

# Starting With UrbanSim: On the Creation of an Introductory Project

Olivier Gallay

Ecole Polytechnique Fédérale de Lausanne (EPFL)  
TRANSP-OR, Transport and Mobility Laboratory

Internal Seminar  
September 23<sup>rd</sup> 2010



# UrbanSim Basics

- Open source urban simulation system (OPUS)
- Support for planning and analysis of urban development
- Interactions between land use, transportation, the economy, the environment, etc
- Python based modularized architecture to facilitate the insertion of plug-ins
- Microsimulation, agent-based, approach

# Context

- SustainCity: Brussels case study
- Current dataset not yet available
- Proceed to a preliminary study and identify potential problems arising with the software
- Incomplete (and old) dataset for Brussels
- Use the latest Developer version
- Start from the `san_antonio_zone` project and adapt it to the Brussels dataset

# Software Architecture

- Layer construction:
  - `urbansim` (general layer)
  - `urbansim_zone` (zone layer)
  - `brussels_zone` (project layer)
- Strict predominance between layers: child, parent files
- Python code and `.xml` above layer:
  - `urbansim.xml`: model general specifications, database connections,...
  - `urbansim_zone.xml`: model parameters
  - `brussels_zone.xml`: model specific configuration (variable set)

# UrbanSim Data Cache Format

The screenshot displays the UrbanSim software interface. On the left, a file explorer shows a tree structure of data folders under 'Opus Data'. The 'jobs' folder is selected. On the right, a 'jobs' window displays a table view for the year 2005, run name 'base\_year\_data'. The table contains 1353417 records with identifiers in the range 1-1353417. Below the table, a 'Table View' section shows a preview of the data with columns for home\_based\_status, job\_id, sector\_id, grid\_id, building\_id, and zone\_id.

**Jobs Table Data:**

home_based_status	job_id	sector_id	grid_id	building_id	zone_id
0.0	0.01	73	0	1	
676709.26	390698.31	1.04178e+09	1	1.35342e+06	
3.5	1.71	4.73573e+06	1	6	
188175.3	48032.7	1.27658e+09	2948	313667	
76624.3	25425.96	6.25419e+08	36375	111818	
243653.34	1847485.35	-9.47903e+08	110050	2.14748e+09	

**Table View:**

home_based_status	job_id	sector_id	grid_id	building_id	zone_id
0	1	5	302670	36400	110050
0	2	1	305257	36414	110050
0	3	6	296215	36376	110050
0	4	5	301382	36392	110050
0	5	1	302660	36399	110050
0	6	1	302031	36398	110050
0	7	5	296861	36378	110050
0	8	4	304615	36410	110050

# Tools to Manage Data

The screenshot displays the 'Tools' panel in the UrbanSim software, specifically the 'Opus Data' sub-tab. The panel is organized into a tree view under the 'tool\_library' root. The tools listed are:

- synthesizer\_tools
  - import\_raw\_pums\_data\_to\_db
  - delete\_empty\_households\_fro...
  - create\_households\_input\_table
  - create\_persons\_input\_table
  - import\_sf3\_raw\_data\_to\_db
  - import\_pums\_id\_to\_bg\_id\_to\_db
  - create\_households\_marginals\_t...
  - create\_persons\_marginals\_table
  - create\_tables
  - prepare\_data
- opus\_data\_import\_export\_tools
  - sql\_data\_to\_opus\_tool
  - esri\_data\_to\_opus\_tool
  - opus\_data\_to\_sql\_tool
  - opus\_data\_to\_esri\_tool
  - opus\_data\_to\_csv\_tool
- data\_conversion\_tools
  - raw\_sql\_query\_tool
  - csv\_to\_sql\_tool
  - shapefile\_to\_postgis
- spatial\_database\_tools
  - postgis\_polygon\_to\_centroids
  - postgis\_create\_spatial\_index
- esri\_data\_tools
  - esri\_buffer\_tool
  - esri\_feature\_to\_point\_tool
- data\_imputation\_tools
  - missing\_value\_replacement\_tool
  - niv\_run\_from\_configuration
  - outlier\_detection\_tool
  - nd\_run\_from\_configuration

# Database Connection (Importation, Exportation)

The image shows two overlapping windows from the Opus Data software interface. The background window is the 'Tools' panel, and the foreground window is the 'Database Server Connections' dialog.

**Tools Panel (Opus Data):**

Name	Value
tool_library	
synthesizer_tools	
import_raw_pums_data_to_db	
delete_empty_households_fro...	
create_households_input_table	
create_persons_input_table	
import_sf3_raw_data_to_db	
import_pums_id_to_bg_id_to_db	
create_households_marginals_t...	
create_persons_marginals_table	
create_tables	
prepare_data	
opus_data_import_export_tools	
sql_data_to_opus_tool	
esri_data_to_opus_tool	
opus_data_to_sql_tool	
opus_data_to_esri_tool	
opus_data_to_csv_tool	
data_conversion_tools	
raw_sql_query_tool	
csv_to_sql_tool	
shapefile_to_postgis	
spatial_database_tools	
postgis_polygon_to_centroids	
postgis_create_spatial_index	
esri_data_tools	
esri_buffer_tool	
esri_feature_to_point_tool	
data_imputation_tools	
missing_value_replacement_tool	
mv_run_from_configuration	
outlier_detection_tool	
ndt_run_from_configuration	

**Database Server Connections Dialog:**

Database Settings

Name	Value
xml_version	4.2.0
services_database_server	
protocol	sqlite
host_name	localhost
user_name	urbansim
password	*****
scenario_database_server	
protocol	mysql
host_name	localhost
user_name	urbansim
password	*****
estimation_database_server	
protocol	mysql
host_name	localhost
user_name	urbansim
password	*****
indicators_database_server	
protocol	mysql
host_name	localhost
user_name	urbansim
password	*****
mysql_test_database_server	
protocol	mysql
host_name	localhost
user_name	urbansim
password	*****
postgres_test_database_server	
protocol	postgres
host_name	localhost

Buttons: OK, Cancel

# Models Definition

The screenshot displays the 'Submodel Editor' window. On the left, a tree view lists various models, with 'submodel' selected under the 'Estimation Configuration' folder. The main editor area is divided into two tabs: 'Variables' and 'Nests and Equations'. The 'Variables' tab is active, showing a table of variables and their definitions.

variable	definition
(X) avg_hh_income	building.aggregate(household.income, function=mean)
(X) ln_income_sqft_per_unit	ln(household.income*building.sqft_per_unit)
(X) persons_sqft_per_unit	household.persons*building.sqft_per_unit

Below the table, there is a button labeled 'Add/Remove variables...' and a checkbox for 'Show advanced variable parameters'. At the bottom of the window, there are buttons for 'Help on...', 'Save and Close', and 'Cancel'.



# Variable Library

Variable Library

Show variables from dataset: [All datasets]

Name	Dataset	Use	Type	Definition
(X) ln_lm_fer	building	I+M	Exp	ln(building.land_area/building.non_residential_sqft)
(X) ln_pop_zone	building	I+M	Exp	ln(building.disaggregate(zone.aggregate(household.persons)))
(X) ln_non_residential_sqft	building	I+M	Exp	ln(building.non_residential_sqft)
(X) unit_price	building	M	Pri	building.average_value_per_unit
(X) is_office	building	I+M	Exp	building.building_type_id==5
(X) is_warehouse	building	I+M	Exp	building.building_type_id==10
(X) is_school	building	I+M	Exp	building.building_type_id==8
(X) is_retail	building	I+M	Exp	building.building_type_id==7
(X) is_government	building	I+M	Exp	building.building_type_id==1
(X) ln_land_per_unit	building	I+M	Exp	ln(building.land_area/building.residential_units)
(X) ln_avg_hh_income	building	M	Exp	ln(building.aggregate(household.income, function=mean))
(X) land_per_unit	building	I+M	Exp	building.land_area/building.residential_units
(X) ln_jobs_within_30_min	building	M	Exp	ln(building.disaggregate(urbansim.parcel.zone.employment_within_30_minutes_travel_time_hbw_am_drive_alone))
(X) ln_number_jobs_retail	building	I+M	Exp	ln(building.disaggregate(urbansim.zone.number_of_jobs_of_sector_5))
(X) ln_jobs_within_15_min	building	M	Exp	ln(building.disaggregate(urbansim.parcel.zone.employment_within_15_minutes_travel_time_hbw_am_drive_alone))
(X) ln_pop_density_zone	building	I+M	Exp	ln(building.aggregate(household.persons)/building.disaggregate(zone.aggregate(building.land_area)))
(X) cbd_time_sq	building	M	Exp	building.disaggregate(zone.travel_time_to_cbd)**2
(X) ln_cbd_time	building	M	Exp	ln(building.disaggregate(zone.travel_time_to_cbd))
(X) ln_avg_hh_income_zonal	building	M	Exp	ln(building.disaggregate(zone.aggregate(household.income, function=mean)))
(X) ln_avg_val_per_unit	building	M	Pri	ln(building.average_value_per_unit)
(X) ln_residential_units	building	I+M	Exp	ln(building.residential_units)
(X) is_single_family	building	I+M	Exp	building.building_type_id == 9
(X) ln_job_density_zone	building	I+M	Exp	ln(building.disaggregate(zone.number_of_agents(job))/building.disaggregate(zone.aggregate(building.land_area)))
(X) avg_value_per_unit	building	M	Pri	building.average_value_per_unit

# Variable (and/or Indicator) Creation

The screenshot shows the 'Variable Library' dialog box with a list of variables. A 'Dialog' window is open over it, showing the configuration for a new variable named 'zone\_tot\_persons'.

**Variable Library Table:**

Name	Dataset	Use	Type	Definition
zone_mf_unit_price	zone	I+M	Exp	zone.aggregate(where(building.building_type_id==3,building.average_value_per_unit,0), function=mean)
zone_sf_unit_price	zone	I		
zone_hh_mf_units	zone			
zone_avg_cars	zone			
zone_hh_avg_income	zone			
zone_hh_emp_30_min	zone	I		
zone_hh_dev_acre	zone	I		
zone_hh_totacres	zone			
zone_hh_pop	zone			
zone_tot_cars	zone			
zone_tot_income	zone			
zone_tot_pop	zone			
zone_retail_emp_45_min	zone			
zone_retail_emp_15_min	zone			
zone_hh_retail_emp_30_min	zone			
zone_hh_sf_units	zone			
zone_hh_average_value_per_re...	zone			
zone_retail_emp_30_min	zone			
zone_hh_rd_density	zone			
zone_rd_density	zone			
zone_sch_district	zone			
zone_hh_average_value_per_sqft	zone			
zone_services_emp_25_min	zone			

**Dialog Window Configuration:**

- Name: zone\_tot\_persons
- Definition: zone.aggregate(household.persons)
- This variable is an **expression** that will be used as a **Model Variable**.
- Variable settings:
  - Use as a Model Variable
  - Use as an Indicator
  - Use as as both Model Variable and Indicator
- Variable type: expression

Buttons: Check syntax, Check against data, OK, Cancel.

# Model Parameters

The screenshot displays the UrbanSim software interface with the 'Models' tab selected. The left pane shows a hierarchical tree view of model parameters. The 'household\_location\_choice\_model' is selected, and its 'Estimation Configuration' sub-tree is expanded. The right pane is currently empty.

Name	Value
Models	
fertility_model	
mortality_model	
refinement_model	
scheduled_development_events_m...	
scheduled_employment_events_mo...	
employment_relocation_model	
household_relocation_model	
distribute_unplaced_jobs_model	
household_transition_model	
employment_location_choice_model	
household_location_choice_model	
Estimation Configuration	
structure	
run	
init	
import	
estimate	
prepare_for_run	
prepare_for_estimate	
output	(specification, index)
portion_to_unplace	0.333333333333
specification_table	household_location_choice_model_specification
agent_set	household
index_to_unplace	None
specification_storage	base_cache_storage
filter	
join_datasets	<input type="checkbox"/>
agents_for_estimation...	households_for_estimation
agents_for_estimation...	base_cache_storage
data_objects	datasets
name	prepare_for_estimate
specification	
residential_development_project_lo...	
non_residential_development_proje...	
development_project transition m...	

# Model Estimation

The screenshot shows the UrbanSim software interface with the 'Models' tab selected. The 'Models' panel on the left contains a list of model names under the heading 'Name' and 'Value'. The 'employment\_location\_choice\_model' is selected, and a context menu is open over it, with 'Run Estimation' highlighted. Other models in the list include 'fertility\_model', 'mortality\_model', 'refinement\_model', 'scheduled\_development\_events\_model', 'scheduled\_employment\_events\_model', 'employment\_relocation\_model', 'household\_relocation\_model', 'distribute\_unplaced\_jobs\_model', 'household\_transition\_model', 'non\_residential\_development\_project\_location\_choice\_model', 'development\_project\_transition\_model', 'governmental\_employment\_location\_choice\_model', 'employment\_transition\_model', 'real\_estate\_price\_model', and 'add\_projects\_to\_buildings'. The 'Estimation Configuration' section is also visible at the bottom of the list.

# Model Estimation: Errors...

```

an)...2.7 sec
sec      cbd_time = building disaggregate(zone.travel_time_to_cbd)...0.3
      sec      Estimating Real Estate Price Model (from urbansim.models.real_estate
price_model): completed...3.4 sec
      Simulate year 2005: completed.....4.0 sec
Closing log file: C:\opus\data\san_antonio_zone\base_year_data\year_2005
_log.txt
Starting simulation for year 2005: completed.....4.0 sec
Start simulation run: completed.....4.0 sec
Closing log file: C:\opus\data\san_antonio_zone\base_year_data\run_model_system.
log
ERROR: Traceback (most recent call last):
  File "C:\opus\src\opus_gui\models_manager\run\run_estimation.py", line 124, in
run
    self.er.estimate()
  File "C:\opus\src\urbansim\estimation\estimator.py", line 72, in estimate
    self.model_system.run(self.config, write_datasets_to_cache_at_end_of_year=Fa
lse)
  File "C:\opus\src\opus_core\model_coordinators\model_system.py", line 128, in
run
    write_datasets_to_cache_at_end_of_year=write_datasets_to_cache_at_end_of_yea
r)
  File "C:\opus\src\opus_core\model_coordinators\model_system.py", line 289, in
_run_year
    self vardict[outputvar] = self.do_process(locals())
  File "C:\opus\src\opus_core\model_coordinators\model_system.py", line 378, in
do_process
    return eval(eu)
  File "<string>", line 1, in <module>
  File "C:\opus\src\opus_core\model.py", line 51, in logged_estimate_method
    results = estimate_method(*req_args, **opt_args)
  File "C:\opus\src\urbansim\models\real_estate_price_model.py", line 78, in est
imate
    estimate_config=estimate_config, debuglevel=debuglevel)
  File "C:\opus\src\opus_core\regression_model.py", line 226, in estimate
    dataset.compute_variables([outcome_attribute], dataset_pool=self.dataset_pool
L, resources=compute_resources)
  File "C:\opus\src\opus_core\datasets\abstract_dataset.py", line 625, in comput
e_variables
    (versions, value) = self.compute_variables_return_versions_and_final_value(n
ames, dataset_pool, resources, quiet)
  File "C:\opus\src\opus_core\datasets\abstract_dataset.py", line 644, in comput
e_variables_return_versions_and_final_value
    qualified_name = self.create_and_check_qualified_variable_name(name)
  File "C:\opus\src\opus_core\datasets\abstract_dataset.py", line 1804, in creat
e_and_check_qualified_variable_name
    self._check_dataset_name(qualified_name.get_dataset_name())
  File "C:\opus\src\opus_core\datasets\abstract_dataset.py", line 1823, in _chec
k_dataset_name
    raise ValueError, 'different dataset names for variable and dataset'
ValueError: different dataset names for variable and dataset
Error returned from Estimation
Estimation Finished with success = False

```

```

File "C:\opus\src\opus_core\variables\variable.py", line 69, in logged_method
results = compute_method(*req_args, **opt_args)
File "C:\opus\src\opus_core\variables\variable.py", line 142, in compute_with
dependencies
    self._solve_dependencies(dataset_pool)
File "C:\opus\src\opus_core\variables\variable.py", line 206, in _solve_depend
encies
    (new_versions, value) = ds.compute_variables_return_versions_and_final_value
([(depvar_name, version)], dataset_pool)
File "C:\opus\src\opus_core\datasets\abstract_dataset.py", line 652, in comput
e_variables_return_versions_and_final_value
    resources=resources, quiet=quiet, version=version))
File "C:\opus\src\opus_core\datasets\abstract_dataset.py", line 1971, in _comp
ute_if_needed
    return self._compute_one_variable(variable_name, dataset_pool, resources=res
ources, quiet=quiet)
File "C:\opus\src\opus_core\datasets\abstract_dataset.py", line 1915, in _comp
ute_one_variable
    data=variable.compute_with_dependencies(dataset_pool, compute_resources),
File "C:\opus\src\opus_core\variables\variable.py", line 69, in logged_method
results = compute_method(*req_args, **opt_args)
File "C:\opus\src\opus_core\variables\variable.py", line 142, in compute_with
dependencies
    self._solve_dependencies(dataset_pool)
File "C:\opus\src\opus_core\variables\variable.py", line 206, in _solve_depend
encies
    (new_versions, value) = ds.compute_variables_return_versions_and_final_value
([(depvar_name, version)], dataset_pool)
File "C:\opus\src\opus_core\datasets\abstract_dataset.py", line 652, in comput
e_variables_return_versions_and_final_value
    resources=resources, quiet=quiet, version=version))
File "C:\opus\src\opus_core\datasets\abstract_dataset.py", line 1971, in _comp
ute_if_needed
    return self._compute_one_variable(variable_name, dataset_pool, resources=res
ources, quiet=quiet)
File "C:\opus\src\opus_core\datasets\abstract_dataset.py", line 1983, in _comp
ute_one_variable
    index_name=id_name)
File "C:\opus\src\opus_core\variables\variable_factory.py", line 81, in get_va
riable
    % (dataset_name, short_name))
LookupError: Incomplete variable specification for 'building.residential_units'
(missing package name).
Error returned from Estimation
Estimation Finished with success = False

```

# Model Estimation (Continued)

The screenshot shows the 'household\_location\_choice\_model estimation' window in UrbanSim. The window has a 'Start Estimation...' button and a progress bar that is 100% complete. Below the progress bar, it says 'Estimation finished successfully'.

The 'Estimation result log' contains the following text:

```

urbansim_zones.building.vacant_residential_units>: completed...1.7 sec
zone_id>0.....0.1 sec
avg_hh_income = building.aggregate(household.income, function=mean)....1.7 sec
ln(household.income*building.sqft_per_unit).....7.6 sec
household.persons*building.sqft_per_unit.....1.7 sec
Choice set size: 30
submodel: 1
Convergence achieved.
Akaike's Information Criterion (AIC): 659983.476553
Bayesian Information Criterion (BIC): 660012.79445
Number of Iterations: 20
*****
Log-likelihood is:      -329988.738276
Null Log-likelihood is: -440982.24652
Likelihood ratio index: 0.251696092347
Adj. likelihood ratio index: 0.251689289352
Number of observations: 129655
Suggested |t-value| > 3.43112698039
Convergence statistic is: 0.000990795543969
*****

```

Coeff_names	estimate	std err	t-values
avg_hh_income	0.000957247	7.01527e-06	136.452
ln_income_sqft_per_unit	0.0474364	0.00982334	4.82895
persons_sqft_per_unit	-3.00468e-05	3.59862e-05	-0.834952

```

*****
Elapsed time: 124.327111287 seconds
Estimating Household Location Choice Model (from
urbansim.models.household_location_choice_model): completed...2 min, 44.5 sec
Simulate year 2005: completed.....2 min, 57.8 sec

```

# Model Coefficients

The screenshot displays a software interface with a file explorer on the left and a results window on the right. The file explorer shows a tree structure under 'Opus Data' with 'base\_year\_data' expanded to '2005'. The 'household\_location\_choice\_model\_coefficients' folder is selected. The results window, titled 'household\_location\_choice\_model\_coefficients', shows the following data:

Year: 2005 Run name: base\_year\_data

name	mean	sd	sum	min	max
sub_model_id	1.0	0.0	3	1	1
t_statistic	46.82	77.68	140.446	-0.834952	136.452
estimate	0.02	0.03	0.0483636	-3.00468e-05	0.0474364
standard_error	0.0	0.01	0.00986634	7.01527e-06	0.00982334

Size: 3 records  
 identifiers:  
 \_hidden\_id\_ in range 1-3

Table View

sub_model_id	t_statistic	estimate	coefficient_name	standard_error
1	136.452	0.000957247	avg_hh_income	7.01527e-06
1	4.82895	0.0474364	ln_income_sqft_per_unit	0.00982334
1	-0.834952	-3.00468e-05	persons_sqft_per_unit	3.59862e-05

# Model Specification

The screenshot displays the UrbanSim software interface. On the left, a file tree under 'Opus Data' shows a hierarchy of folders, with 'household\_location\_choice\_model\_specification' selected. The main window on the right shows the details for this model specification, including a summary table and a table view.

**Model Specification Summary:**

name	mean	sd	sum	min	max
sub_model_id	1.0	0.0	3	1	1
equation_id	-2.0	0.0	-6	-2	-2

Year: 2005 Run name: base\_year\_data  
 Size: 3 records  
 identifiers:  
 \_hidden\_id\_ in range 1-3

**Table View:**

sub_model_id	equation_id	coefficient_name	variable_name
1	-2	persons_sqft_per_unit	persons_sqft_per_unit
1	-2	ln_income_sqft_per_unit	ln_income_sqft_per_unit
1	-2	avg_hh_income	avg_hh_income = building.aggregate(household.income, function=mean)



# Model Estimation: Facts

- Special attention has to be paid on `xxx_id` (`zone_id`, `building_id`, `household_id`, etc.) variables: links between different tables, aggregations
- Error messages when missing data:
  - Some tables are mandatory for the zone version of UrbanSim:  
`www.urbansim.org`:
  - Some mandatory entries
- Specific `employment_sector_groups`: the submodels of the `employment_location_choice` model have to be adapted

# Model Estimation: Facts (Continued)

- To be able to estimate a submodel, your dataset must contain sufficient data from the considered type
- No capital letters in your table names and entries
- The formatting of your data is important, no entries with empty values
- Your dataset must contain the coefficients and specifications of the used models and some UrbanSim constants (here 14 additional tables):
  - `annual_household_relocation_rates` table for the `household_relocation` model
  - `annual_employment_relocation_rates` table for the `employment_relocation` model
  - ...

# Model Simulation

The screenshot shows the 'Models' tab in the UrbanSim software. The 'models\_to\_run' folder is expanded, showing a list of models with checkboxes for selection. The 'Value' column shows the configuration for each model.

Name	Value
brussels_baseline_test	
brussels_baseline	
models_to_run	
refinement_model	<input type="checkbox"/>
real_estate_price_model	<input type="checkbox"/>
development_project_transition_model	<input type="checkbox"/>
residential_development_project_location_choice_model	<input type="checkbox"/>
non_residential_development_project_location_choice_model	<input type="checkbox"/>
add_projects_to_buildings	<input type="checkbox"/>
household_transition_model	<input checked="" type="checkbox"/>
employment_transition_model	<input type="checkbox"/>
household_relocation_model	<input checked="" type="checkbox"/>
household_location_choice_model	<input checked="" type="checkbox"/>
employment_relocation_model	<input checked="" type="checkbox"/>
employment_location_choice_model	<input checked="" type="checkbox"/>
distribute_unplaced_jobs_model	<input checked="" type="checkbox"/>
scheduled_development_events_model	<input checked="" type="checkbox"/>
scheduled_employment_events_model	<input checked="" type="checkbox"/>
model_system	urbansim.model_coordinators.model_system
base_year	2005
years_to_run	
firstyear	2005
lastyear	2019
cache_directory	san_antonio_zone\base_year_data
creating_baseyear_cache_configuration	
advanced	
travel_model_configuration	

# Model Simulation (Continued)

The screenshot displays the UrbanSim software interface, specifically the 'Scenarios' tab. The main window shows a hierarchical tree structure of configuration parameters for a simulation named 'brussels\_baseline\_test'. The tree is expanded to show the 'brussels\_baseline' sub-configuration, which includes various settings for the model system, years, cache, and datasets. The 'Value' column on the right provides the specific values for these parameters. A large empty white area is visible on the right side of the interface, likely reserved for simulation results or a map view.

Name	Value
brussels_baseline_test	
brussels_baseline	
models_to_run	
model_system	urbansim.model_coordinators.model_system
base_year	2005
years_to_run	
firstyear	2005
lastyear	2007
cache_directory	san_antonio_zone/base_year_data
creating_baseyear_cache_configuration	
advanced	
dataset_pool_configuration	
flush_dataset_to_cache_after_each_model	<input type="checkbox"/>
flush_variables	<input type="checkbox"/>
low_memory_run	<input type="checkbox"/>
datasets_to_preload	
zone	<input checked="" type="checkbox"/>
household	<input checked="" type="checkbox"/>
building	<input checked="" type="checkbox"/>
building	<input checked="" type="checkbox"/>
parcel	<input type="checkbox"/>
person	<input type="checkbox"/>
job	<input checked="" type="checkbox"/>
building_type	<input checked="" type="checkbox"/>
travel_data	<input checked="" type="checkbox"/>
target_vacancy	<input checked="" type="checkbox"/>
development_event_history	<input checked="" type="checkbox"/>
home_based_status	<input checked="" type="checkbox"/>
travel_model_configuration	

# Model Simulation (Continued)

The screenshot shows the UrbanSim software interface with the 'Scenarios' tab active. The 'brussels\_baseline\_test' scenario is selected, and a context menu is open over the 'Run This Scenario' option. The main window displays a list of models and their values for this scenario.

Name	Value
brussels_baseline_test	
Run This Scenario	
Duplicate	
Rename	
Delete	
non_residential_development_project_location_choice_model	
add_projects_to_buildings	
household_transition_model	
employment_transition_model	
household_relocation_model	
household_location_choice_model	
employment_relocation_model	
employment_location_choice_model	
distribute_unplaced_jobs_model	
scheduled_development_events_model	
scheduled_employment_events_model	
model_system	urbansim.model_coordinators.model_system
base_year	2005
years_to_run	
cache_directory	san_antonio_zone/base_year_data
creating_baseyear_cache_configuration	
advanced	
travel_model_configuration	

# Model Simulation (Continued)

The screenshot displays the UrbanSim software interface, divided into two main panels.

**Left Panel: Model Configuration Tree**

The tree shows a hierarchy of models under the 'brussels\_baseline\_test' and 'brussels\_base\_line' folders. The 'models\_to\_run' folder is expanded, listing various models with checkboxes for their execution status. The 'Value' column indicates the path for each model.

Name	Value
brussels_baseline_test	
brussels_base_line	
models_to_run	
refinement_model	
real_estate_price_model	
development_project_transition_model	
residential_development_project_location_choice_model	
non_residential_development_project_location_choice_model	
add_projects_to_buildings	
household_transition_model	
employment_transition_model	<input checked="" type="checkbox"/>
household_relocation_model	<input checked="" type="checkbox"/>
household_location_choice_model	<input checked="" type="checkbox"/>
employment_relocation_model	<input checked="" type="checkbox"/>
employment_location_choice_model	<input checked="" type="checkbox"/>
distribute_unplaced_jobs_model	<input checked="" type="checkbox"/>
scheduled_development_events_model	<input checked="" type="checkbox"/>
scheduled_employment_events_model	<input checked="" type="checkbox"/>
model_system	urbansim.model_coordinators.model_system
base_year	2005
years_to_run	
cache_directory	san_antonio_zone\base_year_data
creating_baseyear_cache_configuration	
advanced	
travel_model_configuration	

**Right Panel: brussels\_base\_line Simulation Progress**

This panel provides details for the 'brussels\_base\_line' simulation. It includes a 'Pause simulation run...' button, a 'Run Name' field containing 'run\_2010\_09\_21\_10\_32', and a 'Cancel' button. The 'Indicator Batch' is set to '(None)'. Below these are tabs for 'Simulation Progress', 'Log', and 'Diagnostics'. The 'Simulation Progress' tab shows:

- Total progress: 11% (indicated by a green progress bar)
- Status: (1/2) 2006
- Show progress within current year
- Year progress: 22% (indicated by a green progress bar)
- Status: (3/10) Household Location Choice Model
- Show progress within current model

# Model Simulation: Errors...

```

OPUS
self.model_system.run(self.config, write_datasets_to_cache_at_end_of_year=False)
File "C:\opus\src\opus_core\model_coordinators\model_system.py", line 128, in
run
write_datasets_to_cache_at_end_of_year=write_datasets_to_cache_at_end_of_year)
File "C:\opus\src\opus_core\model_coordinators\model_system.py", line 289, in
_run_year
self vardict[outputvar] = self.do_process(locals())
File "C:\opus\src\opus_core\model_coordinators\model_system.py", line 370, in
do_process
return eval(ev)
File "<string>", line 1, in <module>
File "C:\opus\src\opus_core\model.py", line 51, in logged_estimate_method
results = estimate_method(*req_args, **opt_args)
File "C:\opus\src\urbansim\models\agent_location_choice_model_member.py", line
47, in estimate
agents_index=agents_index[new_agents_index], **kwargs)
File "C:\opus\src\urbansim\models\location_choice_model.py", line 198, in estimate
debuglevel<debuglevel)
File "C:\opus\src\opus_core\choice_model.py", line 364, in estimate
self.create_interaction_datasets(agent_set, agents_index_for_estimation, estimate_config, submodels=submodels)
File "C:\opus\src\urbansim\models\location_choice_model.py", line 291, in create_interaction_datasets
nchunks=nchunks, chunksize=chunksize)
File "C:\opus\src\opus_core\choice_model.py", line 511, in sample_alternatives_by_chunk
dataset_pool=self.dataset_pool
File "C:\opus\src\opus_core\configurable.py", line 28, in config_run_method
results = run_method(*req_args, **opt_args)
File "C:\opus\src\opus_core\samplers\weighted_sampler.py", line 151, in run
sampling_prob = column_stack([sampling_prob_for_chosen_choices, sampling_prob])
File "C:\Python26\lib\site-packages\numpy\lib\shape_base.py", line 297, in column_stack
return nx.concatenate(arrays,1)
MemoryError

```

```

Select OPUS
File "C:\opus\src\opus_core\model_coordinators\model_system.py", line 289, in
_run_year
self vardict[outputvar] = self.do_process(locals())
File "C:\opus\src\opus_core\model_coordinators\model_system.py", line 370, in
do_process
return eval(ev)
File "<string>", line 1, in <module>
NameError: name 'hwn_index' is not defined
Writing specification and coefficients into C:\opus\data\san_antonio_zone\base_y
ear_data: completed. 2 min, 34.4 sec
ERROR: Traceback (most recent call last):
File "C:\opus\src\opus_gui\scenarios_manager\run\run_simulation.py", line 206,
in run
run_id = run_manager.run_run(config, run_name = run_name)
File "C:\opus\src\opus_core\services\run_server\run_manager.py", line 107, in
run_run
model_system.run_multiprocess(run_resources)
File "C:\opus\src\opus_core\model_coordinators\model_system.py", line 466, in
run_multiprocess
self._run_each_year_as_separate_process(start_year, end_year, seed_array, re
sources)
File "C:\opus\src\urbansim\model_coordinators\model_system.py", line 35, in _r
un_each_year_as_separate_process
urbansim.model_coordinators.model_system', resources, optional_args=['-log
file-name', log_file_name])
File "C:\opus\src\opus_core\model_coordinators\model_system.py", line 563, in
_fork_new_process
self.forked_processes[-1].wait()
File "C:\opus\src\opus_core\fork_process.py", line 79, in wait
self.check_status()
File "C:\opus\src\opus_core\fork_process.py", line 85, in check_status
raise StandardError("Child python process exited with failure.\nCalling modu
le: %s\nSystem command: %s" % (self.module_name, self.python_cmd))
StandardError: Child python process exited with failure.
Calling module: urbansim.model_coordinators.model_system
System command: 'C:\Python26\python.exe', 'C:\opus\src\urbansim\model_coo
rdinators\model_system.py', '-r', 'C:\docume~1\ADMINI~1\locals1\temp\tpa
nkg~1\resources.pickle', '--log-file-name', 'run_multiprocess.log')
Error returned from Model
Traceback (most recent call last):
File "C:\opus\src\opus_gui\scenarios_manager\run\run_simulation.py", line 34,
in run
self.model.element.model.run()
File "C:\opus\src\opus_gui\scenarios_manager\run\run_simulation.py", line 227,
in run
self.finishedCallback(succeeded, run_name = run_name)
File "C:\opus\src\opus_gui\scenarios_manager\run\run_simulation.py", line 77,
in finishedCallback
run_name = run_name)
File "C:\opus\src\opus_gui\results_manager\results_manager_functions.py", line
62, in delete_simulation_run
get_manager_instance('results_manager').delete_run(run_node)
File "C:\opus\src\opus_gui\results_manager\results_manager.py", line 75, in de
lete_run
self.xml_controller.delete_run(run_node, force=force)
File "C:\opus\src\opus_gui\results_manager\controllers\xml_configuration\xml_c
ontroller_results.py", line 67, in delete_run
cache_directory = run_node.find('cache_directory').text
AttributeError: 'NoneType' object has no attribute 'find'

```

# Model Simulation (Continued)

The screenshot displays the UrbanSim software interface, divided into two main panels.

**Left Panel: Model Configuration**

This panel shows a tree view of the model configuration under the name "brussels\_baseline\_test". The "brussels\_baseline" sub-model is expanded, revealing a list of components and their status:

Name	Value
brussels_baseline_test	
brussels_baseline	
models_to_run	
refinement_model	<input type="checkbox"/>
real_estate_price_model	<input type="checkbox"/>
development_project_transition_model	<input type="checkbox"/>
residential_development_project_location_choice_model	<input type="checkbox"/>
non_residential_development_project_location_choice_model	<input type="checkbox"/>
add_projects_to_buildings	<input type="checkbox"/>
household_transition_model	<input type="checkbox"/>
employment_transition_model	<input checked="" type="checkbox"/>
household_relocation_model	<input type="checkbox"/>
household_location_choice_model	<input checked="" type="checkbox"/>
employment_relocation_model	<input checked="" type="checkbox"/>
employment_location_choice_model	<input checked="" type="checkbox"/>
distribute_unplaced_jobs_model	<input checked="" type="checkbox"/>
scheduled_development_events_model	<input checked="" type="checkbox"/>
scheduled_employment_events_model	<input checked="" type="checkbox"/>
model_system	urbansim.model_coordinators.model_system
base_year	2005
years_to_run	
cache_directory	san_antonio_zone/base_year_data
creating_baseyear_cache_configuration	
advanced	
travel_model_configuration	

**Right Panel: Simulation Progress (brussels\_baseline)**

This panel shows the simulation progress for the "brussels\_baseline" model. It includes the following information:

- Start Simulation Run...:** Run Name: run\_2010\_09\_21\_10\_32
- Cancel** button
- Indicator Batch:** (None)
- Simulation Progress:** Log, Diagnostics
- Total progress:** 100% (indicated by a full green progress bar)
- Status:** Simulation ran successfully!
- Show progress within current year
- Year progress:** 100% (indicated by a full green progress bar)
- Status:** Finished
- Show progress within current model



# Model Simulation: Facts

- Some models have to be run together (*i.e.* the `household_relocation` model has to be run before the `household_location_choice` model)
- Some models cannot be deleted because mandatory
- In the current version, control totals are used:
  - `annual_household_control_totals`
  - `annual_employment_control_totals`but fertility and mortality models can be used in addition to that.

## Simulation Results: Data

The screenshot shows the Opus Data software interface. On the left is a file tree with folders for 'base\_year\_data', 'runs', '2005', '2006', '2007', and 'shapefiles'. The 'jobs' folder is selected. On the right, a window titled 'jobs' displays data for the year 2007, run name 'run\_5.run\_2010\_09\_21\_10\_41'. The data is presented as a table with columns: home\_based\_status, building\_id, job\_id, zone\_id, sector\_id, and grid\_id. Below the table, it indicates 'Size: 1353417 records' and 'identifiers: \_hidden\_id\_ in range 1-1353417'. A 'Table View' section below the window shows a detailed view of the data table.

**Table View**

home_based_status	building_id	job_id	zone_id	sector_id	grid_id
1	36421	995	110050	3	307190
1	36421	755	110050	4	307190
1	36420	750	110050	6	306547
1	36419	775	110050	2	306546
1	36418	758	110050	6	306545
1	36417	989	110050	2	305907
1	36417	974	110050	4	305907
1	36417	772	110050	4	305907

# Simulation Results: Indicators

The screenshot shows the 'Batch Indicator visualization' window in a simulation software. The window is divided into several sections:

- Left Panel (Tree View):** Shows a hierarchical structure of simulation results. The 'Results' tab is active, showing a tree with 'Indicator Batches' and 'Simulation Runs'. Under 'Indicator Batches', there is an 'Indicator' sub-item containing a 'New visualization' and two other indicator batches. Under 'Simulation Runs', there are several run identifiers and a 'base\_year\_data' item.
- Top Panel (Batch Indicator visualization):** Contains configuration options for the visualization.
  - Output options:**
    - Visualization name: New visualization
    - Type: Table
    - Dataset name: zone
  - Format options:**
    - Format: Tab delimited file (.tab)
    - Output a single table
    - Output a table for every year
    - Output a table for each indicator
- Table of Available Indicators:** A table listing 14 indicators with their names and definitions.
 

	Name	Definition
1	zone_ln_sf_unit...	ln(zone.aggregate(where(building.building_type_id==9,building.a...
2	zone_ln_mf_unit...	ln(zone.aggregate(where(building.building_type_id==3,building.a...
3	zone_ln_emp_10...	ln(urbansim_parcel.zone.employment_within_10_minutes_travel_...
4	zone_ln_time_cbd	ln(zone.travel_time_to_cbd)
5	zone_ln_res_units	ln(zone.aggregate(building.residential_units))
6	zone_ln_emp_15...	ln(urbansim_parcel.zone.employment_within_15_minutes_travel_...
7	zone_time_cbd	zone.travel_time_to_cbd
8	zone_mf_unit_pri...	zone.aggregate(where(building.building_type_id==3,building.ave...
9	zone_sf_unit_price	zone.aggregate(where(building.building_type_id==9,building.ave...
10	zone_avg_cars	zone.aggregate(household.cars)/zone.number_of_agents(house...
11	zone_ln_emp_30...	ln(urbansim_parcel.zone.employment_within_30_minutes_travel_...
12	zone_ln_dev_acre	ln(zone.dev_acre)
13	zone_tot_cars	zone.aggregate(household.cars)
14	zone_schl_district	zone.schl_district
- Right Panel (Indicators in current visualization):** A panel for managing the indicators currently displayed in the visualization, featuring a '+' button to add indicators and a '-' button to remove them.

# Simulation Results: Indicators (Continued)

The screenshot displays the 'Results' tab in the UrbanSim software. On the left, a tree view shows the hierarchy of simulation data. The 'Indicator Batches' folder is expanded, and a context menu is open over it. The 'Run indicator batch on...' option is selected, and a sub-menu is visible with the following items:

- run\_2010\_09\_16\_10\_48
- run\_2010\_09\_15\_09\_29
- base\_year\_data

The tree view also shows other folders like 'Simulations' and 'Data', and various simulation runs listed under 'run\_2010\_09\_14\_11\_46' and 'run\_2010\_08\_30\_16\_24'.

# Simulation Results: Indicators (Continued)

The screenshot displays the UrbanSim software interface. The 'Results' tab is active, showing a tree view on the left and a data table on the right.

**Tree View Structure:**

- Indicator Batches
  - Indicator
    - New visualization
    - zone\_indicator\_batch
    - untkled\_indicator\_batch
- Simulation Runs
  - run\_2010\_09\_16\_10\_48
  - run\_2010\_09\_15\_09\_29
  - run\_2010\_09\_14\_16\_56
  - run\_2010\_09\_14\_16\_46
  - run\_2010\_09\_14\_14\_54
  - run\_2010\_09\_14\_12\_27
  - run\_2010\_09\_14\_12\_01
  - run\_2010\_09\_14\_11\_58
  - run\_2010\_09\_14\_11\_46
  - run\_2010\_09\_14\_11\_42
  - run\_2010\_09\_14\_11\_40
  - run\_2010\_09\_14\_11\_34
  - run\_2010\_09\_14\_11\_27
  - run\_2010\_09\_08\_13\_49
  - run\_2010\_09\_08\_13\_04
  - run\_2010\_09\_06\_09\_00
  - run\_2010\_08\_30\_16\_55
  - run\_2010\_08\_30\_16\_45
  - run\_2010\_08\_30\_16\_24
  - base\_year\_data

**Data Table: zone\_table-1\_2005\_zone\_\_zone\_tot\_persons**

zone_id	zone_tot_persons_2005
110050	14967.0
110370	13214.0
120050	12025.0
120090	14580.0
120250	70567.0
120290	14047.0
120300	14255.0
120340	6559.0
120350	17022.0
120400	20699.0
210010	89699.0
210020	29270.0
210030	18187.0
210041	40666.0
210042	16397.0
210043	57617.0
210044	21977.0
210045	10342.0
210050	45780.0

# Simulation Results

zone\_table-1\_2005\_zone\_\_zone\_tot\_persons

zone_id	zone_tot_persons_2005
110050	14967.0
110370	13214.0
120050	12025.0
120090	14580.0
120250	70567.0
120290	14047.0
120300	14255.0
120340	6559.0
120350	17022.0
120400	20699.0
210010	89699.0
210020	29270.0
210030	18187.0
210041	40666.0
210042	16397.0
210043	57617.0
210044	21977.0
210045	10342.0
210050	45780.0

zone\_table-1\_2006\_zone\_\_zone\_tot\_persons

zone_id	zone_tot_persons_2006
110050	14256.0
110370	13208.0
120050	13256.0
120090	14788.0
120250	70211.0
120290	14910.0
120300	15034.0
120340	6807.0
120350	18071.0
120400	20356.0
210010	85418.0
210020	28566.0
210030	17332.0
210041	37758.0
210042	15329.0
210043	54006.0
210044	21328.0
210045	10045.0
210050	42830.0

zone\_table-1\_2007\_zone\_\_zone\_tot\_persons

zone_id	zone_tot_persons_2007
110050	14316.0
110370	13570.0
120050	13504.0
120090	14947.0
120250	70846.0
120290	15150.0
120300	15337.0
120340	6944.0
120350	18215.0
120400	20536.0
210010	85613.0
210020	28693.0
210030	17325.0
210041	37772.0
210042	15344.0
210043	54217.0
210044	21524.0
210045	10101.0
210050	42825.0

## Simulation Results (Continued)

zone\_table-1\_2005\_zone\_\_zone\_tot\_cars

zone_id	zone_tot_cars_2005
110050	4076.0
110370	3698.0
120050	3407.0
120090	3950.0
120250	20758.0
120290	3852.0
120300	3891.0
120340	1880.0
120350	4841.0
120400	5637.0
210010	25212.0
210020	8932.0
210030	5269.0
210041	12281.0
210042	5275.0
210043	16941.0
210044	6200.0
210045	3644.0
210050	14335.0

zone\_table-1\_2006\_zone\_\_zone\_tot\_cars

zone_id	zone_tot_cars_2006
110050	3960.0
110370	3737.0
120050	3695.0
120090	4038.0
120250	20809.0
120290	4021.0
120300	4054.0
120340	1938.0
120350	5103.0
120400	5584.0
210010	24590.0
210020	8877.0
210030	5132.0
210041	11792.0
210042	5076.0
210043	16331.0
210044	6115.0
210045	3607.0
210050	13846.0

zone\_table-1\_2007\_zone\_\_zone\_tot\_cars

zone_id	zone_tot_cars_2007
110050	3977.0
110370	3814.0
120050	3765.0
120090	4070.0
120250	20935.0
120290	4140.0
120300	4122.0
120340	1977.0
120350	5138.0
120400	5652.0
210010	24631.0
210020	8896.0
210030	5122.0
210041	11793.0
210042	5081.0
210043	16365.0
210044	6173.0
210045	3594.0
210050	13856.0

# Models that Work

- household\_transition model
- employment\_transition model
- household\_relocation model
- household\_location\_choice model
- employment\_relocation model
- employment\_location\_choice model
- distribute\_unplaced\_jobs model
- scheduled\_development\_events model
- scheduled\_employment\_events model



# What Still Does Not Work

## Development Project Location Choice Model:

- Dependent on your specific `building_types`
- Lack of data...
- Dataset used for estimation: `development_event_history`
- Must have sufficient data for each submodel to be estimated
- `residential_units_capacity > residential_units`

# What Still Does Not Work (Continued)

## Real Estate Price Model:

- Different `building_types` dataset
- Specific submodel for each `building_type`
- Tried to simplify or remove some submodels to cope to the Brussels case and to the specific `building_types`
- However, some specificities of the San Antonio project remain...
- Dataset concerned here for estimation: `buildings` and `building_types`
- Outcome of the model:  
`building.average_value_per_unit`
- The same kind of problem must have appeared in the other case studies (Zurich, Paris)

# What Remains to Do with the New Data

## What Remains to Do with the New Data:

- Results visualization
- Integration of Matsim add-on: more precise transportation simulation aspects
  - Python ↔ Java
  - Communication
  - Data transfer