PECAS - for Spatial Economic Modelling

THE ARC PECAS MODEL: Model Improvements and Integration with the Activity Based Model

System Documentation Technical Note

HBA
specto

Calgary, Alberta

December 2015

Table of Contents

1. Introduction

The ARC PECAS model is a spatial economic model for the Atlanta Region that simulates the location of the households and business in the region, as well as the quantity of development through time from 2005 to 2050. It is based on choice theories and profit motivation theories of behaviour, combined with information on economic relationships. It is used by ARC for forecasting, research, and policy analysis.

The ARC PECAS model has been developed using Agile Process Management, with incremental improvement and delivery of functional versions in iterations. A complete iteration of the ARC PECAS model was delivered in 2014. Enhancements to the model were undertaken during 2015 to better represent the needs of the Atlanta Regional Commission (ARC). The agency had developed an Activity Based Travel Demand Model (ABM), a substantial improvement in the resolution and behavioral representation of travel in the region. Both models, the ARC PECAS and the ABM, worked together in 2015 to assist and support ARC in The Region's Plan Forecast, and both are designed to continue to serve ARC for forecasting, policy analysis, and policy exploration.

During 2015, the ARC PECAS model was improved in several ways. In addition, the integration with the ABM model was made fully operational and tested, so the ARC PECAS generates the socioeconomic outputs for the transport model and receives information on future travel conditions ("skims") from the ABM. The skims are used to calculate transport utilities for economic relationships (e.g. labor market commuting) that affect the exchange location choice and ultimately the location and technology choice of the businesses, households, and other institutions in the Atlanta Region.

The enhancements included in the ARC PECAS model can be organized in two types, some associated with improved model applications and some adding features or improving functionality. Specifically, they are:

Improved model application:

- The Space Development (SD) module was re-calibrated to match updated targets and more consistent statistical distributions for prior knowledge of parameters.
- The growth rates for the model-wide activity amounts (households and industries) through time were updated using data from an updated economic forecast.
- Previously forecasted and observed activity amounts (households and business) were made available (at the TAZ level, as "constraints") to match most recent observations (2005 to 2015), and to facilitate updates of previous forecasts and to facilitate comparisons to previous work (2016 to 2050).

Model functionalities:

- The Activity Allocation module was re-calibrated, specifically at the location choice level, through *floorspace calibration*. This included a new feature that allows the ARC PECAS model to be calibrated to observed quantities of activities at the TAZ level (instead of only at the super district level). This floorspace calibration algorithm was updated, tested and applied for the ARC PECAS model.

- The ARC PECAS model's communication and interaction with the ABM model, reading the new skims and generating new outputs required by the transport model through time, was automated[1] and tested.
- A TAZ level constraint on development in the SD module was added as a feature. This is available for both residential and non-residential space. This is used to add local knowledge to PECAS beyond what's available in the parcel database, and also mitigate the effect of inevitable remaining errors in detailed parcel data.
- A functionality to indicate programmed development as a user input was ported from the SANDAG PECAS model to the Atlanta context. This was tested and documented.

All of these ARC PECAS model improvements are reported in this technical document.

A feature originally planned was to constrain the model to match regional level totals by intermediate spatial units larger than a super district, e.g. counties. But, this feature became less important in The Region's Plan Forecast as the planning emphasis shifted to TAZ level detail. Since TAZ level constraints can also be used to constrain to counties, resources were shifted to TAZ level constraints, consistent with the Agile Process Management strategy of adapting to changing requirements.

2. SD module calibration enhancements

SD was recalibrated with updated targets and more consistent Bayesian prior knowledge than the SD calibration performed during the model development. This is described in the technical report *The ARC PECAS Model Development: model components and calibration*. The recalibration also allowed the behavior of SD to be made consistent with the new space prices simulated by the Activity Allocation module of the ARC PECAS model.

The core of the Space Development (SD) model is based on rational profit maximization – developers are more likely to build where construction costs are low and rents are high. However, some parameters are left unspecified and must be estimated based on observations of real activity. Most of these parameters are market share constants: they add a fixed value to a particular alternative under certain conditions to align the total share of developers choosing that alternative with observations of market share from reality. These constants are designed to represent unknown influences on decisions, especially ones that are not of interest for policy analysis. The objective of SD calibration is to find values for these constants and other parameters that best represent the actual behavior of the land development system.

[1] The ABM requires some manual intervention in the preparation of its inputs, so a fully automated time-series run of PECAS and the ABM together is not yet possible as of December 2015.

SD calibration is a "Bayesian" method – it combines prior information about the likely parameter values with new information. This is important because development displays a high degree of variability due to its discrete nature, as entire office towers or neighborhoods are built or not built based on changing economic conditions. Also, some types of development are rare, so the data quality for those types is lacking. The prior information helps temper these variations, preventing the parameters from being influenced excessively by anomalous data and providing a fallback in place of absent data. The priors represent what we know from behavioral theory, national experience, calibration of other PECAS models, and expert knowledge.

SD calibration is an iterative approach that converges towards the most likely parameter values, which are those that reproduce the target data as well as possible without drifting too far from their prior values. As with floorspace calibration, "tolerance" values are specified that determine the balance between matching different targets and priors. SD itself is non-deterministic (and so outcomes vary from one run to the next due to random events in simulation), but during calibration the expected value of each outcome is calculated, providing consistency between iterations.

A detailed explanation regarding the calibration for the SD module including the setup and the output description is contained in a technical document called Bayesian Calibration Software for the SD Module.

Target data and priors

The most important change to the priors affected the way dispersion parameters were handled. Nested logit models are supposed to obey the rule that the dispersion parameter at one level of a nested logit model should be greater than the dispersion parameter at the level above. To encourage the calibration process to respect this constraint, the priors were assigned neutral values that obeyed this order: from top to bottom, these values were 0.1, 0.14, 0.18, 0.25, and 0.3. Then the dispersion parameters associated with a given space type were specified to be correlated with each other (correlation coefficient of 0.5), so that the calibration process would try to move them all together.

The other priors were still low-information, with neutral values and relatively large tolerances; some of the tolerances were larger than in the previous SD calibration because the previous tolerances were judged to be too small.

The original targets were modified for the 2015 calibration by comparing the PECAS synthetic floorspace quantities in 2005 and 2015, with the differences between these quantities implying a possible range of annual construction. This amount was taken to be the lower bound on the total amount of development; it was possible that more development occurred but some of the space was demolished or replaced, resulting in the same net increase. If the previous target was less than the 2005-2015 difference, then the difference was taken as the target; otherwise, the new target was set between the previous target and the 2005-2015 difference.

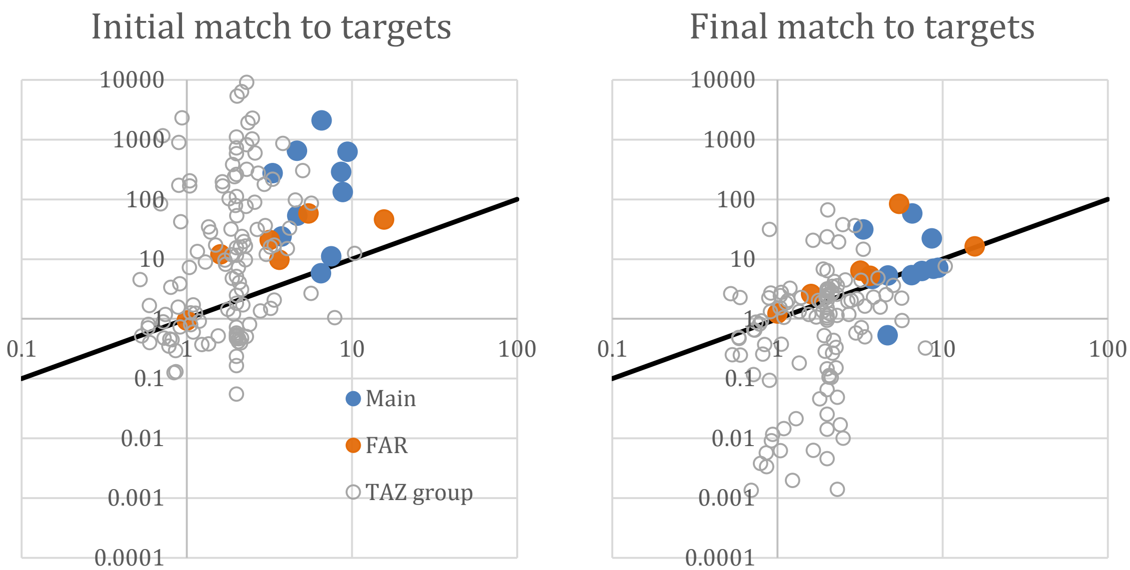
Specifically, the FAR (floor area ratio) targets and the targets for renovation and addition were carried over directly from the F.W. Dodge data. Meanwhile, the new construction targets were compared against the annual change implied by the difference between the 2005 and 2015 floorspace quantities. If the implied value was the larger of the two, then it was used as the new target; but if the implied value was the smaller of the two, then it was possible that some demolition occurred to offset the new construction. For this reason, a weighted average of the two values was used instead. The tolerance for each new construction target was generally half the difference between the implied and Dodge values, but with never less than a minimum value determined by the size of the target value.

Results

Figure 1 shows how the SD calibration process matched the targets. In these graphs, the value on the x-axis represents the target value divided by that target's tolerance; the y-axis represents the corresponding model outcome divided by the tolerance. If an outcome matched its target perfectly, then the target value would be the same as the model outcome, so that data point would lie on the black diagonal line.

For these graphs, the targets have been divided into three types that behave differently: the "main" targets (blue), FAR targets (orange), and TAZ group targets (outline). The main targets are for region-wide total amounts of space affected by some class of development events; for example, the total amount of new construction of detached residential space. FAR targets describe the density of new construction of each space type. The TAZ group targets are similar to the main targets, but at the county level rather than at the region level. Because there is less data available for a given county than for the whole region, the TAZ group targets were assigned larger tolerances, and so they did not match the targets as closely as the main targets and FAR targets.

Figure 1: Fit to the targets before and after SD calibration



The results from the Space Development calibration are contained in a file called **"SDCalibrationResults.xlsx"**. This file contains four spreadsheets relevant to SD calibration: *InitialSDParameters* and *FinalSDParameters* are the parameter values before and after the calibration along with their associated prior values. Similarly, *InitialSDTargetsOutcomes* and *FinalSDTargetsOutcomes* are the model outcomes before and after the calibration along with their associated target values.

3. Updating the growth rates for the households and industries model-wide from 2015 to 2050

Industries as part of the activities grow at a certain rate per year during the 45 years defined for the simulation in the ARC PECAS model.

In the previous version of this model for Atlanta these growth rates were different for every simulation year but they were the same for almost every industry. Basically, the growth rates for each year were defined for two groups: the accounts, which include government accounts and capital; and the rest of the industries. While, for the households the growth rates are different for each simulation year but also they varied for each household category.

As an example, the growth rates calculated for the activities of the ARC PECAS model are presented for the year 2017. In this example the growth rates are shown before and after the update in the current version of this model.

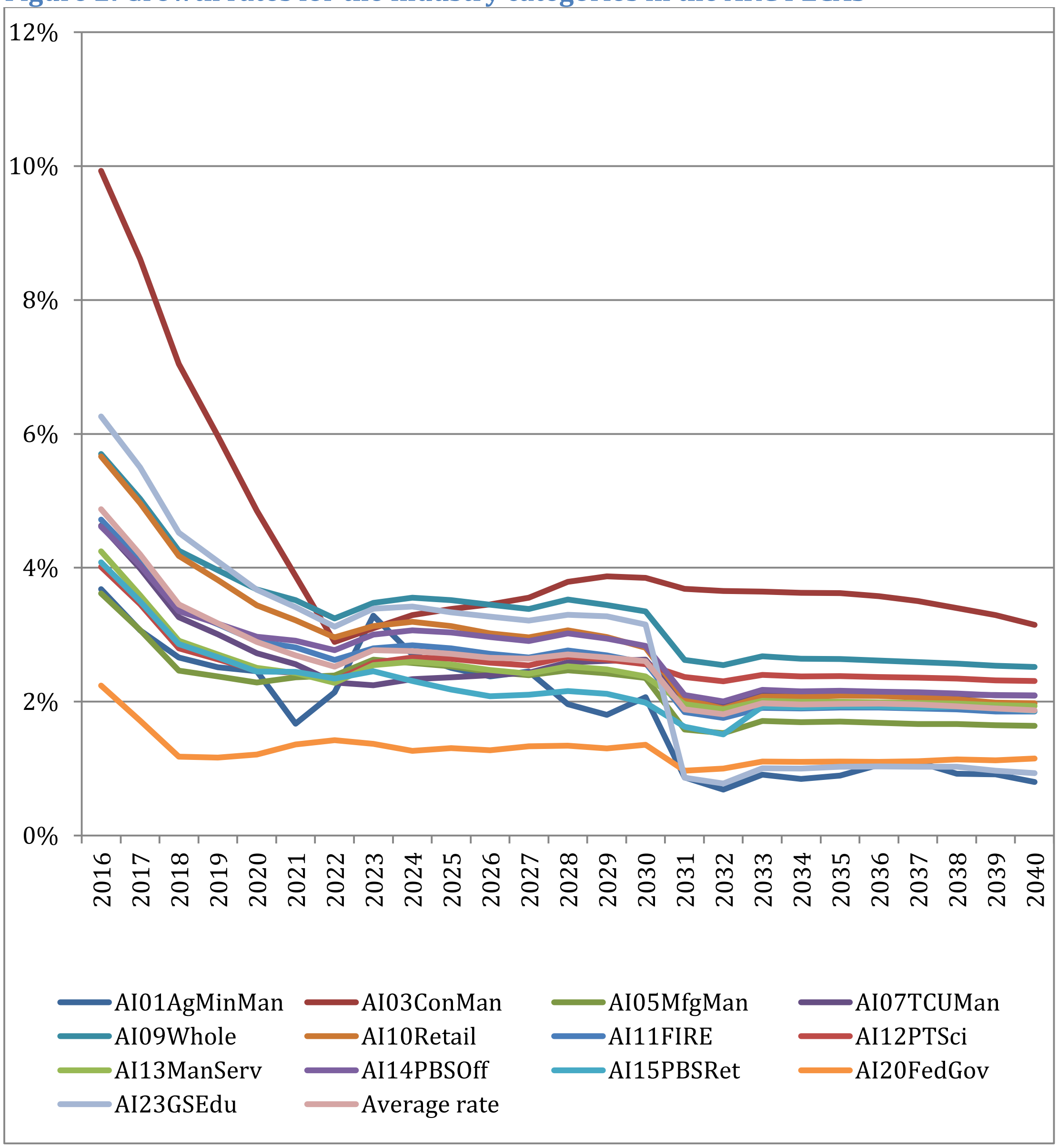
One of the improvements performed in the current version of the ARC PECAS model is that the growth rates for all the activities –households, industry and accounts - were updated using estimations from the ARC economic forecast by county based on NAICS data. These growth rates are different for each simulation year but they are also different among the industries. The growth rates were estimated for each year and industry and are shown in Figure 2.

Some growth rates are quite high, reflecting the current optimistic outlook for certain industries. The PECAS model was modified to enable larger differences in growth rates by calibrating and adjusting the treatment of imports and exports. A tighter linkage between the REMI model (used to produce the growth rates) and PECAS on the assumed local consumption of production, i.e. the "regional purchase coefficient", would be a useful feature to add to the model in the 2016 improvements, especially if forecasts continue to be relatively optimistic for certain industries.

Table 1. Growth rates by industry for the ARC PECAS model

Activities	Growth rates previous model	Growth rates current model
AI01AgMinMan	2.09%	3.06%
AI02AgMinProd	2.09%	3.06%
AI03ConMan	2.09%	8.61%
AI04ConProd	2.09%	8.61%
AI05MfgMan	2.09%	3.07%
AI06MfgProd	2.09%	3.07%
AI07TCUMan	2.09%	3.99%
AI08TCUProd	2.09%	3.99%
AI09Whole	2.09%	5.03%
AI10Retail	2.09%	4.97%
AI11FIRE	2.09%	4.10%
AI12PTSci	2.09%	3.46%
AI13ManServ	2.09%	3.59%
AI14PBSOff	2.09%	4.04%
AI15PBSRet	2.09%	3.50%
AI16PSInd	2.09%	3.50%
AI17Religion	2.09%	3.50%
AI18BSOnsite	2.09%	3.50%
AI19PSOnsite	2.09%	3.50%
AI20FedGov	2.09%	1.71%
AI21StLocGov	2.09%	1.71%
AI22Military	2.09%	1.71%
AI23GSEdu	2.09%	5.50%
AI24HiEdu	2.09%	5.50%
AI25Health	2.09%	3.50%
AA26FedGovAccounts	2.46%	4.20%
AA27StLocGovAccounts	2.46%	1.71%
AA28CapitalAccounts	2.78%	8.61%
AH29HHlt20_12	2.16%	1.63%
AH30HHlt20_3p	1.33%	0.76%
AH31HH2050_12	2.06%	1.49%
AH32HH2050_3p	1.74%	1.14%
AH33HH50100_12	2.11%	1.49%
AH34HH50100_3p	1.96%	1.37%
AH35HHge100_12	2.11%	1.48%
AH36HHge100_3p	1.68%	1.11%

Figure 2: Growth rates for the industry categories in the ARC PECAS



Note that the plot does not include all the industry categories defined for the ARC PECAS model, the ones missing were assumed from other industries. The growth rates applied for each industry in the model is shown in Table 2.

An average rate was estimated and used for the Federal Government Account, while the estimated rate from the Federal Government was applied to the Local Government; the estimated rate for the management of construction was applied to the capital account; and the estimated rate for retail was used for several industries as indicated in light blue. In cases which the industry is split in the production and the management, the growth rate estimated for each industry is applied to both categories (Table 2).

Table 2. Mapping of growth rates by industry to the ARC PECAS model

PECAS activity	Category of the rate applied
AA26FedGovAccounts	Average rate
AA27StLocGovAccounts	AI20FedGov
AA28CapitalAccounts	AI03ConMan
AH29HHlt20_12	AH29HHlt20_12
AH30HHlt20_3p	AH30HHlt20_3p
AH31HH2050_12	AH31HH2050_12
AH32HH2050_3p	AH32HH2050_3p
AH33HH50100_12	AH33HH50100_12
AH34HH50100_3p	AH34HH50100_3p
AH35HHge100_12	AH35HHge100_12
AH36HHge100_3p	AH36HHge100_3p
AI01AgMinMan	AI01AgMinMan
AI02AgMinProd	AI01AgMinMan
AI03ConMan	AI03ConMan
AI04ConProd	AI03ConMan
AI05MfgMan	AI05MfgMan
AI06MfgProd	AI05MfgMan
AI07TCUMan	AI07TCUMan
AI08TCUProd	AI07TCUMan
AI09Whole	AI09Whole
AI10Retail	AI10Retail
AI11FIRE	AI11FIRE
AI12PTSci	AI12PTSci
AI13ManServ	AI13ManServ
AI14PBSOff	AI14PBSOff
AI15PBSRet	AI15PBSRet
AI16PSInd	AI15PBSRet
AI17Religion	AI15PBSRet
AI18BSOnsite	AI15PBSRet
AI19PSOnsite	AI15PBSRet
AI20FedGov	AI20FedGov
AI21StLocGov	AI20FedGov
AI22Military	AI20FedGov
AI23GSEdu	AI23GSEdu
AI24HiEdu	AI23GSEdu
AI25Health	AI15PBSRet

4. Activity Allocation: developing Activity Constraints from 2005 to 2015

Since the beginning of the development of the Atlanta PECAS model, the base model year has been 2005. This means that current multi-year runs simulate ten years of history before the forecast begins. Activity constraints for these ten years were developed so that the model's behavior would conform to observed activity distributions, forming a solid basis for simulating future years.

When the original ARC PECAS activity constraints were created for the base year 2005, a spreadsheet was developed to take employee counts by SIC code and allocate them by PECAS activity. This was then adapted to the NAICS system, the North American standard used for more recent data. For some NAICS codes, there is a single clear PECAS activity that corresponded to it; for example, NAICS 42 stands for Wholesale Trade, which directly corresponds to the PECAS activity AI09Whole. In most cases, however, there were several possible activities; such categories were allocated at the regional level in proportion to the availability of workers with the appropriate occupation. Activity constraints Appendix 1 is adapted from a portion of this spreadsheet, and shows which PECAS activities were associated with each NAICS code, as well as the number of workers allocated to that activity.

Based on these allocations, a process was developed in 2013 for converting census data (employment and population) into activity constraints by super-district. This process was used to produce activity constraints for the model years 2010 and 2015, with constraints for 2005 transferred directly from the original model. The constraints for 2006-2009 were created by interpolating linearly between the 2005 and 2010 constraints, while those for 2011-2014 were similarly interpolated between 2010 and 2015. The conversion process is detailed below, with examples from the 2010 conversion.

- The process allocates employees to PECAS activities, at the region-wide level, in proportion to the values from Appendix 1. For example, the 86,808 employees in the construction industry (NAICS 23) are distributed as 18,197 employees in construction management (PECAS AI03ConMan) and 68,611 employees in construction production (PECAS AI04ConProd), since the proportion 18,197:68,611 is as close as possible to the 25,673:96,798 ratio found in Appendix 1.
- It derives a "productivity" value for each PECAS activity by dividing the total output from that activity (from ActivityTotalsI) by the number of employees allocated to that activity in Step 1. The productivity value represents the amount of activity output in dollars generated by each employee working on that activity. It acts as a conversion factor between units of employees and units of dollars. For example, the 68,611 employees allocated to construction production generated $28,364 million of output according to ActivityTotalsI, so the productivity value is $413,413 per employee.
- It divides up the employees in each NAICS category in each LUZ between the associated PECAS activities. For the categories, such as NAICS 42, that have only one associated PECAS activity, all of the employees in that category are allocated to that activity. For the categories with multiple PECAS activities, the allocation is done by solving equations that force the resulting activity amounts to be proportional to ActivityLocations. These equations are described in detail in Appendix 2.
- Even with allocations at the LUZ-level based on proportionality with ActivityTotalsI, the total region-wide employment in each activity generally does not equal the value in ActivityTotalsI. If these allocations were used directly to create ActivityConstraintsI, the mismatch would cause AA to fail. To avoid this, the spreadsheet introduces a set of correction factors that modify the LUZ-level proportions. These correction factors are calculated using Excel's Solver module so as to satisfy the constraints imposed by ActivityTotalsI. The application of the correction factors is described in detail in Activity constraints Appendix 2.

Spreadsheets containing the conversion calculations are attached, under the names "EmploymentAllocation2010.xlsm" and "EmploymentAllocation2015.xlsm". The process involves solving a system of equations using singular value decomposition; as such, the spreadsheets require the Perceptrics LLC linear algebra add-in for Excel (or an equivalent add-in capable of singular value decomposition).

After this conversion, some recalibration was done in 2015. For consistency with the new calibration, the 2015 activity constraints were replaced with TAZ-based constraints. The new constraints were built alongside the SD space limits (see Section 7), which produced activity amounts by TAZ as an intermediate step.

5. Floorspace calibration update to 2015 employment data

Approach

After running the model to simulate years 2005 to 2015, the 2015 floorspace inventory was calibrated to be consistent with the 2015 Census employment and population numbers.

As described in the previous section, the model was constrained to the observed households and employment amounts by LUZ when running from year 2005 to 2015. At model year 2015, these constraints were removed, and the floorspace amounts were calibrated to match population and employment amounts by TAZ as closely as possible without disturbing the overall space prices. Since 2014 was already constrained (and hence accurate) at the LUZ (a.k.a. Super District) level, this adjustment primarily shifted floorspace between TAZs within each LUZ. Thus employment and household information more closely matched observed 2015 TAZ level detail as a result of this calibration process.

The calibration process aimed to match three different types of targets, in order of greatest to least priority:

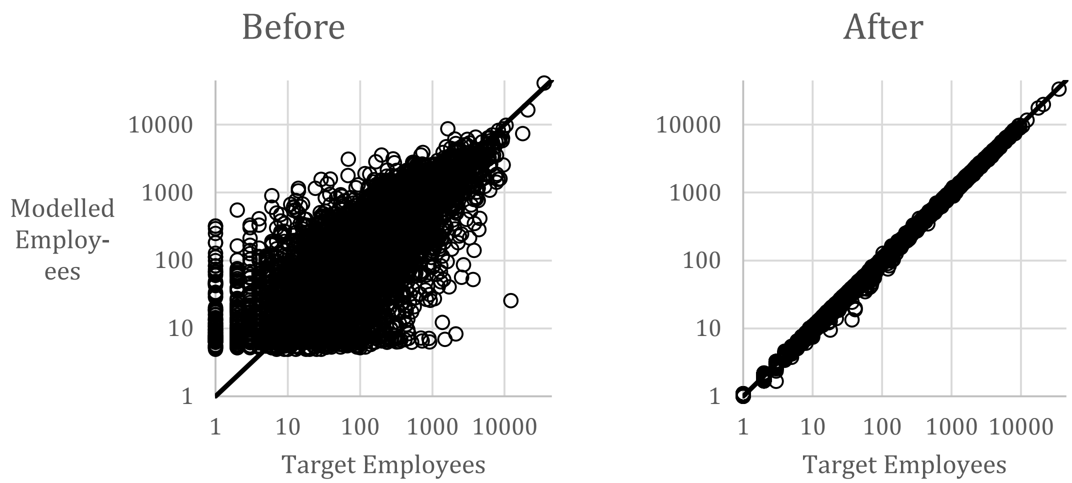
1. Match the observed region-wide **average space price** for each space type. This ensured that zone-by-zone adjustments to match the population and employment targets would not create an overall shortage or surplus of space. (No TAZ-level or LUZ level 2015 space prices targets were available as targets; the model endogenously generated the detailed 2015 prices based its other inputs and the historical simulation.).
2. Match the **total employees** by zone, across all activity types. This was done using the ABM output file as described in the section 6 of this document.
3. Match the **activity amounts** by zone and activity type. It was not possible to match these exactly, since the model has more activity types than the types of space available to adjust. Instead, the process matched the activity amounts as closely as possible without compromising the average space price or total employees by zone.

More detail on this process can be found in the Appendix 1 of this document.

Results

As shown in Figure 3, the employment totals by zone closely matched the targets after calibration. Each circle represents the modelled and target number of employees for one zone, with circles on the diagonal black line indicating a perfect match to the target.

Figure 3: Improvement in the match to zonal employment targets



The fit to the targets for specific industries was much less precise (as would be expected given its lower priority), but it did improve for most activities through calibration. Table 3 shows the total error for each activity across all zones in the model region, as well as the improvement in error (which is negative if the error was worse after calibration). The total error is the sum of the squares of the differences between the modelled and target activity amounts in each zone.

Table 3: Total error by activity before and after calibration

Activity	Total initial error	Total final error	Improvement
Households <20K 1-2 person	1.43×10^7	1.28×10^7	1.51×10^6
Households <20K 3+ person	2.86×10^6	2.43×10^6	4.30×10^5
Households 20-50K 1-2 person	2.24×10^7	1.55×10^7	6.92×10^6
Households 20-50K 3+ person	1.28×10^7	4.29×10^6	8.50×10^6
Households 50-100K 1-2 person	1.62×10^7	5.06×10^6	1.11×10^7
Households 50-100K 3+ person	3.11×10^7	1.63×10^7	1.48×10^7
Households >100K 1-2 person	9.10×10^6	4.91×10^6	4.20×10^6
Households >100K 3+ person	1.19×10^7	1.30×10^7	-1.08×10^6
Agriculture/Mining Production	3.67×10^4	2.98×10^4	6.84×10^3
Agriculture/Mining Management	6.39×10^4	6.98×10^4	-5.87×10^3
Construction Management	1.22×10^5	1.32×10^5	-1.03×10^4
Manufacturing Production	9.87×10^6	1.04×10^7	-5.01×10^5
Manufacturing Management	1.03×10^8	5.28×10^7	5.04×10^7
Transport/Utility Production	1.18×10^8	1.03×10^8	1.47×10^7
Transport/Utility Management	2.20×10^8	3.11×10^8	-9.09×10^7

Activity	Total initial error	Total final error	Improvement
Wholesale	3.23×10^{8}	7.59×10^{7}	2.47×10^{8}
Retail	1.55×10^{8}	7.62×10^{7}	7.86×10^{7}
Financial/Real Estate	6.98×10^{7}	5.64×10^{7}	1.34×10^{7}
Professional/Technical	7.72×10^{7}	6.81×10^{7}	9.02×10^{6}
Management Services	2.84×10^{7}	2.92×10^{7}	-7.36×10^{5}
Personal/Business Services Office	1.54×10^{8}	6.86×10^{7}	8.55×10^{7}
Personal/Business Services Retail	1.64×10^{8}	8.72×10^{7}	7.68×10^{7}
Personal Services Industrial	9.88	3.53	6.35
Religion	2.43×10^{3}	2.71×10^{3}	-278
Federal Government	1.99×10^{5}	1.75×10^{5}	2.37×10^{4}
State/Local Government	1.05×10^{8}	9.33×10^{7}	1.17×10^{7}
Military	1.85×10^{3}	335	1.52×10^{3}
Grade School	2.44×10^{8}	2.56×10^{7}	2.18×10^{8}
Higher Education	3.40×10^{6}	4.56×10^{5}	2.95×10^{6}
Health Services	1.36×10^{8}	1.32×10^{8}	4.10×10^{6}

A plot of the distribution of activity amounts across zones for selected activities is shown in Figures 3 to 5. Again, circles near the diagonal black line are closer to matching the target. While the data points for all three of these activities were still widely spread out, they were more concentrated after the calibration than before.

Figure 4: Improvement in the match to the zonal activity targets – manufacturing production

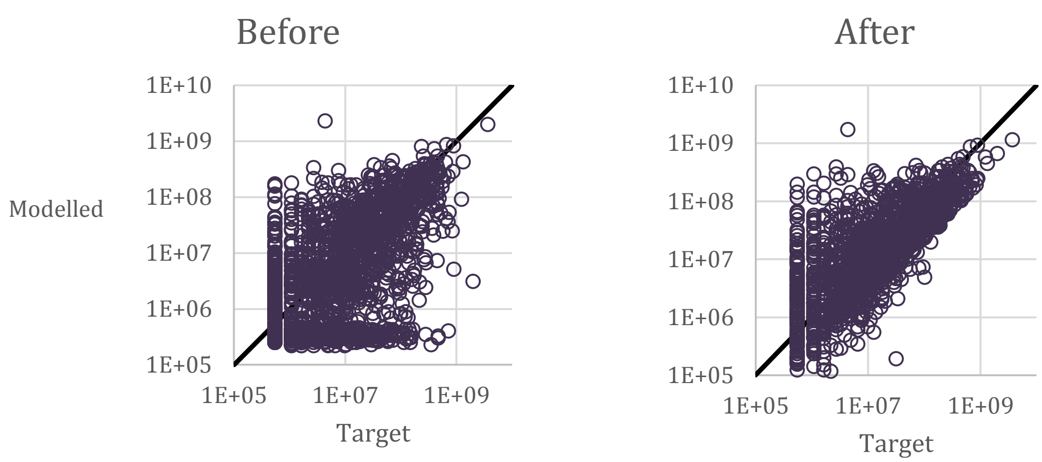


Figure 5: Improvement in the match to the zonal activity targets – professional/technical

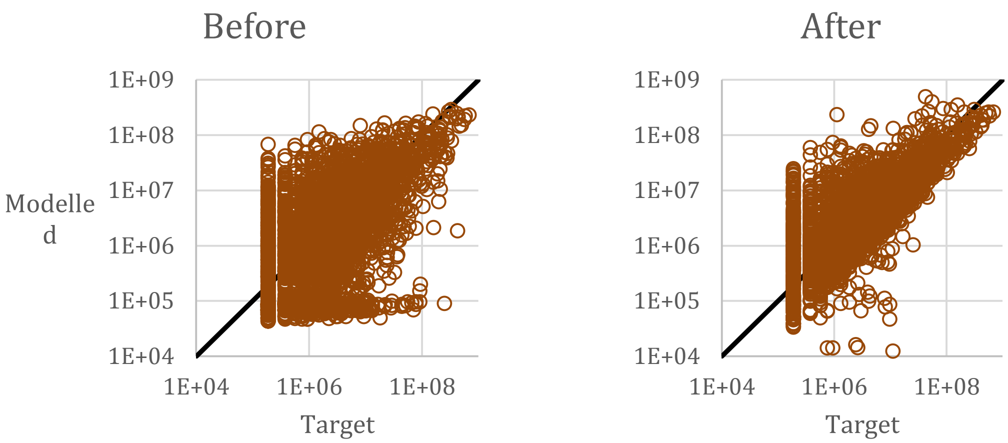
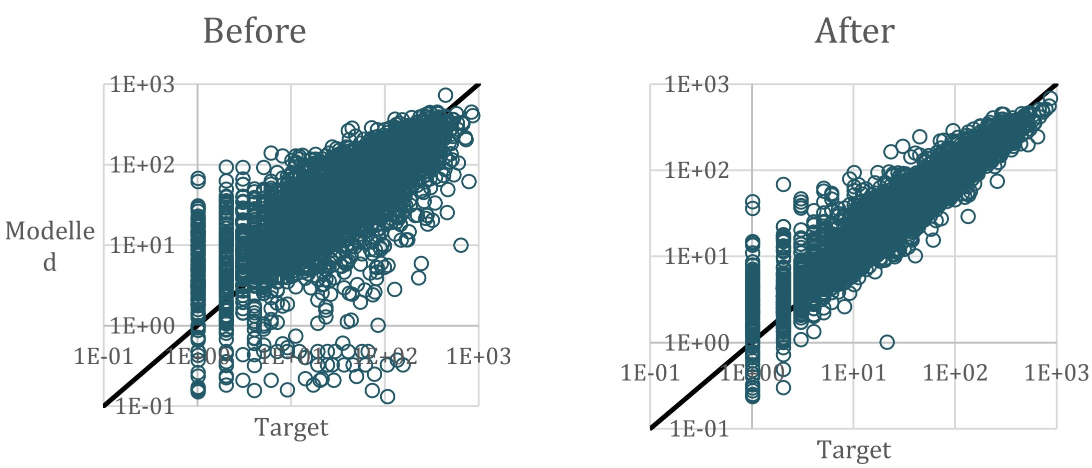


Figure 6: Improvement in the match to the zonal activity targets – 1-2 person households $50-100K



6. Functionality 1: PECAS and ABM integration (JEA)

The primary task undertaken in 2015 for PECAS and ABM integration involved automating the production of ABM inputs from the PECAS outputs.

ABM Totals

The total amount of activity (employment, population, etc.) region-wide to the ABM is a model input to the system, filename **ABModelTotalsI.csv**. The first few lines of this file are shown in the Table 4.

Table 4: Amount of activity by activity type and year run

year_run	abm_code	abm_total_amount
2005	I1H1	132522
2006	I1H1	135923.4
2007	I1H1	139324.8
2008	I1H1	142726.2
2009	I1H1	146127.6

Information is entered for all years in this file because PECAS prepares ABM inputs for each year and the MapIt visualization tool in PECAS allows visualizing of PECAS outputs for any intermediate year. (Intermediate years in which the ABM does not run can be interpolated by the analyst based on ABM run years. A view in the database called *inputs.abm_interpolate_extrapolate* was written to obviate the need to manually interpolate. This view slows down the integration process, so it is recommended to enter travel model total amounts for every year).

Correspondence between ABM categories and PECAS categories

The integration procedure allocates these total amounts, which are specified in this file as an input by the user, amongst the TAZs, based on PECAS outputs.

The system for relating ABM inputs to PECAS outputs is specified in the database table input.pecas_to_abm_categories shown in Table 5. Note that the mapping from NAICS categories to the SIC categories used to build PECAS is implicit in the right side showing the industrial matchings. The weights are based on the number of employees in the combination from the PUMS sample, however some entries were set to zero for small observed combinations and for "OnSite" employee categories who will be allocated in PECAS to the place they deliver services rather than their reported place-of-work.

Table 5: Correspondence table relating ABM inputs to PECAS outputs

abm_code	pecas_activity	weight
I1H1	AH29HHlt20_12	1
I1H2	AH29HHlt20_12	1
I1H3	AH30HHlt20_3p	1
I1H4	AH30HHlt20_3p	1
I1H5	AH30HHlt20_3p	1
I1H6	AH30HHlt20_3p	1
I2H1	AH31HH2050_12	1
I2H2	AH31HH2050_12	1
I2H3	AH32HH2050_3p	1
I2H4	AH32HH2050_3p	1
I2H5	AH32HH2050_3p	1
I2H6	AH32HH2050_3p	1
I3H1	AH33HH50100_12	1

abm_code	pecas_activity	weight
N11	AI02AgMinProd	253
N11	AI01AgMinMan	911
N21	AI02AgMinProd	253
N21	AI01AgMinMan	911
N22	AI08TCUProd	1419
N22	AI07TCUMan	4431
N23	AI03ConMan	2755
N23	AI04ConProd	7692
N313233	AI06MfgProd	3685
N313233	AI05MfgMan	8772
N42	AI09Whole	2650
N4445	AI10Retail	53240
N4849	AI08TCUProd	1419

abm_code	pecas_activity	weight
I3H2	AH33HH50100_12	1
I3H3	AH34HH50100_3p	1
I3H4	AH34HH50100_3p	1
I3H5	AH34HH50100_3p	1
I3H6	AH34HH50100_3p	1
I4H1	AH35HHge100_12	1
I4H2	AH35HHge100_12	1
I4H3	AH36HHge100_3p	1
I4H4	AH36HHge100_3p	1
I4H5	AH36HHge100_3p	1
I4H6	AH36HHge100_3p	1

abm_code	pecas_activity	weight
N4849	AI07TCUMan	4431
N51	AI07TCUMan	0
N51	AI08TCUProd	0
N51	AI14PBSOff	8676
N52	AI11FIRE	4399
N53	AI11FIRE	4399
N53	AI15PBSRet	15764
N54	AI19PSOnsite	0
N54	AI14PBSOff	0
N54	AI15PBSRet	0
N54	AI16PSInd	0
N54	AI18BSOnsite	0
N54	AI12PTSci	8867
N55	AI13ManServ	59
N56	AI22Military	0
N56	AI13ManServ	59
N56	AI18BSOnsite	3775
N56	AI14PBSOff	8676
N56	AI15PBSRet	15764
N61	AI24HiEdu	5056
N61	AI23GSEdu	18604
N62	AI25Health	13817
N62	AI15PBSRet	15764
N71	AI16PSInd	0
N71	AI19PSOnsite	0
N71	AI15PBSRet	15764
N72	AI14PBSOff	8676
N72	AI10Retail	53240
N81	AI16PSInd	567
N81	AI19PSOnsite	3106
N81	AI17Religion	8470
N81	AI14PBSOff	8676
N81	AI15PBSRet	15764
N92	AI22Military	113
N92	AI20FedGov	449
N92	AI21StLocGov	10132

Allocation of ABM totals to TAZs

The database view output.abm_weight joins the above relationship table with the PECAS outputs that are stored in the AA output table output.all_activity_locations2, for the appropriate scenario and year combination, to determine the total weights to apply to each TAZ for each of the ABM categories. The view output.abm_se_taz10 divides the totals entered by the user for each of the ABM categories using these weights.

The model run script writes the table output.abm_se_taz10 to a text file called abm_land_use.csv, containing the amount of each ABM activity in each TAZ.

Labor Make and Use by Occupation

PECAS provides a useful input to the ABM that is not generally available in observed data. This input is the modelled relationship between industry categories and occupation categories. This relationship is used as proportional controls on occupations in the ABM population synthesizer. This information is produced automatically in PECAS in the ZonalMakeUse.csv output file, but the massive amounts of information in ZonalMakeUse.csv is selected and queried using a predefined database view called output.labor_make_and_use. This database view is queried for the scenario and year, and written out to a text file called **"lmau.csv"** by the PECAS run script.

7. Functionality 2: Development maximum capacity by TAZ

Previous scenario runs showed some behavior where space development of certain types would be concentrated into one zone, even though the total amount of space developed across the model region was realistic.

An investigation showed that this was normally due to one of two data-input problems:

- Permissive placeholder zoning for PUD. The "Planned Unit Development" (PUD) zoning generally represents, in reality, land that is not yet zoned for development but will be zoned for development in the future. This was coded into the modelling framework as extremely permissive zoning, since there is no current public information about what will be disallowed. This was corrected by comparing PUD zoned areas with zoning on similar physical locations nearby, to indicate the most likely zoning.
- Incorrectly overlapping parcels, usually for condominiums. In particular, in Fulton County many condominiums along the Peachtree corridor north of downtown are coded as overlapping parcels, each containing the full amount of land of the condominium building. As an example, a 250 unit condominium tower would be miscoded with 250 times the quantity of land. This extra fictional land in the most valuable corridor in the region was an incredible opportunity for the simulated developers in the model. The Fulton county parcel input file was processed extensively to remove most of these issues, although some are known to remain.

Although the above two problems were generally fixed, some issues remain. In modelling, a perfectly accurate parcel-level representation of actual land developability is an unachievable objective (there will always be at least one remaining error in the dataset of almost 2 million parcels.) It was desired to have a general system for preventing parcel coding errors from distorting the simulated development pattern of the entire region.

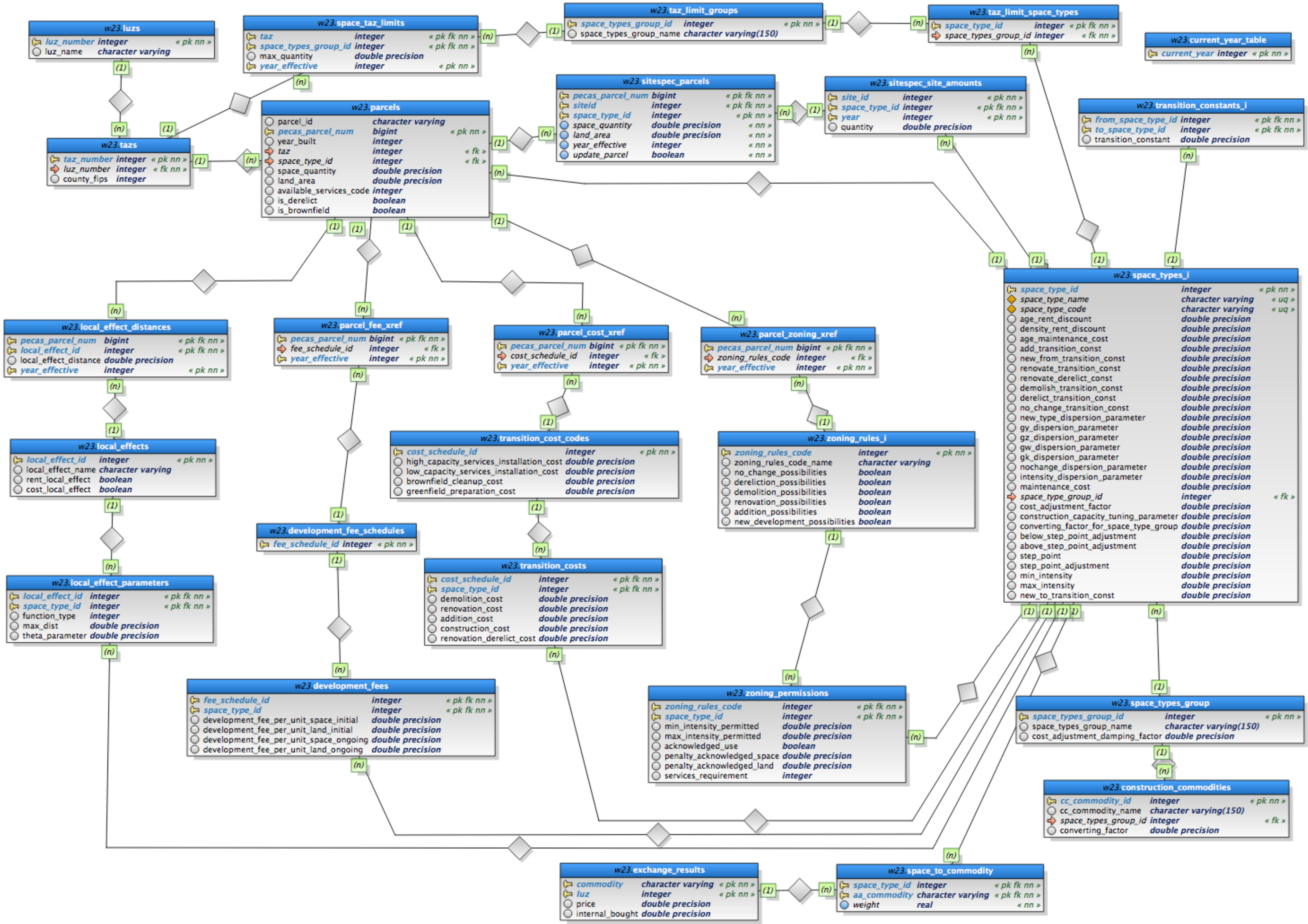
To this end, a new functionality was created in the ARC PECAS model such that a system of TAZ development limits was added to the Space Development (SD) module. The upgrade consisted of changes to the SD Java code (requiring a Pecas.jar version of 2.9 or later) as well as new tables in the structure of the SD database.

These modified SD datatbase input table system is shown in Figure 7.

The limits are monitored as SD runs. If, on a given parcel, it would otherwise be possible to build enough space to exceed the TAZ limit, then the maximum density is reduced to keep the amount of space below the limit.

This functionality requires as input a list of the maximum quantity of space, by type, TAZ to be developed. These quantities were estimated in the ARC PECAS model using, as a baseline (to be factored up as necessary), previously existing employment and household forecasts by TAZ. Two different methods were used to constrain the amount of space by TAZ, one for the residential and another for the non-residential. Both methods are documented in the **Appendix 2** of this report.

Figure 7: Structure of the relational database for the Space Development Module



8. Activating Functionality 3: Adjusting space development based on future planned development by parcel (Site-spec) in the Space Development module

The 'Site-spec' functionality is the term used in PECAS modeling to refer to specified developments on a specific site, planned to occur in a future year. These projects can be known planned developments under construction or with a high likelihood of implementation (e.g. an announced subdivision, hospital, or stadium), or can represent the development outcome of policy (e.g. future inner-city redevelopment), with the outcome provided as input to the model (instead of the more usual case with PECAS of the policy itself as input to the model). The functionality allows the analyst to indicate to the model software that the specified development should be simulated in a specific year, with a specific quantity of space in a group of parcels that make up a specific site in the region.

The 'Site-spec' functionality was already built as part of the PECAS model, but until now it had not been set up and tested for the ARC PECAS model. This functionality is included as part of the PECAS code, in use in different forms in Baltimore and San Diego. The adaptation of code was required because both Baltimore and San Diego use different database systems and have different integrations with their GIS system.

The 'Site-spec' functionality was tested in the Atlanta Region, and 8 steps were defined for its application. These steps are

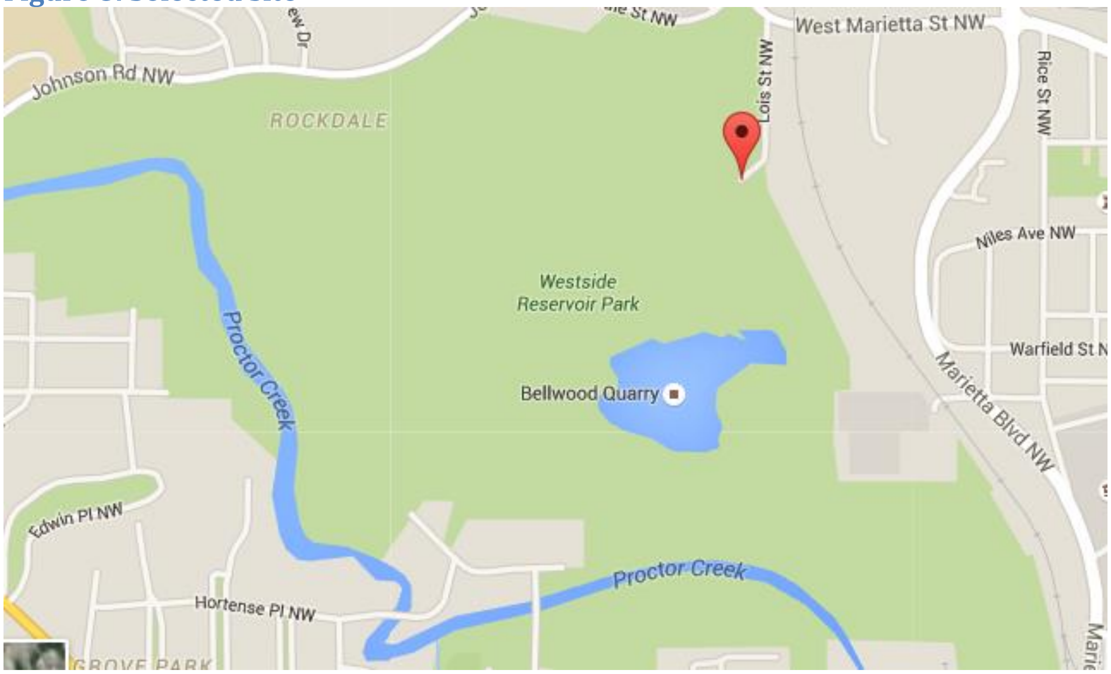
- Select a site
- Choose the parcels to be developed
- Export parcel list
- Define each site-spec
- Update the PECAS SD database
- Prepare the scenario for consistency
- Run the model
- Confirm the development projects were built

The 8 steps were followed in order to test the 'Site-spec' functionality in the ARC PECAS model using as an example a hypothetic development in the area of Bellwood Quarry, a larger vacant area that is not necessarily suitable for development (especially given the topological and water constraints of the quarry itself, see Figure 9) but nonetheless suitable for testing the functionality of specifying a large future development site.

Select a site spec:

A site was selected, as shown in Figure 8

Figure 8: Selected site



Choose the parcels to be developed:

This step includes the selection of several parcels. The one showed in yellow represents one parcel as an example in Figure 9.

Figure 9: Selecting parcels

Export a parcel list:

Export a parcel list as CSV to get list of parcels outside of the GIS system/database, as indicated in Table 6.

Table 6. The parcel list to develop in Bellwood Quarry

Pecas_parcel_num
'17 0190 LL0153'
'17 0190 LL0211'
'17 0190 LL0229'
'17 0190 LL0286'
'17 0225 LL0046'
'17 0225 LL0053'
'17 0225 LL0103'
'17 0225 LL0338'
'17 0225 LL0152'
'17 0225 LL0160'
'17 0225 LL0186'
'17 0225 LL0277'
'17 0225 LL0346'
'17 0225 LL0368'
'17 022600110512'
'17 0225 LL0145'
'17 0225 LL0178'
'14 0144 LL0017'
'14 0144 LL0025'
'14 0145 LL0024'
'14 0145 LL0032'
'14 0145 LL0040'
'14 0145 LL0081'
'14 0145 LL0149'
'14 0145 LL0156')

Define each site-spec:

Decide how much space is required to be developed by space type, and in which years. Enter these data into site_spec_amounts table (Table 7).

Table 7. The site list with the specified attributes

siteid	space_type_id	space_quantity	year_effective
1	72	2022	2,500,000
1	72	2025	2,500,000
1	72	2030	2,500,000
1	72	2035	2,500,000
1	72	2032	2,500,000
1	82	2030	725,000
1	82	2022	500,000

Update the SD database:

Update the PECAS SD database for the corresponding scenario in which is required to specify the development projects in the pre-defined site.

The sitespec_parcels table can be populated initially using the parcel list for the site, but then needs to be modified interactively (e.g. using the PGAdmin3 program) to select different space types (space_type_id attribute) to corresponding space type codes; and to change some of the parcels to be available in a different year (year_effective), if there is space in different years and to update the space_quantity column (Table 8). As well, the hold_until_year column should be populated, to turn off the regular SD simulation possibilities on the parcel until after the specified development on the site as been simulated (or for the entire model run, by leaving “hold_until_year” blank.).

Table 8. The updated parcel list to be developed by space type and year

pecas_parcel_num	siteid	space_type_id	space_quantity	land_area	year_effective	update_parcel	hold_until_year
1.31E+11	1	72	1,156,403	513,768	2032	t	2036
1.31E+11	1	72	1,103,832	490,411	2032	t	2036
1.31E+11	1	72	170,514	75,756	2032	t	2036
1.31E+11	1	72	69,251	30,767	2032	t	
1.31E+11	1	72	1,151,366	698,036	2025	t	2036
1.31E+11	1	72	718,685	435,715	2025	t	2036
1.31E+11	1	72	629,949	381,918	2025	t	2036
1.31E+11	1	72	424,420	432,062	2035	t	2036
1.31E+11	1	72	10,889	11,086	2035	t	2036
1.31E+11	1	72	537,425	547,102	2035	t	2036
1.31E+11	1	72	207,012	210,740	2035	t	2036
1.31E+11	1	72	5	5	2035	t	2036
1.31E+11	1	72	109,058	111,022	2035	t	2036
1.31E+11	1	72	1,211,191	1,233,000	2035	t	2036
1.31E+11	1	72	302,634	364,273	2030	t	2036
1.31E+11	1	72	354,097	426,219	2030	t	2036
1.31E+11	1	72	1,843,269	,218,700	2030	t	2036
1.31E+11	1	72	272,662	467,497	2022	t	2036
1.31E+11	1	72	1,577,983	,705,562	2022	t	2036
1.31E+11	1	72	541,562	928,546	2022	t	2036
1.31E+11	1	72	107,793	184,819	2022	t	2036
1.31E+11	1	82	500,000	453,440	2022	t	
1.31E+11	1	82	725,000	596,756	2030	t	2036

Prepare the scenario for consistency

Predefined queries are then executed to update the "space quantity" field of table 6 (under a constant-FAR assumption for the parcels in the site for the particular year), and update the zoning regulations based on "hold_until_year".

This step updates some of the tables and deletes some inconsistent zoning information. Note that this makes the scenario schema in SD generally unsuitable for cloning for a new scenario with fewer sites specified, as zoning information has been destroyed as irrelevant for parcels on the sites.

If an application has multiple zonings on parcels in different years before the year_effective (e.g. multifamily to 2030, but allow high-rise office in 2031 or later, and a hold_until_year is 2035), it is required to delete any zonings that are earlier than the latest (delete the multifamily zoning since the office zoning would apply in 2035 and later) zoning for the parcel that is before the hold_until_year. For details about how to do this in PostgreSQL see **Appendix 3**.

Run the model:

Run PECAS following the usual procedure.

Confirm that the development was built:

Select the parcels in the SD parcel database and confirm the development projects were built. This step only requires executing a simple query in PostgreSQL to confirm that the space was developed.

The scrips for the queries and updates performed in PostgreSQL for this example can be applied and adjusted when using the 'Site-spec' functionality in any other occasion. The procedure and associated scripts are contained in the **Appendix 3** of this document

Appendix 1: Improvements in the Constraint based Floorspace Calibration theory

Constraint-based floorspace calibration has three main objectives:

- Match the observed region-wide average space price for each space type.
- Match the total employees by zone across all activity types.
- Match the activity constraints by zone and activity type (representing specific industries in specific zones).

Since the first objective is region-wide while the second and third go zone by zone, each iteration proceeds in two stages. First, the floorspace quantities in each zone are adjusted to best match the second and third targets. Second, the quantities of a given space in all zones are adjusted together to recover a region-wide balance in prices.

The second objective is represented by the objective function

$$\Phi_{z2} = \sum_{j=1}^{n} \rho_j^2 \left(a_{jz} - t_{jz}\right)^2$$

where a_{jz} is the modelled amount of activity j in zone z, and t_{jz} is the corresponding target (activity constraint). The squared difference between the target and modelled activity amounts is weighted by the "productivity" factor ρ_j, which is the number of employees required to produce a dollar of activity j.

The third objective is represented by the objective function

$$\Phi_{z3} = k(E_z - T_z)^2$$

where E_z is the modelled number of employees in zone z and T_z is the corresponding target. The modelled employment is calculated using the productivity factors:

$$E_z = \sum_{j=1}^{n} \rho_j a_{jz}$$

In the PECAS Atlanta model, some activities are allowed to choose between two or more types of space to use; for example, financial/insurance services can use either retail or office space, while health services can use either office or institutional space. It is assumed that, when multiple types of space are available, the activity is divided between them in proportion to the amount of each type of space that activity uses (as reported in the MakeUse.csv file). Then the quantity of activity j that uses space of type i is

$$a_{ijz} = a_{jz} \left(\frac{u_{ijz}}{u_{jz}}\right)$$

where $u_{jz} = \sum_{k=1}^{m} u_{kjz}$ is the total amount of space used by activity j in zone z.

When the space quantities are adjusted, the quantity q_{iz} in zone z is replaced with a new quantity q'_{iz} for every space type i. It is again assumed that the change in activity quantity using that space is proportional to the change in space; that is

$$a'_{ijz} = a_{ijz}\left(\frac{q'_{iz}}{q_{iz}}\right) = a_{jz}\left(\frac{q'_{iz}}{q_{iz}}\right)\left(\frac{u_{ijz}}{u_{jz}}\right)$$

The total projected activity is

$$a'_{jz} = \sum_{r=1}^{m} a'_{rjz} = \frac{a_{jz}}{u_{jz}}\sum_{r=1}^{m}\left(u_{rjz}\left(\frac{q'_{rz}}{q_{rz}}\right)\right)$$

The objective function

$$\Phi_z = \Phi_{z2} + \Phi_{z3} = \sum_{j=1}^{n} \rho_j^2 \left(a'_{jz} - t_{jz}\right)^2 + k\left(\sum_{j=1}^{n} \rho_j a'_{jz} - T_z\right)^2$$

can now be differentiated with respect to q'_{iz}. First, the derivative of a'_{jz} is

$$\frac{\partial a'_{jz}}{\partial q'_{iz}} = \frac{a_{jz}}{u_{jz}}\sum_{r=i}\left(\frac{u_{rjz}}{q_{rz}}\right) = \left(\frac{a_{jz}}{q_{iz}}\right)\left(\frac{u_{ijz}}{u_{jz}}\right)$$

Note that the expansion of the derivative does not depend on q'_{iz}; only a'_{jz} itself does.

The derivative of the objective function is

$$\frac{\partial \Phi_z}{\partial q'_{iz}} = 2\sum_{j=1}^{n} \rho_j^2 \frac{\partial a'_{jz}}{\partial q'_{iz}}\left(a'_{jz} - t_{jz}\right) + 2k\left(\sum_{j=1}^{n} \rho_j \frac{\partial a'_{jz}}{\partial q'_{iz}}\right)\left(\sum_{j=1}^{n} \rho_j a'_{jz} - T_z\right)$$

Setting the gradient to zero gives

$$\sum_{j=1}^{n} \rho_j^2 \frac{\partial a'_{jz}}{\partial q'_{iz}}\left(a'_{jz} - t_{jz}\right) + k\left(\sum_{j=1}^{n} \rho_j \frac{\partial a'_{jz}}{\partial q'_{iz}}\right)\left(\sum_{j=1}^{n} \rho_j a'_{jz} - T_z\right) = 0$$

Since there is a value of q'_{iz} for every space type i, this is actually a system of m linear equations. The constant term for equation i is composed of all terms that do not contain a'_{jz}; that is:

$$\begin{aligned} K_i &= \sum_{j=1}^{n} \rho_j^2 \frac{\partial a'_{jz}}{\partial q'_{iz}} t_{jz} + kT_z \sum_{j=1}^{n} \rho_j \frac{\partial a'_{jz}}{\partial q'_{iz}} \\ &= \sum_{j=1}^{n} \rho_j \frac{\partial a'_{jz}}{\partial q'_{iz}}\left(\rho_j t_{jz} + kT_z\right) \\ &= \frac{1}{q_{iz}}\sum_{j=1}^{n} \rho_j a_{jz}\left(\frac{u_{ijz}}{u_{jz}}\right)\left(\rho_j t_{jz} + kT_z\right) \end{aligned}$$

In equation i, the rth term is a coefficient C_{ir} times q'_{rz}. The coefficient is

$$C_{ir} = \sum_{j=1}^{n} \rho_j^2 \left(\frac{\partial a'_{jz}}{\partial q'_{iz}}\right)\left(\frac{a_{jz}}{u_{jz}}\right)\left(\frac{u_{rjz}}{q_{rz}}\right) + k\left(\sum_{j=1}^{n} \rho_j \frac{\partial a'_{jz}}{\partial q'_{iz}}\right)\left(\sum_{j=1}^{n} \rho_j \left(\frac{a_{jz}}{u_{jz}}\right)\left(\frac{u_{rjz}}{q_{rz}}\right)\right)$$
$$= \frac{1}{q_{iz} q_{rz}}\left[\sum_{j=1}^{n} \rho_j^2 a_{jz}^2 \left(\frac{u_{ijz} u_{rjz}}{u_{jz}^2}\right) + k\left(\sum_{j=1}^{n} \rho_j a_{jz}\left(\frac{u_{ijz}}{u_{jz}}\right)\right)\left(\sum_{j=1}^{n} \rho_j a_{jz}\left(\frac{u_{rjz}}{u_{jz}}\right)\right)\right]$$

The coefficients C_{ir} and constants K_i form a matrix equation that can be solved for the optimal values of q'_{rz}. These are the new space quantities in zone z. This process is repeated as many times as desired to approach the true optimum.

It is likely that this zone-by-zone adjustment of space quantities will result in a net gain or loss of space across the whole region. This could cause the price of space to keep increasing or decreasing as the calibration process iterates. Therefore, adjustments to the regional space totals are made in parallel with the zone-by-zone adjustments. The intention is to have the calibration finish with the zonal targets matched as closely as possible without compromising the overall average price of each type of space.

The regional adjustment for each space type has its own objective function, which is similar to that used in "weighted floorspace calibration", but only considering the price. Floorspace targets, which are only necessary for zone-by-zone calibration to prevent extreme changes in small zones, are not used. The objective function for space type j is

$$\Phi_j = \left(\frac{\sum_z D_{jz} \Pi_{jz}}{\sum_z D_{jz}} - \overline{\Pi}_j\right)^2$$

where D_{jz} is the demand for space of type j in zone z, Π_{jz} is the price of type j space in zone z, and $\overline{\Pi}_j$ is the observed region-wide price of type j space (which serves as the target).

The calibration process will make a change to the total quantity of space by adding an amount ΔQ_j. This is done by scaling all zonal quantities up or down by the same factor:

$$Q'_{jz} = Q_{jz}\left(1 + \frac{\Delta Q_j}{\sum_w Q_{jw}}\right)$$

It is assumed that demand will respond proportionately to changes in quantity, so

$$D'_{jz} = \frac{D_{jz} Q'_{jz}}{Q_{jz}} = D_{jz}\left(1 + \frac{\Delta Q_j}{\sum_w Q_{jw}}\right)$$

The new price at the adjusted floorspace quantity is predicted using the linear approximation

$$\Pi'_{jz} = \Pi_{jz} + \frac{\partial \Pi_{jz}}{\partial Q_{jz}} \Delta Q_{jz} = \Pi_{jz} + \frac{\partial \Pi_{jz}}{\partial Q_{jz}} \frac{Q_{jz}}{\sum_w Q_{jw}} \Delta Q_j = \Pi_{jz} - \frac{Q_{jz}}{\sum_w Q_{jw}} \frac{\Delta Q_j}{V_{jz}}$$

where V_{jz} is the derivative reported in ExchangeResults.

To find the optimum quantity adjustment, set the predicted price equal to the target price and solve.

$$\frac{\sum_z D'_{jz}\Pi'_{jz}}{\sum_z D'_{jz}} = \overline{\Pi}_j$$

$$\sum_z D'_{jz}\Pi'_{jz} = \overline{\Pi}_j \sum_z D'_{jz}$$

$$\sum_z \left(D_{jz} + \frac{D_{jz}\Delta Q_j}{\sum_w Q_{jw}}\right)\left(\Pi_{jz} - \frac{Q_{jz}}{\sum_w Q_{jw}}\frac{\Delta Q_j}{V_{jz}}\right) = \overline{\Pi}_j \sum_z \left(D_{jz} + \frac{D_{jz}\Delta Q_j}{\sum_w Q_{jw}}\right)$$

$$\sum_z D_{jz}\Pi_{jz} + \left(\frac{D_{jz}}{\sum_w Q_{jw}}\right)\left(\Pi_{jz} - \frac{Q_{jz}}{V_{jz}}\right)\Delta Q_j - \frac{D_{jz}Q_{jz}}{V_{jz}\left(\sum_w Q_{jw}\right)^2}\Delta Q_j^2 = \overline{\Pi}_j \sum_z \left(D_{jz} + \frac{D_{jz}\Delta Q_j}{\sum_w Q_{jw}}\right)$$

$$\sum_z D_{jz}\Pi_{jz} + \frac{\Delta Q_j}{\sum_w Q_{jw}}\left(\sum_z D_{jz}\Pi_{jz} - \sum_z D_{jz}\frac{Q_{jz}}{V_{jz}}\right) - \left(\frac{\Delta Q_j}{\sum_w Q_{jw}}\right)^2 \sum_z \frac{D_{jz}Q_{jz}}{V_{jz}}$$
$$= \overline{\Pi}_j \sum_z D_{jz} + \frac{\overline{\Pi}_j \Delta Q_j}{\sum_w Q_{jw}} \sum_z D_{jz}$$

Simplify this by defining the following variables:

$$\Pi_j = \sum_z D_{jz}\Pi_{jz}$$

$$D_j = \sum_z D_{jz}$$

$$V_j = \sum_z D_{jz}\frac{Q_{jz}}{V_{jz}}$$

$$\Delta Q_{jr} = \frac{\Delta Q_j}{\sum_w Q_{jw}}$$

Then the equation becomes

$$\Pi_j + \left(\Pi_j - V_j\right)\Delta Q_{jr} - V_j \Delta Q_{jr}^2 = \overline{\Pi}_j D_j + \overline{\Pi}_j \mathrm{D_j} \Delta Q_{jr}$$

$$V_j \Delta Q_{jr}^2 + \left(\overline{\Pi}_j \mathrm{D_j} + V_j - \Pi_j\right)\Delta Q_{jr} - \Pi_j + \overline{\Pi}_j D_j = 0$$

$$\left(\Delta Q_{jr} + 1\right)\left(V_j \Delta Q_{jr} + \overline{\Pi}_j D_j - \Pi_j\right) = 0$$

The solutions to this equation are

$$\Delta Q_{jr} = -1 \text{ or } \Delta Q_{jr} = \frac{\Pi_j - \overline{\Pi}_j D_j}{V_j}$$

The solution of -1 is an extraneous root, since it implies that $\sum_z D'_{jz} = 0$, and the original equation includes $\sum_z D'_{jz}$ in the denominator. The other root is the only true solution, and is used to scale each space type.

Appendix 2: Calculating the constraints for the maximum development capacity by TAZ for the ARC PECAS model

Introduction

This document describes the calculation of forecast PECAS space quantity limits using the employment and household data from the ABM (Activity-Based Model) for each TAZ. The quantity limits are used as constraints when running the PECAS model to restrict development to reasonable levels. Two separate methods were used to calculate residential and nonresidential quantity limits. The data, calculation steps and Python scripts which were used are described here.

The maximum development constraints were calculated based on the previously existing 2040 forecast of employment and households by TAZ. These were scaled up by 25% for use up until 2040 (so that the PECAS model could, in the most general sense, forecast any zone to be up to 25% higher, as measured in space consumption, than what was implied by the previous existing forecast), and scaled up by a further 15% for the period 2040 to 2050.

Calculation of nonresidential space quantities

To calculate the nonresidential space quantities in each TAZ, ABM employment data were used. There are 20 employee categories in ABM data which are available at the TAZ level. However, there is no source which has the PECAS space use rates with respect to the ABM employee categories. Therefore, existing ABM employee categories were converted to the PECAS activity amounts as the first step. Then, space use rates per dollar of PECAS activities were used to calculate the space quantities in each TAZ.

The number of employees in each employee category is available at the TAZ level from the ABM forecasts. Therefore, the dollar value of PECAS activity per employee was used to calculate the quantities of PECAS activities in each TAZ. The process of calculating the dollar amount of activity per employee is described below.

Table 9 how many employees are working in each PECAS activities according to the 2040 forecast. From the ActivityTotalsI file, total amounts of activity for a given year were obtained. Eq. 1 was used to calculate the dollar amount of PECAS activity per employee using the data obtained from Table 9 and ActivityTotalsI file.

$$DE_A = \frac{AT_A}{\sum_{i=1}^{n} ABM_{i,A}} \qquad (1)$$

where:

DE_A = dollar amount of PECAS activity A per employee,

$ABM_{i,A}$ = number of type i employees in the PECAS activity A,

ATA = PECAS activity total.

It can be seen from Table 9 that in some situations one employee category is shared between two or more PECAS activities. Therefore, it is required to find out the proportions of employees work in each PECAS activities in a given employee category.

Eq. 2 was used to calculate the PECAS activity amounts by using the ABM employee data:

$$PA_{A,i} = DE_A \times E_i \times P_{i,A} \tag{2}$$

where:

$PA_{A,i}$ = amount of activity A by employee category i,

DE_A = dollar amount of activity A per employee,

E_i = number of employees in category i,

$P_{i,A}$ = proportion of employee i in activity A

According to the equation, it can be seen that proportions of employees belongs to each PECAS activities are required to calculate the PECAS activity quantities. Moreover the proportions which are used also should minimize the difference between the calculated activity totals using the ABM data and PECAS activity totals. Eq. 3 shows the minimization function which was used to estimate the optimal parameters of proportions of employee categories which will minimize the difference between the calculated activity totals using the ABM data and PECAS activity totals:

$$\text{minimize } AT_{dif} = \sum^{A} \left| T_A - \sum_{i=1}^{n} PA_{A,i} \right| \tag{3}$$

subject to

$$\sum_{j=1}^{a} P_{i,j} = 1 \tag{4}$$

AT_{dif} = sum of differences of calculated activity totals using ABM data and PECAS activity totals,

A = PECAS activity,

$PA_{A,i}$ = amount of activity A by employee category i,

T_A = total amount of activity A from PECAS,

n = number of employee categories in a PECAS activity category,

i = employee category,

j = number of PECAS categories uses the employee category i.

Using Excel's Solver add-in, optimal values for the proportions were found; these values are shown in the Table 10. Then those proportions and dollar value per employee were used to calculate the PECAS activity amounts in TAZ level using Eq. 2. The calculated activity amounts in TAZ level then multiplied by the use rates in the ZonalMakeUse database to calculate the space quantities in TAZ level.

Calculation of residential space quantities

To calculate the residential space quantities in each TAZ, ABM household data were used. There are 24 and 8 household categories in ABM data and PECAS respectively. Therefore, more than 1 ABM household category belongs to a one PECAS household category. To match the household totals in the PECAS, it is required to multiply each ABM household categories by some factors. To calculate those factors Eq. 5 was used.

$$F_{i,H} = \frac{TP_H}{\sum_{i=1}^{n} H_i} \qquad (5)$$

$F_{i,P,H}$ = factor for ABM household category i belonging to PECAS category H,

TP_H = total PECAS households in category H,

H_i = number of ABM households in category i belonging to PECAS category H

n = number of ABM households in categories belonging to PECAS category H

Using the calculated factors for each ABM household, the number of PECAS households in each TAZ was determined. Then the residential space quantities in each TAZ were found by multiplying the rates in ZonalMakeUse by the number of PECAS households.

Table 9: Number of employees in each PECAS category

	AI01 AgMinMan	AI02 AgMinProd	AI03 ConMan	AI04 ConProd	AI05 MfgMan	AI06 MfgProd	AI07 TCUMan	AI08 TCUProd	AI09 Whole	AI10 Retail	AI11 FIRE	AI12 PTSci	AI13 ManServ	AI14 PBSOff	AI15 PBSRet	AI16 PSInd	AI17 Religion	AI18 BSOnsite	AI19 PSOnsite	AI20 FedGov	AI21 StLocGov	AI22 Military	AI23 GSEdu	AI24 HiEdu	AI25 Health	Total
N11	685	947	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1632
N21	1142	1580	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2722
N22	0	0	0	0	0	0	3870	6422	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10292
N23	0	0	15720	223471	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	239191
N313233	0	0	0	0	41705	107723	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	149428
N42	0	0	0	0	0	0	0	0	178760	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	178760
N4445	0	0	0	0	0	0	0	0	0	233823	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	233823
N4849	0	0	0	0	0	0	63835	105917	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	169752
N51	0	0	0	0	0	0	0	0	0	0	0	0	0	73721	0	0	0	0	0	0	0	0	0	0	0	73721
N52	0	0	0	0	0	0	0	0	0	0	206983	0	0	0	0	0	0	0	0	0	0	0	0	0	0	206983
N53	0	0	0	0	0	0	0	0	0	0	86553	0	0	0	65652	0	0	0	0	0	0	0	0	0	0	152205
N54	0	0	0	0	0	0	0	0	0	0	0	342466	0	0	0	0	0	0	0	0	0	0	0	0	0	342466
N55	0	0	0	0	0	0	0	0	0	0	0	0	48786	0	0	0	0	0	0	0	0	0	0	0	0	48786
N56	0	0	0	0	0	0	0	0	0	0	0	0	434	210223	131322	0	0	18707	0	0	0	0	0	0	0	360686
N61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	273657	31976	0	305633
N62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	228473	0	0	0	0	0	0	0	0	0	289339	517812
N71	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45530	0	0	0	0	0	0	0	0	0	0	45530
N72	0	0	0	0	0	0	0	0	0	203915	0	0	0	38005	0	0	0	0	0	0	0	0	0	0	0	241920
N81	0	0	0	0	0	0	0	0	0	0	0	0	0	79662	49763	145	1230	0	243	0	0	0	0	0	0	131043
N92	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9893	226264	580	0	0	0	236738
Total	1826	2528	15720	223471	41705	107723	67705	112339	178760	437738	293536	342466	49220	401610	520740	145	1230	18707	243	9893	226264	580	273657	31976	289339	

Table 10: Estimated proportions of employee categories for PECAS activities

	AI01 AgMinMan	AI02 AgMinProd	AI03 ConMan	AI04 ConProd	AI05 MfgMan	AI06 MfgProd	AI07 TCUMan	AI08 TCUProd	AI09 Whole	AI10 Retail	AI11 FIRE	AI12 PTSci	AI13 ManServ	AI14 PBSOff	AI15 PBSRet	AI16 PSInd	AI17 Religion	AI18 BSOnsite	AI19 PSOnsite	AI20 FedGov	AI21 StLocGov	AI22 Military	AI23 GSEdu	AI24 HiEdu	AI25 Health
N11	0.443	0.557																							
N21	0.405	0.595																							
N22							0.495	0.505																	
N23			0.066	0.934																					
N313233					0.279	0.721																			
N42									1.000																
N4445										1.000															
N4849							0.369	0.631																	
N51														1.000											
N52											1.00														
N53											0.569				0.431										
N54												1.000													
N55													1.000												
N56													0.001	0.514	0.433			0.052							
N61																							0.895	0.105	
N62															0.441										0.559
N71															1.000										
N72										0.843				0.157											
N81														0.797	0.190	0.001	0.009		0.002						
N92																				0.042	0.956	0.002			

Appendix 3: Procedures and scripts for the 'Site-spec' functionality

Identifying parcels for site

Select parcels in QGIS or other GIS e.g. ESRI ArcMAP.

Overlay base layer (e.g. Stamen toner and Google Earth) to see location and whether it's vacant, save screenshots of imagery/maps.

Export parcel list as CSV to get list of parcels outside of the GIS system/database.

Check address in google street view.

Decide how much space and what type, and what years, enter into site_spec_amounts e.g.

Site	SpaceType	Year	Amount
1	72	2022	2500000
1	72	2025	2500000
1	72	2030	2500000
1	72	2035	2500000
1	72	2032	2500000
1	82	2030	725000
1	82	2022	500000

Decide which parcels will be used for which space types, and what year they will be made available. Figure out their PECAS Parcel number with a query like this:

```
select p.pecas_parcel_num, t.county_fips from wXX.parcels_backup p
join wXX.tazs t on p.taz = t.taz_number
where parcel_id in (
'17 0190  LL0153',
'17 0190  LL0211',
'17 0190  LL0229',
'17 0190  LL0286',
'17 0225  LL0046',
'17 0225  LL0053',
'17 0225  LL0103',
'17 0225  LL0338',
'17 0225  LL0152',
'17 0225  LL0160',
'17 0225  LL0186',
'17 0225  LL0277',
'17 0225  LL0346',
'17 0225  LL0368',
'17 022600110512',
'17 0225  LL0145',
'17 0225  LL0178',
'14 0144  LL0017',
'14 0144  LL0025',
'14 0145  LL0024',
'14 0145  LL0032',
'14 0145  LL0040',
'14 0145  LL0081',
'14 0145  LL0149',
'14 0145  LL0156')
```

Update the PECAS SD database for the scenario to specify the sitespec

Insert the sitespec_parcels with a query similar to this

```
insert into wXX.sitespec_parcels
select p.pecas_parcel_num, 1, 72, 0, p.land_area, 2022, 2036 true
from wXX.parcels_backup p
join wXX.tazs t on p.taz = t.taz_number
where parcel_id in (
'17 0190  LL0153',
'17 0190  LL0211',
'17 0190  LL0229',
'17 0190  LL0286',
'17 0225  LL0046',
'17 0225  LL0053',
'17 0225  LL0103',
'17 0225  LL0338',
'17 0225  LL0152',
'17 0225  LL0160',
'17 0225  LL0186',
'17 0225  LL0277',
'17 0225  LL0346',
'17 0225  LL0368',
'17 022600110512',
'17 0225  LL0145',
'17 0225  LL0178',
'14 0144  LL0017',
'14 0144  LL0025',
'14 0145  LL0024',
'14 0145  LL0032',
'14 0145  LL0040',
'14 0145  LL0081',
'14 0145  LL0149',
'14 0145  LL0156')
and t.county_fips = 121
order by parcel_id;
```

This makes every one of these parcels available to sitespec in 2022, and also makes them unavailable to the regular SD development process until 2036, after all of our programmed development has occurred.

Modify the resulting table using pgAdmin3 to select different years and/or space types, we are using parcels `17 0190 LL0211; 131210096515` and `17 0225 LL0160"; 131210096701` for retail, so we interactively set the space_type_id to 82 in wXX.sitespec_parcels.

We also needed to change some of the parcels to be available in a different year, if we have amounts in different years in sitespec_site_amounts, some parcels needed to be available in those years.

The space_quantity can be 0 for now as we have a script (below) to update it under a constant FAR assumption.

Prepare the scenario for consistency

These steps update some of the tables and delete some inconsistent zoning information. Note that this makes the scenario schema in SD generally unsuitable for cloning for creating a new scenario with less sites specified, as zoning information has been destroyed as irrelevant for parcels on the sites.

If we had multiple zonings on parcels in different years before the sitespec year (e.g. multifamily to 2030, but allow highrise office in 2031 or later, and hold_until_zoning is 2035) we need to delete any zonings that are earlier than the latest (delete the multifamily zoning since the office zoning would apply in 2035 and later) zoning for the parcel that's before the hold_until_year: zoning_to_apply = zoning with year = (max year of zoning with zoning.years <= hold_until_year) delete zonings on parcel where year < zoning_to_apply.year

We need to delete any zoning on parcels with hold_until_year = null, to *permanently* disable SDs normal development functionality on those parcels.,

```
delete from wXX.parcel_zoning_xref pzx
where pecas_parcel_num in (select pecas_parcel_num from
wXX.sitespec_parcels where hold_until_year is null);
```

Update the zoning, to disallow standard SD development until after the sitespec is complete, with a query like this.

```
with t as (select pecas_parcel_num, max(hold_until_year) as
first_free_year
from wXX.sitespec_parcels group by pecas_parcel_num)
update wXX.parcel_zoning_xref pzx
set year_effective = t.first_free_year
from t where pzx.year_effective < t.first_free_year
and pzx.pecas_parcel_num = t.pecas_parcel_num;
```

Apply the total amounts of development to the individual parcels (optional, the user could also specify the development on each parcel, but this script assumes a constant FAR on the available parcels for the space type in the year.

```
with t as
(select siteid, space_type_id, year_effective, sum(land_area) as
total_land
from wXX.sitespec_parcels group by siteid, space_type_id,
year_effective order by space_type_id, total_land)
update wXX.sitespec_parcels sitep
set space_quantity = sams.quantity*sitep.land_area/t.total_land
from t, wXX.sitespec_site_amounts sams where
t.siteid = sitep.siteid and
t.space_type_id = sitep.space_type_id and
t.year_effective = sitep.year_effective and
t.siteid = sams.site_id and
t.space_type_id = sams.space_type_id and
t.year_effective = sams.year;
```

Running the model

The PECAS run script performs this query before each SD run:

```
update wXX.parcels p
set year_built = s.year_effective,
space_type_id = s.space_type_id,
space_quantity = s.space_quantity,
land_area = s.land_area,
is_derelict = false,
is_brownfield = false
from wXX.sitespec_parcels s
where p.pecas_parcel_num = s.pecas_parcel_num
and s.update_parcel = true
and s.year_effective = 2022
```

Note that the core PECAS SD module only looks at these tables to remove the sitespec amounts from its total *Construction_Capacity* target, so that SD doesn't also build the amount of space that has been manually specified and applied. The application of the sitespec to parcels is done in the run script, not in the PECAS SD module.

To confirm that the sitespec worked, use this query:

```
select * from wXX.parcels
where pecas_parcel_num in (select pecas_parcel_num from
wXX.sitespec_parcels)
```