SEMANCO Semantic Tools for Carbon Reduction in Urban Planning

Deliverable 5.1
Building Extraction and Classification Tools

Revision: **6**
Due date: **2013-5-31 (m21)**
Submission date: **2013-7-31**
Lead contractor: **UoT**

Dissemination level		
PU	Public	X
PP	Restricted to other program participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Deliverable Administration & Summary	
No & name	**D5.1 Building Extraction and Classification Tools**

Status	Final review	Due	m21	Date	31/5/2013

Author(s)	Martin Carpenter, Agnieszka Knera, Amit Mhalas, Nashwan Dawood, Tracey Crosbie, (UoT), Leandro Madrazo, Alvaro Sicilia (FUNITEC), Tomas Karlsson (Agency 9), Gonzalo Gamboa (CIMNE), Thomas Rønn, Nadeem Niwaz (Ramboll)
Editor	Martin Carpenter (UoT)
DoW	**T5.1 is described in the following way in the DoW:** "Building stock energy modelling tool: integrating. This task will integrate Geographical Information Systems (GIS) with map digitisation and photogrammetry to develop a tool for identification and classification of buildings within a given geographic area. HALCON image processing software will be integrated in the GIS platform 3DMaps by Agency 9. The result of this integration will be packaged as a tool. The input too of this tool will require the purchase and/or conducting of aerial photogrammetry of at least one of the case study areas and will therefore utilise the work subcontracted as detailed in Appendix A2. As far as it is feasible, an automated building identification tool will be developed. The identified buildings will then be classified based on their type, age, size, fuel used, etc. using the most relevant classification for each of the case studies identified in WP2. The classified data will be stored in the Semantic Energy Information Framework (SEIF)**. D5.1 is described in the following way in the DoW:** "Building extraction and classification tools: A prototype building stock energy modelling tool that will integrate image processing software with the GIS platform. This tool will as far as it is feasible automate the identification of building stock for energy modelling."
Comments	

Document history			
V	**Date**	**Author**	**Description**
1	2012/11/07	Martin Carpenter (UoT), Tracey Crosbie (UoT)	Preparation of original draft, internal discussion as to how to deal with delayed LiDAR data.
2	2013/04/02	Martin Carpenter (UoT)	Preparation for inclusion of LiDAR data.
3	2013/06/11	Martin Carpenter (UoT)	Further edits in response to comments from the coordinator
4	2013/06/12	Tracey Crosbie (UoT)	Final edits to proposed draft before internal review and the addition of further referencing/discussion. Decision taken to postpone delivery for a month to allow for LiDAR and software from all three case studies.
5	2013/06/18	Martin Carpenter (UoT)	Edits in response to the final delivery of the LiDAR data, and restructuring to prepare for inclusion of all case studies
6	2013/07/29	Martin Carpenter (UoT), Gonzalo Gamboa (CIMNE), Thomas Rønn, Nadeem Niwaz (Ramboll), Leandro Madrazo (FUNITEC)	Inclusion of sections relating to North Harbour and Manresa. Final edits and proof reading before internal review.

Disclaimer

Table of Contents

EXECUTIVE SUMMARY

This report presents a description written of the work undertaken as part of Task 5.1 *Building stock energy modelling tool* within the SEMANCO project, and a description of the Deliverable 5.1 *Building extraction and classification tool*. Both D5.1 and its description contained in this report are public.

The work conducted during T5.1 focused on two principal issues:

- The development of 3D building models for each of the three case study areas;

- The development of 'building stock energy modelling tools' for each case study area.

The development of the 3D city models formed a major component of the work conducted during T5.1 and covered both the development of the models and work on supporting specific features within the 3D geographic information software used (3DMaps, from Agency9). In developing the map for the Newcastle use case it proved necessary to conduct a LiDAR survey of the relevant area. The potential need for such a survey had been anticipated within the original description of work.

The remainder of the document is devoted to the presentation of the building stock energy modelling tools which have been developed for each of the three case studies involved. Each of these tools is designed around the primary purpose of identifying the baseline energy consumption and CO_2 emissions of a given urban area. Such baseline measurements provide crucial knowledge for targeting potential interventions, and in particular serve as the basis underlying the work in Task 5.3: *Energy simulation and optimisation tool*.

The differing requirements of the three case studies, as outlined in Deliverable 2.1 *Report on the case studies and analysis* and Deliverable 8.1 *Implementation Plan*, in turn require differing tools.

- For Manresa the focus is on evaluating the overall energy efficiency of an area of housing, both newly built and pre-existing, with a focus on how the shadows from the buildings affect the cooling requirements of those buildings,

- For North Harbour the primary purpose of the tool is establishing the energy demand and energy supply of an urban area of newly built city,

- For Newcastle the focus is on identifying the energy efficiency of individual units of already existing housing.

Chapter 2 is devoted to a description of each of these tools and in particular on introducing and justifying their distinct purposes. The following chapters are then devoted to a more detailed discussion of the individual aspects of each tool.

Chapter 3 discusses how the various parameters required by each tool are derived from a combination of open access data and expert knowledge. This includes both open access data and the three dimensional maps developed during the project.

Chapter 4 closes the deliverable by presenting a discussion of how feasible it might be to use publically accessible data in order to move towards a fully automated method for rating the energy efficiency of housing stock.

Appendix A reports on the work involved in developing the three dimensional urban models for each case study. Appendix B reports on some of the work involved in developing the 3DMaps software to cater for the specific needs of the SEMANCO project.

1 INTRODUCTION

1.1 Purpose and target group

The purpose of this report is to provide a written description of the work undertaken as part of Task 5.1 *Building stock energy modelling tool[1]* within the SEMANCO project, and a description of the Deliverable 5.1 *Building extraction and classification tools[2]*. Both D5.1 and its description contained in this report are public.

This report focuses on two principal issues:

- The work undertaken to develop the 3D urban models required for each of the three case study areas.

- A detailed description of the development of 'building stock energy modelling tools' for each case study area.

This report contains the following sections:

- Chapter2 contains a description of the building stock energy modelling tools for each of the three case study areas. This describes both how the specific features of each case study mandate the nature of each tool there in described and provides a brief overview of the ultimately intended functionality of each of these tools.

- Chapter 3 contains a description of how each of the tools currently derive their input data, which includes the use of open access data of various forms and the three dimensional maps developed for each case study area.

- Chapter 4 contains a theoretical discussion regarding how feasible it would be to move from the remote, but human mediated, evaluation process described in chapter 3 to a fully automatic one.

- Appendix A describes the technical work conducted on the development of 3D maps.

- Appendix B describes some of the technical developments that Agency9 have made to their 3DMaps software in order to support further work within the SEMANCO project.

1.2 Contribution of partners

The two principal partners within involved in T5.1 were University of Teesside and Agency9, while Ramboll, CIMNE and FUNITEC played an important part.

Agency 9 led the creation of the 3D city models using their software 3DMaps and were responsible for making certain technical additions to their 3DMaps software to support both this work task and other tasks within WP4 *Semantic Energy Information Framework* and WP5 *Integrated Tools*.

Ramboll, CIMNE and UoT were each responsible for work relating to the development of the 3D urban models relating to, respectively, the North Harbour, Manresa and Newcastle case studies. In the case of Newcastle this involved the commissioning of a LiDAR survey.

---

[1] Henceforth T5.1
[2] Henceforth D5.1

Ramboll, CIMNE and UoT were additionally responsible for producing detailed specifications of the energy modelling tools for their corresponding case study areas. UoT were additionally responsible for the implementation of the Newcastle tool in software.

In the case of Manresa the decision was made to use an existing tool, URSOS. The integration of this tool within SEMANCO has been outsourced and FUNITEC are/were responsible for arranging and coordination this work. This is expected to complete within Task 5.5.

In the case of the Copenhagen tool, Ramboll produced a tool which ran in Excel and FUNITEC were responsible for transforming this into software appropriate for integration within the integrated platform.

FUNITEC has coordinated the integration of the tools developed for each case study in the SEMANCO platform which will be the outcome of Task 5.4 *Prototype of the integrated platform*.

1.3 Relations to other activities in the project

The baseline stock energy modelling tools produced in Task 5.1 will provide baseline estimates of the energy consumption and CO_2-emissions of existing housing stock or new urban development within each case study area. This data is required to quantify the effect of the baseline in each case study area which will form the basis for interventions to reduce CO_2-emissions in the further development work undertaken within Task 5.3. In addition it forms a major component of the integrated platform being developed within the SEMANCO project and so into tasks 5.4 *Prototype of the integrated platform*, and 5.6 *Integrated platform*.

The work relating to the development of the city models is a vital step within the project. In particular these models are a basic requirement for the development of the integrated platform and so for Task 5.4 and Task 5.6. In a similar manner the technical enhancements made to Agency9's software are required to support this work.

The task received input regarding requirements from deliverables D2.1 *Report of the case studies and analysis*, D2.2 *Strategies and indicators for monitoring CO_2 emissions* and preliminary work arising during work in T6.1 and T8.1. Most specifically, D6.1 provided input regarding the choice of modelling methodology and feedback regarding the design of the interface. As part of the software demonstrators being developed during SEMANCO it provides a contribution to work packages 6, 7 and 8. Most specifically it helped to provide inputs for D8.2.

Finally the production of high quality city models for each case study will, by providing visually convincing demonstrators, form a vital contribution to both the dissemination and user validation activities within the SEMANCO project.

2 ENERGY MODELLING TOOLS

2.1 Introduction

Before considering any form of intervention for improving energy efficiency or reducing CO_2-emisisons from energy supply it is crucial to first establish the existing, or baseline, energy efficiency and energy supply on which you are aiming to improve. This is crucial for not only establishing what level of improvement has been achieved but also for targeting improvement measures at those areas most likely to fully benefit.

To this end, tools for measuring the baseline energy efficiency and energy supply for the three case study areas within SEMANCO have been developed. While the tools all address similar underlying goals, the differing nature of the three case studies – outlined in deliverables 3.1 *Report on the accessible energy data* and 8.1 *Implementation Plan* – means that the concrete expressions of the tools within each case study naturally differ.

The current chapter focuses on describing the conceptual basis of each of these tools and in addition focuses on describing the basic input and output of each tool. In summary the three tools have the following form:

- The Manresa tool focuses on evaluating the baseline energy efficiency of a group of both newly constructed and existing buildings. In addition to examining aspects relating to the energy efficiency of each specific building it calculates the collective effects of the building shadows. Since it is applied within the context of a warm climate, where summer cooling is a major source of energy consumption, this can be a crucial factor to consider. An interface between with the URSOS software will be developed based on the work conducted in this task.

- The North Harbour tool focuses on planning the construction of new areas of city with an emphasis on their energy efficiency and energy supply. The tool produces two major outputs. The first is a "baseline" measurement of the current state of the city. This "baseline" is defined as the quantification of CO_2 emissions based on energy consumption within the city's defined geographical area for a given year. CO_2 emissions are mapped from energy fuel sources and the respective reduction potentials.

  Furthermore, a "baseline projection" is produced. This represents the development in energy consumption and CO_2 emissions towards for example 2020 or 2030 (in principle for any future year supported by data) if the municipality/city does not carry out additional activities that could affect the climate. Already adopted national and local initiatives and policies are usually included in the baseline projection. This tool is embedded in the SEMANCO platform.

- The Newcastle tool focuses on evaluating the baseline energy efficiency of individual, pre-existing houses. It also focuses on establishing a baseline of data allowing for evaluating the effectiveness of potential improvements to those buildings. This tool is based on the Standard Assessment Procedure (SAP) and it has been embedded in the SEMANCO platform.

The current level of integration of each one of the tools in the integrated platform differs. The Newcastle and Ramboll tools have been embedded in the SEMANCO platform. The interface with the URSOS based tool will be provided by Task 5.5 *Interoperable tools with SEIF*.

2.2 The Manresa Tool

2.2.1 Motivation

As mentioned in D2.3 *Impact validation*, one cannot study and analyse the behaviour of an urban element in isolation from its context. There is the need to perform simultaneous assessments at different scales in order to make decisions in the urban planning domain. Basically, it may happen that an urban plan may have good results at one scale, but may not be adequate at a different one.

For instance, consider the case of an urban project aimed at increasing the number of storeys of a building in order to increase the solar heat gains and to consequently reduce the energy consumption for heating. At building level, the plan would produce a reduction of energy consumption, and therefore, of CO_2 emissions. However, raising the height of the building may increase the shadows cast on neighbouring buildings (i.e. it competes with other buildings to capture solar energy) and so might increase the over energy consumption of the neighbourhood.

The need to consider the synergistic effects of the shadows cast by buildings is especially prominent in hot, sunny climates such as those encountered in Manresa and motivated the original development of the URSOS tool, and its selection for integration within the SEMANCO platform.

URSOS is a pre-existing piece of software aimed at providing an energy analysis method that allows urban planners to optimize energy demand of a group of buildings. As the basic input to its calculations, URSOS takes the following:

- A detailed set of parameters for determining the stand alone energy efficiency of each individual building,

- A three dimensional model of the modelled urban area.

URSOS then uses the three dimensional model to take into account the microclimate, location (road networks, building area), building types and shadows in order to calculate the energy performance of the overall urban area. Its main outcomes are the following:

- Energy demand for heating and cooling in the buildings in the modelled urban area.

- Total energy demand of the urban development, for heating and cooling.

- Information on built areas surfaces and buildable plots.

- Detailed calculations of shadows for any closure and time during the year.

URSOS is oriented to use by municipalities, energy utilities, consultancy companies and academic organizations in order to perform energy calculation of urban plans at the level of derivative plans. By catering for both the static efficiency of the buildings involved and the effects of their shadows on the solar gains, and so heating and cooling requirements, it allows for the holistic evaluation of plans for entire urban areas.

The work relating to this tool within the SEMANCO project and T5.1 concerns the integration of the calculation engine of an existing tool into the integrated platform and the

investigation of methods for using the SEIF to link pre-existing data to the calculation model within URSOS.

2.2.2 Basic Methodology

The principal idea underlying URSOS is that it should consider both the effects of the static efficiency of housing stock and the effects of their shadows on the heating/cooling requirements for other buildings within the given urban area. This is of particular relevance in hot, sunny, areas such as Manresa.

In order to calculate the individual efficiency of buildings, URSOS works with a statistically simplified method. The calculation method is especially suitable in urban planning stages where the specificities of the plan are not fully defined. For instance, in the so-called Derivative and/or Special plans within the Spanish urban planning framework. In this stage of the urban planning process the user knows the street layout and the basic shape of the buildings. However the detailed type of windows, doors, woodwork or heating system are not yet defined. Therefore the method works with estimated properties of enclosures, energy system and comfort conditions of the building.

URSOS is based on the Method 5000 to calculate the energy demand for heating. The method has been modified in order to consider the internal gains and dynamic of solar radiation for cooling calculations. It therefore meets norm UNE-EN832.

In order to calculate the energy demand for cooling, the software uses an adaptation of the ASHRAE 1997 method. This modification uses a static calculation procedure that considers average monthly figures of temperature and solar radiation. In general terms, the method considers the following heat losses:

- Through enclosures that separate the building from the exterior,
- Heat loss due to air renewal.

The method considers the following types of heat gain:

- Those resulting from solar installations.
- Internal gains - those due to the metabolism of occupants, lighting, thermal dissipation of electric appliances.

In addition to this static evaluation of the energy efficiency of the building, URSOS performs a situational calculation of the effects of shadows on it. This especially affects the calculation of the expected solar gains, and so the need for heating/cooling the building.

The calculation of the effects of shadows is detailed, allowing for differing seasonal effects, and two principle elements are considered. The first is shadows from the elements of the same building (Figure 1) and the second is shadows arising from other urban elements (Figure 2).
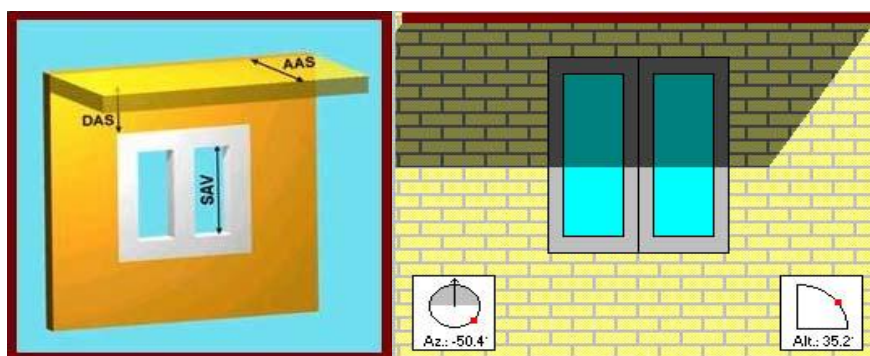
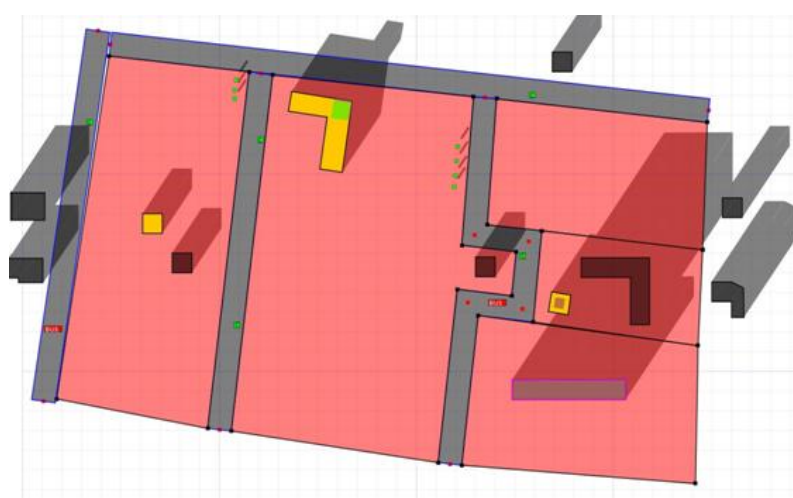*Figure 1. Simulation of shadows from elements of the building (e.g. overhangs)*

Figure 2. Simulation of shadows from urban elements (e.g. other buildings)

2.2.3 Data Input

In order to perform an URSOS calculation for an urban area, two distinct elements of data must be entered:

- The building parameters which allow for the evaluation of the static energy efficiency of each building.

- A three dimensional model of the urban area in question. In principle this can contain a mixture of pre-existing housing and proposed new construction.

The interface of the URSOS software within the SEMANCO integrated platform will thus have two elements. The first concerns the input of the parameters for defining the static energy efficiency of each individual house. These forms have been developed and a screen shot can be seen in Figure 3 below.

The second concerns the need to have a developed three dimensional model for the Urban area under consideration. For pre-existing housing, this work was covered by the development of the three dimensional map for the Manresa case study which is described in Appendix A. For new build housing the interface of the URSOS tool relies on work being developed as part of the SEMANCO integrated platform to be described in D5.4.

This platform will feature tools for both drawing new building shapes directly within the platform and for the import of drawings from existing CAD/GIS systems. In this way it will be possible to construct models of potential new buildings for evaluation with URSOS and to cater for the effects of their shadows.

Figure 3. Screen shot of data entry form, Manresa tool

2.3 The North Harbour Tool

2.3.1 Motivation

Energy efficiency and CO_2-reductions have been subject for thorough discussion throughout the last couple of decades. The focus has led to political action, both in the EU and in Denmark. Energy efficiency and renewable energy is a central part of Denmark's energy policy. The government actively seeks significant energy reductions within all fields and to replace energy supply based on fossil fuels with renewable energy. The CO_2 reduction goal for 2020 is 40 % compared to 1990, and for that reason the Danish government has subscribed to a 12 % reduction of the gross energy use in 2020 compared to 2006 and at least 30% of the final energy consumption should be based in renewable energy (Danish Energy Agency, 2012).

In order to meet these emissions reduction targets it is important that methods are developed for precisely identifying the most efficient ways of reaching the end goals. One of the ways to reach these goals is to make sure that future cities are as energy efficient as possible from the very beginning. The planning of the city is vital, as the building stock, in particular the housing sector, plays a major role in achieving the abovementioned targets. Reducing the

need of future renovations and measures can save money and time that can be used elsewhere.

The above mentioned reasons have led to increased focus on energy strategies for new urban development projects as the North Harbour project. The Urban Energy Planning tool is based on an excel spreadsheet which compiles the methods and calculations Ramboll has undertaken in the energy strategy for the development of urban areas in general. This methodology has been applied before by Ramboll to develop energy scenarios using a LEAP-based model.

Hence, even though the tool focuses on the North Harbour area, it can be applied to other urban development projects in order to develop energy strategies for those specific projects (e.g. SEAP's, Climate Change Plans).

The tool is able to make projections of energy demand and CO_2-emisisons based on data describing for example the characteristics of an urban environment (e.g. the amount and types of buildings), specific energy intensities for different building typologies, CO_2-emission factors for electricity and heating.

The results produced by this tool will serve as the baseline for the projected energy demand and CO_2-emissions. The main items of work relating to the development of this tool within SEMANCO are the development of a software version of this tool from the Excel based prototype and the integration of this software component into the integrated platform.

The tool produces two major outputs. The first is a "baseline" measurement of the current state of the city. This "baseline" is defined as the quantification of CO_2 emissions based on energy consumption within the city's defined geographical area for a given year. CO_2 emissions are mapped from energy fuel sources and the respective reduction potentials.

Furthermore, a "baseline projection" is produced. This represents the development in energy consumption and CO_2 emissions towards for example 2020 or 2030 (in principle for any future year supported by data) if the municipality/city does not carry out additional activities that could affect the climate. Already adopted national and local initiatives and policies are usually included in the baseline projection. There is of course some uncertainty in a baseline projection, since it is dependent on various forecasts and assumptions which may prove to be incorrect. For example, the future fuel prices are important for the development of energy consumption and choice of energy supply technologies. But in spite of the uncertainties, it is still essential to compare the effect of the various future plans and projects against the baseline projection. An example of such a project can be seen in Figure 4 below.

2.3.2 Basic Methodology

One of the key challenges of future energy planning for new urban areas is to implement high energy efficient standards for new buildings and energy supply based on primarily renewable energy technologies.

The Urban Energy Planning tool uses specific data, methods and calculations for determining the energy demand for a new urban area based on certain building characteristics, building typologies and specific energy intensities.

The basic inputs needed to define an urban area to be rated by the tool are twofold. The first is a basic three dimensional model of the area concerned, with features such as the height, floor area and conditioned space of each of the buildings within the urban area defined.

The second set of inputs relates to the further data required in order to make a realistic projection of the future energy use, and CO_2 emissions , of each building. In brief these are given below:

- The use to which the building will be put (e.g. the building typology).

- Specific Energy Demand – This is of importance as the age and type of a building influences its use of energy.

- Net Energy Demand – by using the Building Characteristics (such as the amount of conditioned space) and the specific energy demand of a building, the net energy demand can be calculated for a specific building.

- Energy Supply - By knowing the type of energy supply technology and the efficiency of that technology (Coefficient of Performance) and considering distributions losses in the energy system the total energy supply can be calculated.

- Emission Factors (EMF) – When the total energy use (e.g. electricity and heating) is known, EMF's are needed to "translate" the energy use in MWh/year into tonnes CO_2/year.

These inputs provide the basis for a projection of the energy use and CO_2 emissions for each individual building and so, when the buildings are aggregated, the entire urban area under evaluation over time. There is of course some uncertainty in such a baseline projection, since it is dependent on various forecasts and assumptions which may prove to be incorrect. For example, the future fuel prices are important for the development of energy consumption and choice of energy supply technologies. Figure 4 below shows a graph showing a typical baseline projection generated by the tool.
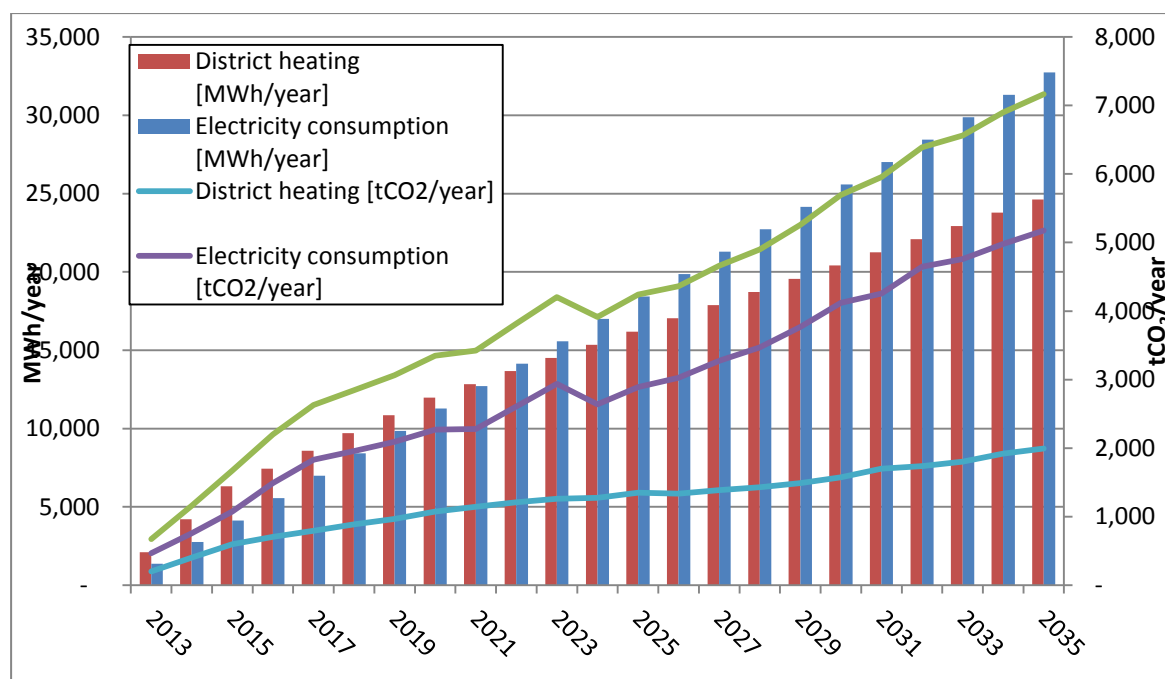

Figure 4: Projections of future Electricity Consumption and amount of District Heating for the North Harbour Area. Furthermore Figure 3 shows projections of total MWh used, and the CO2-emission in Tons/year

Going forward it is planned to build on the foundations of this tool to allow for the analysis of plans and projects to investigate how the energy demand can become more efficient and CO_2-emissions reduced. This further development of the tool will be documented in later deliverables (e.g. D5.3).

2.3.3 Data Input

The screen-shot below in Figure 5 is taken from the integrated platform and shows the input data needed (e.g. building characteristics, energy supply) to calculate the net energy demand, energy supply, final energy consumption and CO_2-emisions for an individual building.

Figure 5. Screen-shot from the technological platform showing input data and results for the UEP tool

Building geometric data from the implemented 3D model covering the North Harbour case study area has been made available by Agency9. As described in Appendix A the original CAD 3D of the North Harbour area provided by the architects only shows building volumes (basic geometric data, such as height, width and length). As a first step to apply the tool, it has been necessary to transform the CAD model into a 3dMaps model. Linking data from the 3D-model with the UEP-tool in order to estimate the total floor area (m2) for each building is explained in Section 3.2.2.

In addition to this use of pre-existing data, the SEMANCO integrated platform will offer the ability to create new building shapes using a drawing tool which will form part of the integrated platform and will be discussed within D5.4. This will enable to model and evaluate multiple alternative plans for a given area.

2.4 The Newcastle Tool

2.4.1 Motivation

The prototype tool developed for use within the Newcastle use case focuses on estimating the energy performance of existing housing stock. European Policy aims to reduce domestic greenhouse gas (GHG) emissions by 80% by 2050 (compared to 1990 levels). The building stock, in particular the housing sector, plays a major role in achieving these 20-20-20 strategic targets. Across Europe the need for renovation and energy measures in housing estates particularly those from the 1960s and 1970s is well recognised (Pittini & Laino, 2012). In the UK the problem is somewhat deeper as it has some of the oldest and least energy efficient housing stock in the Europe (Boardman, 2007). In the UK, as in much of the EU there is a lack of data on the quality of the built fabric of existing housing stock (Mhalas et al., 2012).

In the UK buildings contribute almost half of all CO_2 emissions with 26% of these emissions coming from approximately 26 million residential dwellings. Since, even before the current economic troubles across Europe, the building stock renewal rate in the UK was only 1-2 per cent[3] it is clear that many of these existing buildings will still be present for the foreseeable future. Indeed it is expected that about 75% of the existing UK domestic stock will be still present in 2050 (Boardman, 2007; Wright, 2008).

In order to meet the UK's 80% CO_2 emissions reduction target it is thus a primary requirement to develop efficient methods for precisely identifying both those elements of the existing housing stock that are most need in improvement and how these might be improved (Mhalas et al., 2012). The government has legally mandated a standard way of measuring this called the standard assessment procedure [SAP]. SAP provides indicators of energy performance through energy consumption per unit floor area, an energy cost rating (the SAP[4] rating), an Environmental Impact rating based on CO_2 emissions (the EI rating) and a Dwelling CO_2 Emission Rate (DER).

Currently the normal method of assessing the SAP rating of existing property is to send a qualified engineer to visit the property to conduct detailed measurements of all of the parameters required by the SAP model. This then enables a calculation to take place.

The drawback with such an approach is that such engineer visits require notable amounts of time and money. As such only a limited number of the pre-existing housing stock in the UK has been classified using the SAP model. Since the rate at which houses are replaced with new ones within the UK is between one and two percent per year and SAP is the legally mandated standard by which energy efficiency must be measured, this represents a substantial obstacle to evaluating and so to improving the energy efficiency of the UK's housing stock. User validation work conducted during Task 6.1 clearly identified SAP as the calculation methodology of choice within the UK context.

---

[3] See: http://www.bis.gov.uk/assets/foresight/docs/energy/energy%20final/ravetz%20paper-section%205.pdf;
http://www.eci.ox.ac.uk/research/energy/downloads/40house/40house.pdf
[4] http://www.bre.co.uk/sap2009/page.jsp?id=1642

The goal of the tool developed in Task 5.1 is to make the process of rating houses both quicker and cheaper, while ensuring that the results produced remain aligned with the legally mandated SAP methodology. In doing so the tool will supply the data required for targeting houses in need of energy efficiency interventions. It will also supply the base line energy performance data required to support the work of Task 5.3 in evaluating the types of improvements that will be most beneficial in terms of cost and CO_2 reduction for particular houses.

2.4.2 Basic Methodology

When developing the tool for the Newcastle use case, there were two main requirements:

- That the software should produce results that closely match the SAP rating that would be produced by an engineer visiting the house in question,

- That this happens much more cheaply and easily than at present.

In order to understand how this was done it is helpful to first consider the SAP model itself. The SAP model is designed to be used for both new build and existing housing stock. As such the fullest version of the SAP, which is designed for use in new build housing, expects to take as input a detailed description of the geometry and fabric of a building. This includes factors such as:

- The materials used for construction of the main aspects of the dwelling such as walls, windows and roofs.

- The size of the building together with that of any windows or floor space.

- Thermal insulation characteristics of the building fabric.

- Ventilation characteristics of the dwelling and ventilation equipment,

- Efficiency and control of the heating system(s).

- Gains into the dwellings from solar, metabolism, cooking and lights and appliances.

- Fuel used to provide space and water heating, ventilation and lighting.

- Energy for space cooling, if applicable.

This is then put through a detailed energy physics model before results for that specific building are produced. In total the full SAP spread sheet requires the input of several hundred pieces of information.

While the accurate gathering of such a quantity of data is reasonable when developing new housing, it is not a realistic goal to expect to gather all of it when rating existing buildings. Indeed, this need has been recognised by the designers of SAP who created a reduced version of the model – RDSAP. This model uses precisely the same calculation engine underneath as SAP, thus ensuring that its results are broadly comparable. The difference lies in the quantity of data which must be entered to use the approach. RDSAP is described over 45 pages of the SAP specification and contains detailed charts saying how values that would otherwise be gained by detailed measurement should be approximated using other, more easily measured, values. For example:

- Instead of measuring each individual window in detail, including its U and g5 values, an evaluation is made as to whether the dwelling has a 'typical' quantity of windows

---

[5] Measures of thermal insulation and solar energy transmittance respectively.

or less/more than this. In addition the average glazing quality of the windows is chosen from a small, fixed list of options. The estimate of how typical the area is, is then used alongside the floor area of the dwelling to estimate the total area of windows. This is then combined with estimated U/g values derived from the glazing quality estimates to produce the heat loss and solar gains expected from the windows of the dwelling.

- Instead of providing detailed values for the U values of the walls, they are estimated using the type of construction, age of construction for the dwelling and the quantity of insulation added.

RDSAP thus requires the input of rather fewer input values than SAP, and also ones which require less detailed measurement. The basic assumption underlying the model however remains that an engineer will visit each property which must be assessed in accordance with the technique. This represents a considerable cost and uses a lot of time, and thus represents a major obstacle to rating existing housing stock within the UK.

The purpose of the work underlying this tool, which builds directly on earlier research (Mhalas et al., 2012), is to further reduce the scope of RDSAP to produce a methodology allowing for the remote evaluation of the SAP values of houses. In performing this research the principle goals were:

- To identify which of the RDSAP input parameters could be intelligently estimated without visiting the dwelling in person and which could not.

- To develop methods for estimating those parameters which cannot be directly estimated.

- To demonstrate that the SAP calculations conducted using these estimates approximate those produced by engineer visits as accurately as possible.

Therefore this work included validation of the results produced by the prototype tool using SAP values calculated in the traditional way based on input data provided by site visits (Mhalas et al., 2012). A detailed description, backed by screenshots, of which parameters are used for this calculation and how the values for these parameters are calculated is included in Section 3.2.4, while Section 4 discusses the extent to which it might be possible to automate the process of inputting these parameters into the tool.

The tool develops and refines a SAP calculation tool developed in earlier research (Mhalas et al., 2012). In particular the following steps were taken:

- The original SAP calculation tool prototype was developed in an Excel sheet this has been translated into a Java based calculation engine that has been integrated into the SEMANCO platform;

- An interface has been developed and refined in response to direct user feedback;

- The set of data elements which must be inputted to generate an estimated SAP result has been further streamlined through further sensitivity analysis.

The further streamlining of the input data has been considerable and has reduced the number of distinct data points which must be entered from around 70 in the spread sheet prototype to just fewer than 30 in the Java application. The accuracy of this further simplification has been verified by checking that the results of calculations conducted using both versions of the SAP tool align.

The streamlining of data input involved two principal elements. The first was the removal of some input elements which were functionally dependent on each other, or were constant for every dwelling rated. As an example the spread sheet mirrored RDSAP in containing entries for the total number of low energy light bulbs and the total number of standard light bulbs. Since the flats weren't being visited, these numbers were being generated as a function of the total floor area and the proportions of lighting were always being taken as a fifty/fifty mix of each. This assumption had been found to give accurate answers and so was built directly into the SAP prototype tool that is integrated in the SEMANCO platform.

Overall, the software tool delivered as D5.1 represents a considerably streamlined and more user friendly version of the original excel prototype reported on in Mhalas et al. (2012). The ideal would be to further automate the process and so reduce the number of items which must be entered each time. One major potential area for this will be deriving the geometric measurements from the results of the LiDAR survey. This has potential to reduce the number of data inputs required considerably. A detailed discussed of how further automation might be possible is given in Chapter 4.

2.4.3 Data Input

The Newcastle case study tool has been implemented and deployed a standalone tool which can be found running at: http://arcdev.housing.salle.url.edu/saptool/. The screen shots which follow (Figures 6, 7) are taken from this software and illustrate the interface used in order to fill in the information relating to the SAP calculation for a building. As can be seen multiple items of information must be entered, some of which can be seen. A more detailed discussion of each of these elements, together with how they are derived can be found in Chapter 3.

Figure 6. Screen shot of data entry related to the house fabric

Figure 7. Types of result arising from using the Newcastle case study tool

The prevalence of drop down menus is due to the nature of the RDSAP model wherein many of the parameters take values from a given, specified, list. A good example of this is seen with the wall construction type drop menu presented in Figure 6. These parameterised lists serve much the same purpose that building typologies are used for in some other approaches but allows for a more individual approach to each dwelling.

Certain additional features were added as the direct result of user feedback resulting from the testing of an early version of this software. These include the explanatory text at the bottom of each page and the provision of help text when they hover over certain windows.

The results shown in Figure 7 represent a summary of the full set of data generated by the calculation. The full set of data generated during the calculation is stored and will be used within the tool developed within Task 5.3.

While the Newcastle case study tool is, for clarity, presented here as a standalone entity it has already been integrated within the overall platform being developed within SEMANCO. This will be fully described in the forthcoming Deliverable 5.4 *Prototype of integrated platform.*

3 DETERMINING INPUT PARAMETERS – CURRENT STATUS

3.1 Introduction

For each of the tools described in Chapter 2 above, multiple parameters must be considered before an individual house or building can be assigned an energy rating. Since the tools will typically be used to rate large areas of housing or buildings, there is a clear need to try and make this process as easy and fast as possible. The need for this was indeed anticipated in the description of work where it states that: 'as far as possible, an automated building identification tool will be developed'.

This chapter, and the one following it, contain a discussion of how far the process of attaching the parameters required to perform an energy evaluation can be automated. The focus of the current chapter is on discussing how much of this process is, or will be during the current project, automated. The following, and final, chapter contains a discussion of how this process might be further automated in the future.

The focus of the current chapter is on providing intelligent support for the process of manually assigning parameters. While some are fully automated, for many parameters the procedure is that the user uses a combination of freely available information sources and expert knowledge to manually derive a value. This still represents a considerable saving in terms of time and effort over sending an engineer to visit the property concerned.

3.2 Potential Data Sources

In addition to the expert knowledge of the user of the tool, there are three basic sources of information which can be used to aid the classification of buildings:

- General purpose open access data regarding the building.
- Data directly derivable from the three dimensional model.
- Photographic images of the building.

The extent to which each of these is currently being used within SEMANCO will be discussed in turn.

3.2.1 Open Access Data – the SEIF

One of the primary goals of the SEMANCO project, and so the integrated platform, is to demonstrate how the semantic framework developed in work packages 4 and 5 allow for the generic information relating to a building to be automatically attached to it. The principal limitation to its use in the current energy evaluation tools has been the discovery of appropriate open access data sources. A discussion of these can be found in Deliverables 2.1 and 3.1.

For the North Harbour case study the focus is specifically on evaluating the efficiency and energy supply of areas of newly built city and so no open access information exists. In the case of Newcastle, data protection concerns meant that no usable building level data was found.

However the HEED (the Homes Energy Efficiency Database) and EHCS (the English House Conditions Survey), while only providing aggregated information, remained useful. Firstly they provide useful input when the user is deciding on which value to assign to certain parameters. Secondly they are used to enable the targeting of those areas of the city with

high levels of fuel poverty, which are consequently most likely to benefit from refit programs. This functionality relates to the SEMANCO integrated platform as opposed to D5.1 and more details will be given in the forthcoming D5.4.

The situation in the final case, that of Manresa, is more positive as the Manresa census and land registry provide several items of very useful information. These include:

- The numbers of occupants and the percentage of occupation can be retrieved from the Manresa census,

- The use of the ground floor of the building can be retrieved from the land registry data.

In addition to this form of data that can be used directly additional data sources exist which, as in the case of Newcastle, can be used to make intelligent estimates of certain parameters. For example, the household incomes for an area can be retrieved and is used, in combination with expert knowledge, to deduce the likely comfort temperature of the house and the internal heat gains within it.

Additionally, IDESCAT (2008)[6] contains statistical information regarding the likely heating/cooling and hot water systems for given types of houses. When combined with the incomes for an area, this data can be used to support expert decisions regarding which of these systems are likely to be present within a given house.

As a final note, it should be said that any company mandating the deployment of SEMANCO within a specific context may well have access to their own, private, sources of information. For example a company owning a set of rental houses might keep a record of what hot water systems they each use and what fabric refits have been applied to them. The design of the integrated platform and the SEIF would permit such a data source to be used.

3.2.2 The Three Dimensional City Model

Crucial to any energy efficiency calculations for a dwelling is accurate data regarding its size, footprint and the exposed surface area of its walls. The use of three dimensional models within the SEMANCO platform means that such data can, and will, be automatically supplied to the models without requiring the user to enter it at the interface. The only remaining concern relates to situations with multiple flats within a single building, where the distinct floor areas cannot be worked out without visiting to measure. This is a major limitation for systems - such as the SAP model underlying the Newcastle tool – which deal with individual dwellings rather than entire buildings.

In the case of North Harbour, certain adaptations had to be made. Building geometric data from the implemented 3D model covering the North Harbour case study area has been made available by Agency9. The original architects CAD 3D model however only shows building volumes (basic geometric data, such as height, width and length).

In order to estimate total floor area (m^2) for each building a configurable parameter of storey height has been implemented in the technological platform (default value: 3.5 meter for offices and 2.9 for residential buildings) which is divided by the height of the building and rounded into an integer describing the number of complete storeys. This makes it possible to

---

[6] IDESCAT (2008). Social Survey 2008. Households and environment. Accessed on July 5th, from http://www.idescat.cat/pub/?id=lma&n=4.1.1.5

estimate floor area (m^2) by multiplying the built-up (width x length of the building) area with the number of complete storeys.

The extraction of geometric data from the 3D model was compared to the plot ratio and floor area given in the official master structure plan of the developing area and the approximation, an example of which can be seen in Figure 8, was found to be acceptable.

Figure 8. Mock-up of the technological platform showing geometric data from the 3D model

A further use of the three dimensional models concerns the calculation of the effects of building shadows in the case of the URSOS tool. Again certain adaptations were required.

As mentioned in appendix A, the 3D model of Manresa was created from the GIS of the city, whose polygons were extruded according to the number of storeys of each polygon. In the Manresa GIS, buildings are made up by more than one polygon, as the example of Figure 9. In the example, the building is made up of four polygons with different number of storeys.
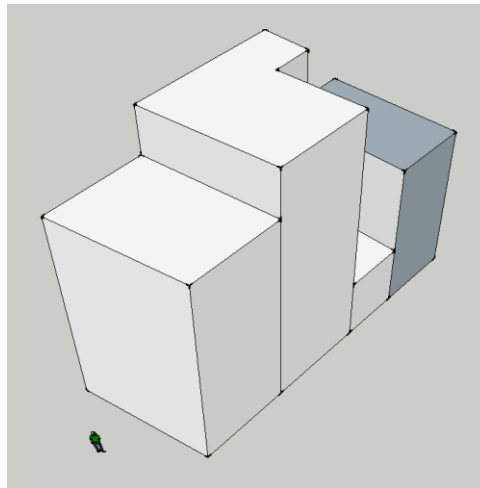
Figure 9. 3D model of a building made up from four extruded polygons

In order to apply the URSOS calculation engine, the building of 9 should be defined using certain standard URSOS structural elements (i.e. towers and inner courtyards). However, this option entails that certain building geometries are nearly impossible to define, which complicates the process significantly. Therefore, and after testing different ways of simplifying building geometries, it was decided to keep the extruded polygons as such and simulate their energy performance as if they are different buildings. Then, the overall energy performance of the real building is the result of the aggregation of its parts. In order to do so, the different extruded polygons making up a building are identified using a similar ID

(Building-Code_Polygon-Code). Also, some energy related parameters of the buildings (e.g. U-values of dividing walls) are defined proportional to area in contact between extruded polygons.

Finally, the 3D model itself can also be used to help to identify the types of housing involved, giving information relating to the likely energy efficiency of those houses. Figure 10 contains the blocks representing each house in the Newcastle case study area, with aerial photography indicating what type of housing this relates to. Again this information provides a useful data source to complement the data derivable from the use of street level photography.

Tyneside Flats
(pre-1915 mix of private rental / owner occupied 'duplex' and single family terraced residential)

Terraces
(mid 1960s mix of local municipality owned / ALMO managed rental with limited owner occupied Right To Buy properties)

Courtyard Blocks
(1970s municipality owned / ALMO managed and owner occupied houses / multiple occupancy corner flats)

Tower Blocks
(2008 refurbished multiple occupancy residential, municipality owned / ALMO managed)

Figure 10. Types of housing, Newcastle, Three dimensional model

A final, and major, consideration relating to any use of these models concerns the considerable effort required to produce a model, and to highlight the individual houses within it. This was true for each of the case studies within SEMANCO. When considering the potential use of such models, these costs must be carefully weighed against the significant benefits arising from their use. In some cases cities might already possess appropriate three dimensional models, thus removing many of the costs involved.

3.2.3 Open Access Photography

The third source of information lies in photographic information relating to the building. In this case there is high quality, relevant, open source data – Google street view for example. In the case of SEMANCO the aerial photography resulting from the Newcastle LiDAR survey provides an alternative source of information.

The sorts of information that can be gathered from an open access street view photograph, such as Google street view, is covered in more detail in Section 3.2.4 below, but basically includes facades features and construction materials. The accurate interpretation of this requires some expert knowledge. This data source is expected to be used in both Newcastle and Manresa.

In the case of Newcastle there is an additional source of photography because the LiDAR survey also captured high quality aerial photography of the area. This can be seen in Figure 11 below. In addition to illustrating the basic quality of the aerial photography, this figure highlights certain structural features of the houses which can be derived from an expert examination of this photography. This information can complement the use of open source street level photography.
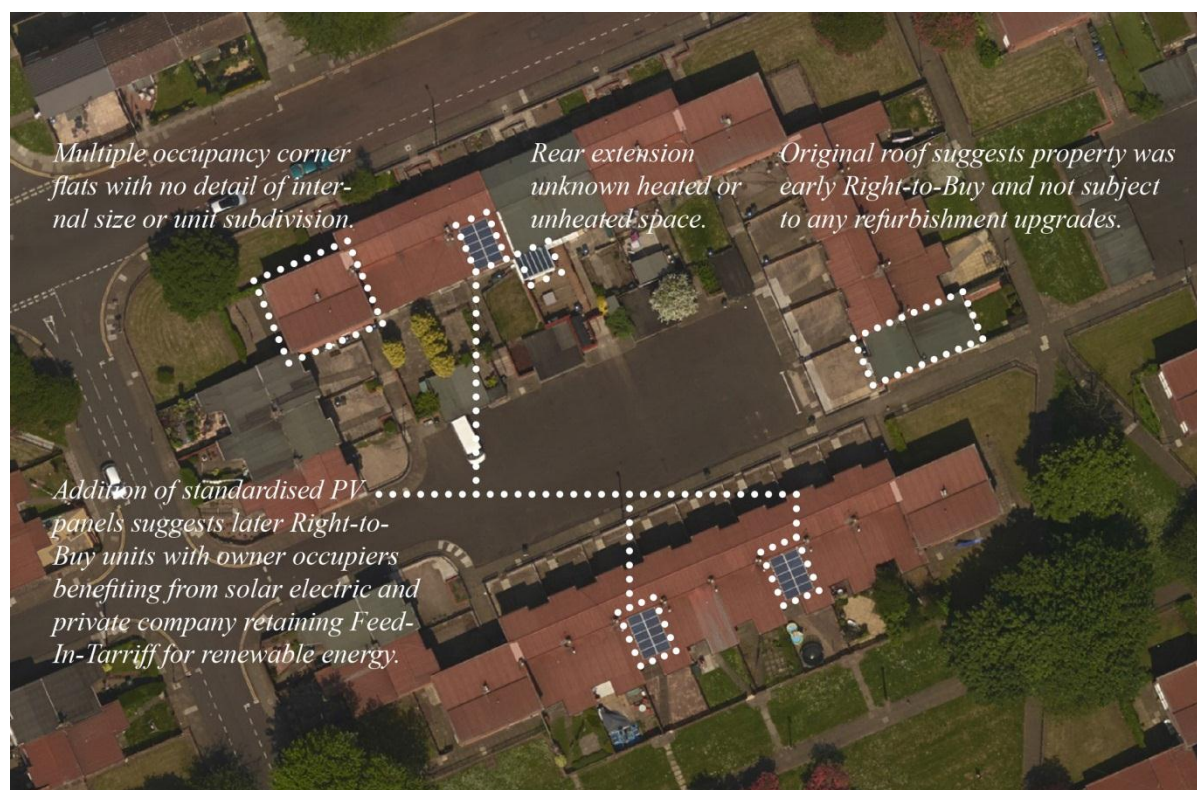

Figure 11. Aerial photography and data derived

3.2.4 Detailed Data Type Derivation, Newcastle

In order to summarise how each data type is used in a specific case, a detailed listing of the parameters involved in using the tool from the Newcastle case study is now given.

- Property type – Number of sides sheltered, Fraction of windows in each direction – each of these values can be easily manually derived from the use of open access photography. No specialist knowledge is required to do this.

- Number of rooms – This uses expert knowledge to derive an approximate figure from the overall floor area and type of property. Since the effect of this figure within the energy calculations is quite small, it should be possible to use a purely automated calculation without major problems arising.

- Window Area – This requires the evaluation of two pieces of information. The first is the quantity of windows in the dwelling concerned. The second is the 'typical' amount of windows in a dwelling of that type. The information regarding the quantity of windows on the actual dwelling can be evaluated using photographic evidence. This is easy to do manually and possible, if not straightforward, to imagine automating using image recognition. The second involves expert knowledge regarding what quantity of windows are in fact typical. While the SAP standard makes no effort to quantify or standardise this figure, doing this could potentially yield a fully automated system.

- Floors of dwelling – This data, in the form required by the 5.1 tool above, could in theory be automatically provided through the use of appropriate vector floor plan information. Without such detailed information it will not be possible to give distinct areas to each floor in the dwelling, or to rate situations with multiple dwellings in one building. Instead it is envisaged that a combination of the two dimensional floor area, the overall height of the dwelling and the number of floors in the dwelling will be combined to provide this information, with the assumption that every floor within the dwelling has equal floor area being made. Of these items, the height of the dwelling will be provided from the LiDAR data. The number of floors will be manually deduced from the photographic evidence and especially the number of windows. Finally the two dimensional floor area of each dwelling will be automatically provided from the model – this will be present as a result of adding the information required to identify the distinct dwellings within the model.

- Roof orientation, Exterior perimeter, Exterior Wall Area, Roof area, Roof tilt – These values are expected to be directly obtainable from the results of the LiDAR survey. If not they will be straightforward to manually estimate from the three dimensional model produced by the LiDAR survey. If the LiDAR survey was not present then the roof area, orientation and tilt would have to be estimated from photographic evidence. The exterior perimeter, and wall area, would come from the floor plans being used to determine which individual buildings were present within the model.

- Wall type, Window glazing type, Roof type, Age of dwelling, Roof window area, Roof window glazing type – each of these takes values from a list given in the SAP specification. Filling in these values requires an examination of the photographic evidence together with knowledge regarding how each of the potential values of these parameters would look.

- Added Wall Insulation, added roof insulation Efficiency of main system, Water heating boiler type, Water storage insulation, Main Boiler Type, Ventilation – if the company using the SEMANCO platform is the formal owner of the houses in question and has refitted them at least once then these values may well already be known. If they are not then they must be approximated based on expert knowledge of typical values for such homes.

The accuracy of the results of this tool against those produced by engineer visits and official SAP assessments been investigated (Mhalas et al., 2012). The results of using the tool closely approximates the official SAP assessments. The major remaining question for validating this tool is how reliable the results are when it is used by multiple different users. This area remains a matter for future exploration.

4 TOWARDS FULL AUTOMATION

A component of the work description for T5.1 is to "*...as far as it is feasible automate the identification of building stock for energy modelling.*". While the work described in Chapters 2 and 3 of this document goes a considerable distance to meeting this goal, and represents a genuine increase in the amount of automation involved, the question of how close it is possible to get to full automation is open.

The process of fully automating the calculation of energy efficiency for a set of houses directly corresponds to that of fully automating the calculation/approximation of the various parameters discussed in the previous chapter. Such a process only makes sense when evaluating the energy efficiency of existing houses – for a proposed new build the parameters must be attached manually. The discussion within this chapter is therefore most relevant to the Newcastle and Manresa tools.

The very first step in automating the process of automatically evaluating the energy efficiency of a set of houses is identifying the actual set of dwellings to be evaluated. The basic solution to this involves first obtaining the list of post codes and house numbers from the ordnance survey. These can then be used to link to freely available street level photography[7] to identify their features.

This data is typically freely available and for the Manresa case, as can be seen in Section 2.2.3 of D3.4, is integrated into SEMANCO using the SEIF. In the case of Newcastle this data will be integrated into the platform as part of the work involved in highlighting which elements of the three dimensional map correspond to which dwelling.

This naturally leads on to the next major problem – namely identifying the position and geographical extent of each of these buildings. The ideal solution to this problem is to have a detailed ground floor plan map of the area which can then be overlaid onto whichever visual map is being used for the user interface[8]. Unfortunately such maps are not always freely available – while the Ordnance survey in the UK does produce such maps they charge substantial license fees for their use, making their use in SEMANCO impossible.

None the less a specific company deploying SEMANCO might already possess such a license, in which case the data could be incorporated into their deployment of SEMANCO via the SEIF. Otherwise visual clues must be used. In the case of Newcastle the LiDAR data survey has supplied a set of floor print data, which can be accurately verified by using the aerial photography.

Any approaches using purely external photography or ground floor maps share a fundamental limitation in that they cannot cater for either blocks of flats or town houses which have been internally converted into multiple flats. The SAP methodology embedded within the tool for the Newcastle case study deals with individual dwellings and not individual buildings, and so detailed information regarding the floor plans of each individual flat is required. This is not as problematic in the case of the UK as might be expected due to the fact that the largest percentage of the housing stock are single family dwellings[9].

---

[7] Such as Google's street view imaging.

[8] In the case of SEMANCO the three dimensional city models.

[9] There are about 24 million homes in the United Kingdom. Of these 21.8 million are in England, comprising 29% terraces, 27% semi-detached, 17% detached, 9% bungalows, 3% converted flats and 14% purpose-built flats (Communities and Local Government, 2009).

An easy solution to this dilemma would arise if the organisation deploying SEMANCO were to have detailed floor plan information of their buildings which contain multiple distinct dwellings. If they did then this information could be incorporated into the system via the SEIF and automatically used. If such data does not exist then either such buildings must either not be treated at all or the share of them allocated to each flat must be approximated. This might for instance be done by assuming each individual flat in a building to comprise an equal share of the volume.

Having determined the identity, floor plan and location of each property the next step in automating the process would be to determine values for each of the parameters needed for the energy evaluation of the dwelling. As was discussed in Section 3.2.4 above, these are determined through a combination of directly usable open access data, expert knowledge and the interpretation of photographic evidence.

The SEMANCO integrated platform is specifically designed to take advantage of any available open access data and, where it exists, to allow it to be used both directly and easily. However which properties will be linkable to such data and which won't will vary for each organisation deploying SEMANCO and no useful general statement can be made. In the case of the Newcastle case study the LiDAR survey has given rise to detailed three dimensional data for each building.

For each property for which such data does not exist a combination of photographic evidence and expert knowledge must be used. In order to do this the following steps must each be carried out:

- The area of openly available street level photography corresponding to the dwelling in question must be identified.

- The area of this photography which corresponds to the facade(s) of the house must be stripped from the overall photograph.

- The features of interest must be identified on this area of photography.

- They, potentially combined with certain pre provided knowledge, must be used to derive an appropriate value for the parameter.

None of these steps are straightforward. The first one may seem to be relatively easy due to the ability to search Google street view by post code and house number, but reliably facing precisely the right house without human intervention is not a trivial task.

The basic problem involved in the second stage is of a type well known within image recognition. However, since the provision of massive data sets of this type is relatively new, the application and adaption of these techniques to the precise details of this problem is still on-going. One paper examining this is (Recky et al., 2011).

Once the facade has been isolated from the remainder of the photograph, and so trees etc. have been excluded, this would have to be aligned with and combined with floor plan data to isolate the portion of the facade corresponding to the specific house in question. This could be especially problematic in the case of multiple dwellings within a single building.

Once the portion of photography corresponding to a specific building has been isolated it must then be analysed. The process for doing this would typically involve the following steps – a set of characteristic features is identified, means of detecting these on the photographs is developed and an automated method of mapping from vectors of the characteristic features to the corresponding values is developed. This latter process might typically use machine learning to develop an appropriate mapping.

Each of these individual steps would require concerted work, a large number of practical examples and would produce country, or even region, specific results. An example of the potential individual complexity of this work relates to the detection of windows from photography of facades. While this has long been an area of study, open issues remain and new techniques such as those discussed in (Recky & Leberl, 2010) are still being proposed.

While this may be discouraging, examples such as that of (Zamir & Shah, 2010) who used image recognition and pattern matching to determine which city a Google street view image came from confirm that the basic problems could in principle be solved.

The overall conclusion to be drawn regarding the fully automated, remote energy evaluation of houses is that, besides the issues with multiple dwellings in one house, the remaining issues are all soluble in principle but would require a major research effort to do. Earlier research (Mhalas et al., 2012) and the work conducted to produce the D5.1 prototype tool, demonstrates both the principle and accuracy of remote evaluation techniques and there is no theoretical reason that much, or all, of the remaining manual interpretation could not be automated.

One final idea involving automation would be to drastically simplify the initial problem – i.e. to find a method of calculating an estimated energy efficiency which retained plausible accuracy while requiring many fewer parameters to be specified. This would then simplify the amount of research required to automate the process. This is an on-going topic of research within Task 5.2.

5 CONCLUSIONS

5.1 Contribution to overall picture

This report presented the work conducted within Task 5.1 of the SEMANCO project. This comprised two major elements of work:

- The development of the models required to give 3D mapping of the three case studies selected for use within the SEMANCO project.

- The development of specifications for software tools for measuring the baseline energy efficiency within each case study and the implementation of these within software.

- The development of the D5.1 prototype tool.

As described in this document this work has been completed in line with the expectations in the description of work. Three-dimensional city models have been developed for each of the case study areas, with work ongoing to improve these for incorporation in the final integrated versions of the SEMANCO platform. The development of this mapping in the case of Newcastle required the conduct of a LiDAR survey which was substantially late and in turn caused a considerable delay in the production of this deliverable but ultimately provided excellent results.

The software tools for measuring baseline energy consumption and CO_2-emissions within each case study area have been specified and plans for their development and integration within the overall SEMANCO platform developed. They are all expected to be included within the prototype of the integrated platform which is due to be delivered as D5.4.

Finally the Newcastle and North Harbour case study tools have been implemented in the platform. Besides, the Newcastle tool is available as an independent tool.. This tool provides the ability to rapidly and efficiently produce SAP ratings for houses without a visit by an engineer. As such it represents a major step towards managing to rate the large numbers of existing houses which are currently without such ratings. This will allow both the effective targeting of houses in need of improvements and the intelligent fitting of improvements suiting the specific characteristics of those houses.

5.2 Impact on other WPs and Tasks

The work done on the 3D models for the case study sites is the foundation for every piece of software being developed to work on those sites, and as such its success was a prerequisite for many other work packages within the project. This is especially true for the other software based tasks within work package 5.

The work conducted in developing each of the energy efficiency tools will form a crucial building block for the energy simulation and visualisation tools to be developed in T5.3. Finally data provided as a result of the work within T5.1 will form an input to the data mining work envisaged within T5.2 and a crucial part of the integrated platform currently being developed in tasks 5.4 and 5.6.

The delay to the production of this deliverable resulted from the delay in the receipt of the results of the LiDAR survey for the Newcastle area due to persistently adverse weather. This has not adversely affected the schedule of work on other tasks and in particular D5.3 is expected to be delivered on time.

5.3 Contribution to demonstration

The principal contribution of T5.1 to the demonstration of the project is the work done on developing the 3D maps of the three case study areas. Without such maps no graphically convincing demonstrators would be possible and as such this forms a crucial contribution to the demonstration effort within the project. The results of the LiDAR survey provide a very convincing map for the Newcastle area, and thus form good demonstration material.

Finally the tools for calculating energy efficiency will form a major component of the SEMANCO integrated platform, to be described in Deliverables 5.4 and 5.6, and so will also contribute substantially to any demonstrations of that platform.

6 REFERENCES

Boardman, B. (2007) Examining the carbon agenda via the 40% house scenario. *Building Research and Information*, 35(4), 363-378.

BRE (2011) *The Government's Standard Assessment Procedure for Energy Rating of Dwellings*, Watford, UK: BRE on behalf of DECC.

Communities and Local Government (2009). *English House Condition Survey 2007* Annual Report, London: Communities and Local Government.

Gamboa, G., Cipriano, X., Oliveras, J., Niwaz, N., Hvid, J., Lynch, D., Madrazo, L., Sicilia, A. (2012) *SEMANCO Deliverable 2.1: Case Study Description.* The SEMANCO project is Co-funded by the European Commission within the 7th Framework Programme Project ICT 287534. Retrieved July 31, 2013 in Web site: http://semanco-project.eu/index_htm_files/SEMANCO_D2.1_20120330_r.pdf

Mhalas, A., Crosbie, T. & Dawood, N. (2012) Development of a Visual Domestic Energy Assessment Tool presented in *CONVR2012, the twelfth international conference of construction applications of virtual reality,* November 1/2, 2012. Retrieved June 2013, in proceedings from website: http://convr2012.com/download/CONVR2012-Proceedings.pdf

Recky, M., Wendel, A. & Leberl. F. (2011) Façade Segmentation in a Multi-view Scenario. In *Proceedings of the 2011 International Conference on 3D Imaging, Modelling, Processing, Visualization and Transmission* (3DIMPVT '11). IEEE Computer Society, Washington, DC, USA, 358-365. Retrieved July 31, 2013 in Web site: http://dx.doi.org/10.1109/3DIMPVT.2011.52

Recky, M. & Leberl, F. (2010) Windows Detection Using K-means in CIE-Lab Color Space, *20th International Conference on Pattern Recognition (ICPR)*, pp.356-359, 23-26 Aug.

Wright, A. (2008). What is the relationship between built form and energy use in dwellings? *Energy Policy*, 36, 4544-4547.

Zamir, A. R. & Shah, M. (2010) Accurate Image Localization Based on Google Maps Street View, Computer Vision – ECCV 2010, *Lecture Notes in Computer Science Volume 6314*, pp 255-268

7 APPENDICES

Appendix A: Creation of 3D city models

A.1 Introduction

The work within SEMANCO targets three chosen case study sites – North Harbour , Manresa and Newcastle. Detailed descriptions of aspects of each of these case studies can be found in SEMANCO Deliverable 2.1 (Gamboa et al., 2012). The current appendix focuses on the development of usable 3D mapping data for each of these case study areas.

The work presented in both this chapter and appendix B was completed within the originally planned timescale for T5.1 and so by month 18 of the project. Both areas of work have been the subject of ongoing work under other tasks and the results of this work will be reported on in D5.4 and 5.6.

Such mapping data forms an essential prerequisite for the development of any software relating to the building stock within the case study areas, and is thus foundational for the software development within the SEMANCO project. This chapter examines each case study in two distinct contexts:

1. The technical work involved in developing the 3D modelling data, and putting it into a format usable by Agency9's 3DMaps software;

2. The prototype work conducted to integrate semantic data into this mapping data.

A.2 Development of data for 3D mapping

To create a useful 3D city model the following items data are required:
- A **DTM** (Digital Terrain Model) – this defines the topography, representing the bare ground surface without any objects like plants and buildings.
- A **DSM** (Digital Surface Model) – this includes all of the permanent features which are above ground. The heights of buildings are commonly extracted from the elevation data within this model.
- A **Shapefile** (.shp) – this is a georeferenced file which contains the footprints and heights of individual buildings. (Geospatial vector data format).

Further, depending on the amount of detail required, orthorectified aerial photography can add a much greater level of detail, bringing the texture and colour. It can also be used to create models. A final challenge in the current case was the need to ensure that all of the data was in a format usable with the 3DMaps software of Agency9.

Manresa

A basic map for a large area of Catalonia was created using a set of aerial images streamed from a public WMS. The performance of the available WMS server was however not adequate for the sort of high throughput streaming required within SEMANCO. A cached collection of images suitable for use with Agency9's 3DMaps was therefore assembled. This combines low resolution images over the majority of Catalonia with high resolution images over Manresa.

A set of shape layers has also been assembled and added to the map and shape files containing building and property outlines exist. A method for using this data to extrude 3D blocks representing the buildings is currently under development, and will be used to

populate the model of Manresa with such information. The current two dimensional map of Manresa can be seen in Figure A1.

Figure A1. The Manresa city map

This model has recently been developed into a three dimensional one. The Agency9 3DMaps SDK can automatically create a map from a set of raster images and a DTM model. A 3DMaps map consists of a set of 3D meshes structured as a quad tree. The 3DMaps SDK toolkit automatically handles projection of the raster images on a DTM to create 3D meshes as well as the creation of the quad tree structure.

The Manresa 3D city model has been created from the cadastre dataset containing building footprints. The dataset lacks the actual building height but contains information about the number of floors per polygon which has been used to approximate the building height.
The Municipality of Manresa has delivered a single shapefile, named PLCONSTRC53D, which contains the basic geometric information, such as a unique ID for each polygon, the date the polygon was created, a code defining the type of polygon (e.g. a construction, pool, garden), centroid of the polygon and number of storeys.

To create the 3D model a Java tool has been created that takes the building footprint in shape format and extrudes the polygon using the number of floor to create a 3D volume. The

information is then converted and saved as a COLLADA 1.4.1 document with one building per file and as CityGML document. Tessellation of complex polygons is done using GLU. Using the COLLADA 1.4.1 is the optimized into a tile layer using the 3DMaps SDK. The tile layer can then be used with 3DMaps.

One important issue is that, in the case of the Manresa GIS, one building is usually made up by more than one polygon. Therefore, the first version of the Manresa 3D map contains cases such as the one presented in Figure A2. This figure shows the 3D map fed with fake data on energy consumption in order to test the visualization features of the platform. As the reader can see, different extruded polygons are considered as different buildings with different energy performances.
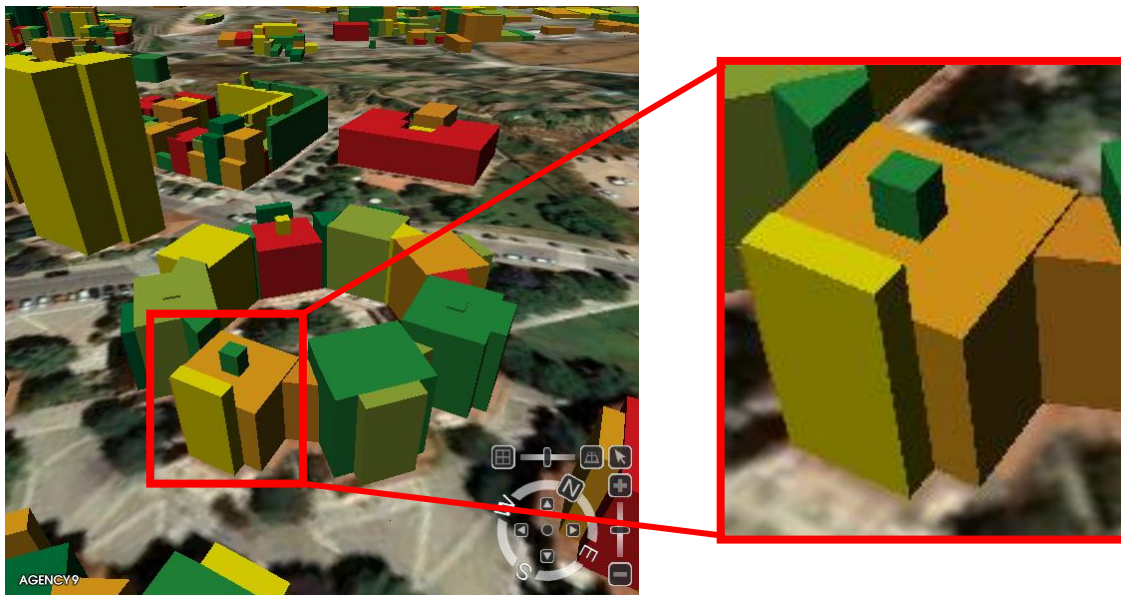

Figure A2. Example of a building made up of several polygons in the Manresa 3D map.

In order to overcome this issue, it was necessary either to merge polygons belonging to a single building or to assign adequate identification codes to the polygons to identify them to the building they belong to. Due to the complications found when simplifying geometries to be entered to the energy performance calculation tool, it has been decided to follow the second option. In a first moment, it was decided to assign adequate ID to the extruded polygons of the current 3D map. However, it was found easier to reconstruct the shapefiles of the GIS and to extrude the polygons again to build an updated 3D map. In order to do so, the following procedure has been performed:

a. Take the shapefile CN_TOP_16, which contains information on the centroids of the constructed polygons. Each centroid has information of the area and perimeter of the polygon, the type of construction and the number of floors.

b. Take the shapefile PLPARCE, which contains information of the urban plots. Each urban plot is represented by a polygon identified with unique land registry codes (i.e. FINREFCAD1 and FINREFCAD2)

c. To add information of land registry codes to the centroid of each constructed polygon. The result is a new shapefile of centroids with upgraded information (i.e. area and perimeter of the polygon, the type of construction and the number of floors + FINREFCAD1 and FINREFCAD2)

    d. Take the shapefile PLMERGED_TOP_16, which contains information about blocks (i.e. set of semidetached buildings). Each block as an unique ID.

    e. To assign a block ID to each centroid. The result is a new shapefile of centroids with upgraded information (i.e. area and perimeter of the polygon, the type of construction and the number of floors + FINREFCAD1 and FINREFCAD2 + BLOCK_ID)

    f. Take the layer LMCOBSTRC5C, which contains information of the lines making up each constructed polygon, and create a new shapefile of polygons. Each polygon is created with the lines surrounding an upgraded centroid. The new polygons contains all the information of the upgraded centroids.

    g. Assign a building code to the constructed polygons. All polygons with the same FINREFCAD1, FINREFCAD2 and BLOCK_ID belong to the same building. The building code has the following format: FINREFCAD1-BLOCK_ID)

Then, the new shapefile of constructed polygons can be extruded according to their number of floors as explained previously. The result is the same than the one presented in Figure A2, in the sense that one building is made up of more than one extruded polygon. However, in this case, the extruded polygons belonging to the same building have a common building code.

Additionally, it was delivered a Digital Terrain Model (DTM) with a resolution of 5x5 meters, which was downloaded from the Catalonian Cartographic Institute website. In this way, it is possible to construct the 3D map over the modelled topography of the area.

The real estate registry (i.e. the cadastre) is an administrative register of the Ministry of Finance and Public Administration. It is regulated by the Law of the Real Estate Cadastre. according to which the registration of properties is mandatory (and free of charge). City councils are responsible, in front of the Ministry, to keep the information of the registry updated. Therefore, when thinking in future incorporation of other case studies or the develop a business model, it has to be considered that in Spain all municipalities would be able to follow the steps described above.

North Harbour

The basic 3D model of North Harbour was created from a set of geotiff aerial images. To optimise the real-time behaviour of the data set, the data was converted from a set of raster images to an Agency 3DMaps A3M quad-tree using the tools in the Agency9 SDK. The map is located in ETRS 1989 UTM Zone 32N.

The area of interest in North Harbour is the case study area. Since the area has not been developed there is neither existing GIS data nor a virtual city model. However there are architect based CAD drawings showcasing the proposed development. The architect's CAD drawing was created using the Rhinoceros (Rhino) modelling software for windows, the data format of which is based around the NURBS primitive. NURBS, Non-Uniform Rational B-Splines, are mathematical representations of 3-D geometry that can accurately describe any shape, including potentially the most complex 3-D organic free-form surface.

However the graphical hardware requirements involved in providing real time rendering of data models described in NURBS are currently prohibitive, meaning that the Rhinoceros models in their original form were not suitable for use within SEMANCO. To remedy this CAD models were triangulated and exported as a set of triangle meshes. Several methods

were then investigated in order to get this data into one of the open standards usable by Agency9's 3DMaps software, namely COLLADA or CityGML.

The content pipeline that proved to be the most efficient and created the best visual result was to first export to .obj or .fbx and then to import the result into Autodesk Maya. Since the built in COLLADA support in Maya proved to be buggy when working with the Rhino meshes, the meshes were exported using the Feeling software COLLADA plugin.

The final requirement before this model could be used with the 3DMaps GIS software was that these models should be georeferenced. Since the architect models provided did not use a useful coordinate system for this manual adjustments were made with reference to aerial photography of the area. This involved both attaching georeferencing data and ensuring that the elements of the model within 3DMaps were appropriately scaled.

A demo of the final virtual city model can be found at http://cityplanner.agency9.se/project/119NnyIk and a screen shot is shown in Figure A3.
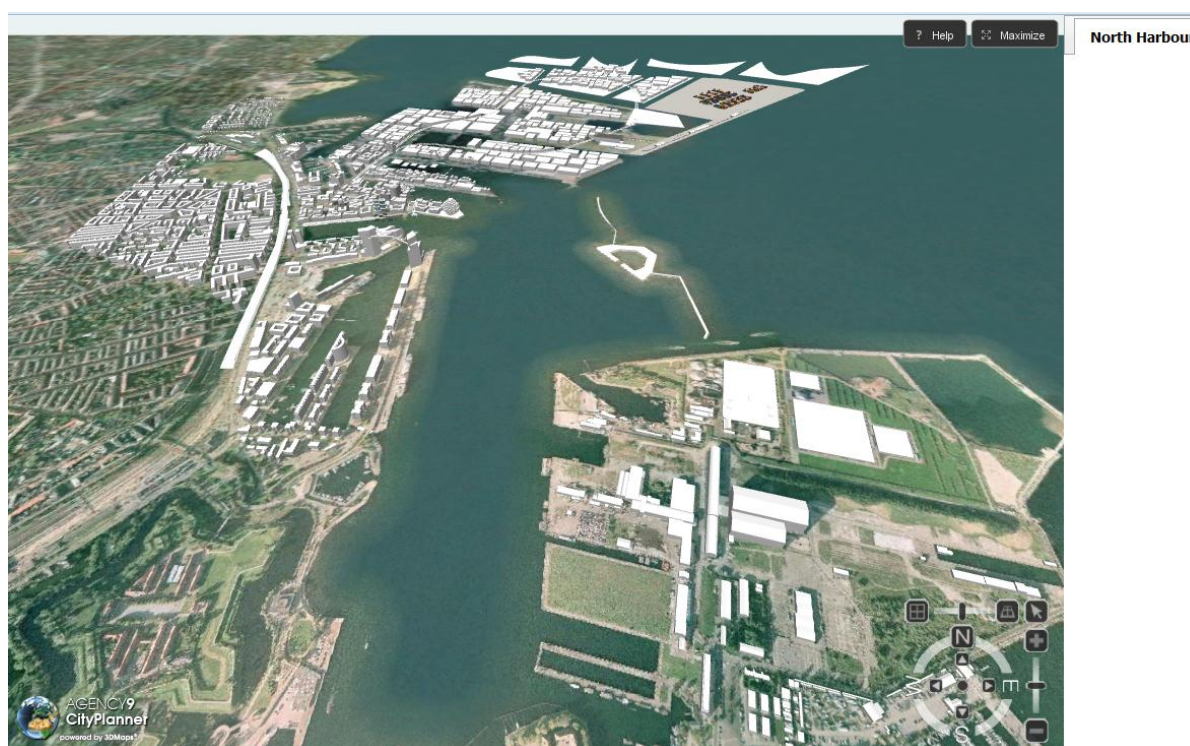

Figure A3. North Harbour case study 3D map

Newcastle

The base map of Newcastle was created from the combination of a set of geotiff aerial images and a DSM dataset. Software from Agency9 was then used to superimpose these aerial images on a triangle mesh created from the DSM to create a 3D map over the area. The result of this process was a quad-tree based 3D map that could be used with Agency 3DMaps.

The original virtual city model of Newcastle was created by extruding the 2D polygons forming the base of the building to height of the building. A tool which takes a shape file containing a set of 2D polygons and a dbf file with a height attribute was used to process the input and create a georeferenced set of COLLADA meshes. A demo showcasing the current

virtual city of Newcastle can be found at http://cityplanner.agency9.se/project/v95DGnvX and a screen shot is shown in Figure A4.


Figure A4. The original Newcastle 3D city map

One principle requirement placed on the city model developed for the Newcastle use case demonstrator is that it should support the demonstration of the baseline energy modelling software tool within Newcastle. Since this tool involves judging the energy efficiency of individual buildings there is a basic requirement that any city model used for it should enable the identification of individual buildings.

Unfortunately, due to the resolution and also potentially the nature of the conveyed survey (data processing regime, date of data capture), some buildings were found to be missing from this map, the height of some buildings to have been judged incorrectly and some of the extruded 3D blocks did not even correspond to buildings at all. Additionally knowledge of the site, site and images visual inspection analysis allowed the recognition of other discrepancies. A graphical depiction of these issues can be found in Figure A5 where each coloured block represents a specific form of problem in the model.

In order to allow the software of D5.1 to precede, a better model of the Newcastle use case was required. The potential for such problems had been foreseen when the DoW was drawn up and is mentioned within the text of T5.1: "The input of this tool will require the purchase and/or conducting of aerial photogrammetry of at least one of the case study areas and will therefore utilise the work subcontracted as detailed in section B2.4.1.", and detailed provision for it being described in Section B2.4.1 of the DoW. The decision was thus made to conduct a LiDAR survey of part of the Newcastle case study area.
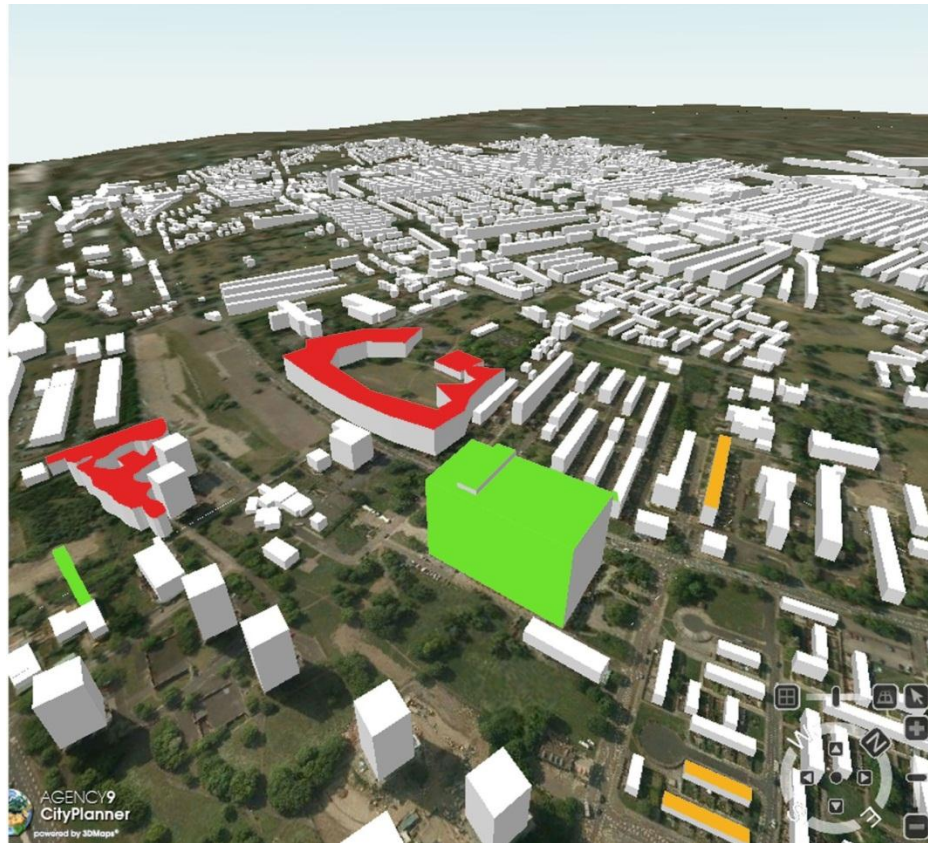
Non-building (mostly trees) assigned with a height attribute

Individual dwellings not identified within the block of terraced houses

Incorrect height information attached to buildings' polygons

Figure A5. Problems with the initial Newcastle 3D city model

Results arising from the LiDAR survey

The principal result arising from the LiDAR survey was an improved fully three dimensional model of the area surveyed within the case study. This can be seen in Figures A6 and A7. In particular, the blocks within the three dimensional model relate to houses and do not include trees. This can be clearly seen in Figure A6 where the basic blocks of the model arising from the LiDAR model are illustrated.
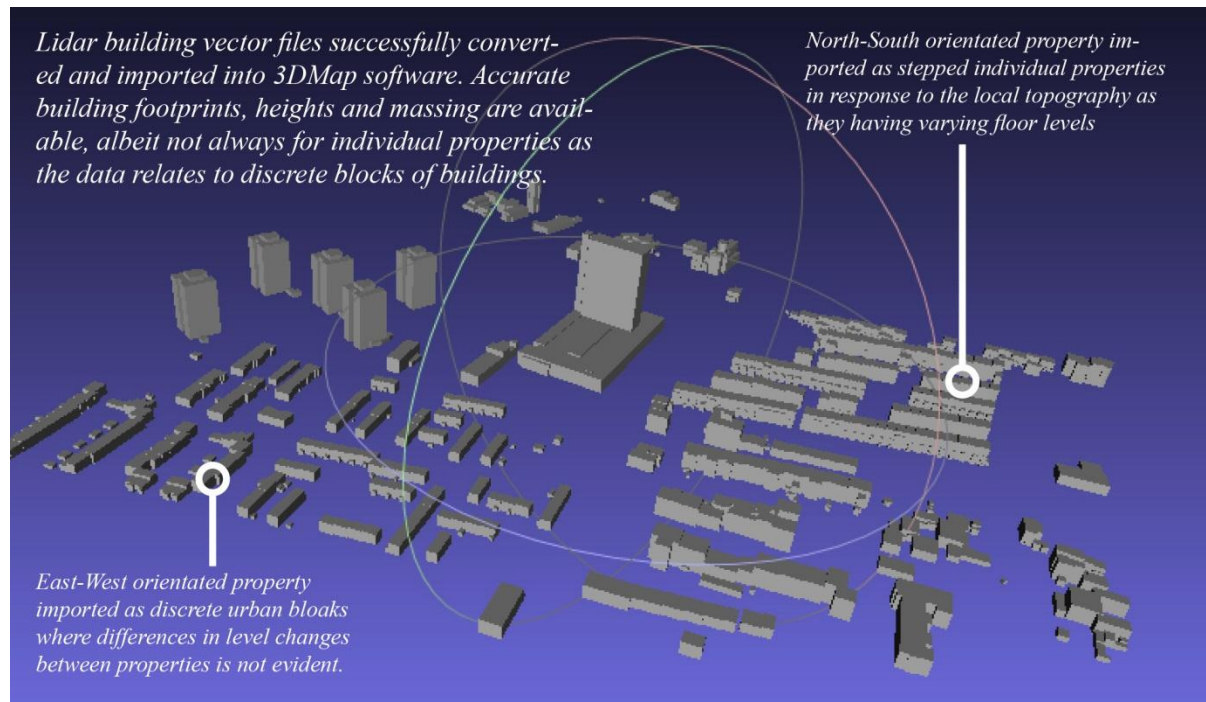
Figure A6. Demonstration of LiDAR data for the Newcastle upon Tyne Case Study area converted into 3DMaps

In addition to the basic geometry representing the houses, the LiDAR survey produced geographical references for the blocks. That this data relating to the placement was very accurate as can be seen in Figure A7, where the blocks arising from the LiDAR data are shown overlain onto aerial photography of the area.
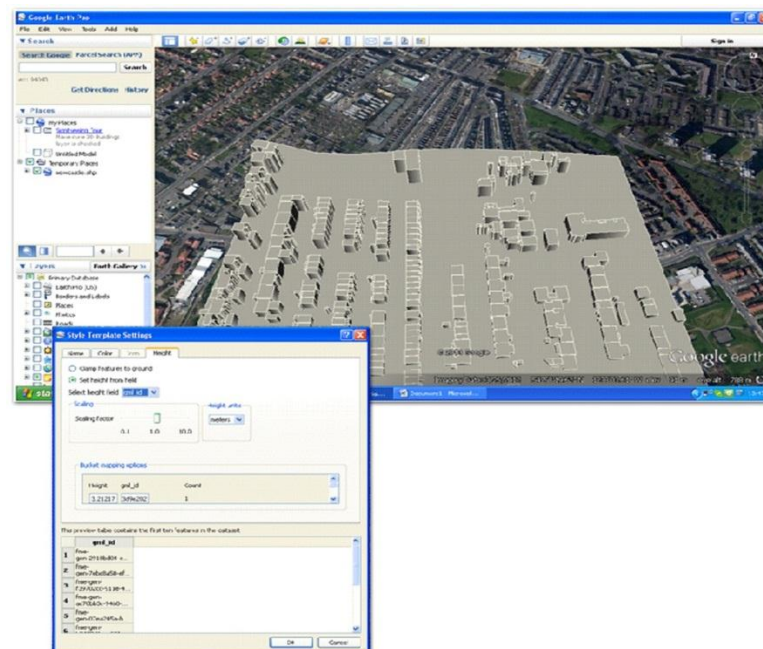

Figure A7. LiDAR data overlain on aerial photography

In order to produce Figure A7, the LiDAR data from the Newcastle case study area was directly imported as individual Collada files into GoogleEarth Pro. This data was georeferenced and located as separate terrain layers imported as CityGML files with selected

individual shp vector files for the buildings blocks. Both are superimposed on GoogleEarth base aerial map.

A final major advantage arising from the LiDAR survey can be seen in Figure A7 - the blocks representing the individual houses within a terrace of houses can be individually identified. Since the Newcastle tool described below focuses on evaluating the energy efficiency of individual houses, the ability to identify individual dwellings within the model was a crucial feature.

Appendix B: Technical Enhancements to Agency9's 3dMaps Software

During the course of the development of the demonstrators within T5.1, and SEMANCO in general, it has become clear that the current 3D mapping software of Agency9 must be further enhanced in a few technical ways. The purpose of the enhancements described in this appendix was to support the specific needs of work tasks in work packages four and five as they arise. Since these work tasks are on-going, some portions of this development work are also anticipated to be on going during the project.[10]

B.1 Building extrusion tool

The building extrusion tool takes a shape file containing a list of polygons together with their height, uses the polygons as the base of the buildings and extrudes them to create a 3D building. This tool was used with publically available data in order to create the building layer for Manresa and the initial data for Newcastle.

B.2 Volume editor

A volume editor that supports the basic editing of volumes has been created and integrated within the 3DMaps software. This editor supports such elements as the drawing of polygons, extrusion and the insertion and removal of points. The features are very close to the behaviour of the basic modelling tool inside Trimble Sketchup. The editor also supports undo and redo functionality and export to COLLADA. Work is on-going to store the data within the SpatialDB.

The user interface of the volume editor is not predefined and can be altered either by using HTML and JavaScript or directly within 3DMaps using the map canvas elements. All editing features and tools are also exposed as a part of the Agency9 3DMaps JavaScript API's. This enables the volume editor to be used in order to create dynamic and procedural buildings. The main anticipated future use of this volume editor is within WP5.4

B.3 Shadow support

The Agency9 3DMaps SDK has had experimental support for shadows rendering for some time. This implementation used a methodology that partitions the shadow space into several shadow maps based on the distance from the viewer. While this approach gave good looking

---

[10] The work reported in this appendix was up to date as of month 18 of the project. Work on developing the 3D maps software to meet the goals of the project is ongoing and will be reported in subsequent deliverables.

results the process was hardware intensive which presented an obstacle to its use in lower power systems (Figure B1).
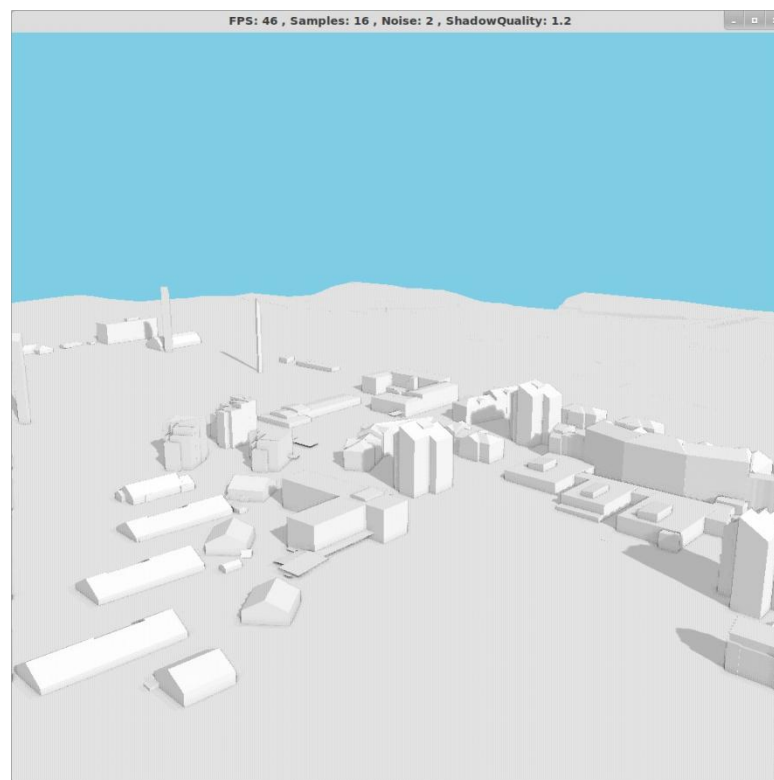

Figure B1. Illustration of new shadow rendering technique

Research has therefore been conducted to enable the support of real-time shadows that only use one shadow map while retaining good image quality. A test implementation of a variance shadow map has been implemented with support for soft edges. This implementation supports both inferred shading and deferred shading and implements soft shadows by blurring the shadow map edges using the variance between two depth values. The implementation is still research quality and has not yet been implemented within the Agency9 3DMaps SDK.

B4 CityGML reader

A reader that serializes CityGML to Java classes has been implemented. The reader is created by using the JAXB XJB compiler to create a java class structure from an xsd schema. The process is mostly automatic except for some minor tuning of the xjb compiler using xml and tuning of the generated classes using Java annotation. The result can read and write the full CityGML 1.4 specification. Efforts to add full CityGML support to the Agency9 3DMaps software itself are on-going.

B5 Spatial Database

In order to allow for the persistent storage of the case study data, a database capable of running on top of PostGIS and serving as a Spatial database had to be located and evaluated. Two different techniques for storing this data within a database were explored.

The first technique investigated was to first use the CityGML reader described in Section B4 to map the CityGML structures into Java classes and then to use standard methods to persist those Java classes within a database. This approach has the substantial advantage of not relying on a specific database implementation, but does require a lot of manual editing of annotations to make the storage of the Java classes operate.

The second technique explored was to directly implement CityGML as a schema inside a db. This approach has been explored within the 3DCityDB project which was a German funded research project whose principal focus was on the creation of a free available data for the storage of 3D city information. During the summer of 2012 this project released a beta version of its database solution which supports PostGIS. Work is currently on going to evaluate this implementation as a potential Spatial database solution for use within the SEMANCO project.