

Combining Engineering and Qualitative Models to Fault Diagnosis in Air Handling Units

Peter Struss², Raymond Sterling¹, Jesús Febres¹, Umbreen Sabir², Marcus M. Keane¹

Abstract. This paper presents a methodology for model-based fault localization and identification that exploits both numerical (Modelica) models and a qualitative model-based approach to diagnosis. It has been applied to diagnosis of an air handling unit based on data recorded by a building management system. The main steps from model development to diagnosis based on the recorded data are discussed.

1 INTRODUCTION

Heating Ventilation and Air conditioning (HVAC) systems are known for being very inefficient for different reasons, one of the most common causes being the presence of undetected failures in one or more of its components. Faults can remain undetected for long periods due to different factors: automated control may mask faults through compensation of symptoms by other elements of the system; lack of proper maintenance, improper timing of flow of energy to/from the building, etc. Though the faulted system may produce proper conditions in the building, this may happen at the cost of unnecessary energy consumption, and, hence, energy saving policies enforce the need for early detection of faults. Even when systems are known to operate in a suboptimal way, the presence of faults may be difficult to localize and identify manually and a costly task for human operators who usually only act when targeted indoor conditions are not met. All this raises the need for developing automated fault detection and diagnosis methods and technologies that assist the building operator.

The focus of this paper is on a model-based diagnostic solution that uses a **Modelica model of plant components**, a **qualitative diagnostic model** derived from it, and **consistency-based diagnosis** [3] for a part of the HVAC system, the **Air Handling Unit** (AHU). This solution starts from a general first-principle Modelica model to generate a qualitative diagnostic model. It exploits a general diagnosis algorithm (consistency-based diagnosis) that feeds deviations of actual sensor measurements from the values predicted by the numerical model to the diagnostic model. The application system isolates and identifies the most common faults that can cause significant loss of system performance and waste of energy in AHUs: defective valves in heating- and cooling-coil, and stuck dampers. It has been applied to real data of several days of an AHU with different faults inserted and successfully evaluated and compared to an existing rule-based solution.

The paper is structured as follows: section 2 presents the plant used in the case study. Section 3 outlines the targeted entire work flow, whose different steps are then presented in section 4. Then we discuss the results of the evaluation of the system on real data.

2 THE AIR-HANDLING UNIT

The case study comprises a constant air volume air-handling unit (AHU) whose schematic is shown in Figure 1. The AHU serves a facility consisting of an audio laboratory of around 50 m², where strict conditions of temperature and humidity have to be met due to the presence of highly sensible musical instruments. The building is located in Cork city in the Republic of Ireland.

An Air Handling Unit is a hybrid system that operates in multiple different modes depending on the prevailing external environmental conditions and the requirements of the area being served by the unit at a given time. Each different mode of operation may utilize the components in the AHU in a different manner to satisfy required conditions in the rooms served by the unit under the given environmental conditions.

An AHU is composed of a number of components that function independently of each other, but are controlled by a central system. As a result, the impact of a fault in one component (e.g. erroneous pre-heating of the incoming air) may be compensated for through the control system by the use of another component (subsequent cooling), causing energy to be wasted (actually in both components).

A typical AHU, as presented in Fig. 1, comprises the following components:

1. Mixing Box (MB): serves to recover heat from exhaust air by mixing a certain (controlled) fraction of it with fresh air from outside;
2. pre-heat coil (preHC): is an emergency heat exchanger to prevent frost in the unit when outside air condition are below freezing point;
3. Cooling Coil (CC): to control both temperature and humidity by cooling the air;
4. Re-Heat Coil (HC): to control temperature by heating the air;
5. Humidifier (H): to control humidity by adding water vapor to the air.
6. SupplyFan and ReturnFan: produce the air flow.
7. Temperature sensors (circles 15, 19, 5, 1, 8, 14, 29).

¹ Informatics Research Unit for Sustainable Engineering, Ryan Institute, NUI Galway, Ireland, email: raymond.sterling@nuigalway.ie

² Computer Science Department, Technische Universität München, email: struss@in.tum.de

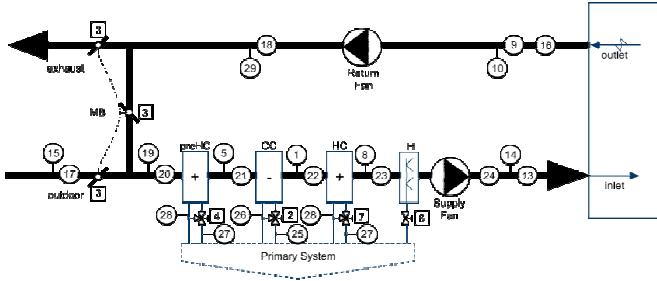


Figure 1. Air Handling Unit Schematic.

The data available for diagnosis are the actuator signals (controlling the mixing box and the coils by opening and closing valves to control the flow of hot or cold water) and the sensor data received by the building management system every minute. Not all AHUs are equipped with such a complete set of sensors, which leads to the requirement that the diagnostic system should produce results for whatever set of sensors is available. In the current solution, we do not include humidity in the model and, hence, do not exploit the respective sensors.

3 THE TARGET WORK FLOW

In our work, we produced a complete workflow for generating model-based diagnosis systems tailored to particular plants and implemented a major part of the respective tool chain in order to realize the case study on the AHU. Fig. 2 presents the three task layers, whose steps will be discussed in more detail in the following sections.

1. Producing the **domain-specific foundation** involves
 - the production of a library of Modelica models and
 - its transformation into a qualitative diagnostic model library.
2. Producing a **plant-specific application system** based on the libraries requires
 - The configuration and calibration of a Modelica plant model of the correct behavior (“OK model”);
 - the composition of the diagnostic model based on the diagnostic library and the component structure of the plant, which can be extracted from this Modelica system model. This composition step is part of the functionality of the tool used in this work, *Raz'r* (from OCC'M Software GmbH [9]). The extraction of the component structure from Modelica has not yet been realized, but is expected to be straightforward, provided that the Modelica models are compliant with the requirements stated below in section 4.
3. For **on-line diagnosis**,
 - qualitative deviations are generated by computing the difference between the real data (currently for steady state only) and the predictions generated by the Modelica OK model, and determining qualitative deviations based on given thresholds, which are then processed by
 - the runtime system, which employs *Raz'r's* diagnosis engine.

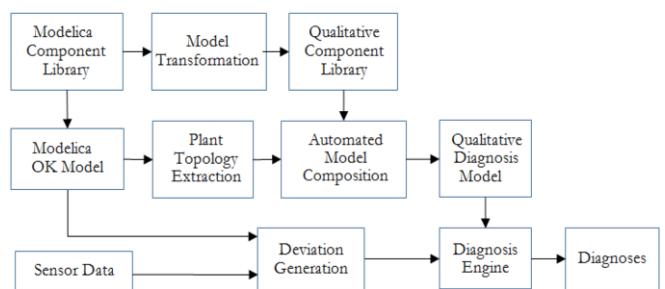


Figure 2. From model to diagnosis, the MBD chain

4 MODELING FOR MODEL-BASED DIAGNOSIS

In order to support component-oriented model-based diagnosis, the diagnosis models and, hence, also the numerical models have to satisfy particular requirements:

1. Strictly **component-oriented** modelling: the library has to be organized around the component types (with models that can be parameterized) that constitute the plant and that are units subject to diagnosis, e.g. heating/cooling coils and mixing box.
2. **Fault models** should be represented (perhaps with a parameter characterizing the fault, such as the opening of a passing valve)
3. The plant model has to be configured strictly according to the **real physical interconnections** in the plant. It must not include computational artefacts that link certain variables that are not really interacting directly via a physical connection. This includes using the concept of connectors in Modelica to reflect the channels of physical interactions between components (rather than connections via single variables as, for instance, in Matlab/Simulink).
4. The models in the library have to be formulated in a **context-independent manner** and must not rely on implicit assumptions about the presence and correct functioning of other components, even though they may exist in most standard configurations. This is relevant for two reasons: it enables the re-use of the component models for different plants, and it is a precondition for the adequacy of the models in fault situations.

4.1 Model Development

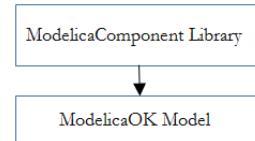


Figure 3. Modelica model development

Besides the above requirements, low development efforts and best use of manufacturer's data have been main guiding principles for the model development. They are closely related since the manufacturer's data (provided for certain operation points) is the

first source of information a model developer will have at hand and forms the input to a first calibration.

In the following, we use the heating coil model for illustration. It calculates the outlet steady-state conditions in both, water and air, using equations derived from the conservation of energy and mass principles and the definition of effectiveness in the classical eff-NTU method which given by equations (1), (2) and (3)[4]:

$$Q = C_a * (T_{ao} - T_{al}) \quad (1)$$

$$Q = C_w * (T_{wl} - T_{wo}) \quad (2)$$

$$Q = eff * min(C_a, C_w) * (T_{wl} - T_{al}) \quad (3)$$

The effectiveness 'eff', depends on the coil configuration (parallel flow, counter flow, or cross flow with both streams unmixed) [5]. A snippet from the Modelica code is shown below to illustrate the match between the above equation formulation and Modelica code:

```
Qflow = Cflow_a*(To_a - Ti_a);
Qflow = Cflow_w*(Ti_w - To_w);
Qflow = eff*min(Cflow_a,Cflow_w)*(Ti_w-Ti_a);
```

For the heating-coil component, there are inputs and outputs for flow of air through the ducting, and flow of hot water through the heating coil. Hence, mass- and energy-balance equations hold for airflow and water-flow. They complement the Modelica model equations.

4.2 Qualitative Diagnostic Models

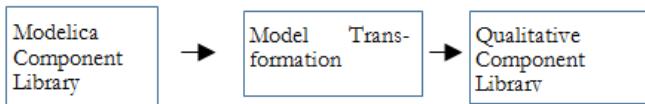


Figure 4. From numerical model to qualitative models.

The diagnostic model library is obtained from the Modelica library (Fig. 4) by transforming the each component model. The models used in our diagnostic approach are stated in relative, rather than absolute terms: they capture the deviation of variable values from the respective under nominal behavior.

Following [6], [7]; the qualitative deviation of a variable x is defined as:

$$\Delta x := sign(x_{act} - x_{nom}) \quad (4)$$

Equation (4) captures whether an actual (observed, assumed, or inferred) value is greater, less or equal to the nominal value. The latter is the value to be expected under nominal behavior, technically: the value implied by the model in which all components are in OK mode.

Qualitative deviation models can be obtained from standard models stated in terms of (differential) equations by canonical transformations, such as equations (5) and (6). We use \oplus , \ominus and \otimes , to denote addition, subtraction and multiplication on signs.

$$a + b = c \Rightarrow \Delta a \oplus \Delta b = \Delta c \quad (5)$$

$$a * b = c \Rightarrow (a_{act} \otimes \Delta b) \oplus (b_{act} \otimes \Delta a) \ominus (\Delta a \otimes \Delta b) = \Delta c \quad (6)$$

It is important to note that these equations do not contain and require values for the reference values x_{nom} and, hence, can be applied to different plants and under distinct operating modes. The qualitative deviation models, obtained from the Modelica models, reflect current modeling assumptions, (steady state, and no deviation in airflow) and become very compact due to their qualitative nature and because constants can be dropped and just replaced by their signs. Internally, this model is automatically transformed into an efficient data structure representing finite relation.

In the following, we illustrate how this transformation can be done by manipulating the equations. According to energy balance equations (equations (1), (2) and, (3)), and assuming no losses, the energy balance in equation (7) can be reformulated in terms of deviations (Δ) as in equation (8).

Assuming that the air flow and the water temperature (drop) are positive and not deviating and replacing the capacity flow by the mass flow $mflow_w$ (which differ only by a constant factor), we obtain equation (9) which applies to all modes of the coil.

$$0 = C_a * (T_{ao} - T_{al}) - C_w * (T_{wl} - T_{wo}) \quad (7)$$

$$0 = \Delta (C_a * (T_{al} - T_{ao})) \oplus \Delta (C_w * (T_{wl} - T_{wo})) \quad (8)$$

$$0 = \Delta T_{al} \ominus \Delta T_{ao} \oplus \Delta mflow_w \quad (9)$$

Following equation (4), each of the variables used for diagnostics can have a deviation of the measured value from the simulated one as follows:

1. positively ('+'), when the actual (measured, predicted, or assumed) value is above the simulated plus a threshold;
2. negatively ('-'), when the actual value is below the simulated minus a threshold;
3. or not deviate ('0'), when the actual value is within the simulated value plus/minus the threshold.

Table 3 depicts the resulting relation on the three deviation variables for equation (9), i.e. all solution tuples. For instance, the first three rows of the table indicate the intuitive fact that, if the mass flow shows no deviation, a deviation of the incoming air temperature will simply be propagated to the output air temperature.

On the other hand, a positive deviation of the output air temperature in combination with no deviation in the input air temperature, is only consistent with a positive deviation in the mass flow rate of the water (last-but-one row). From the diagnostic perspective, this reveals a fault in the coil (e.g. a passing valve), because a correct coil will not produce a deviating water flow. A valve stuck closed may lead to a negative deviation “-”, if the command Cmd to the valve is “open” (to some non-zero position, “+”). If the control commands the valve to be shut, anyway, a stuck-closed valve would cause no deviation in the water flow. This is captured by the model fragment in Table 5, which actually, is the complete fault model. Table 4 and Table 6 show the models of the OK mode and the passing valve, respectively. The table expresses that this mode may coincide with the nominal behavior for a certain range of opening commands, but deviate positively for smaller valve positions.

With respect to their use for diagnosis, tables 4 – 6 jointly with table 3 capture which tuples of temperature and water flow deviations are consistent with which behavior modes. Note that this does not require that the deviations can be observed directly. They

may also be **predicted** by the system model based on observations for a particular system health assignment.

Table 3. Relation on temperature deviations and water flow deviation.

Δm_{flow_w}	ΔT_{al}	ΔT_{ao}
0	-	-
0	0	0
0	+	+
-	-	-
-	0	-
-	+	*
+	-	*
+	0	+
+	+	+

Table 4. Qualitative representation of the OK mode.

Cmd	Δm_{flow_w}
0	0
+	0

Table 5. Qualitative representation of the stuck closed valve mode.

Cmd	Δm_{flow_w}
0	0
+	-

Table 6. Qualitative representation of the passing valve mode.

Cmd	Δm_{flow_w}
0	+
+	0
+	+

Bear in mind that a qualitative representation of one mode does NOT exclude that any other mode can be reached with the same combination of inputs/outputs.

While the above illustrates how to obtain the diagnostic models from the equations underlying the Modelica model, we also developed a tool that automatically **calculates** the abstract model using the respective Modelica model. Since this is beyond the scope of this paper, we refer to [8], which applies the approach to Matlab models.

4.3 Runtime Deviation Generation

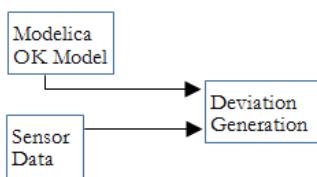


Figure 5. Generating Deviations

At runtime, the system calculates deviations (Fig. 5) as follows:

Table 7. Deviations between sensor data and model data.

	T _{al} (°C)	T _{ao} (°C)	T _{al} (°C)
Sensor Data	18.32	20.87	18.32
Model Prediction	18.44	18.44	18.44
Resulting Deviation	0	+	0

1. A steady state filter is used to extract steady state data from the operation data. (Note that, in this application, exogenous conditions do not change drastically and frequently, such that enough periods will be available);
2. For each recorded “snapshot”, the data vector with the sensor and actuator signals is read;
3. From these, the values of exogenous variables (external temperature, damper commands, and valve commands) are fed as an input to the Modelica model of nominal behavior;
4. For the values predicted by this model for the variables that correspond to the rest of the observables, their deviation from the actual sensor data is computed the deviations. In the current solution, this is done by using a chosen threshold (which can be different for different variables). For the exogenous variables, the deviation is always zero.

In our case study, a threshold of 2°C was chosen in order to produce deviations in the domain of signs ('+', '-','0'). In future solutions, different orders of magnitudes of the deviations could be generated by the abstraction module, which can take arbitrary sets of interval boundaries as an input.

Table 7 shows both the sensor data and the predicted values, highlighting the temperature before and after the heating coil. Using the 2°C threshold, the inflow air temperature is determined as nominal, while the outflow air temperature is higher than expected.

4.4 Diagnosis Inference

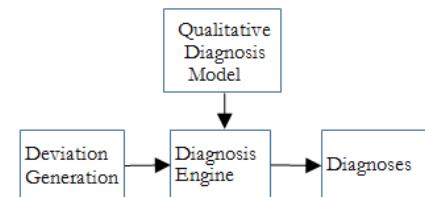


Figure 6. From deviations and qualitative model to diagnosis

In our solution, we use Raz'r [9], a commercial tool that implements component-oriented consistency-based diagnosis [3]. The essence of this is the following: every component can have different behavior modes assigned that correspond to nominal or faulty behavior. These behavior modes have associated models in the diagnostic library, LIB (in our approach capturing how deviations are created or propagated). A mode assignment MA selects one behavior mode for each component and, based on the topology of the system, STRUCTURE, specifies a deviation model of the entire system. The core of the diagnosis engine is to check whether such a model is consistent with the observations OBS. If a mode assignment is consistent

$$\text{LIB} \cup \text{STRUCTURE} \cup \{\text{MA}\} \cup \text{OBS} \not\models \perp$$

then it is a candidate for describing the actual state of the system, i.e. a diagnosis. The initial hypothesis is that all components operate correctly (and, hence, the system behaves without deviations): in the mode assignment MA_{OK} , OK modes are assigned to all components. If the model of the correctly behaving system is inconsistent with the observations:

$$\text{LIB} \cup \text{STRUCTURE} \cup \{\text{MA}_{\text{OK}}\} \cup \text{OBS} \models \perp.$$

a **fault** has been **detected**. If faulty components need only to be **localized**, i.e. separating correctly operating components from faulty ones, it often suffices to restrict the modes of a component C to $\text{OK}(C)$ and $\neg\text{OK}(C)$. To perform fault identification and/or refine fault localization, fault modes of components and their respective behavior models can be defined, as previously mentioned. The combinations of fault models span an entire space of models.

In our solution, the observations OBS are the deviations generated as described in section 4.3 (including the zero deviation of the exogenous variables), and the diagnosis engine searches for mode assignments whose resulting system (deviation) model is consistent with the observation vector. Obviously, a mode assignment can contain several fault modes, and, hence, this technique is not limited to localize and identify single faults. The search generates diagnoses containing minimal sets of faulty components.

What happens for the entire system, can be illustrated for the trivial case of one component: the observations presented in Table 7, namely input/output temperature deviations (0, +) match with only one row in Table 3 that holds for all behavior modes, which fixes $mflow_w$ to be positively deviating. This positive deviation is consistent with the valve passing mode (Table 6), but neither with the OK mode, nor the stuck closed mode. Note, that this result can actually be concluded without information about the command to the valve.

The slow processes and sparse observations in our case study do not establish a challenge to real-time behavior. For applications where this is the case, Raz'r's code generator can be used to compile the consistency-based diagnosis algorithm together with the diagnostic plant model into very compact and efficient C-code.

5 EVALUATION AND COMPARISON

In this section, we present the results of testing the developed qualitative diagnosis methodology on a real facility. First, we describe the experiments carried out, followed by a comparison of the results against the tradition APAR rule based approach [10].

Five experiments were developed, with the initial experiment having the objective of obtaining operational data for when humidity control on the unit has been disabled (Humidity control is not a focus of the research at the present time). Four further experiments then emulated faults with the mixing box, pre-heating coil, cooling coil, and heating coil. These experiments require the control valve of the individual component to be manipulated into a fault condition through the control system. Therefore, it was proposed to fix the damper/control valve position on the individual component for a set period of time, observe the system response and then adjust the position again. A brief outline of the procedure

for one of these experiments is outlined below and the others follow a similar pattern.

In the experiment to simulate a stuck mixing box damper position, the damper is initially set at the minimum fresh air position and the reaction of the system is observed for ten minutes. The damper position is then opened by 10% and again the system is left to settle for a period of ten minutes. This step is performed incrementally with ten-minute settling periods until a damper position of 100% is achieved. The procedure is then reversed going from 100% back to the minimum position in steps of 10% with a 10-minute settling period between changes. The whole procedure is then repeated a second time giving two sweeps through the applicable damper positions to give sufficient data to enable the model calibration and diagnostics analysis on the mixing box component of the AHU.

As a result of the experiments, four 24-hour data sets were compiled from real AHU data and including each one of the experiment scenarios described above.

Table 8. Evaluation results

Fault	APAR	QMBD	Comments
Nominal	No fault identified	No fault identified	No fault identified by each of the two approaches.
Passing Cooling Coil	No fault identified	2 possible faults identified during 5 separate time periods	No fault identified by APAR as the cooling coil being 0% made the engine think that the unit was in heating mode and therefore rules pertaining to the cooling coil were not applied.
			QMBD both correctly identified an issue with the cooling coil.
Passing Heating Coil	6 possible faults identified during 3 separate time periods	4 possible faults identified during 3 separate time periods	APAR identified a number of possible faults including an issue with the heating coil.
			QMBD correctly identified a fault on the heating coil and also correctly identified a fault in the mixing section of the AHU.
Stuck Mixing Damper	4 possible faults identified during 5 separate time periods	2 possible faults identified during 6 separate time periods	APAR identified a number of possible faults including an issue with the mixing dampers.
			QMBD correctly identified a fault on the mixing dampers.

In order to provide a baseline for comparison, the traditional APAR rule set was also applied to the data and results were used for comparison with the qualitative model-based diagnosis. The four data sets chosen for demonstrative purposes come from experiments 1 to 4. **Error! Reference source not found.** presents and comments the comparison results obtained for each experiment.

In contrast to [11], our solution is snapshot-based and does not require temporal reasoning. The work described in [12] is closely related, but focuses on exploiting Modelica models for qualitative simulation.

6 DISCUSSION

In this paper a tool chain from model development to on-line diagnosis and its application to air handling units has been presented. The development tool of choice for the model was Modelica since it provides all the necessary tools to comply with model requirements for component-oriented model-based diagnosis discussed in section 4. Modelica models bear the perspective of becoming a de-facto standard in energy modeling of building components as shown by the recently established International Energy Agency Annex 60.

One of the main advantages of the model-based approach is the adaptability to different plants and to changes in the same plant. A brief description of the steps involved in adapting the qualitative model based diagnosis is presented below.

1. **Structural changes:** These changes will have to be reproduced in the model, which would need to be compiled and recalibrated. The diagnosis model structure is a 1:1 mapping of the model and as such only minor adaptation is needed. However, if the change involves variables considered for diagnosis, the variable mapping between model and diagnosis framework has to be modified and tested with new data sets.
2. **Parameter changes:** recalibration of the models is in principle the only requirement. In the case these parameter changes impact the accuracy of the model, the tolerances of the diagnosis framework might have to be adjusted.
3. **Sensor changes:** similar consideration to the case of structural changes should be taken in the case of adding new sensors or modifying position of existing ones. In the case that existing sensors are to be replaced with new ones with different precision, the steps described in the parameter changes are to be followed.
4. **Changes in control:** plant model and diagnosis framework is, in principle, not affected by changes in the control strategy.

This adaptability makes model-based diagnosis a viable approach to fault detection and diagnosis in air handling units. Since they have a fairly standard structure and comprise a small set of different component types, the current library of qualitative diagnostic models needs only limited extensions in order to be applicable to a broad class of AHUs.

Taking into account that heating ventilation and air conditioning systems are rarely critical systems, the benefits are more economic and environmental rather than a safety issue. And hourly fault detection and diagnosis frequencies are more than acceptable in building applications.

7 ACKNOWLEDGEMENTS

This work was supported by the International EnergyResearch Centre and Enterprise Ireland under project n. CC-2011-4005B and by the Irish Research Council – D’Appolonia enterprise partnership scheme. Special thanks to Dominik O’Sullivan John McCarthy for their invaluable support and help in providing datasets for testing the developments presented in this research work. We are also indebted to Oskar Dressler of OCC’M Software for providing Raz’r and supporting its integration.

8 REFERENCES

- [1] S. Katipamula, P. Michael, and R. Brambley, “Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems— A Review, Part II,” *HVAC&R Res.*, vol. 11, no. 2, 2005.
- [2] R. Isermann, “Model-based fault-detection and diagnosis – status and applications,” *Annu. Rev. Control*, vol. 29, no. 1, pp. 71–85, Jan. 2005.
- [3] P. Struss, “Model-based problem solving” in van Harmelen, F., Lifschitz, V., and Porter, B. (eds.) *Handbook of Knowledge Representation*, Elsevier, 2008, pp. 395 – 465.
- [4] ASHRAE, *ASHRAE Handbook: Fundamentals (SI edition)*. Atlanta, GA: American Society of Heating, Refrigerating and Air-conditioning Engineers, 2009.
- [5] M. Wetter, “*Simulation Model: Finned Water-to-Air Coil without Condensation*”. LBNL-42355,” 1999.
- [6] P. Struss, “Models of Behavior Deviations in Model-based Systems” in *European Conference on Artificial Intelligence*, pp. 883–887.
- [7] P. Struss and A. Fraracci, “Automated Model-Based FMEA of a Braking System,” *8th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes* (Safeprocess 2012), Mexico City, 2012
- [8] Struss, P., Fraracci, A., Nyga, D., "An Automated Model Abstraction Operator Implemented in the Multiple Modeling Environment MOM". In: *25th International Workshop on Qualitative Reasoning*, Barcelona, Spain, 2011.
- [9] OCC’M Software GmbH, www.occm.de 2014
- [10] House, J., Vaezi-Nejad, H., & Whitcomb, J. (2001). An expert rule set for fault detection in air-handling units. *ASHRAE Transactions*, 107(1).
- [11] M. Behrens, G. Provan, Temporal Model-Based Diagnostics Generation for HVAC Control Systems, Proc. *ICMT2010 - Intl. Conference on Model Transformation*, 28 June - 2 July, 2010
- [12] M. Klenk, D. Bobrow, J. de Kleer, and B. Janssen; Making Modelica Applicable for Formal Methods. In: *10th International Modelica Conference*, 2014