^{-2})	MSE (W^2 m^{-4})	RMSE (W m^{-2})	MaxAE (W m^{-2})	Interval (W m^{-2})
Calibration in winter-spring	19.72	1736.24	41.6	271.56	[0, 890]
Validation in winter-spring	29.37	3592.86	59.94	496.81	[0, 910]
Validation in summer-autumn	54.68	7826.46	88.4	299.73	[0, 530]

B. Model sensitivity analysis

In this section, a sensitivity analysis is performed for the air temperature model. Firstly, the sensitivity of the model using constant parameters is investigated during the diurnal and nocturnal periods as shown in Fig. 30. It can be noticed in the diurnal period analysis that $C_{asw,a}$ (unitless) is the most influential parameter on the system, which is logical since it is related to solar radiation.

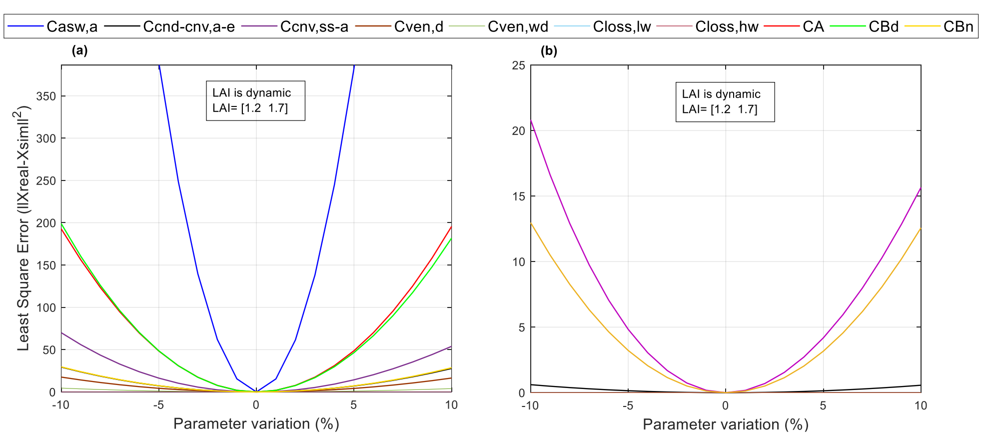

Figure 30. Sensitivity analysis of the inside air temperature model. (a) the diurnal period; (b) the nocturnal period

The transpiration parameters C_A (unitless) and C_{B_d}($kg\,m^{-2}\,h^{-1}\,kPa^{-1}$) have also noticeable relevance, which is also logical according to the fundamental effect of the crop transpiration process. Apart from the parameter $C_{cnv,ss-a}$ ($W\,m^{-2}\,K^{-1}$) related to the important effect of the soil surface temperature, the rest of the parameters mostly have a non-relevant influence. In the nocturnal period, it can be observed that only two parameters mainly affect the system: $C_{cnv,ss-a}$, which explains the role played by the soil, as a heat accumulator during the day and as a heat releaser during the night, and C_{B_n}($kg\,m^{-2}\,h^{-1}\,kPa^{-1}$), which represents the effect of crop transpiration at night.

Furthermore, the model sensitivity using time-varying parameters has been investigated. This was achieved by performing three sensitivity analysis processes with a real dataset of 1440 samples (1 day). The results are shown in Fig. 31 for three different sets of the main parameters presented in Table 11. It can be concluded that different sensitivity responses can be obtained when using time-varying parameters. Besides, it is noticed that increasing or decreasing the parameters $C_{cnv,ss-a}$ and $C_{cnd-cnv,a-e}$ ($W\,m^{-2}\,K^{-1}$) radically affects the model sensitivity, and they also alter the order of the most influential parameters. This also can be seen as a logical influence since the convection and conduction processes depend on temperature differences which are indirectly affected by all of the climate variables. In contrast, changing the values of the rest of the parameters does not affect the model sensitivity as much as $C_{cnv,ss-a}$ and $C_{cnd-cnv,a-e}$ which are considered as the most influencing time-varying parameters according to this test.

The rest of the time-varying parameters $C_{ven,d}$ (unitless), $C_{ven,wd}$ (unitless), $C_{loss,hw}$ and $C_{loss,lw}$ ($m\,s^{-1}$) representing ventilation coefficients have a less relevant influence whether on the model output or the order of the most influential parameters. It is physically logical since it is known that

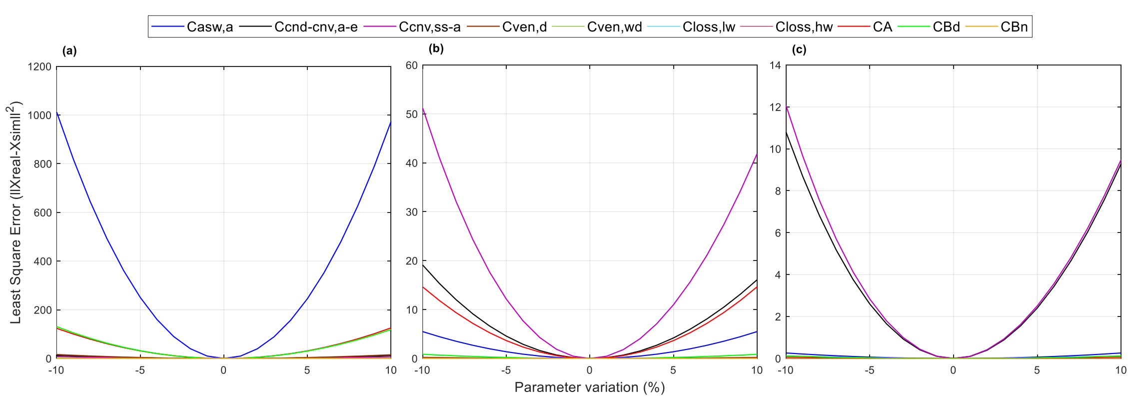

Figure 31. Sensitivity analysis of the inside air temperature model according to different sets of parameters. (a) set 1; (b) set 2; (c) set 3

Table 11. Different sets of parameters used for the model sensitivity analysis with time-varying parameters

Parameters	C_{asw}	$C_{cnv,ss-a}$	$C_{cnd-cnv,a-e}$	C_A	C_{B_d}	$C_{ven,d}$	$C_{ven,wd}$
Set 1	0.59	3.88	1	0.42	14	0.0021	0.23
Set 2	0.2	20	17	0.65	9	0.0024	0.6
Set 3	0.35	35	22	0.35	11	0.0027	0.4

the heat loss due to ventilation has less effect on the inside air temperature compared to the rest of heat fluxes, leading to conclude that the ventilation coefficients could be considered constants for simplicity. However, since the ventilation is a rapidly varying flux due to the wind effect, the ventilation parameters were considered time-varying to exploit all chances of model adaptation. The air leakage coefficients $C_{loss,lw}$ and $C_{loss,hw}$ were compensated by only one time-varying parameter C_{loss} to be online estimated representing the heat loss due to air leakage in all scenarios.

IV.2.2.2 Online parameter estimation stage

The settings of the online estimator are chosen as follows. The execution time of the parameter estimation process for each model is:

- 1 minute for the estimation of the parameters affecting the air temperature model. This time is equal to the simulation step time of the model (sample-by-sample parameter estimation).
- 20 minutes for the estimation of the parameter affecting the solar radiation sub-model. It should not be changed to a faster rate than this because it could generate an undesirable divergence in the solar radiation simulation, in turn, negatively affecting the air temperature prediction.

The number n_X of the previous time instants (last scenario) at $[t - n_X,\ t]$ to be evoked by each model is selected as:

- $n_{T_{in}} = 3$ for the air temperature model adaptation. It is found that beyond this value, undesirable fluctuations could be generated in the predicted variable, and below this value, the information would not be sufficient for an efficient model adaptation process, leading to less accurate prediction.
- $n_{V_{sr,cr}} = 60$ for the solar radiation sub-model adaptation because its parameter varies very slow in time.

The dynamic search ranges $[LB_j,\ UB_j]$ used in the RSBA are adaptively updated based on the defined variation ratio for each parameter as presented in Table 12. The search range of the time-varying

Table 12. Variation ratios for each parameter representing the adaptivity rates of the search ranges

Time-varying parameters	Variation ratios	Physical characteristics and effect on the air temperature model
$C_{asw,a}$	$\pm 2\%$	- Medium variation ratio affected by external climate and covering material. - Model sensitivity is high due to the direct effect of solar radiation on the inside air, soil surface and crop.
$C_{cnv,ss-a}$ and $C_{cnd-cnv,a-e}$	$\pm 10\%$	- Very fast variation ratio affected directly by soil surface, inside and outside air temperature differences and indirectly by radiation, ventilation and transpiration. - Model sensitivity is very high because they are the most influential parameters.
C_A and $C_{B_{d/n}}$	$\pm 0.2\%$	- Very slow variation ratio affected by crop transpiration process which follows the slow crop growth evolution (LAI). - Model sensitivity is medium, essentially affecting the inside air but with less influence than solar radiation.
$C_{ven,d}$, $C_{ven,wd}$ and C_{loss}	$\pm 7\%$	- Fast variation ratio affected by the opening of vents and wind velocity which varies quickly. - Model sensitivity is low but it has a fast effect, highly dependent on wind velocity.

parameter $C_{tsw,cv}$ in the solar radiation sub-model is not adapted and is maintained constant at [0, 1]. This allows the estimator to directly reach the lowest or the greatest values in the search range in case the shade screen or the cover whitening process are applied to or removed from the greenhouse. Moreover, when the greenhouse cover is deteriorated or becomes stained, the radiation inside the greenhouse can be correctly simulated thanks to the online estimation of $C_{tsw,cv}$ in real-time. The defined constraints to restrict each adaptive search range are presented in Table 13. They differ from those used in the offline calibration for two parameters: $C_{cnv,ss-a}$ and $C_{cnd-cnv,a-e}$. These constraints are enlarged because they are empirical parameters compensating several physical dynamics, and also because they get saturated each time a smaller upper boundary is defined during trial and error tests.

Table 13. Restrictions of the adaptive search ranges of each parameter

Parameters	$C_{asw,a}$	$C_{cnv,ss-a}$	$C_{cnd-cnv,a-e}$	C_A	C_{B_d}	C_{B_n}	$C_{ven,d}$	$C_{ven,wd}$	C_{loss}	$C_{tsw,cv}$
Range restriction	[0.1, 0.9]	[1, 100]	[1, 300]	[0.2, 0.7]	[4, 26]	[4, 26]	[15, 35].10^{-4}	[0.1, 1]	[0.1, 1]	[0.1, 1]

A. Simulation study

In this study, simulations of the online parameter estimation are conducted to assess the adaption performance under changing climatic conditions and with different crop states. For this purpose, recorded experimental datasets of two periods of different agri-seasons are used.

Fig. 32 and 33 present the graphical results after using the online parameter estimator with the internal air temperature and internal solar radiation models using the dataset of the winter-spring period. The corresponding heat fluxes evolution, the evolution of the estimated parameters and the

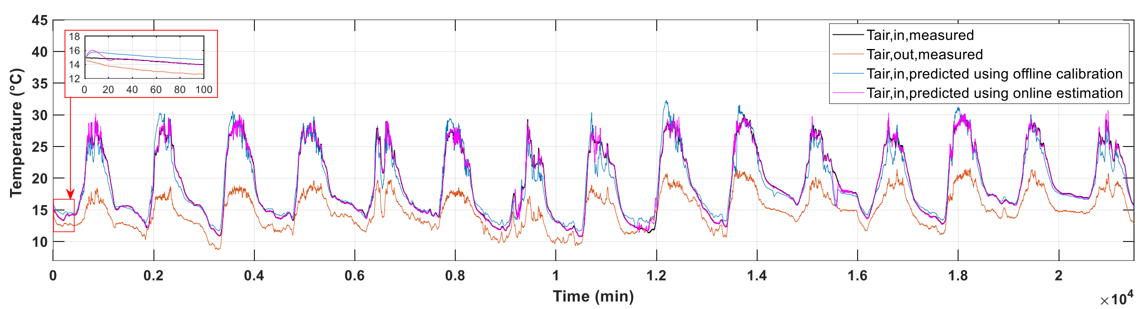

Figure 32. Internal air temperature prediction using the online parameter estimation with winter-spring dataset

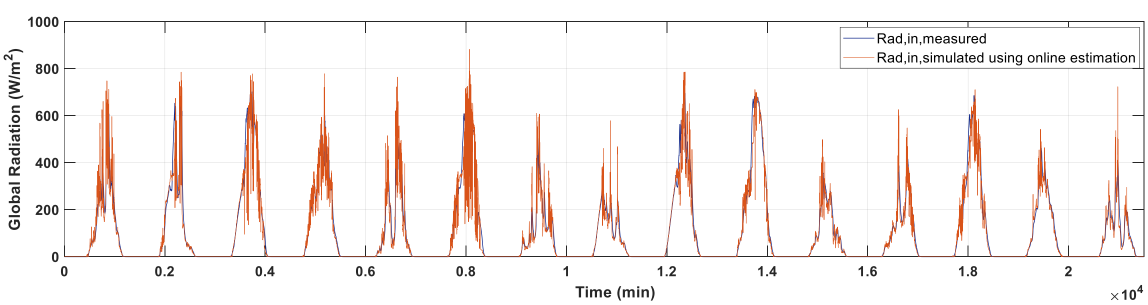

Figure 33. Internal radiation simulation using the online parameter estimation with winter-spring dataset

vents control signal are presented in Fig. 34, 35 and 36 respectively. The same test was performed with the dataset of summer-autumn and its results are presented only numerically in Table 14. In general, the graphical results show a very promising model performance based on the remarkable accurate fit between the measured and the predicted or simulated variables in comparison to the results of the offline calibrated model.

Statistical indices to evaluate the performance of both models are presented in Tables 14 and 15, and the evolution of the residual error for air temperature prediction with both datasets is shown in Fig. 37. Regarding the comparison of these results with the ones obtained with the offline calibrated model, it can be observed that the performance with the adaptive model has highly improved thanks to the online estimation of the parameters. For the air temperature model, the prediction using the online parameter estimation with the winter-spring dataset presents a MAE = 0.22 °C, meaning that the average error is decreased by 77.5%, which proves the high efficiency of the estimator, an MSE = 0.21 °C^2, R^2 = 0.992 and a MaxAE = 4.68 °C as a sporadic value, since it surpassed 4 °C only once in 15 days. The online parameter estimator succeeded very quickly in adapting the model to suit the

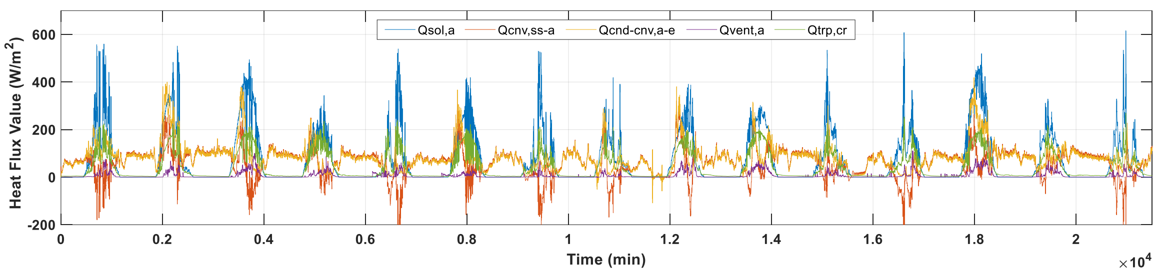

Figure 34. Heat flux evolution using online parameter estimation with the dataset of winter-spring period

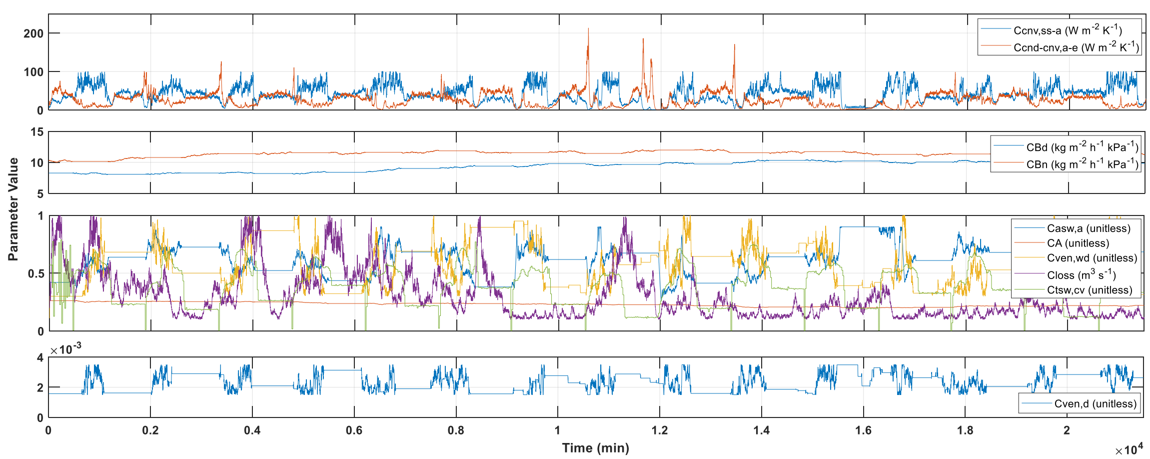

Figure 35. Variation of the online estimated parameters with the dataset of winter-spring period

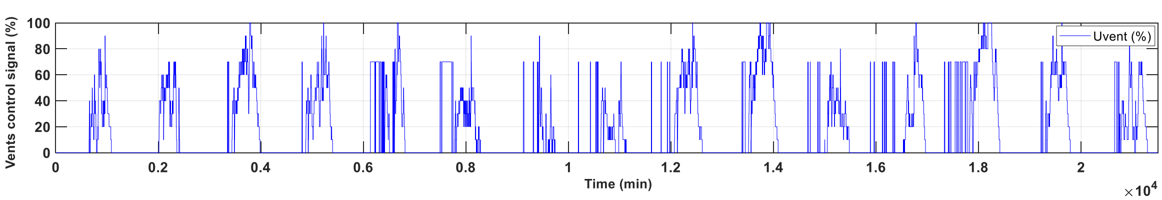

Figure 36. Vents control signal (opening percentage) from the dataset of winter-spring period

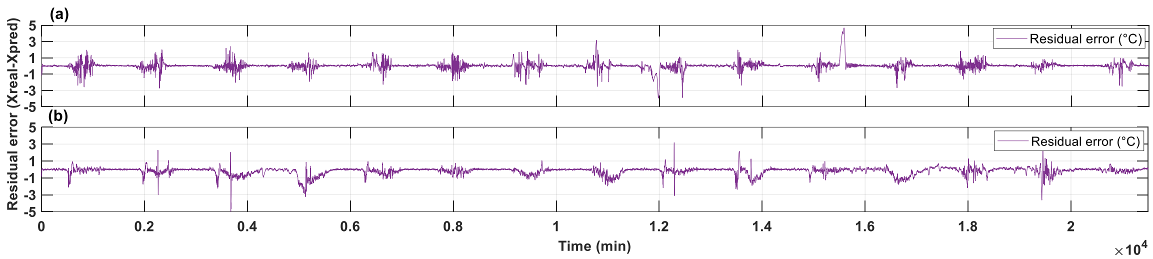

Figure 37. Evolution of the residual error of the air temperature prediction using online parameter estimation. (a) with winter-spring dataset; (b) with summer-autumn dataset.

Table 14. Statistical evaluation of internal air temperature prediction using online parameter estimation in different agri-seasons

	MAE (°C)	MSE (°C^2)	RMSE (°C)	MaxAE (°C)	Interval (°C)
Winter-spring dataset	0.22	0.21	0.46	4.68	[10.8, 30]
Summer-autumn dataset	0.27	0.23	0.48	5.57	[19.7, 39.3]

Table 15. Statistical evaluation of internal solar radiation simulation using online parameter estimation in different agri-seasons

	MAE (W m^{-2})	MSE (W^2 m^{-4})	RMSE (W m^{-2})	MaxAE (W m^{-2})	Interval (W m^{-2})
Winter-spring dataset	19.81	2065.70	45.45	467.02	[0, 910]
Summer-autumn dataset	8.15	318.7	17.85	201.05	[0, 530]

different climate conditions in less than 40 prediction steps (40 minutes) at the nocturnal period (see Fig. 32). Furthermore, the air temperature prediction using the summer-autumn dataset presents a MAE $= 0.27$ °C, meaning that the average error is decreased by 83.6%, a $MSE = 0.23$ °C^2, $R^2 = 0.995$ and a MaxAE $= 5.57$ °C as an acceptable sporadic value that does not surpass 3 °C in most of the 15 days. The residual error evolution for the summer-autumn dataset (see Fig. 37b) shows in some days a decrease in prediction accuracy compared to the residual error obtained with the winter-spring dataset. Nonetheless, it is still considered much better than the result obtained with the offline calibrated model. In this sense, this is a very promising response of the estimator, highlighting a powerful capability which is that the user might be able to avoid the offline model calibration process by directly applying the online estimator for such similar greenhouse facilities under similar climate conditions. A resembling results enhancement is observed for the simulation of the internal radiation with the online parameter estimator presenting an $R^2 = 0.97$. The statistical results present a decrease in the average error by 32.56% with the winter-spring dataset which is considered as the harshest one (especially in terms of solar radiation) and by 85.1% with the summer-autumn dataset.

Concerning the evolution of the heat fluxes with the online estimated parameters, Fig. 34 shows logical amplitudes and variations according to the modelled physical behaviour and the physical nature of each heat flux. Regarding the variation of the estimated parameters, it shows a good tendency in terms of respecting the pre-defined search range constraints (search limits) and the variation ratio of search ranges (Tables 12 and 13). It is highly interesting and important to investigate the dynamics of the estimated parameters to understand the model responses from a physical point of view. Analysing the graphs of their evolution also helped in enhancing the proposed online estimation mechanism and its specified settings and constraints. Furthermore, it helped in determining the best settings with the continuous observation of parameters' variations through trial-and-error processes.

As for the evolution of the estimated parameters in Fig 35, it can be observed that: (i) The parameters vary while respecting the restrictions defined for each parameter range. (ii) The values of the parameters respect the defined variation neighbourhood $\pm R_j\%$ according to their physical nature. (iii) The values of the parameters change only when the corresponding physical process is active, as explained in Section 1.2.2. (iv) There are no significant saturations along with the variations of most parameters which explains the appropriate selection of the constraints of the search ranges. (v) The evolutions of the convection and the thermal loss parameters $\mathrm{C_{cnv,ss-a}}$ and $\mathrm{C_{cnd-cnv,a-e}}$ show large variations up to 100 and 200 $\mathrm{W\ m^{-2}\ K^{-1}}$, respectively. It is physically logical because they are coefficients of the heat fluxes $\mathrm{Q_{cnv,ss-a}}$ and $\mathrm{Q_{cnd-cnv,a-e}}$ that are calculated based on simplified empirical equations compensating several physical dynamics which depend mainly on the state of the fluid (laminar or turbulent regime) and other variables that all vary in time (Rodríguez et al. 2015). Mathematically, those equations consist of temperature differences multiplied by the corresponding coefficients that directly affect the amplitude of the heat flux. Hence, since those heat fluxes are the most influential ones besides $\mathrm{Q_{sol,a}}$, the values of their coefficients should be large in some periods to achieve the necessary balance between the different heat fluxes according to Eq. III.31. Furthermore, it can be proven that the evolutions of both parameters show two different behaviours in the nocturnal and diurnal periods and also in the different seasons, which highlights a possibility of calculating a set of different constant values for each period in each season. However, considering them as constants is not recommended since it is proven that their evolution is fast and largely varying and because it urges on performing a laborious periodical offline model calibration demanding a minimum of a one-year dataset and leading in all cases to a loss of information and consequently, less accurate model performance as observed in the offline calibration results. (vi) The online estimation has proven that transpiration parameters could also be considered as constants for simplicity due to their slow evolution. The rest of the parameters should be time-varying to ensure the correct adaptation of the model.

Regarding the computational burden of the developed online estimator, it was found that the average time consumed by one step of the online parameter estimation process with the air temperature model is 2.03 seconds, and the average time consumed by one step of the online parameter estimation process with the radiation sub-model is 0.0038 seconds. Thus, both estimation processes are performed at the same time instant, in which the average total time consumption is 2.04 seconds, which only represents 3.4% of the total step time (60 seconds). The time consumption of the developed parameter estimator scheme is suitable for real-time application. Moreover, it leaves a sufficient time gap for the online parameter estimation of more microclimate models and the online optimisation of controllers for greenhouse control applications.

B. Experimental implementation in real time

The real-time implementation of the developed online estimator was performed on the greenhouse system in the winter season. Specifically, the estimator was tested in the period starting from 07 January 2021 to 22 January 2021 (15 days, 22000 samples). The real evolution of the climate variables registered in this period is shown in Fig. 24. The application period presented some cloudy, rainy and windy days, which are different from the usual weather in the region, and thus, they are considered as a challenging microclimate scenario to be reproduced by the model due to the strong variation of the external weather variables. Another detail to be considered as a challenge for the developed estimator is that a second polyethene cover was installed inside the greenhouse on top of the tomato crop (see Fig. 6b) to offer the plants more favourable climate conditions. The impact of this second cover was not highly relevant since it did not cover all the greenhouse surface, only the crop area but not the corridor, so, it did not create a second isolated environment inside the greenhouse. Thus, its effect from a physical point of view was assumed to cause an additional attenuation in solar radiation reaching the crop and a slight reduction in internal ventilation flux.

The estimator was executed in real time to online estimate the model time-varying parameters sample by sample to adapt it to the real changing conditions. The aim of the test performed at the real greenhouse is to investigate the capability of the online estimator in adapting the microclimate model in real time, moreover, to test its robustness considering the effect of the second cover on the internal air temperature and solar radiation.

Fig. 38 presents the graphical result of the inside air temperature prediction in real time using the developed online parameter estimator. It can be observed that the fit between the predicted and measured variables is impressive. The results are very satisfactory for all days: calm ones and even for the rainy, cloudy and windy days as highlighted in Fig. 39 and Fig 40. Table 16 presents the statistical evaluation results of the inside air temperature prediction using the online parameter estimator in real time. It presents a $\mathrm{MAE} = 0.22\ ^{\circ}\mathrm{C}$, a $\mathrm{MSE} = 0.18\ ^{\circ}\mathrm{C}^2$, $\mathrm{R}^2 = 0.992$ and a $\mathrm{MaxAE} = 3.49\ ^{\circ}\mathrm{C}$, which does not surpass $3\ ^{\circ}\mathrm{C}$ on most days. The residual error evolution for the estimation in real time shown in Fig. 45a presents a better evolution than the corresponding to the air temperature prediction obtained in Section 1.2.3.2. A with the winter-spring dataset. It can be noticed that there are some peaks in the residual error during the transition from night to day (or vice versa). It happens when the inside air temperature starts to increase (or decrease) rapidly due to the solar radiation effect. It is also related to the change in $\mathrm{C_{cnd-cnv,a-e}}$, and $\mathrm{C_{asw,a}}$ values as a response of the parameter estimator while adapting the model.

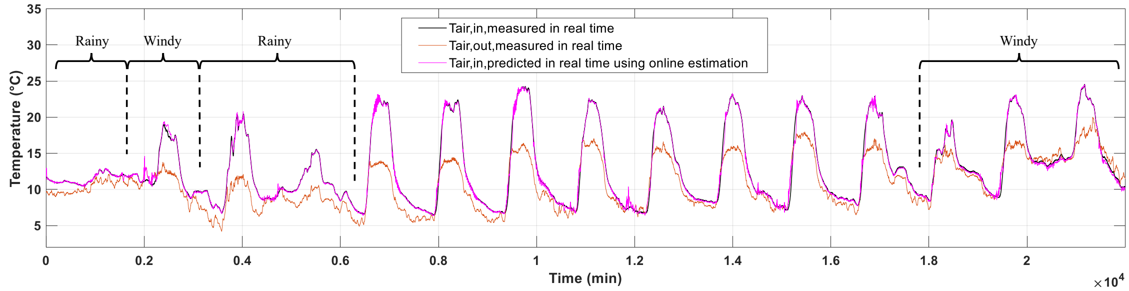

Figure 38. Internal air temperature prediction using online estimation in real time from 07 January 2021 to 22 January 2021

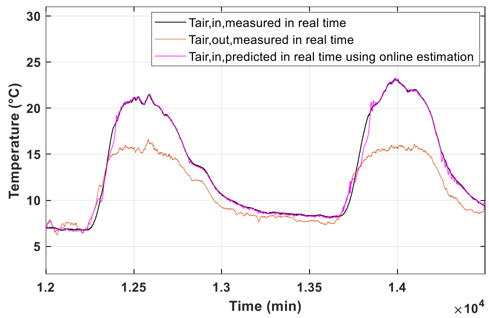

Figure 39. Air temperature prediction using online parameter estimation in real time: two days with calm climate conditions (zoomed in)

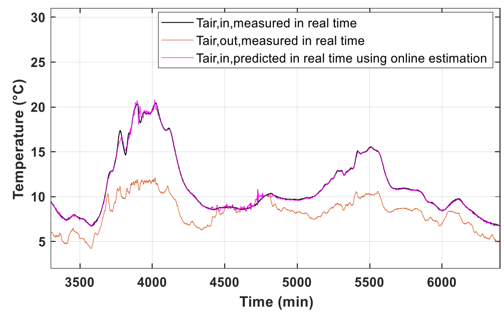

Figure 40. Air temperature prediction using online parameter estimation in real time: of two days with turbulent climate conditions (zoomed in)

Fig. 41 shows the graphical result of the inside solar radiation simulation in real time exhibiting the successful performance of the adaptive sub-model. The fit between the measured and simulated variables is very satisfactory and radiation variations are well fitted presenting an $R^2 = 0.99$. It also proves the estimator efficiency in adapting more than one model as a multi-objective task. The corresponding statistical results are presented in Table 17. The evolution of the residual error using the online estimator in real time is presented in Fig. 45b, showing a very promising result.

The real-time evolution of the heat fluxes is shown in Fig. 42. The first important aspect to be mentioned is that it was assumed previously that the physical effect of the second cover could mean more attenuation mainly on the radiation reaching the crop, and partly on the heat loss due to natural ventilation. Thus, as a confirmation of the assumption, in comparison to the evolution of the heat fluxes corresponding to the online estimation tests in the previous Section, it can be noticed that: (i) The amplitude of the solar radiation heat flux $Q_{sol,a}$ is decreased averagely by 39.3% as a main effect of the second cover. (ii) It can be graphically noticed that the amplitude of the heat loss

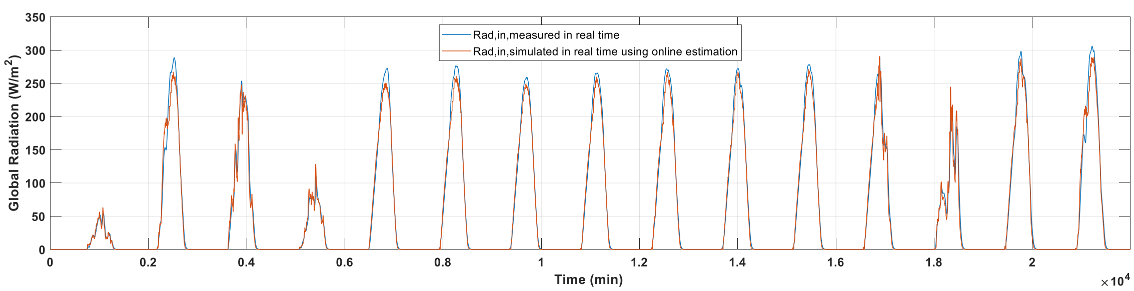

Figure 41. Internal radiation simulation using the online estimation in real time from 07 January 2021 to 22 January 2021

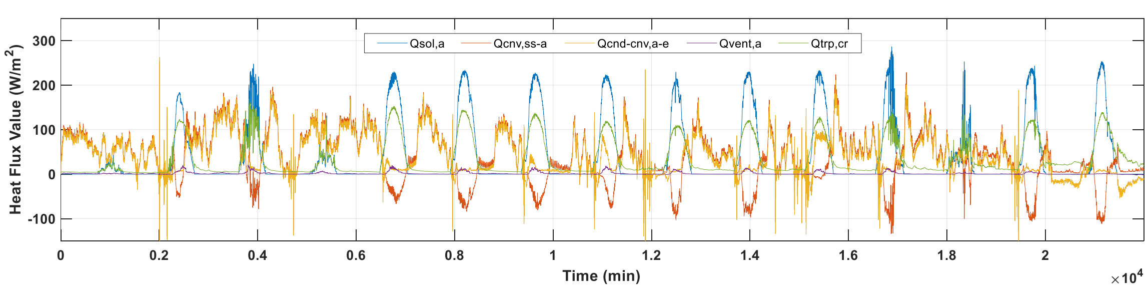

Figure 42. Heat flux evolution with parameter estimation in real time from 07 January 2021 to 22 January 2021

flux due to natural ventilation $Q_{vent,a}$ is also decreased, however, this is also dependent on wind velocity.

As for the evolution of the estimated parameters in real time in Fig. 43, it shows a remarkable resemblance with the evolution of parameters discussed in the previous section. Thus, these real-time results can be considered as a validation for the conclusions in the previous section. The vents control signal is presented in Fig. 44.

In this last stage, the real-time model adaptation was successfully achieved without changing any settings of the online estimator, neither re-programming its algorithm nor applying new mechanisms or restrictions. This proves the efficiency and robustness of the implemented online parameter estimator against the uncertainties for the adaptation of the greenhouse microclimate model. These results obtained for the real-time implementation can also be briefly analysed

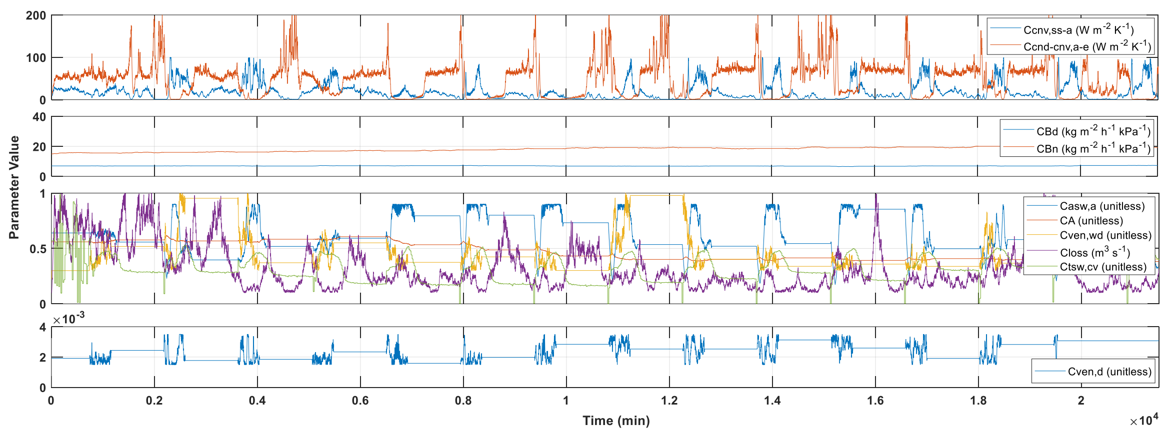

Figure 43. Variation of the estimated parameters in real time from 07 January 2021 to 22 January 2021

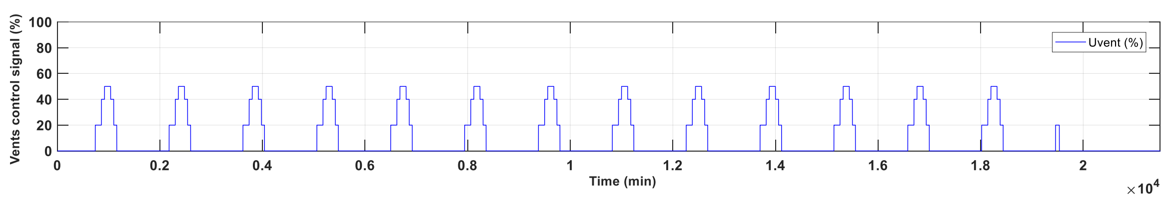

Figure 44. Vents control signal (Vents opening percentage) recorded in real time from 07 January 2021 to 22 January 2021

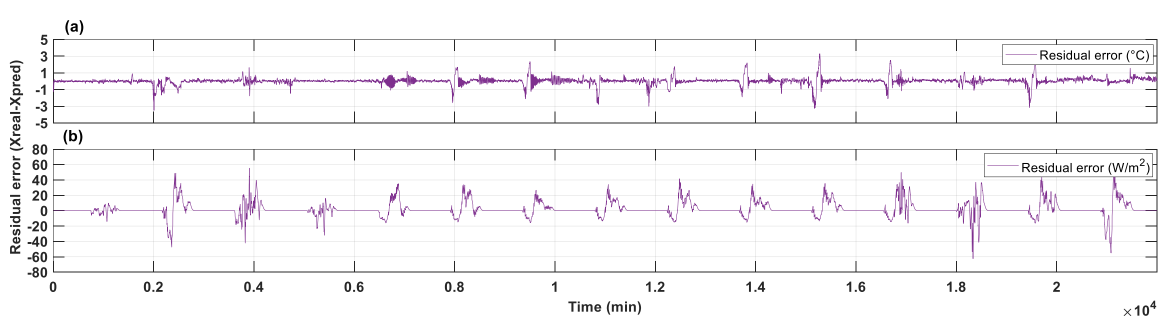

Figure 45 Evolution of the residual error with parameter estimation in real time. (a) air temperature prediction; (b) solar radiation simulation

Table 16. Statistical evaluation of internal air temperature prediction using the online parameter estimator in real time

	MAE (°C)	MSE ($°C^2$)	RMSE (°C)	MaxAE (°C)	Interval (°C)
Winter season (validation in real time)	0.22	0.18	0.43	3.49	[6.4, 24.5]

Table 17. Statistical evaluation of internal solar radiation simulation using the online parameter estimator in real time.

	MAE ($W\ m^{-2}$)	MSE ($W^2\ m^{-4}$)	RMSE ($W\ m^{-2}$)	MaxAE ($W\ m^{-2}$)	Interval ($W\ m^{-2}$)
Winter season (validation in real time)	4.62	89.56	9.46	62.47	[0, 309]

comparatively with those previously published in other works. Among the different works cited in this chapter including similar adaptation methods for first principles-based or pseudo-physical models, very satisfactory results for air temperature and relative humidity prediction were published in Pérez-González et al. (2018). It presents a $\mathrm{MSE} = 5.64\ °C^2$ using a variant of PSO algorithm and a $\mathrm{MSE} = 9.42\ °C^2$ using a variant of the DE algorithm which were considered as the best reported results for a real-time test. They were obtained in a short period of 3 days (calm and windy days) with a sampling time of one sample per one second (259200 samples). The experiments were done in an empty greenhouse (absence of crop) and without ventilation (closed vents and fans turned off). In comparison, as a substantial contribution and a distinctive aspect of this work, the evolutions of the time-varying parameters are graphically presented and thoroughly explained from a physical point of view. It is important to highlight the relevance of the achieved results since the real-time experiment have been executed under more dynamic conditions, with a grown tomato crop and with active actuators to regulate greenhouse natural ventilation. Accordingly, superior results are obtained in this chapter in terms of internal air temperature prediction with a $\mathrm{MSE} = 0.18\ °C^2$ using the developed RSBA-based online estimator in real time during 15 days (calm, windy and rainy days) with a sampling time of one sample per one minute (21500 samples).

IV.3 Virtual sensor for ventilation flux estimation

In this section of this chapter, a virtual sensor for greenhouse ventilation flux has been developed based on PCA-NARX modelling following the methodology illustrated in Fig. 46. Two datasets have been generated from a Mediterranean multi-span greenhouse located in Almería, Spain (described in Chapter II) including a combination of measured microclimate variables and the evolutions of model heat fluxes. The model heat fluxes were calculated using an adaptive air temperature model due to its capability of providing their optimal estimations (error of <5% between a test and another under the same conditions) and reliable for estimating the greenhouse ventilation flux without installing expensive sensors (described in Chapter III). All the obtained variables were firstly processed by: signal filtring, centralisation, reduction and standardisation. Secondly, the treated dataset was used to generate the PCs for data reduction using PCA. These PCs are then considered the new inputs of the neural network. Thus, the network was trained based on the PCs to fit the target which is the estimated heat loss using the opening percentage of roof and side vents based on

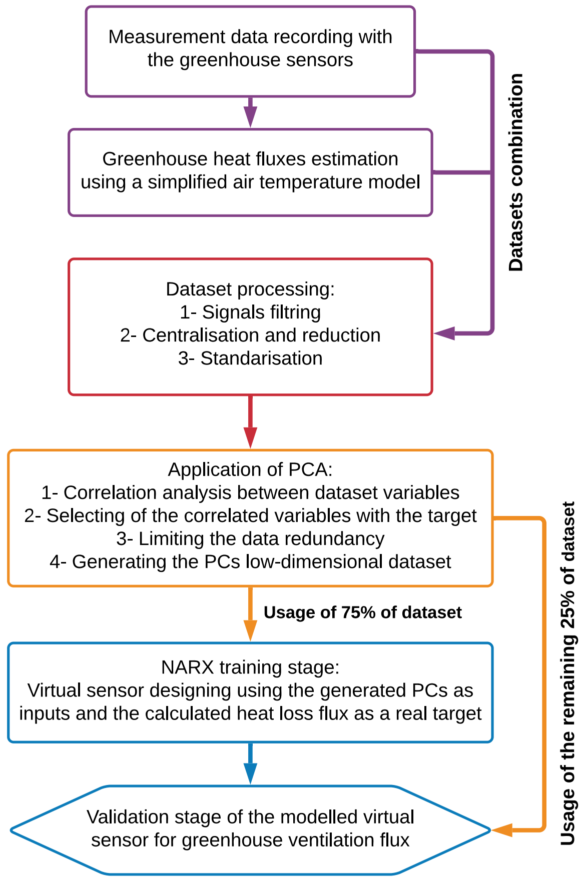

Figure 46. The methodology for developing the PCA-NARX-based virtual sensor for greenhouse ventilation flux

the pre-mentioned explicit approach (Kittas., 1997). Finally, the estimated heat loss flux was used to inversely calculate the ventilation flux representing the ultimate objective in the proposed virtual sensing method. The validation of this developed virtual sensor has shown promising preliminary results and it has to be more investigated.

IV.3.1 Materials and Methods

IV.3.1.1 Greenhouse microclimate dataset

The greenhouse microclimate dataset used in this work to design the virtual sensor of ventilation flux consists of two combined parts which are:

- An experimental dataset collected from the greenhouse containing measured variables is presented in Fig. 47. It includes: the wind velocity $D_{ws,e}$, the difference between the internal and external measured air temperature $T_{air,diff}$ and as well as relative humidity $H_{air,diff}$.

- A simulated dataset consisting of the estimated evolutions of the greenhouse heat fluxes is presented in Fig. 48. It includes: $Q_{sol,a}$, $Q_{cnv,ss-a}$, $Q_{cnd-cnv,a-e}$, $Q_{trp,cr}$ and $Q_{ven,a}$ that were estimated based on the equations of the adaptive air temperature model in which the

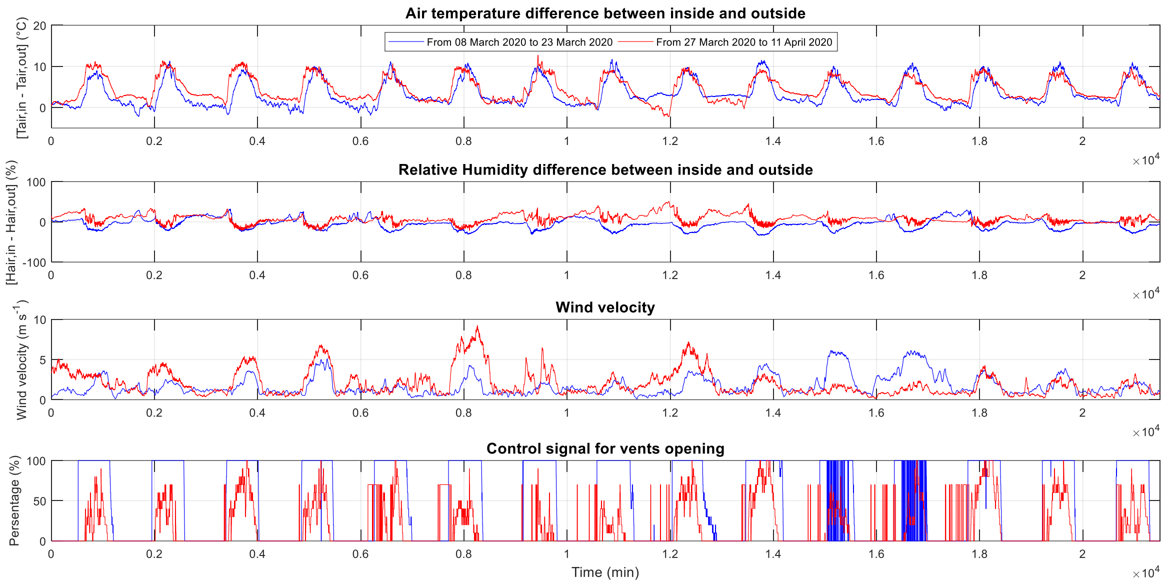

Figure 47. Dataset of the measured microclimate variables in the transitional period between winter and spring seasons

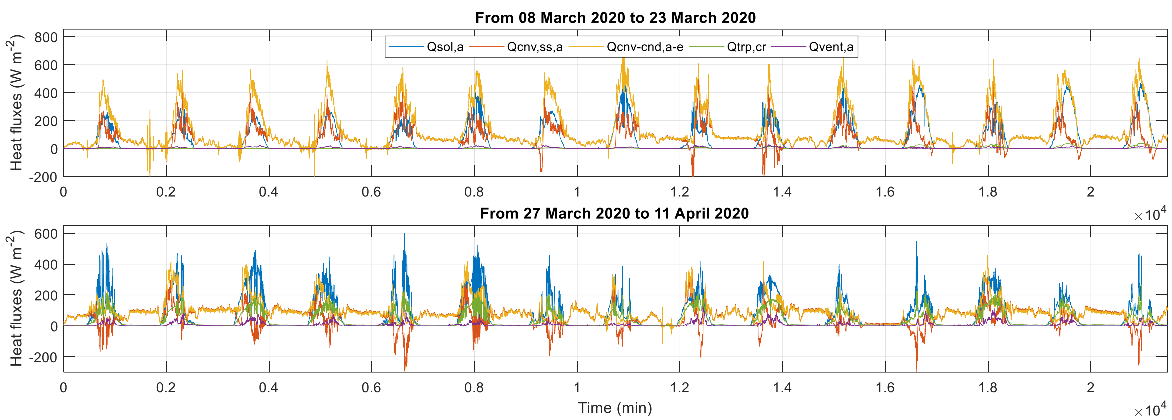

Figure 48. Dataset of the calculated greenhouse heat fluxes in the transitional period between winter and spring seasons

measured variables were supplied as inputs. Substantially, this was performed by executing the model in two scenarios:

a. Using the available vents control signal as one of the model inputs to consider the explicit effect of ventilation flux (calculated based on Eq. III.31) on the heat loss flux (calculated based on Eq. III.32), and in turn, on the model heat balance. Thus, the calculated variables of internal ventilation and heat loss fluxes are considered the real targets in designing the virtual sensor of ventilation flux.

b. Without using the vents control signal which means that the ventilation flux and the heat loss flux are not influencing explicitly (Indirect effect) heat balance of the air temperature model (using the same Eq. III.32 but with $Q_{ven,a} = 0$). Since the used air temperature model is adaptive thanks to the RSBA-online parameter estimator, the effect of the heat loss flux will be adaptively compensated by certain behaviours in the evolutions of the other heat fluxes. This has led to an accurate prediction of internal air temperature which is considered the criterion for successful estimation of heat fluxes. Thus, the estimated heat fluxes in fact hold implicitly the ventilation flux information (ventilation effect). Accordingly, the heat fluxes are considered useful data to be used as inputs for designing the virtual sensor of ventilation flux.

All variables were processed by centralisation, reduction and standardisation and some of them were filtered.

IV.3.1.2 Nonlinear autoregressive network with exogenous inputs

A discrete-time NARX model is used in this work. It consists of a recurrent dynamic network with feedback connections including numerous layers of the network (Wang et al., 2017). This model is defined by the following function:

$$y(k) = f(y(k-1), \dots, y(k - d_{output}), \mathrm{u}(k-1), \dots, u(k - d_{input})) \qquad \text{(IV.55)}$$

where $y(k)$ is the greenhouse ventilation flux estimated by the model at the discrete-time step k, u is the column vector of inputs, d_{input} and d_{output} are the orders of the past inputs and outputs, respectively, to be used for producing $y(k)$. The NARX was chosen for its advantage in relating the output to the past inputs and outputs.

IV.3.1.3 Principle components analysis

PCA is a multivariate technique that transforms a set of correlated variables into a smaller set of uncorrelated variables, called principal components (PCs). The power of PCA is more apparent for a larger number of variables. Several PCA applications in the field of greenhouse data analysis have been investigated. In (He and Ma., 2010) the PCA was used to simplify the data samples and to optimise the model learning speed for internal air humidity modelling. In (Pessel, N. and Balmat, J.F., 2008), neural networks were combined with PCA to explain mathematically the redundancy between variables in order to simplify the complex model by keeping the efficient one. In this chapter, the PCA is used to choose the relevant information from the greenhouse microclimate variables of the available dataset.

The starting point for PCA is the sample covariance matrix S. The covariance matrix for the p-variable is calculated as follows:

$$S = \frac{1}{1-n} X_c^T X_c \quad \in \ \mathbb{R} \qquad \text{(IV.56)}$$

$$S = \begin{pmatrix} s_1^2 & \cdots & s_{1p} \\ \vdots & \ddots & \vdots \\ s_{1p} & \cdots & s_p^2 \end{pmatrix} \qquad \text{(IV.57)}$$

$$X_c = X - X_{mean} \qquad \text{(IV.58)}$$

where n is the number of samples and X_c is the centred data around the mean value and it is calculated as follows:

Secondly, the obtained covariance matrix S is reduced to a diagonal matrix L by pre-multiplying it and post-multiplying it by a particular orthonormal matrix U. It is calculated as follows:

$$U^T S U = L \qquad \text{(IV.10)}$$

where s_i^2 is the variance of the ith variable and s_{ij} is the covariance between the ith and j^{th} variables.

The diagonal elements of L are called the eigenvalues of S. The cumulative variance contribution of each principal component is calculated with the next expression:

$$W_{PC}(i) = \frac{l_i}{\sum_{j=1}^{m} l_j} \qquad \text{(IV.59)}$$

The columns of U are called the eigenvectors, which represent the direction of maximal variability for each variable. These vectors are used for analysing the correlation between variables. Finally, the uncorrelated components Z are calculated as follows:

$$Z = U^T X_c \qquad \text{(IV.60)}$$

The PCA correlation circle is also used to study the relation between the inputs and the target and to reveal any possible redundancy of data and avoid it by eliminating the variables that approximately hold the same information.

IV.3.2 Results and Discussion

This section presents the results and observations obtained from each stage of the proposed methodology (see Fig. 46) and their corresponding discussion toward designing and validating a virtual sensor for greenhouse ventilation flux.

IV.3.2.1 Principal component analysis application

The application of the PCA on the available dataset was performed for two purposes as described in the following sub-section.

A. Data analysing

Initial PCs were generated based on the full dataset including the inputs (generated without considering $V_{ven,flux}$ effect) and the target (generated with considering $V_{ven,flux}$ effect). Fig. 49 shows the correlations between the PC1 and PC2 that hold the largest amount of information from the original data via coordinates in a 2-dimensional plot. Based on the fact that the correlations are proportional to the angle between vectors, it can be noticed that: (I) $Q_{ven,a}$ which represents the target has a strong positive correlation with $Q_{trp,cr}$, $D_{ws,e}$, $Q_{sol,a}$ and $T_{air,diff}$ since their vectors are grouped in the neutral quadrant; (II) $Q_{cnv-cnd-a-e}$ and $H_{air,diff}$ have much lower correlation since they are orthogonal with the target; (III) $Q_{cnv,ss-a}$ has a low but noticeable correlation with the target which is negative since it is in the opposing quadrant.

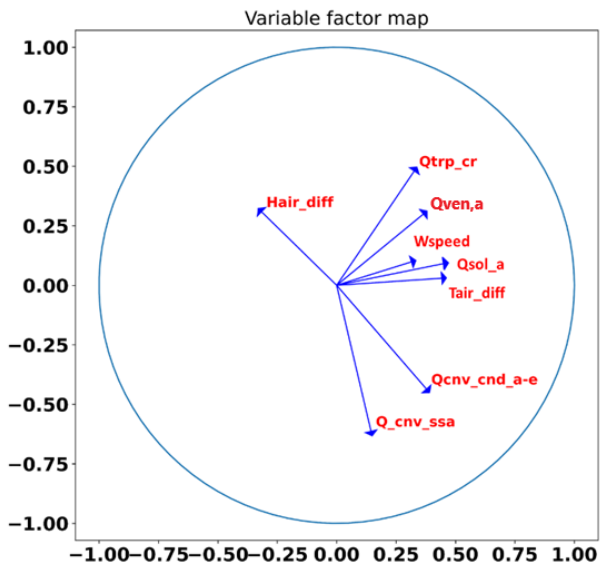

Figure 49. Correlation circle between PC1 and PC2

The relation between the inputs and the target was also analysed with the Pearson correlation coefficient using datasets of different climate conditions as illustrated in Table 18. It can be noticed that: (I) $Q_{trp,cr}$ appears to be the most correlated variable in the different calm, windy and cloudy days. This could be physically explained as a result of the direct effect of internal air temperature and the indirect effect of relative humidity through vapour pressure deficit on crop transpiration. (II) Both $T_{air,diff}$ and $H_{air,diff}$ also have high positive and negative correlations, respectively, with heat loss since ventilation flux affects them directly. (III) $Q_{cnv-cnd,a-e}$ could also be considered as one of the most correlated variables with the target based on the shown correlation values because it is calculated based on $T_{air,diff}$. (IV) $D_{ws,e}$ is known to be the main driving force of natural ventilation in greenhouses but it shows a strong correlation only on the calm day, however, it is still considered a useful input. (V) $Q_{sol,a}$ has an interesting noticeable positive correlation especially on the calm day. This could be considered an effect as a consequence of the model adaptation using the RSBA-based online parameter estimator. $Q_{sol,a}$ has prominently compensated the implicit effect of heat loss due to ventilation, because the model heat balance is influenced by $Q_{sol,a}$ more than the other heat fluxes. It can be considered as the driving flux manipulated by the online estimator according to the target (with respect to the physical constraints). (VI) $Q_{\mathrm{cnv,ss-a}}$ has the weakest correlation due to its relation with the soil surface temperature that changes slowly differently than the rapidly varying ventilation flux, thus, it could be eliminated to lower the computational cost.

Table 18. Pearson-based correlation coefficients between the input variables and the target in different climate conditions

	Climate	$D_{ws,e}$	$T_{air,diff}$	$H_{air,diff}$	$Q_{sol,a}$	$Q_{cnv,ss-a}$	$Q_{cnv-cnd,a-e}$	$Q_{trp,cr}$
$Q_{ven,a}$	Calm	0.656	0.663	- 0.627	0.736	0.069	0.476	0.809
	Windy	0.367	0.397	- 0.325	0.472	- 0.256	0.264	0.516
	Cloudy	0.381	0.040	- 0.138	0.475	0.025	0.325	0.708
	All	0.357	0.669	- 0.610	0.547	0.299	0.729	0.342

B. Generating of principal components

Four new essential PCs are generated using the full dataset combining the two datasets from the winter and spring seasons. In this case, the target $Q_{ven,a}$ has not been included because these PCs are the ones to be used as inputs of the developed PCA-NARX model. Fig. 50 shows the specific amount of information maintained by each PC and the total amount of information maintained by all the PCs is 93.8% which is considered sufficient. It can be noticed that PC1 include the largest amount of maintained information. The dimension of the dataset has been decreased from 7 to 4 variables.

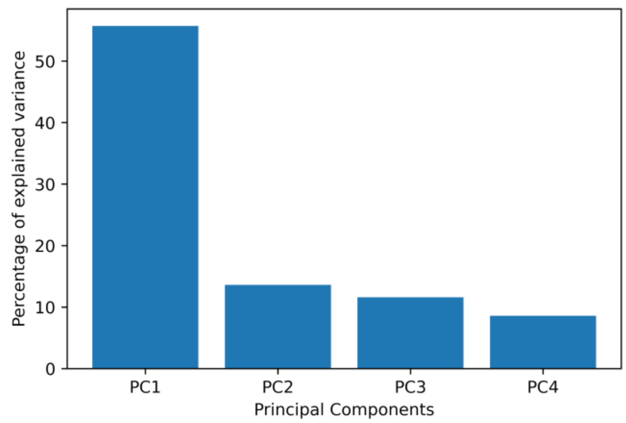

Figure 50. Amount of information maintained by the principal components constituting a low-dimensional dataset

IV.3.2.2 Virtual sensor design

The designed NARX model was trained using the generated PCs as inputs and the calculated heat loss due to ventilation as a target. Several training processes were performed using 75 % of the dataset for training and the remaining 25% for validation. The NARX model was trained using different orders of feedback output parameters using the Scaled conjugate gradient backpropagation (trainscg) as a training function and a tapped delay line with a delay from 0 to 5 samples at the input and also from 1 to 5 samples at the output. The best structure of the NARX model included: 6 layers in total and 2 hidden layers, they consist of 4, 30, 50, 20, 1 and 1 neurons, respectively, as is shown in Fig. 51. The simulation results after training are very satisfactory where

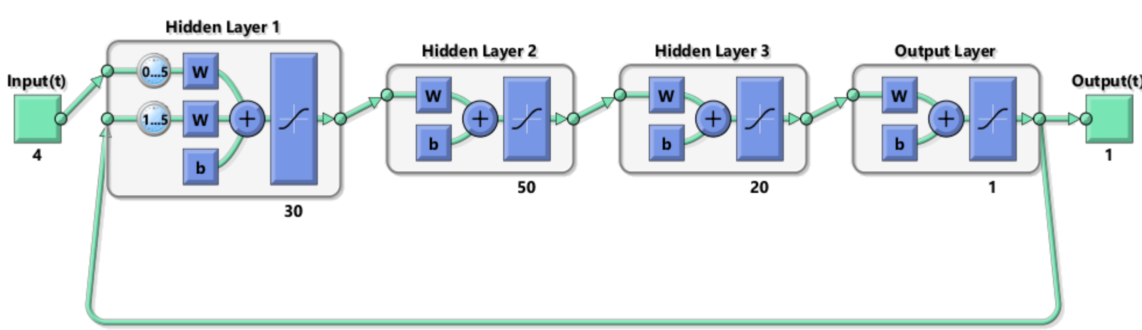

Figure 51. NARX network structure

the estimated variable follows the variations efficiently as it is qualitatively shown in Fig. 53. Settings of the NARX model design were chosen in a way that helps to avoid the overtraining of the network and enhances the generalisation in its dynamics, especially because only a short dataset was used in this preliminary study of this problem. The quantitative result of this training process present a mean absolute error $\mathrm{MAE} = 1.88\ \mathrm{W\ m^{-2}}$ which represents a percentage of $1.34\ \%$ in a variation interval of [-2.98, 137.26], which is considered sufficient although the variations of the target are not fully captured by the trained model.

IV.3.2.3 Virtual sensor validation

The best PCA-NARX model among the tested ones is selected to be the preliminary developed virtual sensor for greenhouse ventilation flux in this work. The estimation of ventilation flux was based on the calculated heat loss flux using the inverse formula of $Q_{ven,a}$ (inverse of Eq. III.32). The best PCA-NARX model was resulted after training and validating the network through a set of trial and error processes in two cases:

1- Using a dataset combining both datasets obtained from different periods in winter and spring seasons.
2- Using only one dataset of the first period recorded in the winter season.

As presented in Table 19, the preliminary quantitative results using the combined dataset show better performance than only using a separated dataset of one period presenting a $\mathrm{MAE} = 3.26\ \mathrm{W\ m^{-2}}$ for simulating the heat loss flux and leading to a $\mathrm{MAE} =\ 0.41\ \mathrm{m^3 s^{-1}}$ which are both considered very promising as preliminary results. Fig. 52 shows the qualitative results and interesting performance of $Q_{\mathrm{ven,a}}$ heat loss simulation can be observed in both diurnal and nocturnal periods and for both different datasets affected by the different vents control signals. Fig. 53 shows

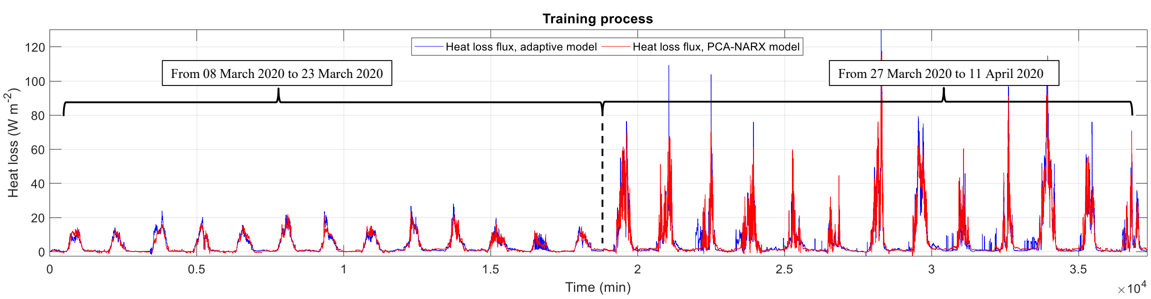

Figure 52. Estimated heat loss flux $Q_{\mathrm{ven,a}}$ as a target after training of the PCA-NARX model using the PCs as inputs

Table 19. Quantitative results of the validation process of the PCA-NARX virtual sensor

		MAE	Error persentage	Variation interval
Using the combined dataset	$Q_{\mathrm{ven,a}}\ (W\ m^{-2})$	3.26	3.93 %	[0.08, 83.11]
	$V_{ven,flux}\ (m^3 s^{-1})$	0.41	4.57 %	[0.1, 5.96]
Only one dataset: from 27 March 2020 to 11 April 2020	$Q_{\mathrm{ven,a}}\ (W\ m^{-2})$	6.27	5.47 %	[0.14, 114.75]
	$V_{ven,flux}\ (m^3 s^{-1})$	0.95	14.23 %	[0.1, 12.74]

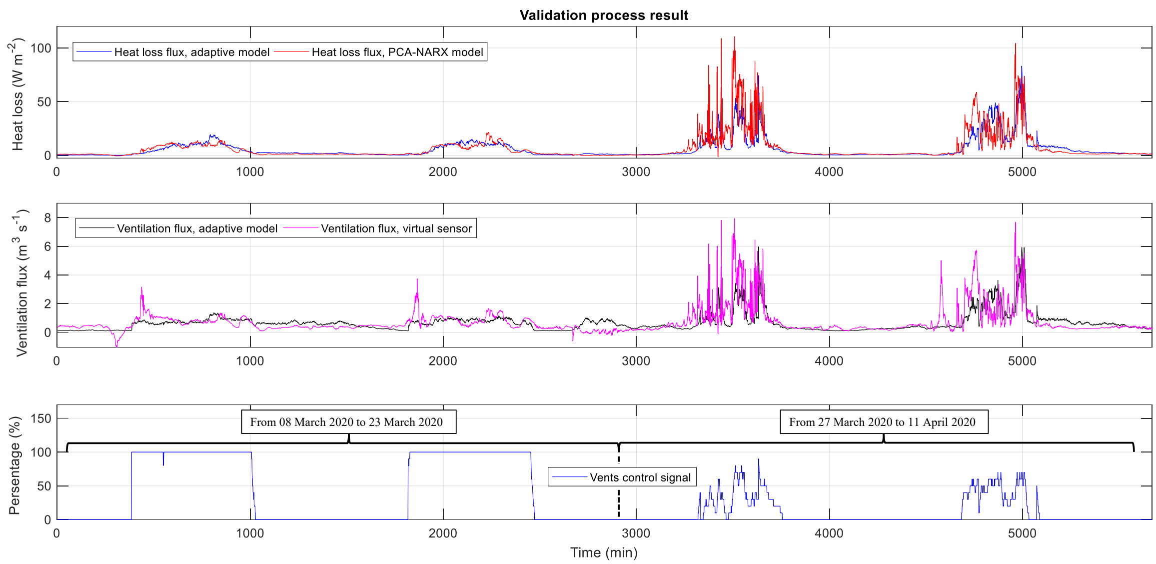

Figure 53. Validation of the virtual sensor for the estimation of ventilation flux $V_{\mathrm{vent,flux}}$ based on the PCA-NARX model

the final qualitative results of $V_{ven,flux}$ simulation as the targeted output of the virtual sensor that exhibits an interesting performance. This supports the fact that the more diverse and large the dataset used for training is, the more the dynamics of the resulting model are driven by the supplied data rather than the bias parameters added by the ANN, and the more accurate the output is in terms of capturing the variations in the target evolution. Simulating such a strongly nonlinear phenomenon with such fast dynamics requires more than the provided information in this work. However, it is still considered a preliminary result to help in leading to a better version of this solution.

IV.4 Conclusions

Greenhouse cultivation is always in need of continuous development and optimisation of automatic systems. This work proposes a methodology for online parameter estimation for greenhouse microclimate model adaptation as one of those possible greenhouse system optimisations. It is proposed as an alternative to the periodic offline calibration of the time-varying parameters, which is commonly considered a laborious procedure that consumes time and computational resources. Thus, an online parameter estimator is developed in this chapter to achieve the real-time adaptation of a greenhouse microclimate model and intends to thoroughly study the time-varying parameters, aiming for optimal prediction performance. The more accurate the model performance is, the more accurate the yield control, and the better the economic profits are, quantitatively and qualitatively. The online estimator works based on the RSBA as an enhanced variant of the nature-inspired BA algorithm. It has been developed in four phases: Firstly, an offline model calibration using a real experimental dataset was achieved. Secondly, a sensitivity analysis to investigate the influence of each parameter on the model outputs was performed, and it could be concluded that the ventilation parameters can be considered constants for simplicity. Thirdly, an online parameter estimation using real datasets of different agri-seasons was performed. The performance of the developed online parameter estimator in adapting the greenhouse microclimate model has been evaluated from both physical and statistical points of view. The evolution of the estimated time-varying parameters also proved that transpiration parameters could be considered constants for simplicity. However, the parameters of convection and conduction processes should be time-varying since they compensate for several physical dynamics, and their evolution is fast and vastly varying according to their corresponding search ranges. Research works that include a graphical illustration and a detailed discussion of the evolution of the time-varying parameters have not been encountered in literature to be compared to the results presented in this chapter.

Qualitative and quantitative evaluation results show a very satisfactory performance of the adaptive microclimate model and the online estimator in terms of:

- The accuracy in predicting the internal air temperature and simulating the internal solar radiation is due to the successful microclimate model adaptation.
- The efficiency of the online parameter estimation mechanism respects the defined constraints, in turn, the physical sense of the time-varying parameters.
- The robustness of the online estimator against:
    - The changing weather conditions (calm, cloudy, rainy and windy days)
    - The uncertainty after installing the second plastic cover.
- The limited total time consumption of every parameter estimation process allows for a future adaptation of more microclimate models and controllers in real time.
- The successful adaptation of both models supports the possibility of using the developed online estimator for the adaptation of a set of connected or interconnected microclimate models representing a greenhouse MIMO model, for example, taking into account air temperature, air humidity, solar radiation and CO_2 concentration models, among others.

Finally, the real-time implementation of the proposed online estimator was tested in an experimental greenhouse under Mediterranean climate conditions. The results exhibited an outstanding performance of the estimator in adapting the models in different agri-seasons, presenting an average error of less than $0.28\ °C$ for air temperature prediction and $20\ \mathrm{Wm}^{-2}$ for solar radiation simulation. It proves the successful adaptation methodology and the efficiency of the developed online estimator for greenhouse microclimate model adaptation.

As a future perspective, the estimation mechanism can be improved by automating the selection of the estimator settings in real-time. The methodology proposed in Section 3 could be applied to adapt other models for different microclimate variables considering the greenhouse as a MIMO system. Preliminary results show that the estimator could be applied without an initial offline calibration stage to other similar greenhouses in different locations to study the variability of the parameters and their dependence on the weather conditions, the type of greenhouse structure and the dimensions of its actuators. In addition, its application could highly be possible in other fields, to other time-delay systems and might also be possible with real-time operating systems. Furthermore, a potential application for the developed online parameter estimator is related to adaptive control and nonlinear control since these kinds of strategies usually trust the equations of a model to calculate a control signal. Therefore, the solution presented in this work could help to improve the performance of automatic control strategies for greenhouses.

In greenhouses, the absence of measurements for the opening percentage of vents (vents control signal) complicates the calculation of the ventilation flux. For this reason, a virtual sensor for greenhouse ventilation flux has been proposed as the second contribution in this chapter. It is designed based on a PCA-NARX model following the methodology illustrated in Fig. 46. A measurements dataset generated from a Mediterranean multi-span greenhouse located in Almería, Spain, has been used. A complete dataset has been prepared by combining the measured microclimate variables and the calculated heat fluxes of an adaptive air temperature model, all from different periods of winter and spring seasons. A set of PCs was generated using PCA with the dataset, leading to a significant data dimensional reduction. The resulting components are considered the new inputs of the neural network, and the calculated heat loss flux due to natural ventilation was considered a real target. Thus, the network was trained using the prepared inputs to fit the real target to estimate the ventilation flux as the primary objective.

Preliminary quantitative and qualitative results have been obtained in this work. The validation process of the developed virtual sensor has presented a $\mathrm{MAE} = 0.41\ \mathrm{m^3s^{-1}}$ which represents an error percentage of 4.57% between the estimated ventilation flux without considering the vents control signal and the calculated ventilation flux using the vents control signal. The qualitative result shows that the variations of ventilation flux are captured acceptably by the virtual sensor even though the vents control signal included different patterns.

As conclusions and future perspectives, the preliminary results obtained in this work support the fact that developing a superior and reliable virtual sensor for long-term applications calls for some essential requirements and potential procedures that could be highlighted as follows:

- A large database of at least one year or several datasets from different seasons have to be provided, consisting of the needed measured and calculated variables representing the inputs and the target for the proposed PCA-NARX model.
- More greenhouse environment dynamics (e.g., the difference between internal and external CO_2 concentration) should be analysed to obtain more correlated variables to the greenhouse ventilation flux if they exist.
- The correlation between the target and the provided measured and calculated data has to be more investigated from different perspectives using other data analysis techniques.
- Different structures of the NARX model and other data-driven modelling methods could be investigated.

Based on the fact that the correlation between the inputs and the target can change depending on the climate conditions (calm, windy and cloudy days, etc.), a classification technique and multiple data-driven models can be obtained for each climate scenario for more accurate performance.

CHAPTER V

CONCLUSION AND FUTURE PERSPECTIVES

V.1 Conclusions

Following the defined objectives in the primary work plan, a set of contributions have been successfully achieved in this thesis. They are briefly described as follows:

1. A gable-shaped small-scale greenhouse prototype was constructed to be used as a nursery in an arid region, Meziraa, Biskra, Algeria. It included designing a low-cost microcontroller-based data acquisition system for the wireless monitoring of the prototype. They have been used to acquire a modest dataset including climate variables from inside and outside environments of the greenhouse prototype under moderate desert climate conditions. This project was chosen as the best project in the exhibition of the International Symposium of Technology and Sustainable Industry Development 2019 (ISTSID), El-Oued, Biskra, Algeria. The successfully acquired dataset was used in the investigations of two international conferences.
2. A nonlinear grey-box model of greenhouse air temperature is proposed. It describes the inside air temperature as a set of heat exchange processes generated by the differences in energy content between the inside and outside air. This model was derived by reformulating a physical-based model. The reformulation includes having new static parameters linearly dependent that have to be identified based on offline parameter estimation using the acquired dataset from the greenhouse prototype. A less complicated model has been derived for greenhouse air temperature prediction. The lack of information on the system parameters is considered a kind of uncertainty in greenhouse models. For this issue, the commonly known metaheuristic bio-inspired algorithm called the *Random inertia weight particle swarm optimisation algorithm* (RIWPSO) was chosen to be implemented in the proposed model for model calibration as an offline parameter estimation method.
3. An enhanced variant of the *Bat Algorithm* (BA) is proposed in this thesis and also implemented in the proposed grey-box temperature model. It is called the *Random scaling-based bat algorithm* (RSBA). It was used to identify the unknown values of the proposed static parameters of the model through an offline estimation process. The RSBA was proven to have superior performance to the standard BA in terms of accuracy and speed of convergence. Finally, another comparative study has been carried out on the performance of RIWPSO and RSBA in identifying the parameters of the grey-box model and the prediction accuracy of the different obtained greenhouse air temperature models. The results have shown the superiority of the RIWPSO over the RSBA in solving the problem at hand. However, the RSBA could still be more useful against other different problems, such as the online parameter estimation in real-time, where the advantage of the early convergence to optimality can be necessary due to time constraints.
4. A methodology for online parameter estimation is proposed for the adaptation of a greenhouse climate model. It is proposed as an alternative to the laborious periodical offline calibration of the time-varying parameters, which is commonly considered a laborious

procedure that consumes time and computational resources. Specifically, an online parameter estimator is developed to achieve the real-time adaptation of a greenhouse microclimate model and intends to thoroughly study the time-varying parameters, aiming for optimal prediction performance. The online estimator works based on the RSBA as an enhanced variant of the nature-inspired BA algorithm. The performance of the developed online parameter estimator in adapting the greenhouse microclimate model has been evaluated from both physical and statistical points of view. The evolution of the estimated time-varying parameters has proven that transpiration parameters could be considered constants for simplicity. However, the parameters of convection and conduction processes should be time-varying. Qualitative and quantitative evaluation results have shown a very satisfactory performance of the adaptive microclimate model and the online estimator in terms of:

- The very low prediction error, which is less than $0.28\,°C$ for air temperature prediction and $20\ Wm^{-2}$ for solar radiation simulation.
- The efficiency of the online parameter estimation mechanism respects the defined constraints, in turn, respects the physical sense of the time-varying parameters.
- The robustness of the online estimator against:
- The changing weather conditions (calm, cloudy, rainy and windy days)
- The uncertainty after installing the second plastic cover.
- The limited total time consumption of every parameter estimation process allows for a future adaptation of more microclimate models and controllers in real time.
- The successful adaptation of both models supports the possibility of using the developed online estimator to adapt a set of connected or interconnected microclimate models representing a greenhouse MIMO model, for example, considering air temperature and air humidity, solar radiation and CO_2 concentration models, among others.

Research works that include a graphical illustration and a detailed discussion of the evolution of the time-varying parameters have not been encountered in literature to be compared to the results presented in this paper. The more accurate the model performance, the more accurate the yield control and the better the economic profits quantitatively and qualitatively.

5. A virtual sensor for greenhouse ventilation flux has been developed based on PCA-NARX modelling following the methodology explained in detail in the folds of this chapter. A dataset has been generated from a Mediterranean multi-span greenhouse located at “Las Palmerillas” Experimental Station, which is a property of the Cajamar Foundation (36.79316 latitudes, -2.72014 longitude). The dataset includes a combination of measured microclimate variables and the evolutions of greenhouse heat fluxes. The heat fluxes were estimated using an adaptive air temperature model due to its capability of providing their optimal estimations

with an error of <5% between a set of the same tests and its reliability in estimating the greenhouse ventilation flux without installing expensive sensors. All the obtained variables were first processed by: signal filtring, centralisation, reduction and standardisation. Secondly, the treated dataset was used to generate the PCs for data reduction using PCA. These PCs are then considered the new inputs of the neural network. Thus, the network was trained based on the PCs to fit the target, which is the estimated heat loss using the opening percentage of the roof and side vents based on the previously mentioned explicit approach. Finally, the estimated heat loss flux was used to inversely calculate the ventilation flux representing the ultimate objective of the proposed virtual sensing method. The validation of this developed virtual sensor has shown promising preliminary results.

V.2 Future perspectives

The proposed grey-box modelling methodology can be used for energy balance studies of greenhouses. It could also be adopted as a tool to study different climate conditions or for practical applications of control systems. This study paves the way for future investigations on applying a cooling system to overcome the harsh summer climate of arid regions. The proposal of the RSBA paves the way for future investigation on developing a novel mechanism that will make the scaling parameter changes adaptatively with the closeness to finding the optimal solution.

Concerning the development of the RSBA-based online parameter estimator, the estimation mechanism can be improved by automating the selection of the estimator settings in real-time. The methodology proposed in Section 3 could be applied to adapt other models for different microclimate variables considering the greenhouse as a MIMO system. Preliminary results show that the estimator could be applied without an initial offline calibration stage to other similar greenhouses in different locations to study the variability of the parameters and their dependence on the weather conditions, the type of greenhouse structure and the dimensions of its actuators. In addition, its application could highly be possible in other fields, to other time-delay systems and might also be possible with real-time operating systems. Furthermore, a potential application for the developed online parameter estimator is related to adaptive control and nonlinear control since these kinds of strategies usually trust the equations of a model to calculate a control signal. Therefore, the solution presented in this work could help to improve the performance of automatic control strategies for greenhouses.

Regarding the developed ventilation flux virtual sensor, the preliminary obtained results support the fact that developing a superior and reliable virtual sensor for long-term applications calls for some essential requirements and potential procedures that could be highlighted as follows:

- A large dataset of at least one year or several datasets from different seasons has to be provided, consisting of the needed measured and calculated variables representing the inputs and the target for the proposed PCA-NARX model.

- More greenhouse environment dynamics (e.g., the difference between internal and external CO_2 concentration) should be analysed to obtain more correlated variables to the greenhouse ventilation flux if they exist.
- The correlation between the target and the provided measured, and calculated data has to be more investigated from different perspectives using other data analysis techniques.
- Different structures of the NARX model and other data-driven modelling methods could be investigated.

Based on the fact that the correlation between the inputs and the target can change depending on the climate conditions (calm, windy and cloudy days, etc), a classification technique and multiple data-driven models can be obtained for each climate scenario for more accurate performance.

It could be very interesting to model other climate variables like humidity, CO2, and even crop growth variables. Concerning the application of control techniques, the developed adaptive model in this thesis can be of very efficient and reliable use as a greenhouse digital twin to test many control strategies and methods.

References

Ali, R.B., Bouadila, S. and Mami, A, 2018. Development of a Fuzzy Logic Controller applied to an agricultural greenhouse experimentally validated. *Applied Thermal Engineering,* Volume 141, pp. 798-810.

Atia, D. a. E.-m. H., 2017. Analysis and design of greenhouse temperature control using adaptive neuro-fuzzy inference system. *Journal of Electrical Systems and Information Technology,* 4(1), pp. 34-48.

Boulard, T., Baille, A., 1995. Modelling of air exchange rate in a greenhouse equipped with continuous roof vents. *Journal of Agricultural Engineering Research,* 61(1), pp. 37-47.

Čekon, M., Slávik, R., Juráš, P, 2016. Obtainable Method of Measuring the Solar Radiant Flux Based on Silicone Photodiode Element. *Applied Mechanics and Materials,* October , Volume 824, pp. 477-484.

Choab, N., Allouhi, A., El Maakoul, A., Kousksou, T., Saadeddine, S. and Jamil, A., 2019. Review on greenhouse microclimate and application: Design parameters, thermal modeling and simulation, climate controlling technologies. *Solar Energy,* Volume 191, pp. 109-13.

Cunha, J. B., Couto, C., and Ruano, A. E. B., , 1997 . Real-time parameter estimation of dynamic temperature models for greenhouse environmental control. *Control Engineering Practice,* Volume 5, pp. 1473-1481.

Fourati, F., 2014. Multiple neural control of a greenhouse. *Neurocomputing,* Volume 139, pp. 138-144.

Frausto, H.U., Pieters, J.G., Deltour, J.M., 2003. Modelling greenhouse temperature by means of auto regressive models. *Biosystems Engineering,* 84(2), pp. 147-157.

Ghosal, M.K., Tiwari, G.N., Das, D.K. and Pandey, K.P, 2005. Modeling and comparative thermal performance of ground air collector and earth air heat exchanger for heating of greenhouse. *Energy and buildings,* 6(37), pp. 613-621.

Guesbaya, M., García-Mañas, F., Megherbi, H., Rodríguez, F, 2021. Real-time adaptation of a greenhouse microclimate model using an online parameter estimator based on a bat algorithm variant. *Computers and electronics in agriculture ,* Volume 192.

Guesbaya, M., García-Mañas, F., Rodríguez, F., Megherbi, H., Ouamane, M, R.,, 2021. *Virtual sensor for ventilation flux estimation in greenhouses.* Valedeloualid, Spain , XI CONGRESO IBÉRICO DE AGROINGENIERÍA.

Guesbaya, M., Megherbi, H. and Megherbi, A.C., 2019. *Random Scaling-Based Bat Algorithm for Greenhouse Thermal Model Identification and Experimental Validation.* Singapore, Springer, pp. 47-62.

Guesbaya, M., Megherbi, H, 2019. *Thermal modeling and prediction of soilless greenhouse in arid region based on particle swarm optimization: experimentally validated.* Algiers, Algeria, s.n.

Hasni, A., Taibi, R., Draoui, B., Boulard, T, 2011. Optimization of Greenhouse Climate Model Parameters Using Particle Swarm Optimization and Genetic Algorithms. *Energy Procedia,* Issue 6, pp. 371-380.

He, F. and Ma, C.,, 2010. Modeling greenhouse air humidity by means of artificial neural network and principal component analysis. *Computers and Electronics in Agriculture,* Volume 71, pp. 19-23.

Holden, Z. A., Klene, A. E., Keefe, R. F., Moisen, G. G, 2013. Design and evaluation of an inexpensive radiation shield for monitoring surface air temperatures. *Agricultural and Forest Meteorology,* October, Volume 180, pp. 281-286.

Horsey, J, 2016. DIY Arduino Wind Speed Meter Anemometer Project.

Hoyo, A., Moreno, J.C., Guzman, J.L. and Rodríguez, F., 2019. Robust QFT-based feedback linearization controller of the greenhouse diurnal temperature using natural ventilation. *IEEE Access,* Volume 7, pp. 64148-64161.

Huang, X., Zeng, X., Han. R, 2017. Dynamic inertia weight binary bat algorithm with neighborhood search. *Computational intelligence and neuroscience.*

Iga, J.L., Iga, J.L., Iga, C.L. and Flores, R.A, 2008. Effect of air density variations on greenhouse temperature model. *Mathematical and computer modelling,* 47(9-10), pp. 855-867.

Imran, M., Hashim, R. and Abd Khalid, N.E, 2013. An overview of particle swarm optimization variants. *Procedia Engineering,* Volume 53, pp. 491-496.

Khan, K. a. S. A., 2012. A comparison of BA, GA, PSO, BP and LM for training feed forward neural networks in e-learning context. *International Journal of Intelligent Systems and Applications,* 4(7), p. 23.

Kherif, F., Latypova,, 2019. Principal component analysis. In: *Machine Learning: Methods and Applications to Brain Disorders.* s.l.:Academic Press, pp. 209-225.

Kittas, C., Boulard, T., Papadakis, G., 1997. Natural ventilation of a greenhouse with ridge and side openings: sensitivity to temperature and wind effects. *Transactions of the ASAE,* 40(2), pp. 415-425.

Laktionov, I.S., Vovna, O.V., Kabanets, M.M., Derzhevetska, M.A. and Zori, A.A., 2020. Mathematical model of measuring monitoring and temperature control of growing vegetables in greenhouses. *International Journal of Design and Nature and Ecodynamics,* 15(3), pp. 325-336.

Lamrani, M.A., Boulard, T., Roy, J.C. and Jaffrin, A, 2001. SE—Structures and Environment: AirFlows and Temperature Patterns induced in a Confined Greenhouse. *Journal of Agricultural Engineering Research,* 1(78), pp. 75-88.

Li, K., Sha, Z., Xue, W., Chen, X., Mao, H. and Tan, G, 2020. A fast modeling and optimization scheme for greenhouse environmental system using proper orthogonal decomposition and multi-objective genetic algorithm. *Computers and Electronics in Agriculture,* Volume 168, p. 105096.

Lyu, S., Li, Z., Huang, Y., Wang, J., Hu, J, 2019. Improved self-adaptive bat algorithm with step-control and mutation mechanisms. *Journal of computational science,* Volume 30, pp. 65-78.

Ma, D. Carpenter, N. Maki, H. Rehman, T. U. Tuinstra, M. R. Jin, J, 2019. Greenhouse environment modeling and simulation for microclimate control. *Computers and Electronics in Agriculture,* Volume 162, pp. 134-142.

Montoya-Ríos, A.P.; García-Mañas, F.; Guzmán, J.L.; Rodríguez, F. , 2020. Simple tuning rules for feedforward compensators applied to greenhouse daytime temperature control using natural ventilation. *Agronomy,* Volume 1327, p. 10.

Pérez-González, A., Begovich-Mendoza, O. and Ruiz-León, J., 2018. Modeling of a greenhouse prototype using PSO and differential evolution algorithms based on a real-time LabView™ application. *Applied Soft Computing,* Volume 62, pp. 86-100.

Pessel, N., Balmat, J.F., 2008. Principal component analysis for greenhouse modelling. *WSEAS Transactions on Systems,* 1(7), pp. 24-30.

Ramli, M. R., Abas, Z. A., Desa, M. I., Abidin, Z. Z., Alazzam, M. B, 2018. Enhanced convergence of Bat Algorithm based on dimensional and inertia weight factor.. *Journal of King Saud University-Computer and Information Sciences..*

RodrÃguez, F. and Arahal, M.R., 2012. Virtual Sensors for Designing Irrigation Controllers in Greenhouses. *Sensors,* 12(11).

Rodríguez, F., Berenguel, M., Luis Guzmán, J., Ramírez-Arias. R., 2015. *Modeling and control of greenhouse crop growth.* s.l.:Springer International Publishing Switzerland.

Sanchez, J.A., Ming, L., Rodriguez, F., Guzman, J.L., Hui, W. and Xinting, Y., 2017. Development and test verification of air temperature model for Chinese solar and Spainish Almeria-type greenhouses. *International Journal of Agricultural and Biological Engineering,* 10(4), pp. 66-76.

Sánchez, J.A., Rodríguez, F., Guzmán, J.L. and Arahal, M.R, 2012. Virtual sensors for designing irrigation controllers in greenhouses. *Sensors,* 12(11), pp. 15244-15266.

Snyder, R. L, 2005. Humidity Conversion.

Speetjens, S. L., Stigter, J. D. and Van Straten, G.,, 2009. Towards an adaptive model for greenhouse control. *Computers and electronics in agriculture,* Volume 67, pp. 1-8.

Taki, M., Mehdizadeh, S.A., Rohani, A., Rahnama, M. and Rahmati-Joneidabad, M, 2018. Applied machine learning in greenhouse simulation; new application and analysis. *Information processing in agriculture,* 5()(2), pp. 253-268.

Van Straten, G., van Willigenburg, G., van Henten, E. and van Ooteghem, R, 2010. *Optimal control of greenhouse cultivation.* s.l.:CRC press.

Vanthoor, B.H.E., Van Henten, E.J., Stanghellini, C. and De Visser, P.H.B, 2011. A methodology for model-based greenhouse design: Part 3, sensitivity analysis of a combined greenhouse climate-crop yield model. *Biosystems Engineering*, 110(4), pp. 396-412.

Wang, H., Sánchez-Molina, J.A., Li, M., Berenguel, M., Yang, X.T. and Bienvenido, J.F., 2017. Leaf area index estimation for a greenhouse transpiration model using external climate conditions based on genetics algorithms, back-propagation neural networks and nonlinear autoregressive exogenous models. *Agricultural Water Management,* Volume 183, pp. 107-115.

Yang, X. S, 2014. *Nature-Inspired Optimization Algorithms.* 1 ed. s.l.:Elsevier.

Yang, X.S and He, X, 2013. Bat algorithm: literature review and applications. *International Journal of Bio-inspired computation,* 5(3), pp. 141-149.

Yilmaz, S., Kucuksille, E. U, 2013. *Improved bat algorithm (IBA) on continuous optimization problems.* s.l., s.n.

Yu, H., Chen, Y., Hassan, S.G. and Li, D., 2016. Prediction of the temperature in a Chinese solar greenhouse based on LSSVM optimized by improved PSO. *Computers and Electronics in Agriculture,* Volume 122, pp. 94-102.