OpenStreetMap data assessment for extraction of urban land cover and geometry parameters required by urban climate modeling

T.E.Samsonov[1], P.I.Konstantinov[2]

[1]Automation Lab, Dept. of Cartography and Geoinformatics,
Faculty of Geography, Lomonosov Moscow State University,
Leninskiye Gory 1, Moscow, Russia, 119234
Email: tsamsonov@geogr.msu.ru

[2]Meteorological Observatory, Dept. of Meteorology and Climatology,
Faculty of Geography, Lomonosov Moscow State University,
Leninskiye Gory 1, Moscow, Russia, 119234
Email: kostadini@mail.ru

1. Introduction

More than a half of world's population lives in cities now, and the urban/rural ratio is increasing (World Urbanization… 2014), which deserves increased attention to heavily populated areas. Contemporary meteorological models assimilate information about urban conditions for implementation of scenarios of physical interaction between atmosphere boundary layer and underlying surface (Kusaka et al. 2001). This allows more precise weather and climate predictions (Konstantinov et al. 2014).

During the last two decades significant progress has been made in description of urban environment for urban climate modeling. Required parameters are extracted from satellite imagery and spatial datasets such as city vector geodatabases (Lindberg 2007). However, expensiveness and unavailability of timely data often limits the possibility of urban climate studies. Recently, volunteered geographic information received great attention as a source of information about land cover (Comber et al. 2013). Our research pioneers in the assessment of OpenStreetMap data for possibility of extraction of main parameters of land cover and urban geometry needed for urban climate research.

2. Land cover classification

Meteorological (climate) models consider physical characteristics of surfaces to model their interactions with atmosphere. For example, WRF model (Skamarock et al. 2008) uses GLCC 1 km resolution land cover database (Loveland et al., 2000) that contains 24 land cover classes. This information should be refined for urban areas.

Meso-scale models use simple land cover refinements. For example, Kusaka et al. (2001) considers urban and vegetation ratio (implemented in WRF), while Trusilova et al. (2013) differentiates fractional area of urban land and fractional artificial area occupied by buildings. At the same time considering water and green area ratios in experimental high-resolution models facilitates better reproduction of temperature effects above those surface types (Konstantinov et al. 2014). This demonstrates the potential of fine-grained classifications of urban land cover for micro-scale modeling.

Guided by availability of various OSM keys (OpenStreetMap... 2014) we developed reclassification scheme that is close to proposed by Lemonsu et al. (2008) and is presented in Table 1. This classification can then be easily reclassified into more simple parameterizations.

Table 1. Extraction of land cover types from OSM Data

#	OSM Key	OSM values	Destination class
1	building	ALL except bunker / cabin / construction / farm_auxiliary / hut / shed / stable	buildings
2	waterway	river / riverbank / stream / canal / ditch	water
2	natural	wetland / water	water
2	landuse	reservoir	water
3	natural	tree / tree_row / wood	tall vegetation
3	landuse	forest	tall vegetation
4	landuse	orchard / vineyard / scrub / farm / farmland / greenfield	low vegetation
5	landuse	grass / meadow / pasture	grass
6	leisure	garden / park	mixed vegetation
7	surface	ground / earth / dirt / mud / sand	bare ground
7	natural	bare rock / mud / sand / beach	bare ground
8	surface	asphalt	asphalt
9	surface	concrete	concrete
10	highway	ALL except track / path / footway / bridleway / steps / proposed	roads
11	landuse	construction / garages, industrial / military / railway	industrial

In current research we focused on the availability of building data as being the most important, keeping assessment of other land cover types for future investigations.

3. Urban canyon geometry

The central concept of urban meteorology is *urban canyon* (Nunez and Oke 1977) which stands for the space between buildings characterized by its width (W), height (H) and length (L) (Figure 1).

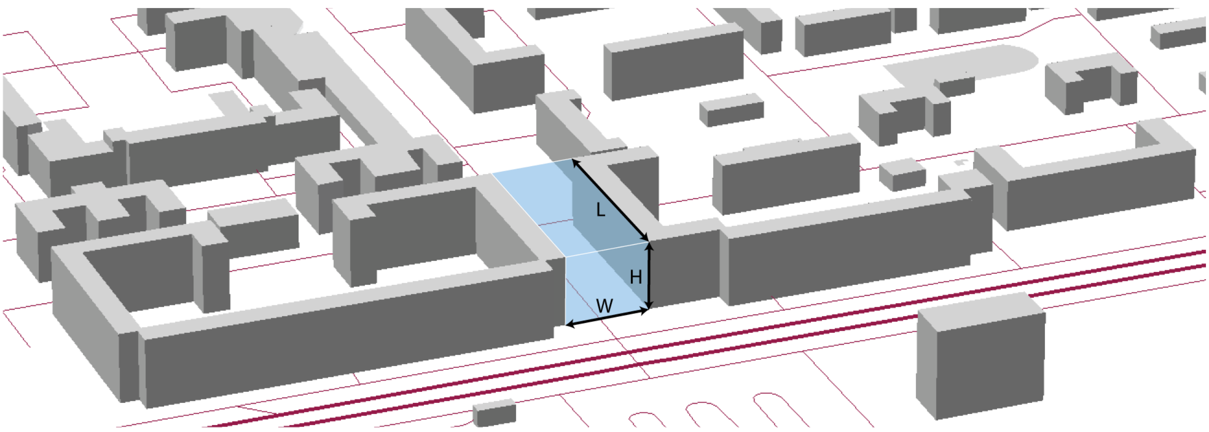

Figure 1. Urban Canyon

Current meso-scale models are not canyon-resolving due to their spatial resolution (~1 km). They assimilate mean parameters of buildings to reconstruct average canyon geometry in every cell. Trusilova's et al. (2013) scheme includes such parameters as building height, height to width ratio of canyons and roughness length for the building–canyon system. Kusaka et al. (2001) uses more sophisticated parameterization that includes street-canyon orientation. Derivation of these parameters requires information about buildings and their heights.

Buildings are coded in OSM data using "building" tag which can be filled by simply "yes" value or the value containing the particular type of the building. There are also two options for coding building heights. The first is "building:height" tag and the second is "building:levels" tag. As detailed information about precise building height is rarely available the second tag is much more common.

We examined the completeness of OSM building data in 29 largest world urban areas that have more than 10 mln inhabitants as of March 2014 (Demographia... 2014). Results are summarized in Table 2. L-ratio reflects how many buildings are attributed with levels data. B-ratio is synthetic index that shows how many buildings are digitized in relation to number of inhabitants and thus reflects the completeness of building geometry.

Table 2. OSM building data availability for world's largest urban areas (> 10 mln people)*

#	Rank	Country	Name	Population	Area	Density	Buildings	Levels	L-ratio	B-ratio
1	**27**	**France**	**Paris**	**10 975**	**2 845**	**3,9**	**2415331**	**2333**	**0,10%**	**220,08**
2	**8**	**United States**	**New York**	**20 661**	**11 642**	**1,8**	**1103982**	**692**	**0,06%**	**53,43**
3	**29**	**United Kingdom**	**London**	**10 149**	**1 738**	**5,8**	**532550**	**11240**	**2,11%**	**52,47**
4	**15**	**Russia**	**Moscow**	**15 885**	**4 662**	**3,4**	**337229**	**58386**	**17,31%**	**21,23**
5	**28**	**Japan**	**Nagoya**	**10 238**	**382**	**2,7**	**152337**	**768**	**0,50%**	**14,88**
6	**14**	**Japan**	**Osaka-Kobe-Kyoto**	**17 234**	**3 212**	**5,4**	**249902**	**20692**	**8,28%**	**14,50**
7	**1**	**Japan**	**Tokyo-Yokohama**	**37 555**	**8 547**	**4,4**	**516715**	**6611**	**1,28%**	**13,76**
8	**5**	**Philippines**	**Manila**	**22 710**	**158**	**14,4**	**218464**	**699**	**0,32%**	**9,62**
9	16	United States	Los Angeles	15 250	6 299	2,4	51341	416	0,81%	3,37
10	2	Indonesia	Jakarta	29 959	3 108	9,6	86109	11084	12,87%	2,87
11	20	Bangladesh	Dhaka	14 816	337	44	19510	14269	73,14%	1,32
12	18	Thailand	Bangkok	14 910	2 461	6,1	17120	210	1,23%	1,15
13	26	Brazil	Rio de Janeiro	11 723	202	5,8	9882	791	8,00%	0,84
14	23	Turkey	Istanbul	13 187	1 347	9,8	10995	54	0,49%	0,83
15	21	Argentina	Buenos Aires	13 913	2 642	5,3	10931	123	1,13%	0,79
16	11	China	Beijing	19 277	3 756	5,1	14099	92	0,65%	0,73
17	6	China	Shanghai	22 650	3 626	6,2	14749	182	1,23%	0,65
18	19	India	Kolkota	14 896	1 204	12,4	9674	0	0,00%	0,65
19	10	Brazil	Sao Paulo	20 273	2 849	7,1	12632	2549	20,18%	0,62
20	13	India	Mumbai	17 672	546	32,3	10524	100	0,95%	0,60
21	17	Egypt	Cairo	15 206	1 761	8,6	6069	49	0,81%	0,40
22	4	South Korea	Seoul-Incheon	22 992	2 266	10,1	8418	119	1,41%	0,37
23	9	Mexico	Mexico City	20 300	2 072	9,8	3648	119	3,26%	0,18
24	3	India	New-Delhi	24 134	2 072	11,6	4219	41	0,97%	0,17
25	24	China	Shenzhen	12 860	1 748	7,4	2073	37	1,78%	0,16
26	25	Nigeria	Lagos	12 549	907	13,8	795	0	0,00%	0,06
27	22	Iran	Tehran	13 429	136	9,9	844	23	2,73%	0,06
28	7	Pakistan	Karachi	21 585	945	22,8	1082	17	1,57%	0,05
29	12	China	Guangzhou-Foshan	18 316	3 432	5,3	824	0	0,00%	0,04
				X10³ people	km²	x10³ people / km²	count	count	Levels / buildings, %	Buildings / 10³ people

*The list of urban areas and data about population, area and density is taken from (Demographia... 2014).

Results show only 8 cities with high values of B-ratio (bold font) — those having relatively full information about built-up. And only 4 cities (highlighted in green) have significant value of L-ratio. The most satisfactory results are shown by Moscow city, however even there the completeness of information is not enough for its usage in urban studies, as only 17% of buildings are attributed with levels.

We assessed the quality of OSM building levels in Moscow city using Geocentre Consulting Ltd. database. OSM levels (L) were reduced to heights using H = 4L formula. Heights from both databases was averaged for 262 cells with 200 m resolution. Figure 2 presents scatterplot with reference heights along Y axis and OSM heights along X. Coefficient of determination is $R^2 = 0.77$ for this dependency. This shows satisfactory level of dependency and proves that data can be potentially used in urban climate tasks. However, the similar verification should be done for all other major urban areas in the future.

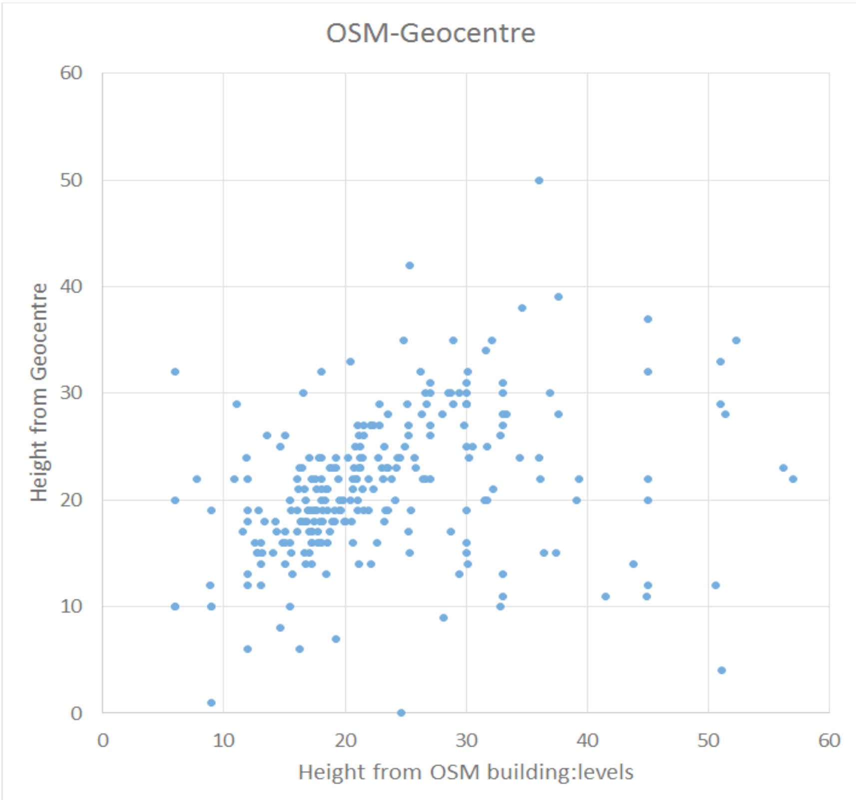

Figure 2. Scatterplot of OSM building heights and referential heights from Geocentre Consulting database (262 cells with 200 m spatial resolution).

4. Discussion

In this paper, OSM data is assessed in terms of land cover and urban geometry characterization for urban climate modelling for the first time. A mapping of OSM tags to land classes is proposed. The completeness of OSM building data is estimated over 29 largest world urban areas. Results showed that OSM data is nor ready yet for urban canyon estimations due to incompleteness of buildings and/or their levels attribute. The quality assessment of Moscow OSM building levels show satisfactory correspondence with reference data and thus potential applicability of OSM data in extraction of urban canyon geometry.

Acknowledgements

This work was supported by RFBR grant 13-05-41306-RGO_a.

References

Comber AJ, Brunsdon C, See LM, Fritz S, McCallum I, 2013. Comparing Expert and Non-expert Conceptualisations of the Land: An Analysis of Crowdsourced Land Cover Data. LNCS 8116: Proceedings of 11th International Conference COSIT 2013, Springer, p 243-260.

Demographia World Urban Areas (Built-Up Urban Areas or Urban Agglomerations) 10th Annual Edition: March 2014, Wendell Cox Consultancy, Belleville, IL, USA, 127 p.

Konstantinov PI, Varentsov MI, Malinina EP, 2014. Modeling of thermal comfort conditions inside the urban boundary layer during Moscow's 2010 summer heat wave (case-study). Urban Climate (in press). DOI: 10.1016/j.uclim.2014.05.002. 10 p.

Kusaka H., Kondo H, Kikegawa Y, Kimura F, 2001. A simple single-layer urban canopy model for atmospheric models: Comparison with multi-layer and slab models. Boundary-Layer Meteorology, vol. 101, pp. 329–358.

Lemonsu A, Leroux A, and Belair S., 2008. A general methodology of urban land cover type classication for atmospheric modelling. Technical report, Meteo France, Centre National de Recherches Meteorologiques

Lindberg F, 2007. Modelling the urban climate using a local governmental geo-database. Meteorological Applications, 14 (3), 263-273.

Loveland TR, Reed BC, Brown JF, Ohlen DO, Zhu, J, Yang, L., and Merchant, J.W., 2000, Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1-km AVHRR Data:International Journal of Remote Sensing , v. 21, no. 6/7, p. 1303-1330

Nunez M and Oke TR, 1977. The energy balance of an urban canyon. *Journal of Applied Meteorology*, 16, 11-19.

OpenStreetMap Wiki, 2014. http://wiki.openstreetmap.org [accessed 11 Aug 2014]

Skamarock W, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, et al. (2008) A description of the advanced research WRF version 3: NCAR technical note TN–475+STR. National Center for Atmospheric Research Boulder, Colorado, USA, 125 p.

Trusilova K, Früh B, Brienen S, Walter A, Masson V, Pigeon G, Becker P, 2013. Implementation of an Urban Parameterization Scheme into the Regional Climate Model COSMO-CLM. J. Appl. Meteor. Climatol., vol. 52, no. 10, p. 2296–2311.

World Urbanization Prospects, the 2014 Revision: Highlights, 2014. United Nations, Department of Economic and Social Affairs, Population Division: New York, 32 p.