CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Why and how to assess the quality of Building Performance Data
Authors	Menzel, Karsten;Hryshchenko, Andriy;Mo, Kai
Publication date	2014-09-17
Original Citation	Menzel, K.,Hryshchenko, A. and Mo, K. (2014) 'Why and how to assess the quality of Building Performance Data', 10th European Conference on Product & Process - ECPPM 2014, Vienna, Austria, 17-19 September.
Type of publication	Conference item
Link to publisher's version	http://info.tuwien.ac.at/ecppm/sites/default/files/conftool-StatischeVersion/20141212161803/default_047.html
Rights	© 2014 the authors
Download date	2026-09-08 03:17:46
Item downloaded from	https://hdl.handle.net/10468/10448

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Why and how to assess the quality of Building Performance Data

K. Menzel, A. Hryshchenko, Kai Mo

IRUSE, School of Engineering, University College Cork, Ireland

ABSTRACT: Over the past decade buildings have become much more complex technical artefacts due to the fact that numerous energy co-generation and storage capabilities can be integrated into buildings. The integration of complementing operation of generation, storage, and building services systems is not a trivial task. Numerous data streams need to be analysed to ensure that the operation of the aforementioned systems contributes to overall energy savings and CO_2-reductions without undermining the tenants' comfort and wellbeing in those buildings. One could assume that through the emergence of wireless monitoring systems an improved knowledge base can be quickly and efficiently build-up to support the decisions which must be made by building automation systems or systems' operators. However, the pure availability of additional data compiled from sensors, meters, access control systems and local weather stations can worsen the situation in case the compiled data is incomplete or of insufficient quality. Therefore, this paper discusses the question of why and how to

1. WHAT IS BUILDING PERFORMANCE?

Building operations is one of the most significant expenses to any organisation, and these operating costs can be more than 30% of the annual corporate spend. Another additional factor to consider is fulfilling the requirements of many different European Legislations and Regulations demanding reduction of CO_2 emissions and the energy consumption of buildings. As a result, in order to ensure the ongoing reliability of buildings and the costs to operate and maintain building stock, an interest to obtain the most accurate performance data looks set to rise even further.

Building Performance can be defined from different perspectives, (i) the tenants perspective – focusing on user comfort, (ii) the operators perspective – usually focusing on minimal cost (fixed and running), and (iii) the owners perspective focusing on the optimal cost of ownership.

Depending on the quality of the building the above interest might considered to be contradictive to each other. It is one of the tasks of building operators to moderate the perspective and to identify a commonly agreed upon denominator for building operation, i.e. to satisfy the building user by optimising the operational cost and convincing the building owner to constantly invest in upgrades and retrofit activities to maintain a building and its systems on a satisfactory standard representing the state-of-the-art of building technology (products) and available building operational strategies (processes). Based on the above the following aspects of building performance can be identified:

Comfort related Performance Indicators

Those indicators are used to evaluate the user comfort. This is usually data compiled from different types of sensors installed in rooms or zones. Examples for values measured are room temperature, humidity-levels, CO_2-levels or illumination levels. Previously, this data was used as control parameters in building automation algorithms. However, the maintenance of such sensors – and subsequently the quality of data provided by these sensors - can be described as insufficient. Commissioning processes usually run over multiple months and sensor calibration becomes often a part of 'errors and omissions management'.

More recently, there is an increasing demand to use comfort data to verify that comfort levels are kept

within certain thresholds which were agreed in Service Level Agreement (SLA) or which are defined in national, European, or international standards.

Systems-related Performance Indicators

Those indicators are used to evaluate the performance of building services systems, sub-systems or individual components. Currently, there is a deficit with reference to commonly agreed standards for the definition of technical performance indicators. Therefore, we have introduced the term "Use Intensity (UI)". Parameters for the performance evaluation can be either compiled from sensors (e.g. supply and return temperature), from meters and sub-meters (e.g. energy consumption of central heating systems or the motor of a Geothermal Heat Pump) or from actuators (e.g. valve or flap is open between 0% .. up to .. 100%).

Finally, also data from burglar alarm systems can be integrated, e.g. indicating that a window is open whilst a heating system is running.

Usage efficiency-related performance indicators

Those indicators provide helpful information to building owners. Through those indicators owners can quickly develop an understanding for how often how many people used a specific room, zone, or part of a building. This data is usually compiled from security or access control systems.

Additionally, this data can be also used to develop a real-time understanding about internal loads from users and required periods for pre-heating and pre-cooling. Such information is extremely valuable if it comes to the development of improved prediction capabilities of Building Automation Systems.

Summary

Building Performance can be analysed from three different stakeholder-perspectives. Certain performance aspects, such as Use Efficiency or Comfort might be included in multiple views for performance analysis.

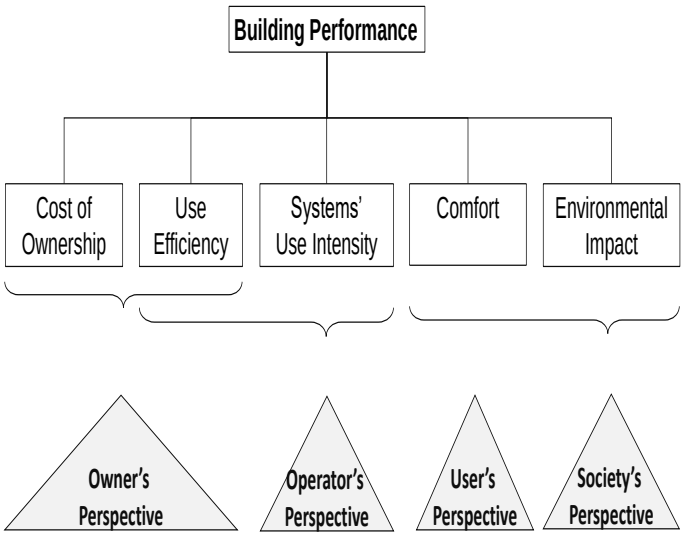

Figure 1: Different Views on Building Performance

2. SOURCES FOR PERFORMANCE DATA ACQUISITION

Performance Data is usually compiled from metering and monitoring systems. This data is compiled frequently, in many cases in 15-minutes reading intervals. Many sources classify this type of data as Dynamic Data, since it documents the dynamic change of either comfort conditions or the status of plant. From a mathematical-physical point of view this data can be also characterised as absolute measurements.

However, for performance evaluation it might be desirable to either relate absolute measurements to other values or to normalise monitored value in order to allow a more generalised evaluation and benchmarking. To achieve this, one needs to acquire further data, so called descriptive data, which can be compiled from Building Information Models. Since this data is not changing frequently, many sources call that data Static Data. The following sections briefly characterise each data type.

Dynamic Data requirements

Dynamic Data is compiled from Building Automation Systems and used as Fact Data in Data Warehouses. This data document the performance history of building systems over time.

We can distinguish two different modes of incoming dynamic data:

1. Regular Time Series - this is data captured or retrieved in set intervals from the data source;
2. Irregular Time Series Data - this data is captured only in case a value changes over a determined threshold. Therefore, this data cannot be retrieved. Vice-versa, the data source must have a "minimum intelligence" allowing the data source to identify the change of the threshold and send the report.

The above explained modes have further implications for the definition of data quality and data consistency checks.

Static Data requirements

This type of data is compiled from Building Information Models and used for Dimensional Modelling in Data Warehouses. This type of data can be characterised as so-called "slowly changing data". Consequently, this data is loaded using an initial load process and does not require frequent updates.

However, the initial load process might be complemented by manual integrity checks and physical inspections of sites, to ensure that the model content is correct, up-to-date and consistent. Product model data may exist on different levels of granularity over the building life-cycle, e.g. for the geometrical representation of a wall a centre-line might be used in the early design phases and for the documentation of the structural system, whereas a 2D-representation might be used in construction documentation and a full 3D-representation could be used for Architectural visualisation or progress monitoring.

As one can see there is little opportunity for automated consistency, integrity or other checks. Therefore, we do not further discuss the problem of uncertainty for static data in this paper. Similar thoughts can be developed for the evolution of the documentation of building services systems. The minimum information that can always be provided for analysis is the topological model (system schematic) of building services systems.

3. Why to assess Performance Data

In Section 1 we have explained that Building Performance Data (BPD) is compiled and used for different purposes, such as:

(1) Operational Support and Building Automation;

(2) Verification that national, European or international legislative requirements are fulfilled

(3) Verifying that contractual obligation defined in SLA are fulfilled.

(4) Decision support of maintenance activities.

It is obvious that low quality performance data can have serious negative impacts on the efficiency of buildings' operation but also on the Customer Relationship Management or the overall implementation and execution of certain business models [Menzel, Sirr; 2014].

Assessment of Data used for Building Operation

In this case wrong data can lead immediately to negative effects. In case building automation and control

algorithms use wrong parameters they will trigger wrong actuation commands and set systems and components in an inefficient operational mode. This might either lead to reduced user comfort or increased (unnecessary) resource consumption.

Potential sources for low data quality are (i) wrongly calibrated sensors, (ii) inappropriately adjusted actuators (e.g. 'slipping' valves), or (iii) uninterrupted data transmission which leads to data losses.

A single wrong or missing value might lead to unwanted negative impacts and a false execution of building automation commands.

Results of low quality comfort data might be that heating or cooling systems continue to operate due to a wrongly measured temperature value. Depending on the wrong offset value this might cause substantial additional energy consumption.

Assessment of Data used for Verification against standards and regulations

In this case wrong data can lead to mid-term negative effects, since the identification of malfunctioning systems and components will lead to a degradation of customer relationships since the documentation of wrong values leads to the impression that a building or a building services sub-system is operated under conditions which are not compliant to legislative requirements.

Since most standards allow for deviations within a given threshold it needs a larger series of wrong or missing dynamic data to create any substantial negative impact.

Assessment of Data used for Verification against Service Level Agreements

In this case wrong data can lead to a mid-term and long-term negative effect. The documentation of inappropriate comfort levels usually results to penalties, e.g. that a building user can withhold payments in total or in part or that the building operator has to pay a fine for breach of contract.

Assessment of Data used for Decision Support for maintenance management

In Figure 2 we present how data from monitoring systems can be used to inform and support the scheduling processes for maintenance activities.

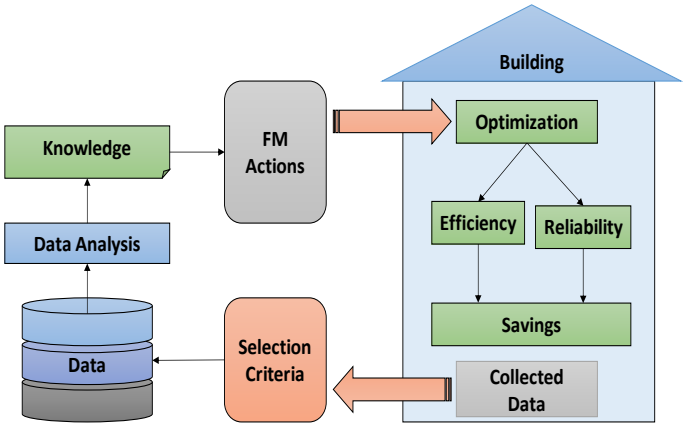

Figure 2: Data selection and analysis will provide new knowledge for when and how to execute maintenance activities

In this case wrong data can lead to long-term negative effects, since the identification of malfunctioning systems and components will lead to incorrect decision support for the scheduling of maintenance activities. However, operators usually go through a series of verifications and integrity checks before systems and components are finally replaced, i.e. operators diagnose systems and components before they decide to replace them. Therefore, low quality data can be identified as a failure source and unnecessary maintenance activities can be avoided.

Summary

In this section we discussed use cases requiring the assessment of BPD. We have identified potential negative impacts which can be triggered by using low quality BPD in different fields.

Figure 3 summarises the findings of section 3 in a graphical format.

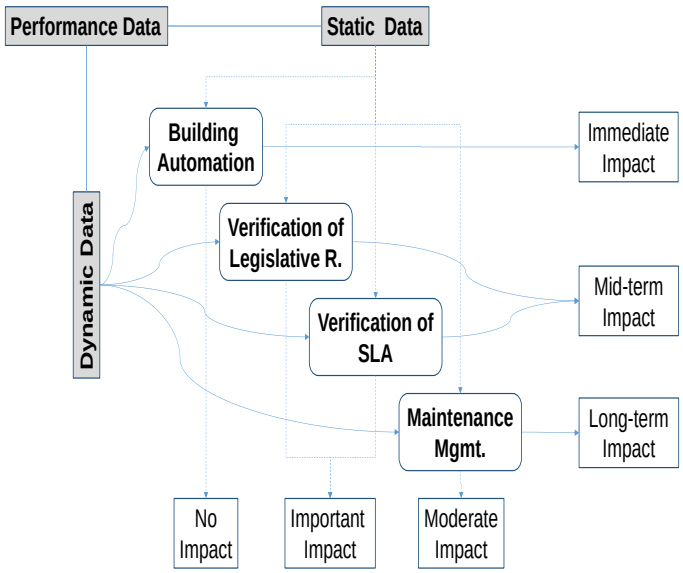

Figure 3: Impact-Assessment due to Low-Quality Data

4. HOW TO ASSESS PERFORMANCE DATA

Building Performance Data (BPD) can be assessed against multiple criteria, such as (1) data generated by Energy Simulation tools, (2) against temporally correlated data from the same or similar sources, (3) against spatially correlated data compiled under comparable conditions, or (4) against systemically correlated data.

The following sections will explain how and against what criteria BPD can be assessed.

Assessing BPD against data generated by Energy Simulation tools

On the first view this methodology seems to be straight forward, since a complete data set generated by the energy simulation tool seems to be available.

However, when setting up energy simulation models many assumptions are made and used as input parameters for those models. In case these assumptions are either incorrect or imprecise the results from the simulation tools are also incorrect. Some authors argue that the inaccuracy of energy simulation models lies between 30% and 100% (Soebarto, et al (2001), Turner, et al (2008); Yudelson (2010)).

Therefore, a calibration of the energy simulation model needs to be executed in order to run the model with as precise as possible values. Such a model calibration must not necessarily be time or labour intensive since many data can be easily compiled from monitoring systems, such as data from local weather stations or data from access control systems. In case of using local weather data it is said that data accuracy can be improved by 7% (Bhandari 2012).

The exploitation of user occupancy data is more complicated since it involves multiple sources for imprecision, such as (1) respecting tenants' or building occupants' privacy, (2) the fact that the functionality of access control systems is constrained by fire-safety regulations in such a way that only the entry to a room can be managed but leaving the room/zone.

Long-term data collected by researchers in the US show variations of occupancy diversity factors in private offices differing as much as 46% from those currently published in ASHRAE 90.1 2004 energy cost method guidelines (Duarte et al (2013)).

The so called static data play also an important role for the calibration of energy simulation models, especially if it comes to models to support renovation or retrofit activities. In these case the precise knowledge about the technical specifications of existing systems is of importance, in order to allow modellers to specify the systems and components to be modelled as precisely as possible. Precise and accurate BIM-data will help in the near future to address this problem.

Assessing BPD using temporal correlations

This approach is of importance for the assessment of BPD used to inform algorithms and control actions in Building Operation. Additionally, this approach might be also used for integrity checks of comfort and meter data used for the verification of SLA-conditions.

The assessment method for the first case is based on the assumption that changes of real-time data must appear in limited increments. Thresholds defining the increments between consecutive measurements can be concluded from multiple sources, such as manu-

facturers' manuals, the specification of allowed tolerances for metering devices in legislative documents, etc.

The assessment method for the second case is based on more detailed correlation analysis. Experimental results executed within the EU-FP7 CAMPUS21 project (Campus21 (2013)) have shown, that comfort levels and meter readings in public buildings show good correlation results on weekly cycles. In many cases the correlation between the same day of the week (e.g. Monday) is better than the correlation between two consecutive week days.

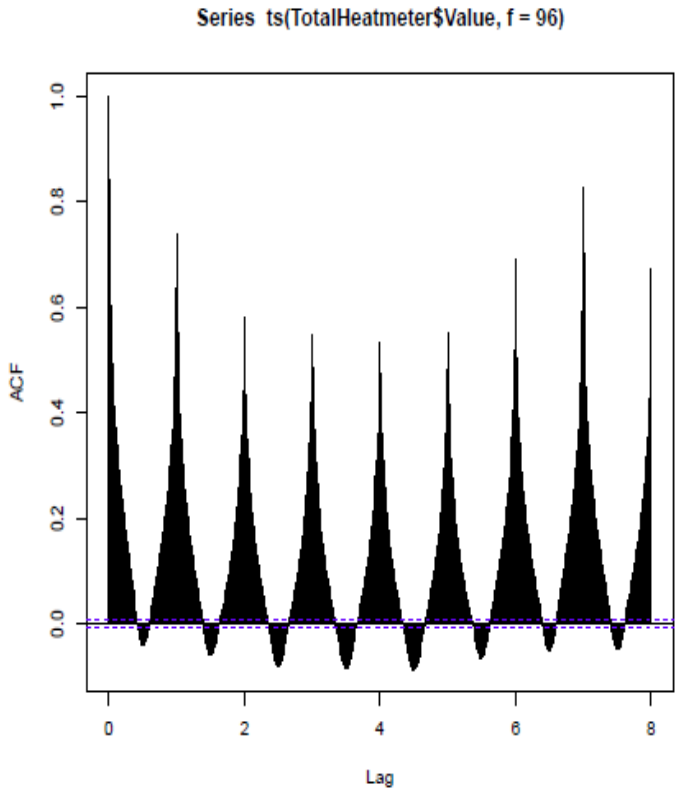

Figure 4: Heat Meter autocorrelation function over 8 days (Campus21 (2013))

Assessing BPD using spatial correlations

This approach needs more effort and thus computing power than the previous one since it needs to determine the spatial dependencies between monitoring devices first. The analysis result can be exploited in a subsequent analysis step.

Exploiting geometrical relationships

Spatial relationships between zones or rooms can be exploited by analysing the geometrical model. Firstly, one needs to identify if rooms or zones are geometrically co-located. This can be done by comparing the Cartesian coordinates of the room representations. An example how to access coordinates of Cartesian Points is presented in Figure 5.

However, the analysis of the geometrical relationship is not sufficient. Additionally, one needs to check if two rooms/zones have a comparable size and a comparable number of occupants and occupation schedules. Finally, one needs to check if rooms or zones have a similar orientation, i.e. if there exist windows within a similar geometrical region. Table 1 provides information how relevant attributes can be identified in a standardised building product model (IFC4).

Only in case all of the above constraints can be satisfied one can exploit spatial correlations for the assessment of user comfort-data.

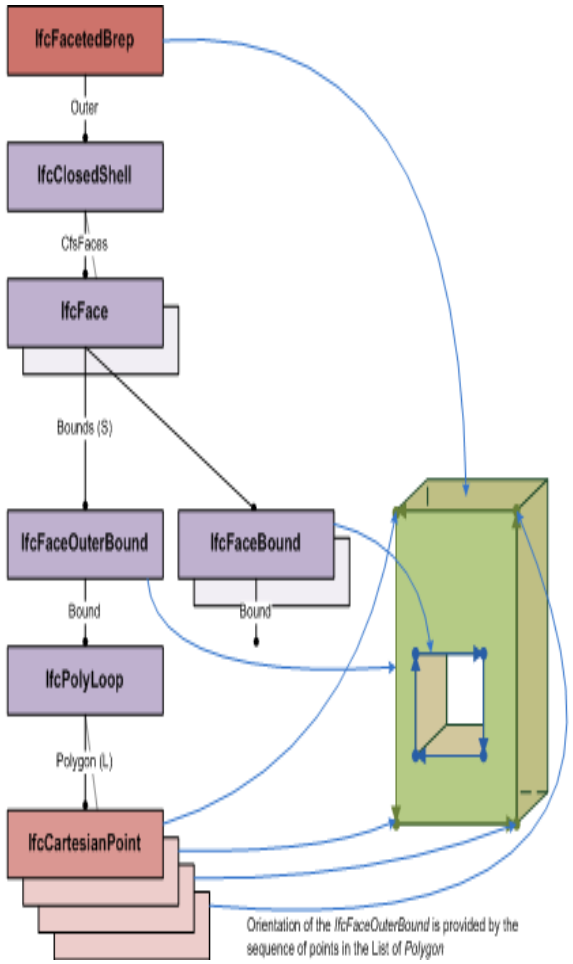

Figure 5: Boundary Representation in IFC4 (Building Smart, 2014)

Table 1: ifc-BIM parameters to check geometrical constraints

Constraint	Ifc-relationship or ifc-property set: attribute
Neighbouring Rooms/zones	ifcProductDefinitionShape → used to identify Cartesian points
Comparable Size Rooms/zones	Pset_SpaceCommon: NetPlannedArea
Comparable Occupation/usage Rooms/zones	Pset_SpaceOccupancyRequire-ments: {all attributes}
Comparable Orientation Rooms/zones	Pset_SpaceCommon: IsExternal

Assessing BPD using topological correlations

Building information models also contain information documenting the topology of building services system. This knowledge can be exploited for the assessment of sensor and actuator data monitoring building services systems. In any thermal energy system one should usually observe a temperature drop between the source and the sink(s).

A similar assumptions holds for the pressure, in case there is no change in the diameter of pipes or due to any turbulence. Table 2 provides information how relevant attributes can be identified in a standardised building product model (IFC4).

Additionally, we provide information in Figure 6 and Figure 7 how these topological relationships are modelled in IFC4.

ENTITY IfcDistributionPort

SUBTYPE OF IfcPort;

FlowDirection: OPTIONAL IfcFlowDirectionEnum;

PredefinedType: OPTIONAL IfcDistributionPortTypeEnum;

SystemType: OPTIONAL IfcDistributionSystemEnum;

END_ENTITY;

Figure 6: EXPRESS notation of the ifcDistribution-Port element in IFC4 (Building Smart, 2014)

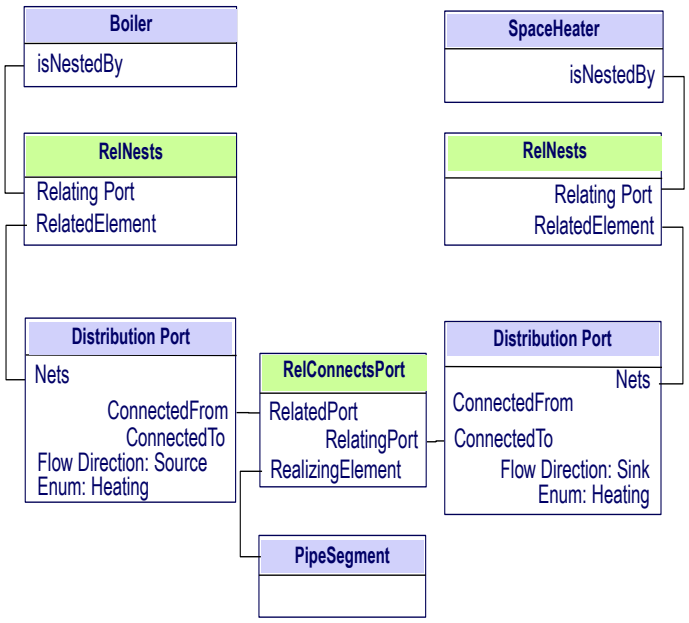

Figure 7: Simplified UML-Class Diagram illustrating the modelling of topological relationships for Building Services Systems (Building Smart, 2014)

Table 2: ifc-BIM parameters to check topological constraints

Constraint	Ifc-relationship or ifc-property set: attribute
Flow Direction	ifcDistributionPort: FlowDirection:
Pipe/Duct Diameter	Pset_PipeSegmentTypeCom-mon: InnerDiameter
Pipe/Duct Pressure	Pset_PipeSegmentTypeCom-mon: PressureRange
Pipe fittings and Control Flow Elements	ifcRelConnectsPorts: RealizingElement

5. ASSESSMENT TYPES

This section defines assessment types based on the classifications developed in the previous sections.

Type 1: This type compares monitored values against values derived from simulation results. An assessment formula is given below.

$Xs - |Xt| < X_m < X_S + |Xt|$ Equation 1

Where:

Xs……value from simulation tool

X_m .....monitored value

|Xt| …specified tolerance of simulated results

Type 2: This assessment type checks if monitored values are compliant to legislative or contractual requirements. An assessment formula is given below.

$L_{min} < X_m < L_{max}$ Equation 2

Where:

L_{min} …minimum value as per requirement

L_{max} ....maximum value as per requirement

X_m …..monitored value

Type 3a: This assessment type checks if monitored values unreasonably deviate from temporally consecutive values. An assessment formula is given below.

$X_{m-1} - X_m < X_D$ Equation 3

Where:

X_{-m-1}…predecessor of monitored value

X_m .....monitored value

X_D …specified allowed deviation between values

Type 3b: This assessment type checks if monitored values unreasonably deviate from temporally correlated values. An assessment formula is given below.

$X_{m-interval} - X_1 < X_D$ Equation 4

Where:

$X_{m-interval}$…value monitored at correlated time

Interval….specification of correlation

X_m .....monitored value

X_D …specified allowed deviation between values

An example for a correlation might be that monitored values always correlate to the value measured on the same weekday. In case we assume a 15 minutes reading interval we can define the value for interval as:

Interval = ((24 * 60/15) *7) = 672

Type 4a: This assessment type checks if monitored values unreasonably deviate from values in neighbouring location, i.e. evaluating the geometrical relationship.

$X_{location-1} - X_m < X_D$ Equation 5

Where:

$X_{location-1}$ …value measured in neighbouring location

X_m ....monitored value

X_D …specified allowed deviation between values

Type 4b: This assessment type checks if monitored values unreasonably deviate from spatially correlated values, i.e. evaluating the topological relationship.

$X_{location\ SC} - X_m < X_D$ Equation 6

Where:

$X_{locationSC}$ ..value measured in correlated location

$X_{m\ ....}$.monitored value

X_D …specified allowed deviation between values

Implementation of Assessment Types

Different assessment types have substantially different requirements with reference to compute intensity and access to other information. Therefore, they can

be implemented in different ways. Simple comparisons, such as assessment types 2 and 3.a can be implemented on filed level in a decentralised way.

In cases of assessment types 4.a and 4.b we need to analyse static data first, before we can identify correlated data points. Those complex checks should be executed and implemented on central components providing sufficient computing power for complex model analysis.

In case of types 1 and 3.b we need to run additional applications (energy simulation tools, or statistical analysis tools) in order to produce the data sets relevant for further analysis. Therefore these assessment types should be also installed and executed on systems with sufficient computing power – preferably at a central location.

Table 3: Compute requirements f. Assessment Types

	Type 1	Type 2	Type 3.a	Type 3.b	Type 4.a	Type 4.b
Low compute power		●	●			
Access to further dynamic data required				●		
Access to further static data required					●	●
Integration with additional software	●			●		

Summary

This section summarises the discussion by mapping the specified assessment types against the potential use cases defined in Section 3. It becomes clear that the compute intensity increases from case 1 (Building Automation) to case 3 (SLA).

This is due to the nature of these use cases; i.e. we need short response times to identify the quality of the current reading in order to provide near real-time feedback to the building automation applications. On the other hand data quality checks for weekly or monthly reports can be executed much longer, since those reports have to be compiled usually on a weekly or even monthly basis.

Table 4: Mapping Data Types to Assessment Types

Data Type	**Assessment Type**
(1) Building Automation	
Comfort Data	Type 3a
Systemic Data	Type 3a
Actuation feedback	n.a.
(2) Verification against legal requirements	
Comfort Data	Type 2
Environmental Impact	Type 2
(3) SLA-verification	
Comfort Data	Type 3.b, Type 1, Type 4.a
Systemic Data	Type 3.b, Type 1, Type 4.b
Consumption Data	Type 3.b, Type 1,
Environmental Impact	Type 3b, Type 2
Occupation Level	Type 3b, Type 2

6. NEXT STEPS

The authors are aware that the formulas presented in chapter 5 are very generic in nature. For the implementation of these formulas the tolerances and "agreed deviation" have to be defined.

As previously discussed, tolerances can be derived from either legislative constraints or manufacturers' manuals. This activity needs intensive literature research and should be complemented by choosing a flexible implementation style, allowing the extension of the scope of these definitions.

Agreed tolerances between spatially or temporally collocated systems or components needs further experimental research in order to guide the definitions of those values. This could be complemented by efforts to categorise systems, components and building types according to climatic zones and observed use-patterns.

7. ACKNOWLEDGEMENTS

This work has been developed with support of European funding provided as part the seventh Research Framework Programme to the projects CAMPUS21 and BaaS.

We would like to especially thank our colleagues from TU Wien, TU Crete, and Honeywell who contributed to the development of Deliverables for work packages 2 in both the CAMPUS21 and BaaS project.

8. REFERENCES

Mahabir Bhandari, Som Shrestha, Joshua New (2012): 'Evaluation of weather datasets for building energy simulation': in: *Energy and Buildings*, vol. 49, pages 109–118.

BuildingSmart International - MSG (2014): 'IFC model specification': *http://www.buildingsmart-tech.org/ifc/IFC4/final/html/*: last accessed May 2014.

CAMPUS21 (2013): 'Campus 21: Deliverable 6.3 - Evaluated Load Balancing Tools': unpublished project report: Cork, Brussels, 2013.

C. Duarte, K. Van Den Wymelenberg, C. Rieger (2103): 'Revealing occupancy patterns in an office building through the use of occupancy sensor data': in: *Energy and Buildings*, vol. 67, pages 587–595.

Goodway. (2013): 'The next frontier for automated FM with Automated Continuous Commissioning': retrieved from: *http://www.goodway.com/hvac-blog/2009/06/the-next-frontier-for-building-energy-management-automated-continuous-commissioning/:* February 24, 2013.

K. Mo, K Menzel, S. Hoerster (2103): 'Development of An IFC-Compatible Data Warehouse for building performance analysis', in: *Proceedings of CIB W78, Beijing 2013*: pages 837 and following.

V.I. Soebarto, T.J. Williamson (2001): 'Multi-criteria assessment of building performance: theory and implementation': in: *Building and Environment* vol. 36, pages 681–690.

C. Turner, M. Frankel (2008): Energy Performance of LEED for New Construction Buildings, 2008.

J. Yudelson (2010): Greening Existing Buildings, Green Source/McGraw-Hill, New York, NY, 2010.