Further Sources for Additional Data and Integration on 3D Models

Felipe Coelho Costa
Lu Lu

April 25, 2013

Contents

List of Tables

List of Figures

1 Abstract

The goal of this report is to describe the procedure performed by the group 3 for the phase 1 and 2 of the GIS-Project in the Winter Semester 2012/13. The main task for phase 1 regards the search for further data and its integration on the data-model using an automatic or semi-automatic process. The operations involved a large number of data exchanging through different formats and a large number of applications, such as SQL, ArcGis, FME and Excel were applied. Though, it is hard to establish an unique and solid procedure which could be systematically repeated. After the integration of the required data, the indicators Apartments/Volume (m3) and Number of Residents/Apartment were calculated and, in the case of the first indicator, different values were obtained since they were grouped by the attribute Age of the Building. The phase 2 comprises the actual estimation of the residential annual electricity consumption on the area of interest. Therefore, all integrated data from phase 1 is gathered in an appropriate algorithm, which relates households and annual electricity consumption. The final result for the households average is 1.584, in contrast to the statistical one: 1.65 (Amt fr Statistik Berlin-Brandenburg, 2011). That led for an average annual electricity consumption range of 2671 - 2931 kWh per apartment. A further research towards a complete automation of the process is recommended and normally requires programmed scripts (e.g in Java) which might deal direct with both the stored data and the data to be integrated from other sources (e.g WMS and WFS), and also perform the final algorithm.

2 Phase 1

2.1 Data Sources

There are many sources of data available which can be integrated either directly by a simple spatial relationship or by joining operations which might become complex depending on the structure differences.

In the case of Berlin, the FIS-Broker database (http://fbinter.stadt-berlin.de) presents a large and spatially located amount of statiscal and survey data regarding social, economical and infra-structural fields. It is the most important source used during the accomplishment of the tasks. The figure 1 elucidates the whole process divided into three steps searching, processing and integrating:

 As figure 1 shows, the sources of data consist basically on the FIS-Broker, the 3D City-Model, which provides all geometric and semantic required in-
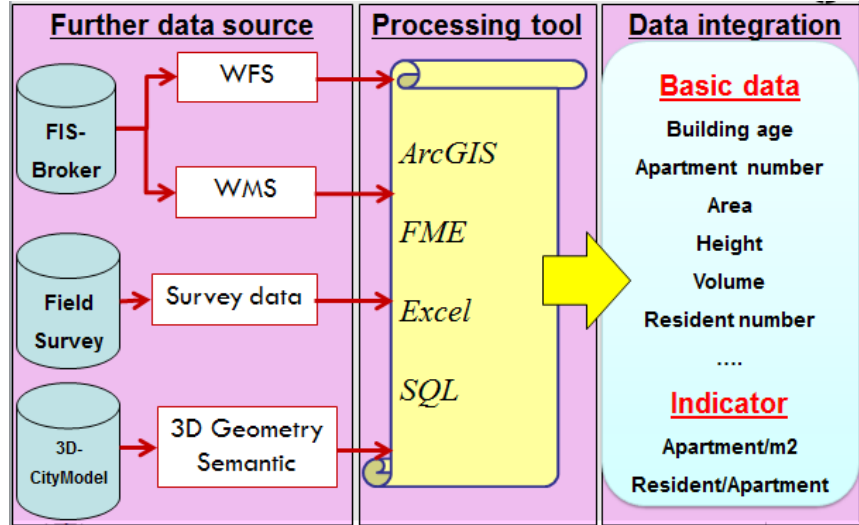
Figure 1: Data Workflow

formation, and the so called "Field Survey", consisting of a survey previously realized in the Moabit area that gathered information for random buildings regarding its number of apartments.

Depending on the application, many other sources can be integrated, enabling such a variety of possible calculations. The exact type of data depends also on the method to be used. The more detailed the algorithm, the more data is required.

2.2 Data Acquisition

2.2.1 WMS

The WMS provides the web map services as a raster map. Normally, only the color information and legend are provided, but not the exact data. However, it is still an important data source and there are plenty of map services provided by WMS. In FIS-Broker database, most of data such as the building age group, proportion of different age group can be accessed by using WMS. In this project, a whole process of getting WMS data in FIS-Broker was created:

1. Get the link to URLfor WMS service from:

www.stadtentwicklung.berlin.de

2. Using ArcGIS connect to the WMS.

3. Set the coordinates system

5

4. Export map by choosing .tiff format in high resolution (300 dpi) and geotiff Tags

5. Combine the export map (Spatial Analyst Tools-¿Local-¿Combine)

6. Re-color raster data and make new legend

7. Zonal statistics as table

8. Join statistics table and new legend to the feature layer

The WMS information is saved in the feature layer after all these steps above. But maybe there are some errors in the boundary due to the data resolution.

Figure 2: Mismatch of Overlap (Carrion, 2010)

2.2.2 WFS

The WFS is the short term for Web Feature Service in which the graphical features can be requested from the services. Both geometry and semantic data are provided by WFS. Therefore, not the color information, but the database can be easily requested by the users compare to WMS. For example, in WMS, the information of building age is given in a time period which is shown in the legend, while in WFS it is provided the exact construction year, as a vector data, which is more specific and useful.

To get WFS data, several steps are needed as follows:

1. Applying FME Intergration Console - add FME extension to ArcGIS

2. Add FME connection

3. Select WFS data Format

4. Set parameter.

Some feature data like building and statistical block as well as semantic data like number of residents in each block are extracted from WFS. However, the WFS services are not commonly open. The FIS-Broker WFS services is not availabe anymore available, due to security reasons.

2.2.3 Field Survey

The energy consumption have a tightly relationship with number of apartments. Therefore a field survey for apartment number is necessary. This survey is done based on different building age group. For each building group, at least 10 buildings were selected to count the number of apartment by counting the ring bells.

2.2.4 CityGML

Also some useful data can be obtain from CityGML, especially 3D geometry data, such as the footprint area and building height. Because only the residential volume is important in this project, the height of the building is calculate by removing the roof height.

2.3 Integration

The aim of Phase I is to get two important indicators: Apartment/Volume(m3) and Resident/Apartment. To calculate these two indicators, not only the data from different data sources is required, but also different tools (FME, ArcGIS, SQL, Excel) should be integrated together.

2.3.1 Age of Buildings

The attribute age of buildings specifies the exact year of construction for the buildings. For cities like Berlin which have a long history, the building age may vary from more than 100 years to recent 1 or 2 years. As architectural properties might not vary so deeply during years or even decades, this attribute is better represented by a range of values.

Therefore, the attribute age class was separated into 6 groups: 1889-1918, 1919-1945, 1946-1961, 1962-1974, 1975-1993 and 1994-2012.

2.3.2 Apartment Number / Volume

The number of apartment is the key value for the next steps. As it is impossible to do the survey for all buildings in the test area, this indicator was introduced. It is calculated from the field survey data and then can be easily gotten the apartment number of other buildings by multiplying this indicator with the building volume.

As a further data to be attached on the calculations, a survey was performed on the area of interest, with the goal of obtaining the number of apartments on the region. Thus, the survey data provides the counting of the ring bells for a selected group of buildings.

The process to join all the information must contain an export of the surveying data on the database. Then, the GMLID have to be joined accordingly this is realized by the address of the buildings, as both the surveying data and the data model provide it.

The following script is an example of a join by the address. It shows basically the final step, once the surveying data is already imported on the data model:

```
select   street, house_number, apartment, a.id,
b.building_id ,co.gmlid
from join_surveying js
inner join address a
on js.strasse=a.street  and js.hnr=a.house_number join
address_to_building b
on a.id=b.address_id join cityobject co
on b.building_id=co.id;
```

As the footprints are to provide the area of the buildings, one must aggregate the height information. A close look on the database highlights the existence of three different heights per building:

- Measured_Height (Hm): representing the total height of the building

- H_Trauf (Ht): the lower height of the roof

- H_First (Hf): the upper height of the roof

So, the real height (Hr) is calculated through the following relationship:

$$Hr = Hm - (Hf - Ht) \qquad (1)$$

The real volume for each building is trivially obtained from the multiplication of the footprint area with the real height. Then, the indicator "

apartment Number / Volume(m3) is performed like:

$$I = \frac{\sum\limits_{k=1}^{n} T_k}{\sum\limits_{k=1}^{n} V_k} \qquad (2)$$

Where:
I = Indicator (Apartment number / Volume(m3))
T = number of apartments of building k
V= volume of building k
n = total surveyed buildings

Afterwards, the above procedure is expanded. Each building age class receives a different indicator, as they present different periods of construction and, therefore, different architectural configurations which renders different results.

It is important to mention that the indicators inherit an expectation "behave". It is well known that older buildings used to have less apartment for the same amount of volume. Another socio-economic issue is related to the period right after the second World War. As the demand for new buildings was very high due to the destruction caused by the war, the short period after 1945 should present the highest indicators.

Also, the integration of the surveying data presents some mismatches from real data to the 3D Model. Some buildings have more than one address and vice-versa. Some garages are integrated on the buildings and, therefore, would be part of the volume calculation.

In order to avoid mismatches and to build reliable indicators, some unclear data were not taken into account. For the next survey missions, it is recommended to take a brief look on the 3D Data Model before going to the field, so that only reliable buildings could be selected.

2.3.3 Number of Inhabitants

A very important data for input in energy simulations is the number of inhabitants of the residential area of interest. Once the city model supports semantic information about the usage of the buildings, it is possible to split

the inhabitants inside the buildings, and so, have an estimation of the house-holds.

For this task, a WFS was provided containing the total amount of inhabitants per block. In order to obtain the total amount of inhabitants per building, a weighting by an approximation of the building volume was performed.
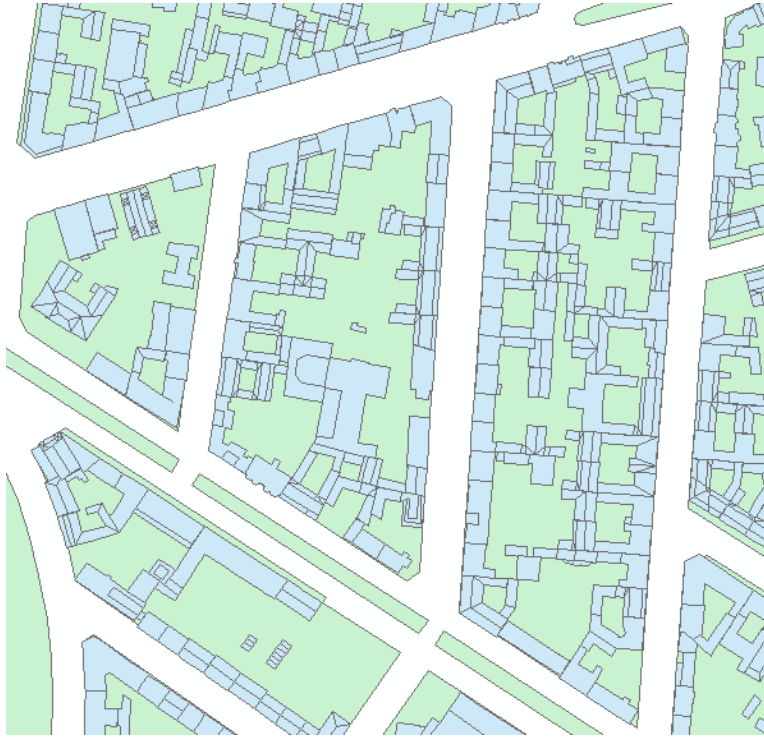

Figure 3: Building Footprints X Block (Self Made)

In order to calculate the number of inhabitants or residents per building, a previous selection is performed on the database, returning only the buildings which present residential purpose. As the buildings are semantically defined also by its function of usage, based on OSKA(Objektschlsselkatalog 2003) the following script was performed:

```
select   b.id ,co.gmlid ,  b.function ,  b.measured_height ,
a.street ,  a.house_number
from cityobject co  ,  building b,  address a,
address_to_building ab
```

```
where b.id=co.id and b.id=ab.building_id and
ab.address_id = a.id and (
b.function='1373' or b.function='1331' or b.function='1311'
or b.function='1372' or b.function='1399' or b.function='1379'
or b.function='1381' or b.function='2199' or b.function='1301'
or b.function='1361' or b.function='1341' or b.function='1221'
or b.function='1321' or b.function='1374' or b.function='1371'
or b.function='1231' or b.function='2131'
or b.function='2121' or b.function='2101'
or     b.function='2141');
```
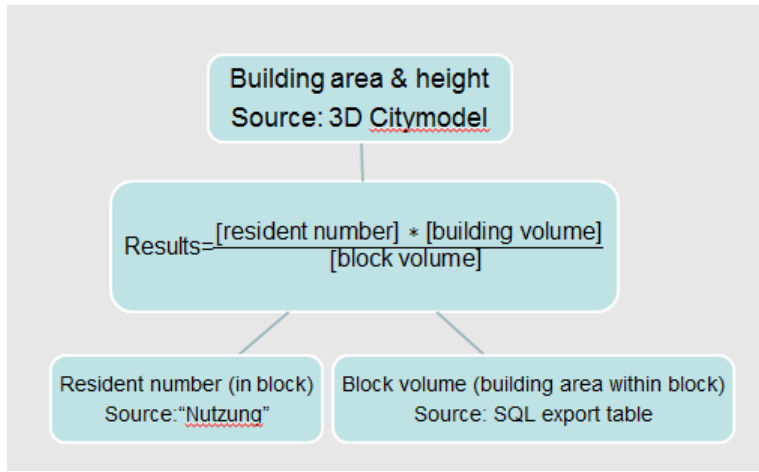

Figure 4: Inhabitants per Building (Self Made)

Figure 4 shows how to calculate the number of inhabitants in each building. In order to group the buildings per block, a spatial join is realized in ArcGis and the results are exported to the data-base. Then, as the buildings contain their BlockId attribute, as well as the sum of inhabitants in that block (for instance called S attribute), the residents shall be split into the buildings according to the next sentence:

$$Ri = (Sk * Vi)/Vk$$

Where:
Ri = Residents on building i
Sk = Total number of inhabitants on block k
Vi = Volume of building i

Vk= Total Volume of buildings on block k

2.4   Results

Table 1 shows the results of Apartment/Volume for different building age group.

Class	Period of Construction	Number of Sample Buildings	Result [apart./m³]
Class 1	1889-1918	20	0.0023247637
Class 2	1919-1945	20	0.0025399709
Class 3	1946-1961	20	0.0035708397
Class 4	1962-1974	20	0.0032491869
Class 5	1975-1993	20	0.0026467066
Class 6	1994-2012	20	0.0026363702

Table 1: Indicator 'Apartments/Volume (m3)' (Self Made)
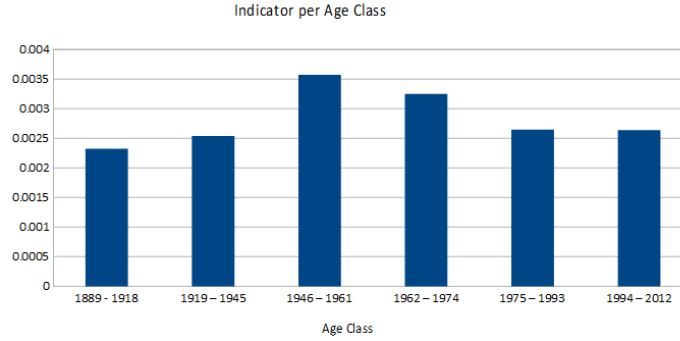

Figure 5: Indicators grouped by Age Class (Self Made)

Applying the indicators above, the apartment number for the whole test area can be calculated by multiplying the indicator with the building volume. Furthermore, the average household can by computed:

$$Hi = \frac{Ri}{Ni} \tag{3}$$

Where:
H = Household average for building i
R = Total number of residents on building i

12

N = Number of apartments on building i (based on the volume)

The average result for the area of interest is: **H=1.584**

As a validation procedure, the result might be compared to to statistical data on Figure 6. Statistically, the Mitte neighborhood presents 1.65 as household average value.

Bezirk	Ins-gesamt	Ein-personen-haushalte	Mehrpersonenhaushalte				Durch-schnittl.
			zu-sammen	davon mit ... Personen			Haus-halts-größe
				2	3	4 und mehr	
			1 000				Personen
Mitte	202,7	126,9	75,9	45,1	15,5	15,3	1,65
Friedrichshain-Kreuzberg	166,7	104,2	62,5	36,8	13,0	12,7	1,64
Pankow	228,2	134,3	93,9	59,4	20,1	14,4	1,64
Charlottenburg-Wilmersdorf ...	194,9	114,3	80,7	52,9	14,9	12,8	1,64
Spandau	127,7	65,0	62,7	39,8	13,0	9,8	1,79
Steglitz-Zehlendorf	159,1	77,3	81,8	51,5	14,3	16,0	1,84
Tempelhof-Schöneberg	192,2	102,5	89,7	56,8	17,6	15,3	1,76
Neukölln	173,0	95,5	77,4	44,6	16,1	16,8	1,80
Treptow-Köpenick	137,6	68,0	69,5	47,0	12,7	9,9	1,75
Marzahn-Hellersdorf	134,3	59,2	75,1	48,5	16,7	9,8	1,86
Lichtenberg	151,9	74,3	77,6	55,0	13,4	9,1	1,73
Reinickendorf	127,1	57,2	69,9	43,6	12,0	14,3	1,91
Land Berlin	1 995,4	1 078,8	916,6	581,2	179,3	156,2	1,74

Figure 6: Household in Berlin (Amt fr Statistik Berlin-Brandenburg, 2011)

3 Estimation of Residential Electricity Consumption Based on 3D City Model

3.1 Define Algorithm

Normally, the algorithms used to calculate the electricity consumption depend on a large amount of variables and factors, as can be seen on figure 7:

According to this sketch, in order to calculate the electricity consumption, except the number of people, it is required to know the penetration of appliances, the power of appliances and the hours of use. Such data are changed from one situation to another. The available data cannot support this algorithm, and, therefore, some assumptions are done in order to simplify the process.
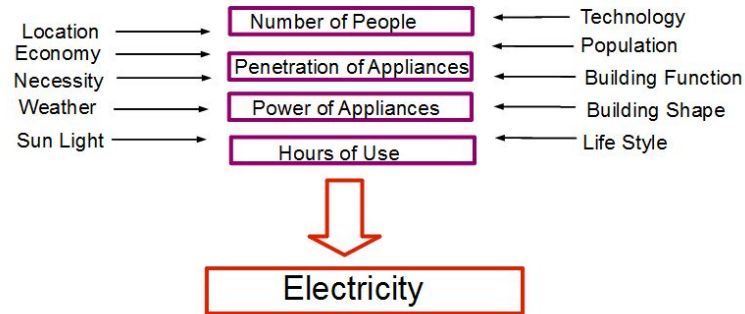
13

Figure 7: Factors for calculating the electricity consumption (Costa, 2012)

According to this situation, a new algorithm is created to estimate the annual consumption based on different household. Figure 9 illustrates the annual consumption for different number of household. Since the number of apartment and residents in each building already calculated in Phase I. If the residents can be distributed into each apartment as household, the electricity consumption can be calculated from it.

Durchschnittlicher Jahresstromverbrauch ausgewählter Elektrogeräte					
	Personen im Haushalt/Verbrauch in kWh				Ihr Jahresverbrauch
Gerät/Anwendung	1	2	3	4	
Elektroherd	195	390	445	575	
Kühlschrank	280	310	330	355	
Gefriergerät	305	350	415	420	
Waschmaschine	70	125	200	265	
Wäschetrockner	125	225	325	465	
Geschirrspüler	120	200	245	325	
Warmwasserversorgung Küche*	245	295	345	415	
Bad	470	780	1.080	1.390	
Fernseher	120	150	190	205	
Hilfsgeräte für die Zentral-/Etagenheizung	250	290	330	370	
Beleuchtung	195	285	330	435	
Sonstiges	340	560	800	860	
Ø Stromverbrauch pro Jahr	2.050	3.440	4.050	4.750	

*ohne Geschirrspüler

Figure 8: The annual electricity consumption for different number of household (Vattenfall 2012)

3.2 Data Source

From Phase I, the results are mainly two indicators: Apartments/volume and Resident/ Building. If these two indicators are applied to all the residential buildings, the number of apartments per buildings will be get. Also the average share of household in Mitte can be found in statistic department. Combining the electricity consumption data from Vattenfall, all the data for estimating the annual consumption is provided.

3.3 Method

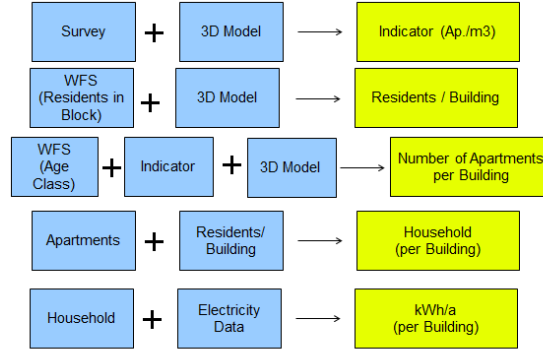Figure 8 illustrates the work flow for calculate the results as well the data sources:

Figure 9: Workflow for Phase II

For the last step, there are two ways to calculate the electricity: 1) Establish the regression equation and apply the average household in the equation (see figure 10). This equation was created from the data available on Figure 8.
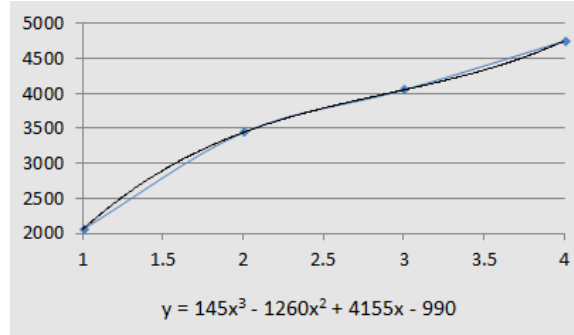
$$y = 145x^3 - 1260x^2 + 4155x - 990$$

Figure 10: The regression equation for electricity consumption

2) Distribute the residents into each apartment to get the number of apartment with different household. And use these numbers to multiply the corresponding electricity and lighting consumption. This distribution is base on the statistical share of households for the giving area.

3.4 Household Distribution

One of the most important parts in the second method is distributing the inhabitants into apartment. There are three inputs already known: the apartment number, the residents number and the average share of different household. To distribute the residents, only two variables above can be fixed with one variable left. Therefore there are three methods:

- Fix the share and residents number, take apartment number as variable

- Fix the share and apartment number, take residents number as variable

- Fix the apartment and residents number, take share as variable

While for the first two methods the variable parameters will be changed in a unreasonable way, and make the results not reliable, a new distribution method is developed from the third method.

3.4.1 Basic Idea

The basic idea of this method is to try to keep the share close to the average share of Mitte, by applying the following steps:

16

1. Spit the total number of apartment with different household based on the percentage
2. Check how many residents left
3. Distribute the left residents
4. Keep the ratio of the 2 4 household: 3:1:1

3.4.2 First Distribution

After the first distribute based on the share, there will be three difference scenes:

- $R > 0$ more people live in one apartment

- $R < 0$ less people live in one apartment

- $R < Apartment$ only household(1) and empty apartment

In the last scenes, all residents will be put into each apartment with only 1 household and the other apartments are empty. For the first and second scenes, the left residents have to be distributed again.

3.4.3 Second Distribution

In the second distribution, as can be seen in figure 11, if only one residents left, it will be put into an apartment with 1 household, which means the number of 1 household will decrease 1 and the number of 2 household will increase 1. If two left, put both of them into 1 household. And the number of 1 household will also decrease 1 but the number of 3 household will increase 1. Do the same as shown in the figure 11 until 8. if there are 9 left, distribute the first 8 as -5:3:1:1 then put the left 1 to 1 household.
By this distribution method, it keeps the ratio of the 2 4 household as 3:1:1, which also means every 8 residents as a loop.

House hold	1	2	3	4	5	6	7	8
1	-1	-1	-1	-2	-2	-3	-4	-5
2	+1			+1		+1	+2	+3
3		+1			+1	+1	+1	+1
4			+1	+1	+1	+1	+1	+1

Figure 11: The second residents distribution method (Self Made)

3.5 ArcPython User Interface

To apply this method to other dataset, an Arctool was designed by using python script which has a good connection with ArcGIS 10.
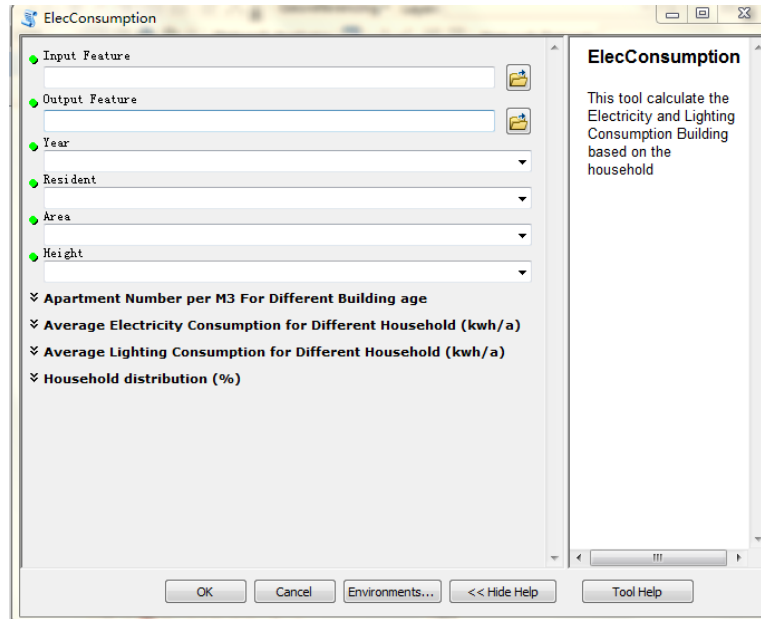

Figure 12: Developed Arctool (Self Made)

Figure 13: Developed Arctool - State Variables (Self Made)

The user can choose the input and output feature class and select the corresponding fields for building age, residents, area and height to calculate the electricity consumption. Besides, the user can also change the optional parameters such as the share of household and the annual electricity consumption for different household.

3.6 Results

After the resident distribution, the light and electricity consumption can be computed by multiplying the number of apartment with different household with corresponding consumption
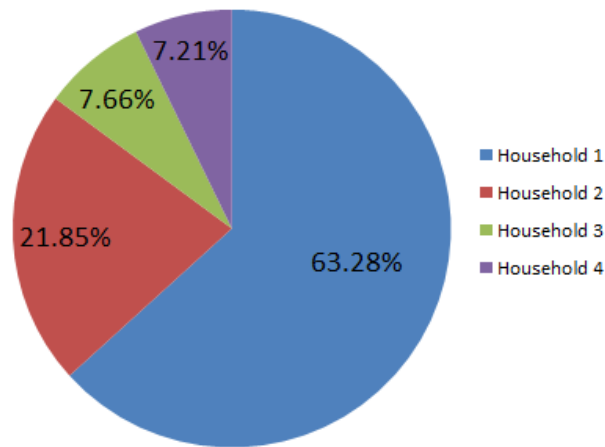

Figure 14: Final Household Share (Self Made)

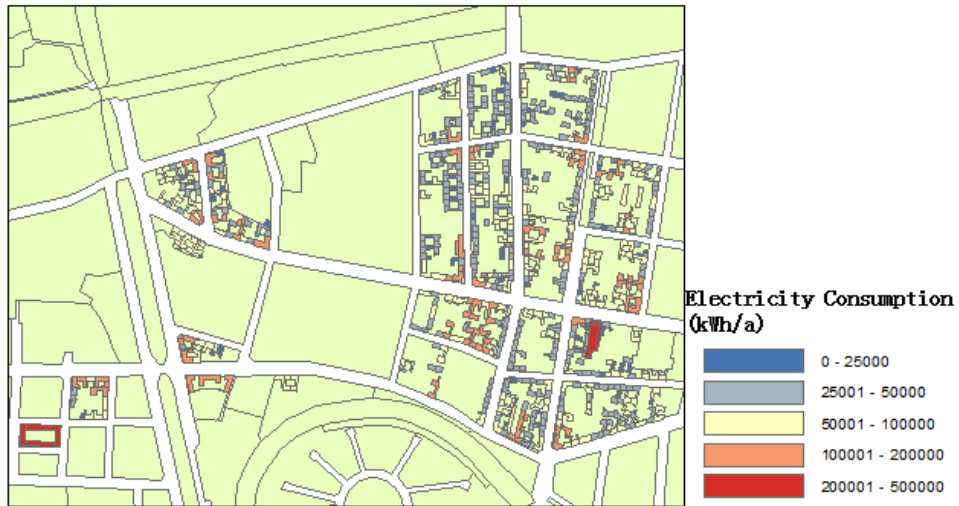Figure 15: Final Lighting Consumption (Self Made)

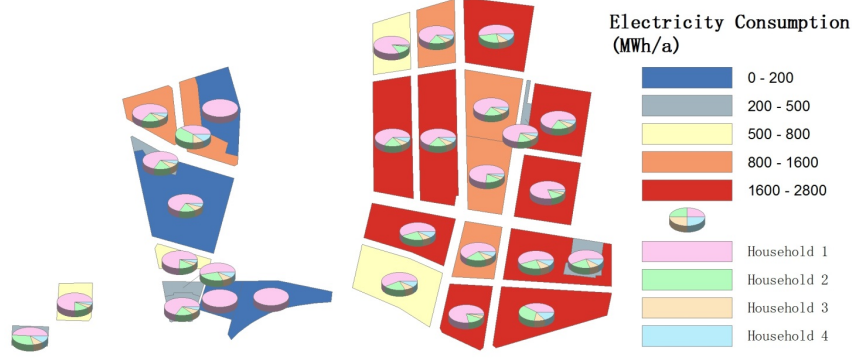Figure 16: Final Electricity Consumption (Self Made)

Figure 17: Final Electricity Consumption per Block (Self Made)

Table 2 shows the results from the two calculation methods. As can be seen from it, using the average household and regression equation the results are smaller than the distribution method. Nevertheless, both estimation methods rendered values within a short interval - as both are based essentially on the residents per building and/or per apartment.

Method	Per building (kWh/a)	Per apartment (kWh/a)
Distribution	50900	2671
Average household	55829	2931

Table 2: Final Electricity Consumption (Self Made)

4   Conclusion

The described method is able to integrate statistical data, surveying data and virtual 3D City Models. Furthermore, it provides a rough tool to calculate the electricity for other area which can be implemented in 3D City Models. As a result, instead of having only average results, the calculations are also locally performed. In other words, as the parameters are first calculated per each building, it is also possible to select the final results for each building. The two distribution methods provide the user two different level results with different data. If there is no household share data, the regres-

sion equation method can give the rough results. While with the household share data, the distribution method can be applied for more reasonable results.

The 3D Model is able to provide all the demanded geometrical factors. Once the electricity consumption estimation is geometrically based on the volume of the buildings, data regarding its footprint area and height are required.

However, there still some parts need to pay attention to during applying this method. Such as the function of the building, the relationship between 3D City Model and the survey address.

5   Bibliography

Carrin, D. **Estimation of the energetic rehabilitation state of buildings for the city of Berlin using a 3D city model represented in CityGML.** Master Thesis. Technische Universitt, Berlin (2010)

Costa, F.C. **Estimation of Electrical Power Consumption Using 3D City Models on Building/Neighborhood Scale.** GIS Seminar. Technische Universitt, Berlin (2012)

Statistik Berlin-Brandenburg. **Ergebisse des Mikrozensus im Land Berlin 2011. Amt fr Statistik Berlin-Brandenburg.** Potsdam, 2012.

Energie sparen-Stromverbrauch prfen. VATTENFALL, September,2012.