Global approach test improvement using a neural network model identification to characterise solar combisystem performances

Antoine Leconte a,b,⇑, Gilbert Achard a, Philippe Papillon b

a LOCIE, CNRS FRE3220, Université de Savoie, Polytech'Annecy-Chambery, 73376 Le Bourget du Lac, France
b CEA LITEN INES, BP 332, 50 avenue du Lac Léman, 73377 Le Bourget du Lac, France

Received 7 April 2011; received in revised form 9 March 2012; accepted 8 April 2012
Available online 4 May 2012

Communicated by: Associate Editor C. Estrada-Gasca

Abstract

Solar CombiSystems (SCSs) are very efficient systems for reducing conventional energy consumption of building but their thermal performances are strongly dependent on the environment where they are installed (type of climate and thermal quality of the building). Currently it is impossible to predict the energy savings generated by a SCS as there is no standard test to characterise SCS performances. Currently, the Short Cycle System Performance Test (SCSPT), based on a 12 days test of the complete SCS on a semi-virtual test bench, is able to predict annual energy savings with a good accuracy, but the performance prediction is limited to only one environment (the building and the climate corresponding with the test).

Based on the SCSPT procedure, this paper proposes an improvement of the method by identifying a global SCS model from the test data. Then, the identified model would be able to simulate the tested SCS in any environment and thus to characterise its performances.

The proposed model to identify is a “grey box” model, mixing a “White Box” model composed of known physical equations and a “Black Box” model, which is an Artificial Neural Network (ANN). A complete process is developed to train and select a relevant global SCS model from such a test.

This approach has been validated through numerical simulations of three detailed SCS models. Compared to those annual results, “Grey Box” SCS models trained from a twelve days sequence are able to predict energy consumption with a good accuracy for 27 different environments. An experimental application of this procedure has been used to characterise a real system.

© 2012 Elsevier Ltd. All rights reserved.

Keywords: Neural network; Solar combisystem; Performance prediction; Test bench; Characterisation

1. Introduction

Solar combisystems are complex solar thermal systems that provide energy for Domestic Hot Water (DHW) preparation and space heating. Solar energy and auxiliary energy are managed in such a way that the thermal loads are covered to fulfil the comfort requirements of the user, but also to save as much as possible auxiliary energy. Each combisystem has its own feature concerning, for example, controller algorithm, hydraulic loops and storage design.

Nowadays, some combisystems can be very efficient and cover up to nearly 50% of the total heat demand with solar energy gains (Letz, 2006). However, such good results are met in very special cases because combisystem performances are very sensitive to climatic conditions and energetic quality of buildings. Besides, it is not unusual to see even poor design or installation mistakes making a

Abbreviations: ANN, Artificial Neural Network; CTSS, Component Testing – System Simulation; DHW, Domestic Hot Water; FSC, Fractional Solar Consumption; SCS, Solar CombiSystem; SCSPT, Short Cycle System Performance Test; SFH, Single Family House.

⇑ Corresponding author at: INES RDI, 50 avenue du Lac Léman, 73377 Le Bourget du Lac, France. Tel.: +33 4 79 44 46 69; fax: +33 4 79 62 13 74.
E-mail address: antoine.leconte@cea.fr (A. Leconte).
Nomenclature

Variables and parameters in the collector model

\( A_{\text{coll}} \) collector area (m\(^2\))
\( c_1 \) heat loss coefficient at \( (T_m - T_a) = 0 \) (W m\(^{-2}\) K\(^{-1}\))
\( c_2 \) temperature dependence of the heat loss coefficient (W m\(^{-2}\) K\(^{-2}\))
\( c_3 \) wind speed dependence of the heat loss coefficient (J m\(^{-3}\) K\(^{-1}\))
\( c_4 \) sky temperature dependence of the heat loss coefficient (–)
\( c_5 \) effective thermal capacity (J m\(^{-2}\) K\(^{-1}\))
\( c_6 \) wind dependence in the zero loss efficiency (s m\(^{-1}\))
\( E_l \) longwave irradiance \( (\lambda > 3 \mu \text{m}) \) (W m\(^{-2}\))
\( F \) collector efficiency factor (–)
\( G \) total solar irradiance on the collector plane (W m\(^{-2}\))
\( G_b \) beam solar irradiance on the collector plane (W m\(^{-2}\))
\( G_d \) diffuse solar irradiance on the collector plane (W m\(^{-2}\))
\( K_{ib} \) incidence angle modifier (–)
\( K_{id} \) incidence angle modifier for diffuse radiation (–)
\( \dot{Q}_{\text{coll}} \) heat flow rate supplied by the collector (W)
\( \dot{Q}_{\text{sol.net}} \) net solar energy available at the bounds of the collector (W m\(^{-2}\))
\( \dot{Q}_{\text{sol.ref}} \) reference solar irradiation (W)
\( T_a \) ambient temperature \((^\circ \text{C})\)
\( T_{\text{coll}} \) mean temperature of the collector \((^\circ \text{C})\)
\( v \) collector surrounding air speed (ms\(^{-1}\))
\( \theta \) angle of incidence (°)
\( \sigma \) Stephan–Boltzman constant (W m\(^{-2}\) K\(^{-4}\))
\( \tau(\alpha)_{\text{en}} \) effective transmittance-absorptance product for direct solar radiation at normal incidence (–)

Variables and parameters in the radiating model

\( C_{\text{em}} \) heat emitter effective heat capacity (W h K\(^{-1}\))
\( K \) heat emitter characteristic coefficient (W K\(^{-n}\))
\( n \) heat emitter characteristic exponent (–)
\( \dot{Q}_{\text{em}} \) heat flow rate supplied for the heat emitter (W)
\( T_{\text{em}} \) mean temperature of the heat emitter \((^\circ \text{C})\)

Variables and parameters in the building model

\( \dot{Q}_{\text{d}} \) standard heat load at design outdoor temperature for chosen location (W)
\( T_{a,d} \) ambient design temperature of heating system at chosen location \((^\circ \text{C})\)
\( T_{\text{room}} \) mean temperature of the room air \((^\circ \text{C})\)
\( T_{\text{set,room}} \) room setpoint temperature \((^\circ \text{C})\)

Variables and parameters in the store model

\( C_{\text{sto}} \) storage tank effective heat capacity (W h K\(^{-1}\))
\( \dot{Q}_{\text{aux,sto}} \) heat flow rate supplied by the auxiliary heater, at the bounds of the storage tank (W)
\( \dot{Q}_{\text{dhw}} \) heat flow rate supplied for the DHW demand (W)
\( \dot{Q}_{\text{dhw,sto}} \) heat flow rate supplied for the DHW demand, at the bounds of the storage tank (W)
\( \dot{Q}_{\text{coll,sto}} \) heat flow rate supplied by the collector, at the bounds of the storage tank (W)
\( \dot{Q}_{\text{em,sto}} \) heat flow rate supplied for the heat emitter, at the bounds of the storage tank (W)
\( T_{a,sto} \) temperature of the air surrounding the storage \((^\circ \text{C})\)
\( T_{\text{tap}} \) temperature of water at input of tap water net \((^\circ \text{C})\)
\( T_{\text{sto}} \) mean temperature of the storage tank \((^\circ \text{C})\)
\( T_{\text{set,dhw}} \) DHW setpoint temperature \((^\circ \text{C})\)
\( (\text{UA})_{\text{aux,loop,hot}} \) heat loss capacity rate of auxiliary loop hot side (W K\(^{-1}\))
\( (\text{UA})_{\text{aux,loop,cold}} \) heat loss capacity rate of auxiliary loop cold side (W K\(^{-1}\))
\( (\text{UA})_{\text{coll,loop,hot}} \) heat loss capacity rate of collector loop hot side (W K\(^{-1}\))
\( (\text{UA})_{\text{coll,loop,cold}} \) heat loss capacity rate of collector loop cold side (W K\(^{-1}\))
\( (\text{UA})_{\text{dhw,loop,hot}} \) heat loss capacity rate of DHW loop hot side (W K\(^{-1}\))
\( (\text{UA})_{\text{dhw,loop,cold}} \) heat loss capacity rate of DHW loop cold side (W K\(^{-1}\))
\( (\text{UA})_{\text{em,loop,hot}} \) heat loss capacity rate of heat emitter loop hot side (W K\(^{-1}\))
\( (\text{UA})_{\text{em,loop,cold}} \) heat loss capacity rate of heat emitter loop cold side (W K\(^{-1}\))
\( (\text{UA})_{\text{sto}} \) heat loss capacity rate of the storage tank (W K\(^{-1}\))

Variables and parameters in the auxiliary boiler model

\( a_{\text{aux}} \) boiler performance coefficient (W\(^{-1}\))
\( b_{\text{aux}} \) boiler performance coefficient (–)
\( c_{\text{aux}} \) boiler performance coefficient (W)
\( Q_{\text{aux}} \) power consumed by the auxiliary heater (W)
\( Q_{\text{aux,nom}} \) auxiliary heater nominal power (W)
\( Q_{\text{aux,out}} \) heat flow rate supplied by the auxiliary heater (W)
\( T_{\text{set,boil}} \) set outlet temperature of auxiliary boiler \((^\circ \text{C})\)

Symbols in system identification

\( H \) Heissian matrix of the cost function
\( J \) cost function
\( N_i \) number inputs
\( N_n \) number of neurones
\( N_{\text{n,max}} \) maximum number of neurones
\( N_o \) number of outputs
\( N_{\text{tp}} \) number of training patterns
\( N_{\text{sw}} \) number of synaptic weights
\( u \) input vector of system
combisystem behave differently as it is supposed to (Thu¨r,
2011). Performances can be reduced significantly in this way. Therefore, it would be important to have an idea of
the performance expected before choosing a solar combi-
system and installing it in a building.

Nevertheless, there is no test to characterise combisys-
tems currently. Due to the large variety of energy manage-
ment strategies and their sensitivity to external conditions,
it is difficult to predict combisystem thermal performance
in a reliable and complete way. Some test methods have been
developed. They can be classified in two categories, depend-
ing on their approach to evaluate primary energy savings:

1.1. Component approach

Component approach method, like the “Component
Testing – System Simulation” (CTSS) (Albaric et al.,
2010; Dr¨uck and Bachmann, 2002), consists in testing the
main components of solar combisystem (heat store, con-
trollers, and collector) according to existing standard tests.
Parameters of each component are identified and then used
within thermal simulation software like TRNSYS (Klein
et al., 1994).

This approach is very flexible as it is possible to test
nearly every system configuration and to evaluate perfor-
mances for different climatic conditions and different kinds
of buildings, through several annual simulations.

On the other hand, it is not possible to check real system
controls and interactions between components since they
are only simulated. Design mistakes and installation diffi-
culties can pass through this kind of test.

1.2. Global approach

Global approach method, like the “Annual Calculation
and/or Direct Characterisation” (Bales, 2002), the “Con-
cise Cycle Test” (Vogelsanger, 2002) and the “Short Cycle
System Performance Test” (SCSPT) (Albaric et al., 2008),
consists in testing each combisystem as a whole. The com-
plete system is set up in an indoor test facility (except the
collectors). Charging and discharging of the system are
done according to a test sequence, peculiar to each method.
For SCSPT method, the annual performance is derived
from a simple extrapolation of the auxiliary energy con-
sumed during the procedure.

This kind of approach ensures that all system functions
and interactions are taken into account during the test.
Furthermore, the time needed with this approach is quite
short. The longer test sequence is 12 days long.

However, in this way, results are available to only one
fixed set of conditions, according to the heating load, the
climatic conditions and the system sizing (especially collec-
tor area) during the test.

The objective of the work presented in this paper is to
further develop one of the global approach methods, the
SCSPT. The improvement proposed here tries to get
through the limitation of the global approach results in
order to have a complete methodology to characterise com-
bisystem thermal performance.

2. The current SCSPT procedure

The new methodology proposed in this paper is based
on the SCSPT that has been developed at INES, the
French National Institute for Solar Energy (Albaric et al.,
2008, 2010). As a global approach method, the SCSPT first
consists in installing the entire tested system on the semi-
virtual test bench.

2.1. The semi-virtual test bench

The semi-virtual test bench located in the INES facilities
(Fig. 1) makes possible to test any thermal system or
process, confronting it with a virtual environment. In the
case of solar combisystem, all the components of the tested
system (store(s), controller(s), pump(s), valve(s), sensor(s)
and auxiliary energy device(s), if needed) are installed on
the test bench as they are usually in a building. Weather
variables, collectors, DHW draw-offs, heat emitter and
building are simulated, in real time, within the TRNSYS
software (Klein et al., 1994).

Communication between the real system and the virtual
environment is done thanks to hydraulic modules. Modules
are plugged in a hot water (180 °C) and a cold water
(-10 °C) network, supplied by a central heating room. Thus, by means of heat exchangers, each module can produce or consume heat quickly from a secondary fluid when they are ordered to.

Each loop of the combisystem (mainly DHW, space heating and collector loops) is connected to a module. At every time step (1 min), modules record the temperature and the flow rate of incoming fluids and send these measurements to the monitoring computer. Those measurements are used as inputs to the TRNSYS simulation software. Then, according to the outputs of the TRNSYS calculation, fluid temperatures at the outlet of the modules are adjusted. This way, each module emulates every virtual component simulated by the software. Data exchanges are represented in Fig. 2.

2.2. The SCSPT virtual environment and the weather test sequence generation

The SCSPT has created a virtual environment to evaluate performance of combisystem with a 12 days test on the semi-virtual bench. The objective of the method is that the tested system use proportionately the same auxiliary energy during the test sequence ($Q_{12d,aux}$), as it would use during a complete year with the selected building and climate conditions ($Q_{y,aux}$). So, the auxiliary energy annual consumption is evaluated by a simple extrapolation, as described by the following equation.

$$Q_{y,aux} = \frac{365}{12} \cdot Q_{12d,aux}$$

The virtual environment is made up of the building model defined in the IEA Solar Heating and Cooling programme Task 32 (Heimrath and Haller, 2007), a standard DHW draw offs profile and a specific 12 days weather data sequence selected from the annual weather data thanks to an optimisation algorithm.

The process for the selection of the weather data sequence is presented in (Albaric et al., 2008) and is the special feature of the SCSPT. It is based on TRNSYS simulations of the reference combisystem modelled within IEA-SHC Task 32 (Heimrath and Haller, 2007). The algorithm selects 12 days from the annual weather data file so that the two following criteria, internal storage energy and space heating demand, have the same evolution day-by-day during the test sequence as month-by-month during the annual sequence.

Once generated for a climate, the same weather data sequence is used to test any combisystem. The performance evaluation is available for the corresponding climate and the selected building model.

2.3. Validation of the SCSPT

The SCSPT approach has been tested within a French project and a European project Combisol (http://www.combisol.eu) for various combisystems available on
the market. The method has been especially studied in (Mette et al., 2010). To sum things up:

- Weather data selection relevancy has been validated through numerous TRNSYS simulations;
- Experimental results are close to CTSS results, when this one is adjusted, considering pipe heat losses and controller defaults;
- Performance evaluations for three tested combisystems are consistent with on-site monitoring results.

However, it seems impossible to characterise thermal performances of combisystems with the current SCSPT procedure. This paper proposes an extension of the SCSPT procedure to make it able to evaluate thermal performances of combisystems for other boundary conditions. The improvement relies on global model identification, analysing completely data recorded during the SCSPT test.

3. Improvement of the procedure with global model identification

The objective is to develop a procedure which is able to learn the behaviour of the tested combisystem from a 12 days test. Considering energy fluxes and temperatures as inputs and outputs of the system, a global model could be identified from an analysis of the test data. Such a model could be used to simulate the combisystem with various boundary conditions and then to evaluate its thermal performance whatever the conditions.

3.1. The SCS global model

The proposed model is represented in Fig. 3. It is mainly composed of one linear dynamic model with known equations (“White Box”), which estimates the states of the system for the next time step, and one non-linear static model to identify (“Black Box”), which represents the real specific behaviour of each system. More explanations are given below.

3.1.1. Inputs and outputs of the global model

The global model must evaluate the energy flows involved in a combisystem according to weather and load conditions. Therefore, inputs and outputs are defined as follows:

- Inputs are defined as external variables, i.e. all the variables that are not controlled by the system. Those are climate variables like ambient temperature and solar irradiations, and DHW loads data like cold-water temperature and DHW draw-offs.
- Outputs are defined as energy flows controlled by the system: on the one hand, heat supplied to space heating loop and DHW draw-offs; and on the other hand, energy supplied by the collector and the auxiliary energy. In this work, the auxiliary energy system is considered as part of the combisystem. Therefore, the energy used by auxiliary system is considered as an output of the model.

This set of variables is also chosen to represent combisystem behaviour as simply as possible, to help the model identification.

3.1.2. Dynamical model representation

One interesting representation of dynamical model in this case is the state space model. The general form of a state space model can be described by the following equation.

\[
\begin{align*}
\dot{x} &= g(x, u) \\
y &= h(x, u)
\end{align*}
\]

where \(y\) and \(u\) are respectively outputs and inputs vectors of the model. \(x\) is the state vector of the model. \(g\) and \(h\) are two functions, linear or not.

Analogy between this representation and the real functioning of solar combisystem has inspired the structure of

![Fig. 3. The proposed global SCS model structure.](image)
the proposed model. Usually, heat flows are controlled depending on the temperature of various components part of the combsystem (collector, store, heat emitter, building rooms). They provide information about the energy needed, stored and available. Therefore, it seems relevant to consider those temperatures as the state of the system. It is even more interesting to use these temperatures as they can be estimated with well-known and simplified equations.

The g function can so be expressed as a “White Box” model. The real special feature of each system to be tested with heating floor. The h function, identified from the test data analysis, is so a "White Box" model. Observing most system controlling, it can be stated that the h function is non-linear. The whole model is a “Grey Box” model.

3.1.3. The white box part

The “White Box” part aims to evaluate mean temperature of the collector, the store, the heat emitter and the room of the building, from one time step to the next one. This must be done only with the set of variables presented in Section 3.1.1 and with standard parameters as far as possible.

The “White Box” part presented below is the one currently used. Nevertheless, it is possible to envisage other models, as long as no other variable is required to provide evaluations of the temperatures needed.

For numerical computing, all equations described in Section 3.1.3.1 are calculated with an explicit discretization scheme and a 30 min time step.

3.1.3.1. Collector model. Eqs. (3) and (4) represent the model used to evaluate the mean temperature of the collector. It is described in Perers (1997). This model is based on the well-known “Hottel–Whillier–Bliss” equation for flat plate solar collectors that is adapted to characterise almost every kind of collectors, except ICS collectors. It has been widely used for simulation (Perers and Bales, 2002) and standards (EN 12975-2, 2006; ASHRAE 93-86, 1986).

\[
c_3 \frac{dT_{\text{coll}}}{dt} = \dot{Q}_{\text{sol,net}} - \frac{\dot{Q}_{\text{coll}}}{A_{\text{coll}}} - c_1(T_{\text{coll}} - T_{\text{a}}) - c_2(T_{\text{coll}} - T_{\text{a}})^2 - c_3(T_{\text{coll}} - T_{\text{a}}) + c_4(E_L - \alpha T_{\text{a}}) \quad (3)
\]

\[
\dot{Q}_{\text{coll,net}} = F'({\text{v}})_{\text{em}}(K_{\text{th}}(0) \cdot G_b + K_{\text{td}} \cdot G_d) - c_5 v G - c_1(T_{\text{coll}} - T_{\text{a}}) - c_2(T_{\text{coll}} - T_{\text{a}})^2 - c_3 v(T_{\text{coll}} - T_{\text{a}}) + c_4(E_L - \alpha T_{\text{a}}) \quad (4)
\]

3.1.3.2. Heat emitter model. So far, tested combsystems have been designed to work with radiators. This part of the model must be adapted to be used with systems working with heating floor.

The radiator model used to evaluate the mean temperature of the heat emitter is based on standard parameters, calculated with (EN 442-2, 1996). A thermal capacitance has been added to take into account the dynamic behaviour:

\[
C_{\text{em}} \frac{dT_{\text{em}}}{dt} = -K(T_{\text{em}} - T_{\text{room}})^n + \dot{Q}_{\text{em}} \quad (5)
\]

3.1.3.3. Building model. The building model used is the one defined in the international standard (ISO 13790:2008(E), 2008). Every heat transfer coefficients and the internal capacity can be calculated out of architectural and physical parameters of the building. The complete set of equation is not described in this paper but it can be found in the mentioned standard. Parameters of this model have been calculated and tuned in order to have similar results as the reference building described in IEA SHC Task32 (Herrmann and Haller, 2007): SFH30, SFH60 and SFH100 (i.e. single family house with space heating loads respectively of 30 kW h m$^{-2}$, 60 kW h m$^{-2}$ and 100 kW h m$^{-2}$ over a year for the Zurich climate).

3.1.3.4. Store model. As the store is one of the central components of a combsystem, this part is not modelled during the SCSP, unlike the other components presented in the above sections.

The goal of this model is to provide information about the energy stored in the tank. Given that only few variables and parameters are available, equations come down to a simple energy balance, completed by a heat capacitance and heat loss parameters:

\[
C_{\text{sto}} \frac{dT_{\text{sto}}}{dt} = \dot{Q}_{\text{aux,sto}} + \dot{Q}_{\text{coll,sto}} - \dot{Q}_{\text{em,sto}} - \dot{Q}_{\text{dhw,sto}} - (U_{\text{a}})_{\text{sto}}(T_{\text{sto}} - T_{\text{amb,sto}}) \quad (6)
\]

In this equation, heat flows take into account heat losses of pipes. Only gross estimations of the heat losses can be done for each loop with the variables available. They consider both hot and cold side of each loop, giving them a heat loss coefficient and a representative temperature. Losses are cancelled when no heat is delivered through pipes.

\[
\begin{align*}
\dot{Q}_{\text{aux,sto}} &= \dot{Q}_{\text{aux,out}} - (U_{\text{a}})_{\text{aux,loop,hot}}(T_{\text{set,boil}} - T_{\text{amb,sto}}) - (U_{\text{a}})_{\text{aux,loop,cold}}(T_{\text{sto}} - T_{\text{amb,sto}}) \\
\dot{Q}_{\text{coll,sto}} &= \dot{Q}_{\text{coll}} - (U_{\text{a}})_{\text{coll,loop,hot}}(T_{\text{coll}} - T_{\text{amb,sto}}) - (U_{\text{a}})_{\text{coll,loop,cold}}(T_{\text{sto}} - T_{\text{amb,sto}}) \\
\dot{Q}_{\text{em,sto}} &= \dot{Q}_{\text{em}} + (U_{\text{a}})_{\text{em,loop,hot}}(T_{\text{sto}} - T_{\text{amb,sto}}) + (U_{\text{a}})_{\text{em,loop,cold}}(T_{\text{em}} - T_{\text{amb,sto}}) \\
\dot{Q}_{\text{dhw,sto}} &= \dot{Q}_{\text{dhw}} + (U_{\text{a}})_{\text{dhw,loop,hot}}(T_{\text{set,dhw}} - T_{\text{amb,sto}}) + (U_{\text{a}})_{\text{dhw,loop,cold}}(T_{\text{tap}} - T_{\text{amb,sto}})
\end{align*}
\]

3.1.3.5. Auxiliary energy model. The “White Box” model takes $Q_{\text{aux}}$, the auxiliary energy used by the SCSP, as input. Yet, the store model needs $Q_{\text{aux,out}}$, the heat supplied by the auxiliary part. Thus, a simplified model describing the auxiliary system is required to evaluate its efficiency.
So far, tested combisystems have been combined with auxiliary gas boilers. This part of the model must be adapted to be used with other kind of auxiliary energy.

The gas boiler model used in the “White Box” part is derived from the model used in the French thermal regulation for building (RT2005, 2005). This model evaluates the energy losses of the boiler according to the heat demanded with a simple second-order polynomial. The relationship between the energy input and output of the boiler can be expressed as follows:

\[
Q_{aux} = a_{aux} Q_{aux, out}^2 + (1 + b_{aux}) Q_{aux, out} + c_{aux}
\]  

(8)

Parameters \(a_{aux}, b_{aux}\) and \(c_{aux}\) are calculated with characteristic powers and losses of the boiler. Calculations are available in the French regulation. \(Q_{aux, out}\) is determined by solving simply the quadratic equation. The nearest root to \(Q_{aux}\) is taken as the solution.

3.1.4. The black box part

The “Black Box” part aims to model the tested combisystem behaviour, i.e. how it manages energy according to external variables and its own states (temperatures described in Section 3.1.3 above). Since combisystem behaviours are complex and non-linear, it seems appropriate to include in this “Black Box” part, a powerful mathematical model such as an Artificial Neural Network (ANN). ANNs are widely used nowadays because, as (Kalogirou, 2001) highlights, they can learn from examples, are fault tolerant and are able to deal with non-linear problems. They are even widely used in solar energy fields for predicting complex phenomenon and modelling solar components as well as complete solar system.

Mellit and Massi Pavan (2010) and Paoli et al. (2010) have used ANNs to forecast daily solar radiation, both to predict performance of grid connected PV plant (at Trieste for the former, Ajaccio for the latter). They have built different neural network architectures but results are relevant for both authors. Paoli has even found that ANNs are more precise compared to conventional prediction methods.

Baccoli et al. (2010) has inferred the steady state performances of a thermal solar collector with evacuated pipes, only from a 4 days experimental data set of the collector in transient regime. He has tried two kind of model: a grey box model (based on known physical model and parameters identification) and a Dynamic Adaptive Neural Network. Both models have shown good results compared to the official EN 12975 efficiency curve.

Medinelli Sanino and Rojas Reischel (2007) has modelled a solar domestic water heating system located near Vina del Mar, Chile. After testing different model structures, the author turns to semi-physical modelling combined with ANN: available system knowledge helps to organise the conceptual form and ANN adjusts that knowledge into a particular non-linear model structure. According to the authors, this incorporation of physical knowledge is advantageous because it shares advantages of each one of the structures, providing a model at the same time parsimonious and flexible. The so trained model predicts with good accuracy the mean water temperature of the storage tank.

Many other examples are reviewed in (Kalogirou, 2001) but those presented above emphasise some interesting points:

- Complex phenomenon, like solar radiation, can be modelled by ANN and can even be more precise than usual prediction method.
- Modelling and predicting global system performances seem possible. It has been done for different kind of solar domestic water heating systems.
- Integrating knowledge of the system as far as possible is advantageous when modelling systems.

So far, nothing has been done to link ANN modelling with SCS. These points encourage to consider the global model structure presented in this paper and especially to integrate ANN in the “Black Box” part of the model.

ANNs are numerical model inspired by biological neural networks. They build their knowledge through a learning process that strengthens connections between neurons, like a brain.

The structure of a mathematical neurone is represented in Fig. 4 and is described in Dreyfus et al. (2002). Neurone inputs \(x_i\) came from other neurones outputs or model inputs. Those signals are transmitted to the neurone through connections called “synapses”. Synaptic weights \(w_i\) are linked to each connection. There are several ways to combine inputs with their corresponding synaptic weights. In this model, a simple linear combination as described in Eq. (9) is used. The result of this combination \(v\) is the argument of a transfer function \(f\). The outcome \(y\) is the activation of the neurone.

\[
v = w_0 + \sum_{i=1}^{n} w_i x_i
\]

\[
y = f(v)
\]

(9)

Fig. 4. Scheme of a mathematical neurone.
ANNs are built linking neurones together and with the model inputs and outputs. When presenting a training data set, an optimisation algorithm searches the best synaptic weights set to minimise a cost function. All functions and parameters selected for the neural network training of this model are presented in Section 3.2.

3.2. Identification of the model

Once the training data set is obtained from the lab test, it is used to train several neural networks. Then, the same training set helps finding the best ANN to make the complete SCS model reproduce the system behaviour as faithfully as possible. The complete process of training and selection is represented in Fig. 5 and described below.

3.2.1. Pre-processing of inputs and outputs

The training set is processed to have data with a 30 min time step. The “White Box” model is applied to the complete data set to have estimations of the system’s states during the lab test. This includes “White Box” part errors as input of the neural network so that it learns to manage them.

Dreyfus et al. (2002) suggest normalising inputs and outputs. Ideally, variables should be centred and reduced. Another way proposed here to come close to this data reduction is to combine variables and parameters of the system and its environment. This way, inputs and outputs of the ANN are not raw temperatures and heat flows but reduced criterions of the interaction between the system, the building and the climate. The current proposed inputs and outputs reduction is described in the following equation.

\[
u = \begin{bmatrix}
\frac{\dot{Q}_{\text{sol,net}}}{\dot{Q}_{\text{sol,ref}}}
\frac{K(T_{\text{em}} - T_{\text{room}})}{\dot{Q}_{d}}
\frac{(T_{\text{room}} - T_{\text{set,room}})/(T_{\text{set,room}} - T_{a,d})}{(T_{\text{room}} - T_{a})/(T_{\text{set,room}} - T_{a,d})}
\frac{(T_{\text{sto}} - T_{\text{set,dlw}})/100}{(10)}
\end{bmatrix}
\]

\[
y = \begin{bmatrix}
\frac{\dot{Q}_{\text{aux}}}{\dot{Q}_{\text{aux,nom}}}
\frac{\dot{Q}_{\text{coll}}}{(A_{\text{coll}}\dot{Q}_{\text{sol,ref}})}
\frac{\dot{Q}_{\text{em}}}{\dot{Q}_{d}}
\end{bmatrix}
\]

\(\dot{Q}_{\text{sol,net}}\), the net solar irradiance on the collector defined in Eq. (4) and \(\dot{Q}_{\text{coll}}\), the heat supplied by the collector are both divided by a reference solar irradiance taken at
3.2.2. Construction of the ANN structure

The ANNs tested for the “Black Box” part of the model are simple static Multi Layer Perceptrons (MLPs). By means of trial and errors, a more precise structure is set to limit the number of ANNs to test. The activation function for every neurone is a hyperbolic tangent. Only one layer of hidden neurone is considered. The structure of tested ANNs is represented in Fig. 6. Finally, tested ANNs differ in the number of neurones in the hidden layer.

3.2.3. Determination of the maximum number of neurones

The main issue of ANN modelling is to find the correct number of neurones. With too few neurones, the ANN will not be able to learn enough the process. However, with too many neurones, the ANN will learn well the training data set but will not be able to generalise for other simulations.

This number is strongly dependent on the process to model.

So far, there is no rule to find the right number of neurones. It is a matter of trial and error. Nevertheless, to develop an automatic way to find the best ANN for the tested SCS, it is possible to refer to the Widrow rule of thumb described in Widrow and Stearns (1985): the number of training patterns \( N_{\text{tp}} \) must be larger than \( N_{\text{sw}} \), the number of synaptic weights to be identified, divided by the error tolerance \( \varepsilon_{\text{tol}} \).

\[
N_{\text{tp}} \geq \frac{N_{\text{sw}}}{\varepsilon_{\text{tol}}} \quad (11)
\]

From this rule, a maximum number of hidden neurones can be calculated. Since only one hidden layer is considered in this ANN (see Section 3.2.2) and that a bias is added to each neurone (including output neurones), \( N_{\text{sw}} \) can be expressed with \( N_o \), the number of outputs, \( N_i \), the number of inputs, and \( N_n \), the number of neurones.

\[
N_{\text{sw}} = (N_i + 1) \cdot N_n + (N_n + 1) \cdot N_o = N_n \cdot (N_i + N_o + 1) + N_o \quad (12)
\]

Therefore, the maximum number of hidden neurones is expressed as follows:

\[
N_{n,\text{max}} = \frac{\varepsilon_{\text{tol}} \cdot N_{\text{tp}} - N_o}{N_i + N_o + 1} \quad (13)
\]

During the selection process, networks with complexities going from 1 to \( N_{n,\text{max}} \) neurones will be trained.

3.2.4. Initialisation of the synaptic weights

Initial synaptic weights of the ANN are determinant for the optimisation algorithm convergence and the solution found. Several initialisations of ANNs are tested in this process \( (I_{\text{max}}) \).

Some weights initialisation methods enhance ANN training. The one used in this process is the Nguyen–Widrow method (Nguyen and Widrow, 1990).

3.2.5. Training algorithm (open loop)

Optimisation algorithms search for the best synaptic weights set to minimise a cost function. The cost function \( J \) chosen here is the sum of squared errors over the training data set.

\[
J = \sum_{i=1}^{N_o} (y_i - \hat{y}_i)^2 \quad (14)
\]

Most training algorithms are based on the calculation of the gradient of the cost function. In the case of feedforward network with differentiable activation function for every neurone (like the MLP considered here), this gradient can be smartly calculated with the backpropagation technique (Rumelhart et al., 1986). To improve algorithm speed compared to a simple steepest descent method, the Levenberg–Marquardt algorithm (Marquardt, 1963) is chosen. It calculates an estimation of the inverse of the Hessian matrix.
$H$ of the cost function and changes the step $\lambda$ at each iteration. The recurrence relation is described by the following equation.

$$w_{j+1} = w_j - (\hat{H}(w_j) + \lambda I)^{-1} \cdot \nabla J(w_j)$$  \hspace{1cm} (15)

To avoid an overtraining phenomenon, a “regularization” strategy has been chosen in the training process using Bayesian regularization in complement of the Levenberg–Marquardt algorithm (Mackay, 1992).

3.2.6. Test of the neural network (closed loop)

It is possible to test the trained network generalisation to a certain extent with the same training data set. The ANN has learnt the SCS behaviour in an “Open loop” way, i.e. with fixed ideal states calculated in the pre-processing step (see Section 3.2.1). It is interesting to check whether the global model is stable and precise in a “Closed loop” simulation of the training data set. In this case, ANN errors affect the state estimation of the “White Box” part of the global model. This way, the network shows if it can handle its own errors. A still overtrained network will be excluded in this step.

3.2.7. Selection criterion

ANNs with different complexities and different initialisations are trained in this process. To select the best one, they are compared according to two criterions.

The Root Mean Square Error (RMSE) assesses the estimation quality of the model at each time step. This is the criterion commonly used to evaluate an estimator in statistics.

$$\text{RMSE} = \frac{1}{N_{tp}} \sum_{i=1}^{N_{tp}} (y_i - \hat{y}_i)^2$$  \hspace{1cm} (16)

The quality expected for this kind of model is above all to estimate precisely the global energy involved by the SCS during any simulation. Even if occasional large errors happen, the model could be able to evaluate with a good accuracy the auxiliary energy used for a complete year. Therefore, it is interesting to check the global estimation error.

$$\text{Err} = 1 - \frac{\sum_{i=1}^{N_{tp}} y_i}{\sum_{i=1}^{N_{tp}} \hat{y}_i}$$  \hspace{1cm} (17)

To observe both expected model quality as only one criterion, a coefficient $d$ is introduced: it combines RMSE and Err as follows

$$d = \sqrt{\left(\frac{\text{RMSE}}{\alpha_{\text{RMSE}}}\right)^2 + \left(\frac{\text{Err}}{\alpha_{\text{Err}}}\right)^2}$$  \hspace{1cm} (18)

where $\alpha_{\text{RMSE}}$ and $\alpha_{\text{Err}}$ are fixed coefficients which assigned a different weight for RMSE and Err. So far, they were arbitrarily set respectively to 5 and 0.9.

Among all networks trained (see Sections 3.2.3 and 3.2.4), the one that gives the lowest $d$ coefficient when simulating the training sequence is selected to be part of the final “Grey Box” model of the tested combisystem.

3.3. Final procedure

Considering the modelling aspect proposed in this paper to improve the SCSPT, the procedure to evaluate SCS performances can be summed up as follows:

- Install the SCS on the semi-virtual test bench as it is installed in a real building.
- Collect parameters needed for the “White Box” (especially storage and auxiliary energy parameters).
- Carry out the test with the virtual environment corresponding to the 12 days Zurich climate sequence, the SFH60 radiator and building parameters and with the collector parameters suggested by the manufacturer. Models are all defined in Section 3.1.3.
- Collect data needed for the model training during the test sequence.
- Apply the process defined in Section 3.2 to get a global model of the tested SCS.
- Simulate the SCS behaviour over a year with different buildings, climates and collector areas.
- Characterise the SCS with the FSC method (Letz et al., 2009).

The FSC method considers that fractional energy savings $f_{\text{sav}}$ of a SCS can be expressed as a quadratic function of the fractional solar consumption FSC, a dimensionless quantity which only depends on the environment of the SCS. Thus, the performances of the tested SCS are characterised by a simple parabola.

4. Results and discussion

4.1. Numerical results

First, to validate this new approach, data from simulations of three different SCS model were used to train global models as described in Section 3.2. Simulations were done with TRNSYS (Klein et al., 1994).

4.1.1. SCS TRNSYS models

The three detailed models used for this numerical evaluation can be briefly described as follows.

- SCS1 is the reference system modelled during the IEA SHC Task 32 (Heimrath and Haller, 2007). It is represented in Fig. 7. The model does not consider DHW loop and space heating loop losses.
- SCS2 and SCS3 are commercial systems models used during the Combisol project (Mette et al., 2010). They can be represented by the generic SCS scheme B1 (Fig. 8) according to (Thür et al., 2010). Both systems differ in their controls and their heat store charging strategies (SCS3 use a special stratification device). No losses are considered in the loops.
Fig. 7. Generic scheme of the reference combisystem model.

Fig. 8. Combisol generic scheme B1: immersed DHW Heat exchanger – auxiliary as return flow increase.
4.1.2. Comparison with TRNSYS calculations

For each SCS model, the SCSPT test (corresponding to the Zurich climate and the SFH60 building) has been simulated within TRNSYS. Then, data from this simulation have been used to train and select one global “Grey Box” model, according to the procedure described in Section 3.2.

Global models have then been used to do the 27 annual simulations as described in Table 1. In Figs. 9–11, results are compared to the corresponding detailed models simulations within TRNSYS. Each simulation is represented by three points: $Q_{aux}$ (circles), $Q_{coll}$ (dots) and $Q_{em}$ (stars). On the whole, identified models and TRNSYS simulations are very close. Differences between both methods to estimate characteristic heats and energy are almost within the $\pm10\%$ range for all simulations. For low heat demand simulations, absolute differences are not excessive but the low energy level makes those differences proportionally higher.

Correlations coefficients ($R^2$) are close to 1 except for $Q_{coll}$. This output seems harder to evaluate precisely. In SCS working, the annual energy coming from the collector depends not only on the quantity of heat stored in the tank (measured by $T_{store}$) but also on the quality of the heat stored (linked to the good stratification in the tank). The store model used in the “White Box” part does not take this aspect into account. Another more detailed store model may enhance $Q_{coll}$ evaluation accuracy. Moreover,

### Table 1

<table>
<thead>
<tr>
<th>Climate</th>
<th>Building</th>
<th>Collector area $^a$ (m$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barcelona</td>
<td>SFH30</td>
<td>10</td>
</tr>
<tr>
<td>Zurich</td>
<td>SFH60</td>
<td>15</td>
</tr>
<tr>
<td>Stockholm</td>
<td>SFH100</td>
<td>20</td>
</tr>
</tbody>
</table>

$^a$ When changing the collector area, the design mass flow-rate in the collector loop and the store volume also change according to fixed ratios specified by the manufacturer (respectively 15 kg h$^{-1}$ m$^{-2}$ and 50 L m$^{-2}$ for the reference SCS for example).

Fig. 9. Results of SCS1 energy estimations.

Fig. 10. Results of SCS2 energy estimations.
Results comparison: SCS3

- Estimated energy (ANN−model) [kJ]
- Calculated energy (TRNSYS) [kJ]

Qaux : $R^2 = 0.99541$
Qcapt : $R^2 = 0.76144$
Qem   : $R^2 = 0.99978$

Fig. 11. Results of SCS3 energy estimations.

SCS1 characterization

- ANN performances estimations
- TRNSYS performances calculations
- ANN characterization
- TRNSYS characterization
- Range of properly working systems

Fig. 12. Results of SCS1 characterisation.

SCS2 characterization

- ANN performances estimations
- TRNSYS performances calculations
- ANN characterization
- TRNSYS characterization
- Range of properly working systems

Fig. 13. Results of SCS2 characterisation.
estimations are less accurate for extreme conditions, i.e. for simulations combining Barcelona climate and a large collector area. In these cases, SCS, facing very low heat load, high solar gains and a storage almost full all year long, can behave very differently compared to their behaviours observed during the training sequence. That is most probably why the global grey box model overestimates \( Q_{\text{coll}} \) for those simulations, resulting in a bad correlation coefficient.

In Figs. 12–14, each combisystem is characterised from TRNSYS results and global model results, according to the FSC method (Letz et al., 2009). Characteristic curves are close for the three combisystems models. Observation made above for energy estimations impacts on the characterisation. Curves are closer when FSC is below 0.6. For simulations with higher FSC, the SCS behaviour start to be different compared to the training set. Low power consumptions enlarge proportional errors. Nevertheless, estimated results are acceptable.

4.2. Experimental results

4.2.1. The tested SCS

A real combisystem has been tested according to this procedure on the INES test bench. As SSC2 and SSC3 numerical models, this system can be represented by the generic scheme B1 (Fig. 8). The store volume is larger (1000 L).

Two tests have been done for this combisystem, with different virtual environments:

1. The standard 12 days test with the Zurich climate, the SFH60 building and 16 m\(^2\) of collector area.
(2) The 12 days test corresponding to the Stockholm climate, the SFH100 building and 10 m² of collector area.

4.2.2. Comparison between the two tests sequence

The former standard test data have been used to train and select a global model of the tested combisystem. The resulting model has been used to simulate the 2nd test sequence. Comparison between the global model simulation and the real recorded power consumption are presented in Fig. 15. It shows that the dynamic behaviour of the global model is in the whole quite corresponding to the real system. Even though model estimation does not fluctuate exactly in the same way as the real system did during the test, power variations are corresponding in average.

Table 2 shows that energy estimations are accurate for $Q_{aux}$ and $Q_{em}$. The error with $Q_{coll}$ is higher. The “Grey Box” model seems to overestimate the solar gains and compensate this overestimation through the two other energies. The difference is acceptable because the level of this output is low compared to the others. Observations made about $Q_{coll}$ estimations in Section 4.1.2 are still valid in this case. Nevertheless, those results are promising for other simulations.

4.2.3. Performance characterisation of the tested combisystem

The same selected model has been used to characterise the performances of the tested system with the FSC method (Fig. 16). So far, there is no more information about the system for comparison but the tests data. The characteristic curve of the tested system seems slightly above the performances estimated by the classic SCSPT method for the two tests. This is somehow in accordance with the performance overestimation observed in the Section 4.2.2. Anyway, Test1 and Test2 points are not that far from the estimated curve in comparison with the general scatter of the other points. The experimental application of the methodology offers a characterisation that seems realistic.

Other real combisystems with better-known performances (for example with several on-site data sets) should be modelled and characterised like this to further validate the methodology. However, this validation project would need a lot of time to find available systems, to collect and to analyse enough data for more complete comparisons.

5. Conclusion and perspectives

The improvement of the current SCSPT methodology is promising. While the former version of those tests evaluate with a good accuracy SCS performances for only one environment (climate and building) and only one system sizing, the “grey-box” model identification of the tested system extend the results to any environment and several sizes of collector. This extent let SCS be characterised only from a 12 days test on a semi-virtual test bench.

The model to be identified is made of two parts: a “White Box” part that gathers characteristics equations of the SCS main components; and a “Black Box” part using artificial neural networks that are taught how the system deals with external variables and its own states to meet the demands.

Comparison with numerical SCS models shows that the methodology is accurate. Identified models estimations of the energy consumed by the auxiliary boiler are closed to the detailed models ones. However, errors are getting larger when the environment become more extreme compared to...
the training one. Some more work about the storage model may improve the accuracy of the model, especially for the collector output energy evaluation. Further studies must be done to check the validity range of estimations.

A real combsystem has been tested according to this methodology. The model trained from a test data set imitates well the SCS behaviour compared to a completely different experimental data set. Its characterisation seems to be in accordance with experimental results but need more information to be further validated. Some other real SCS will be tested soon on the INES semi-virtual test bench.

So far, the methodology is adapted for SCS using gas boiler as auxiliary energy and radiator as heat emitter. Some additional work must be done to find other models (like heating floor for instance) that suit the requirements of the “White Box” part of the model and so to have a methodology adapted for most of existing SCS. If the results still show good prediction performances, it would be interesting to consider this way as a standard test to improve the SCS market development.

Acknowledgement

This work has been supported by the French Agency for Environment and Energy Management (ADEME).

References


