PECAS - for Spatial Economic Modelling

THE ACTIVITY ALLOCATION MODULE CALIBRATION FOR THE ARC PECAS MODEL

System Documentation Technical Note

HBA specto

Calgary, Alberta

Table of Contents

1. Introduction

The Atlanta Regional Commission's PECAS ("Production Exchange Consumption Allocation System") model began development in the year 2008. It is a spatial economic model for the Atlanta Region in the State of Georgia. Its main purpose is to simulate the future location of activities (households and employment), and the development of space by developers, for both forecasting and policy analysis.

The ARC PECAS model includes the two standard PECAS modules: the Activity Allocation module (AA) and the Space Development module (SD). AA follows an aggregate approach and represents how and why industries, government and households choose to locate in different zones or locations in the region. SD follows a deterministic microsimulation approach and simulates development at the parcel level, taking into account developers' profit-motivated behavior as well as land and market characteristics. These two modules interact with each other, and both also interact with the Atlanta transport model by providing it with land use data. The transport model, in turn, provides an indication of travel conditions for use in AA.

The base year for the current ARC PECAS model is the year 2020 and it is run through time to 2050. It was originally developed with a zone system involving 79 super districts – also called land use zones (LUZ) – and 2,024 Traffic Analysis Zones (TAZ). The current resolution of the model is higher than before presenting around 1,000 LUZs and 6,000 TAZs.

The current ARC PECAS model has been enhanced in different ways, but the two major recent improvements include updated parcel data and rent data for residential and non-residential space for 2020, which is the current base year of the model.

To complete the ARC PECAS model development, the two PECAS modules – Activity Allocation and Space Development – are being calibrated to set the value of important behavioral parameters.

The objective of this report is to briefly describe the procedures involved in the calibration of the Activity Allocation Module of the ARC PECAS model, which include new target data for the floorspace calibration, along with updated calibration software and procedures. Most of the procedures involved are contained in spreadsheets, scripts, databases, programing code and other technical documentation. These valuable pieces of information are too extensive to be included in this technical note, but extracts or excerpts of some files are shown in this document, so the reader can have a sense of the data collected, updated, and required to execute calibration procedures in the Activity Allocation module of the PECAS framework.

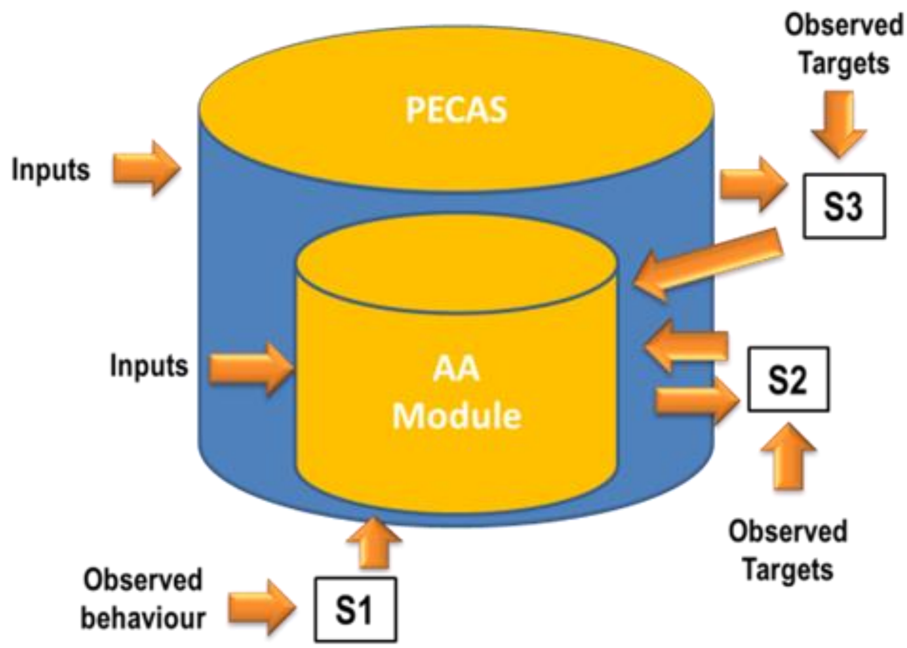
2. Overview of calibration procedure followed in the PECAS framework

Under the PECAS framework the word "calibration" refers to the processes to establish parameter values and constants that better allow the model to reproduce the modelled conditions, without compromising any important piece of the model. For this it is required to have observed target data to represent the conditions to be modelled.

2.1. Parameter types

For the entire PECAS model system three types of parameter values can be identified: Stage 1 (S1), Stage 2 (S2) and Stage 3 (S3). Figure 1 shows the interaction of these parameters in the PECAS framework. A brief description of these parameters and some examples are presented below:

Figure 1. Parameter values based on the three stages of calibration.



Source: G. Fuenmayor. Building a spatial Economic Model for Caracas using PECAS.

Stage 1 (S1) parameters are 'one time' parameters, that are estimated once using data from outside of the model, and therefore are not based on model values instead they come from observed behavior (John Douglas Hunt 1994). At this stage the model is "under construction" so the whole system does not have to run yet.

Examples of S1 parameters are observed behavior of the households or the employment such as rates of consumption of housing for different residential types or rates of space of non-residential space per employee for different industries; or average rents and average vacancy rates associated with different markets of space; or the transport costs associated with the movements of goods, services, or labor in a study area. More details in Table 1.

Table 1. Inputs for S1 parameters

Parameter	Inputs
Fixed and elastic technical coefficients	- Aggregate Economic Flow Table - Synthesize households (residential space consumption rates)
Short-term floorspace supply functions	- Vacancy rates per space type - Average rents per space type
Transport Coefficients for the Goods commodities	- Cost per weight (US Dollar/ton) per commodity - Average use of truck capacity by commodity and by type (capacity) - Cost per minute (US Dollar /min) per type of truck - Cost per kilometer (US Dollar/km) per type of truck - Distribution of trucks by commodity and by type (capacity)
Transport Coefficients for the Household Obtained Services commodities	- Total annual consumption (money value) of commodities by households - Total number of annual visits by purpose - Transport Money Cost Coefficient
Transport Coefficients for the Labour commodities	- Distribution of annual wages by socioeconomic level - Total number of annual trips by socioeconomic level - Transport Money Cost Coefficient
Transport Coefficients for the Worker delivery services commodities	- Total annual production (money value) of commodities by industries - Total number of annual visits by purpose - Transport Money Cost Coefficient

Source: G. Fuenmayor. Building a spatial Economic Model for Caracas using PECAS.

Stage 2 (S2) parameters are 'heuristic parameters' indicating that "they are determined from data synthesized by the model and must be re-estimated each time the model is adjusted in any way that alters these data" (John Douglas Hunt 1994).

In this second stage initial values are established for all the S2 parameters based on runs of components of the whole model system to better match the targets. This considers a single "module" of PECAS, the Activity Allocation or the Space Development Module. For the Activity Allocation Module, we present these S2 parameters in Table 2.

Table 2. parameters for the Activity Allocation Module

Calibration type	Stage 2 parameter	Target
Trip Length (Exchange Location)	Parameter to control the sensitivity of buyers and sellers of a commodity c, to differences in desirability between exchange locations. This parameter is also referred as the **Exchange Location dispersion parameter.**	Average trip length for the 'Puts' or 'Commodities'.
Technology Choice for Households (Option Weight) or Households calibration	This parameter is a term inside of the constant of the utility function for the technology choice. It is also referred as the **Option Weight dispersion parameter.**	Quantities of households by category with labour occupations and average use of residential space Quantities of non-residential space used by industry.
Technology Choice (Option Weight) for Imports and Exports or Import-Export Calibration	This parameter is a term inside of the constant of the utility function for the technology choice. It is also referred as the **Option Weight dispersion parameter.**	Quantities of imports and exports for import providers and export consumers

Calibration type	Stage 2 parameter	Target
Floorspace	There is no parameter associated with this calibration. The model adjusts floors space quantities in order to match floor space prices (rents) or adjust rents in order to match floor space quantity targets.	Quantities of space by type and by zone. Estimated rents by zone.

Source: G. Fuenmayor. Building a spatial Economic Model for Caracas using PECAS.

Stage 3 includes the refinement of parameter values based on running the whole integrated model system, considering the fit of all modules together, and focusing on the long-term model performance and in the sensitivity test. This stage supposes that both PECAS modules, AA and SD are developed and calibrated towards the base year, and that the model runs through time. Parameter values could be adjusted during sensitivity tests to validate the whole system (J.E. Abraham, Weider, and Hunt 2005).

2.2. Activity Allocation Calibration Cycles

In the previous section when referring to the S2 parameters four types of calibration were presented: Trip Length, Households Calibration, Import-Export Calibration and Floorspace Calibration. The calibration process can be run in 2 cycles, each cycle runs under different model conditions, adjusting the corresponding parameters or quantities depending on the calibration type, and depending on the goal and on the objective to achieve during the specific calibration type.

During the first cycle only two calibration types are run, trip length and option weight (Households and Import Export Calibration). At this point the model doesn't need to converge, the focus is on matching the activity targets model-wide (commodity flows and rates of production and consumption of labor and space) while the model uses observed prices. This means that the prices for the commodities are required inputs. Initially, we are setting each parameters and constant within a reasonable range of possibilities, so that the various behaviors are reasonable, but not necessarily realistic (yet), and the software doesn't encounter any numerical overflows (Table 3).

During the second cycle all the calibration types can be run, trip length, option weight (Households and Import Export Calibration) and floorspace. At this point the model needs to converge reaching equilibrium prices, and the focus is on matching the targets by land use zones (LUZs). Trip length and rates of production and consumption should be match model-wide, but the model needs be able to produce simulated prices as model outputs (which need to match target prices) and match constraints (categories of households and employment) by LUZ. The main objective after running this cycle is to generate the Activity Location choice Alternative Specific Constant for each LUZ, which helps the model to simulate the observed base year behavior by LUZ in the future (Table 3).

Table 3. Calibration Cycles

Cycle	Parameters being adjusted	Conditions when running the model	Goal	Objective of the calibration cycle	Calibration status
1st cycle	Exchange Location Dispersion Parameters (DP)	- Using observed prices - Match activities model-wide (unconstrained) - Model does not need to converge	- Match approximately the commodity flow length distributions	- To simulate behaviour using observed prices. This can be achieved by changing the DP in the logistic equation until the resulting probability is in the sensitive region of the cumulative probability distribution - To help in obtaining parameters and constants (option weight) roughly close to what is required for the next cycle	**Is the model ready for the next cycle?** - The model should be sensitive enough (probability in the sensitive range of the cumulative probability distribution) - The model should run without crashing, which means parameter values and constants need to be in the appropriate range of values. - If this is achieved the model is ready for the next cycle.
	Option Weight		- Match approximately the model-wide proportions of each technology option (labour and space) model wide		
2nd cycle	Floorspace quantities and prices	- Using simulated prices - Constrained to match activities by LUZ - Model Convergence is required	- Match targets of floorspace quantities and rents by space type and LUZ	- Simulated equilibrium prices for all the commodities in all the LUZ - To refine parameter values and constants (option weight) until converged model outputs match targets - To provide the activity location choice alternative specific constants for each LUZ	**How much calibration should be performed in this cycle?** - The model should be running without crashing (parameters and constants should be in range) - Converge model values needs to be compared to the targets: in each calibration type: o TLC (Trip Length) o OWC (Technology proportions) o FC (Quantities and Rents) - This cycle can be repeated until converge model outputs match the target
	Exchange Location Dispersion Parameters (DP)		- Match the commodity flow length distributions		
	Option Weights		- Match the model-wide proportions of each technology option (labour and space)		
	Alternative Specific Constant by LUZ		- Match locations of the activities by LUZ		

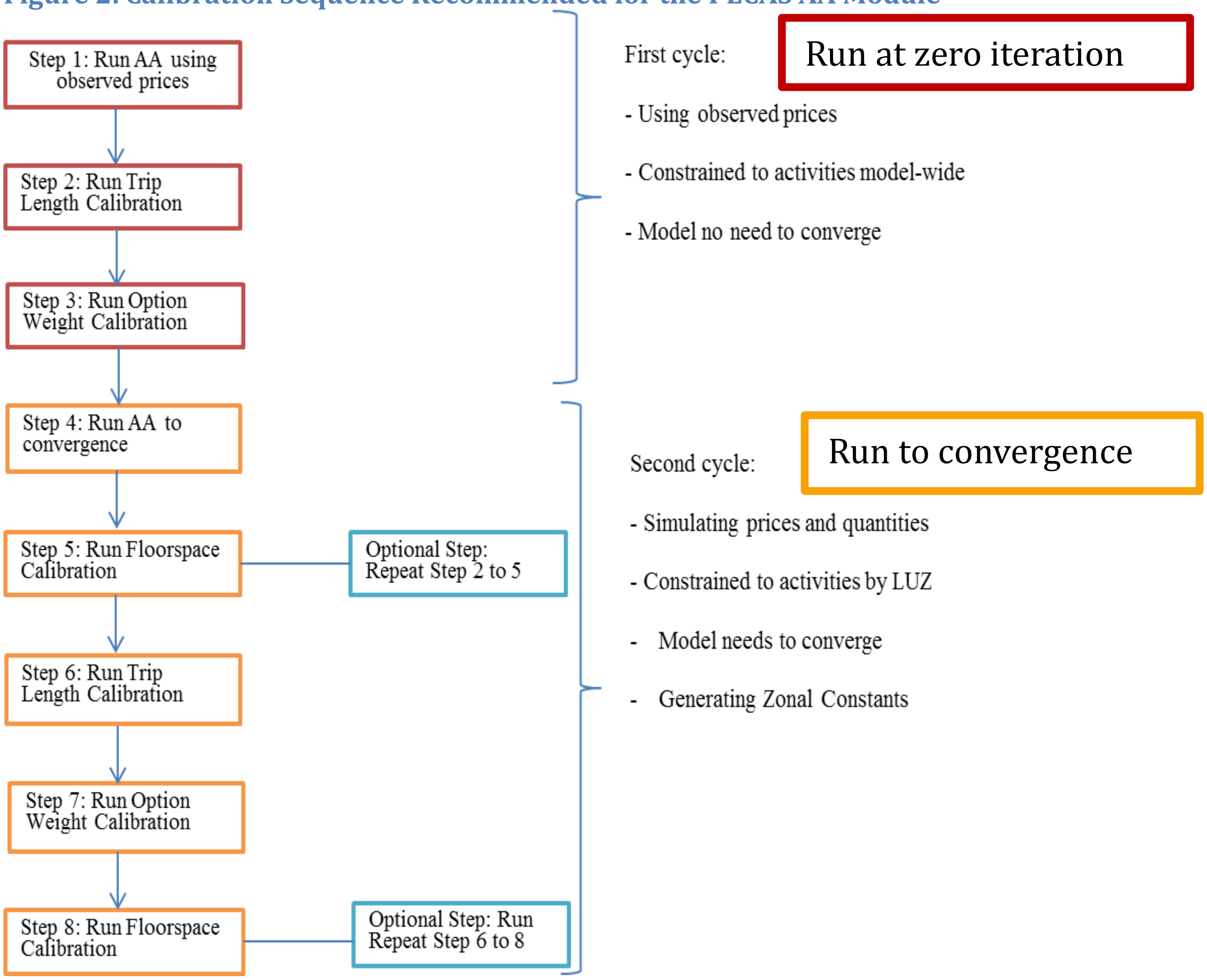
2.3. Activity Allocation Calibration Sequence

Based on experience with several early PECAS models, an AA calibration sequence was developed initially for the PECAS Sacramento model. A series of training videos explain these sequences (J. E. Abraham 2014), and a simplified flowchart is shown in Figure 2. This suggested sequence was used as a general guide for the ARC PECAS model (Figure 2).

Following this sequence allows certain components to be calibrated to a certain degree, where the targets are close to being met, and then improvements can be made in all of the components, revisiting some steps with more restrictive conditions.

During the calibration process each step needs to be tuned until is accomplished to be able to run smoothly, meaning that the parameter values should allow the model to simulate the targets. The recommended calibration sequence provides a guidance to the analyst but shouldn't be used as a constraint.

Figure 2. Calibration Sequence Recommended for the PECAS AA Module



Source: J. Abraham. Recommended sequence from the Sacramento PECAS Model.

The following subsections explain briefly the three types of calibration performed for the AA module. They report the calibration script and other files required, examples of target data, and examples of the results from the calibration process by calibration type.

3. The Trip Length calibration

The aim of trip length calibration is to make the average length of the commodity flows generated by the AA module match the actual average trip length for that type of commodity. This is done by adjusting the buying and selling dispersion parameters (DP) for each commodity until the average trip lengths equal the targets.

These dispersion parameters appear in the formulas for allocating the buying and selling of commodities to the different exchange zones. The desirability of an exchange location includes representation of the sensitivity to size, to commodity price and general transport cost. Size is implicitly adjusted inversely with the DP. Since the DP is multiplied by the utility values for exchange size, commodity price, and generalized cost of transportation, it affects the way that the exchange zone allocation responds to those signals. If the dispersion parameter is large, the allocation becomes more sensitive to transport cost, meaning that AA will exchange more of the commodity in nearby zones, reducing the average length of trips for that commodity.

The targets for the average trip length can be reported in time or distance and the data is normally obtained from the trip length frequency distribution by trip category and income. The targets for the ARC PECAS are obtained from the calibrated trip model for the Atlanta Region, which is the previous version of the current Activity Based Model and are shown in Table 4Table 4.

The trip length calibration script (TLC.py) matches the targets by increasing the dispersion parameters for commodities whose trips are too long and decreasing them for commodities whose trips are too short. The TLC algorithm does this by running the AA model several times (number of iterations indicated by the analyst) while adjusting the dispersion parameters until it matches targets. The calibration algorithm TLCalib.py uses the secant method for root finding. Since the secant method requires two initial guesses, the calibration algorithm uses a given initial scale factor of 1.2 (This can be adjusted by the analyst) to generate the first adjustment to the DP. In the ARC PECAS model, travel length is affected by the transport disutility of the 24 transportable commodities defined for this region. These commodities were divided into 8 groups. The assigned groups of commodities and their travel time targets are shown in Table 4.

The current calibration script is run from a terminal window (e.g. Git Bash under Windows) and the settings are specified in the file: tlcalib.yml. Before the TLC script is run, two files are needed to indicate the setup of the calibration. TLCGroupsI.csv contains the groups of the commodities, while TLCTargetsI.csv has the initial dispersion parameters and target trip length values. Moreover, the number of maximum iterations and allowable error for the trip length should be specified in the script. The skim matrix used to calibrate the parameters of each commodity and intervals of the trip lengths used to calculate the average trip lengths are specified in an input file called HistogramsI.csv.

Table 4. Commodities, Groups and Target distance values used for the trip length calibration in the ARC PECAS

Commodity	Group	Target (Average Trip Time in minutes)
CG11AgFor	CV Heavy	45
CG21Mining	CV Heavy	45
CG22Util	CV Med/Light	35
CG23Constr	CV Heavy	45
CG313233Manu	CV Med/Light	35
CS42Whlsale	CV Med/Light	35
CS4445Retail	CV Med/Light	35
CS4849Trans	CV Med/Light	35
CS51Info	CV Med/Light	35
CS52Finance	CV Med/Light	35
CS53RealEst	Work-Other	18
CS54ProfTech	Work-Other	18
CS55Manag	Work-Other	18
CS56AI56Admin	Work-Other	18
CS61EduServ	HB School	15.5
CS62Health	HB Work	34
CS71Arts	HB Work	34
CS72Accom	HB Work	34
CS81Other	HB Work	34
CL01BlueCollar	HB Work	34
CL02Health	HB Work	34
CL03RetailandFood	HB Work	34
CL04Services	HB Work	34
CL05WhiteCollar	HB Work	34

When TLC.py runs, it creates “TLCOutput_n.csv” where “n” is the iteration number, which has the results of the average trip lengths, estimated parameter and model error for every iteration. The file called “AllTLCOutput.csv” appends all the files from that calibration session (e.g. all the ones specified at that moment in the settings). These files allow the analyst to track the error coming down in each iteration and are created in the scenario folder. An example of how the trip length calibration results is presented in Figure 3.

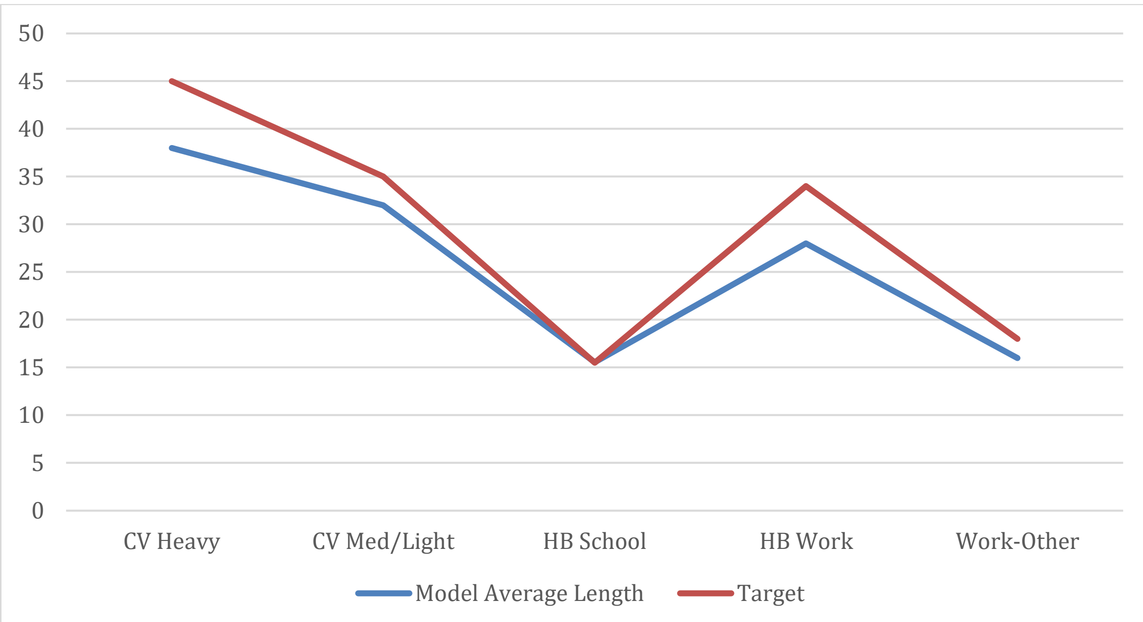
When the trip length error of all commodities is less than five percent, the run will be stopped and the trip length for each commodity for a given trip length intervals are written in the output file called Histograms.csv.

When the system shows a file with a number between parenthesis (n) it means that the calibration system failed to achieve an improvement towards the targets. In this case, two things could be happening:
a) The model cannot improve because it already achieved the targets
b) Under the current conditions (e.g. data influencing the transport utility function) and parameters (DP) the model is not able to achieve the targets

At that point in the calibration process the analyst has the following options:

a) Revisit and check the calculation of the transport cost coefficients associated with each commodity and the skims.
b) Check the range of the DP to verify if it hit the maximum or the minimum of the valid range, usually from 5 to 100 for applications in the US and Canada.

Figure 3. Target versus calibrated model by commodity group caegory



4. The Technology Choice or Option Weight calibration

Option weight calibration is the second step in calibrating the AA module of PECAS, following trip length calibration.

PECAS uses technology options (vectors of technology, called "lifestyles" when used with household activities) to represent the different combinations of commodities that an activity can produce and consume. The aim of option weight calibration is to find weights such that the total amounts of production and consumption for certain commodities match the targets. The option weight is part of the constant, one of the terms in the technology utility for each option. Changes in the constants affect the probability of choosing that technology option and ultimately impact the total quantity chosen (produced or consumed) from that option.

The option weight term for a technology option reflects the “inherent desirability” of that option relative to the others. With all else equal, the odds of choosing a technology option are proportional to that option weight term.

There can be a number of options in the same activity. In PECAS, tags in the TechnologyOptionsI.csv file are used to indicate how each technology option differs from the others under the same activity. With respect to a single commodity, option size calibration relies on three tags for each option describing how it relates to levels of production or consumption of commodities: “more”, “less”, or “zero”.

Option weight calibration script adjusts the weights for each technology option to match production and consumption targets. Each target is the total amount of a given commodity that should be made or used by a given activity across the entire model region.

Option weight calibration uses an iterative approach, each iteration adjusting the weights on “more” and “less” options to move the aggregated production and consumption amounts towards the targets. The adjustment factor for “less” options is smaller than that for “more” option: both “less” and “more” options contribute to the overall production or consumption of a commodity and so they both need to be adjusted. (This means that even if one or more of the options is missing e.g. there is a “more” and a “less” option but no “zero” option, the balance between the options will still shift in the right direction.).

The option weight calibration is usually performed in two stages. In the first stage, calibration is done unconstrained (or constrained model-wide) with observed prices in AA, in order to reach a region-wide balance of production and consumption at prices that are exogenously set to be close to observed or target prices. In the second stage, calibration is done constrained by Transport Analysis Zones (TAZs) with AA run to convergence, which achieves a balance with equilibrium prices zone-by-zone. During this second stage, AA was set to not change the prices of the floorspace commodities, as those are treated in the floorspace calibration step instead.

The option weight calibration algorithm is applied for the calibration of households and for the Imports and Exports. But because the purpose and the targets are different, they are executed as two separate calibration procedures and in consequence the script, the settings, the targets file names, and the calibration output file names are different.

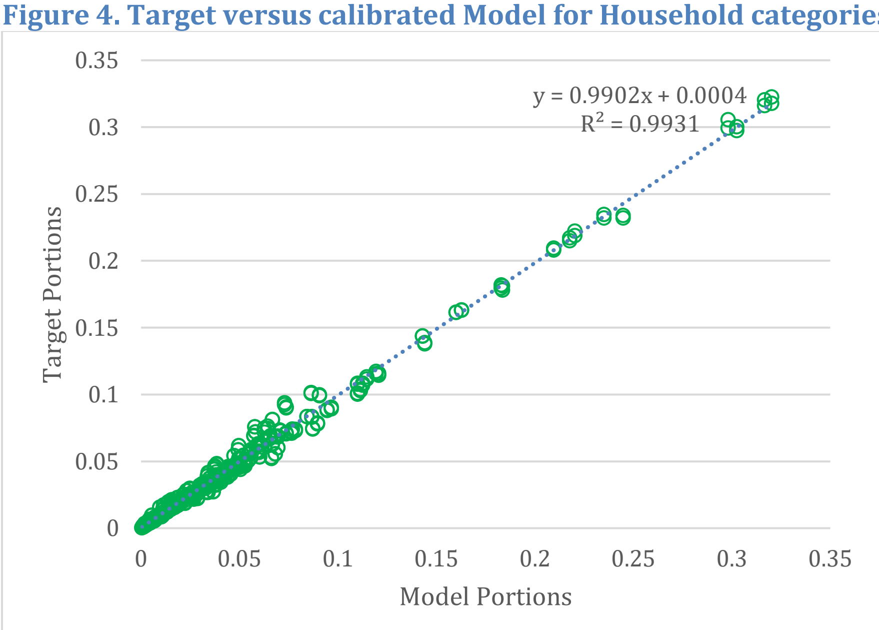
For the Households Calibration, the script is called clustercalib.py, and the settings are specified in the clustercalib.yml file. To run the household calibration one target files containing portions of households by size and income with their labour occupation and living in certain housing type is needed. The portions inside of each household type need to add to 1. An example of an excerpt of the target file is shown

in Table 5. The file is called by default householdClusterChoiceTargets.csv, but the analyst can choose a different file name, and adjust the settings correspondingly. An example of households calibration results are shown in
Figure 4.

Table 5. Excerpt of targets for households' calibration

Activity	Option	Target
I1H1	I1H1|Less CA36HiDenResid|More CL01BlueCollar	0.052796
I1H1	I1H1|Less CA36HiDenResid|More CL02Health	0.014058
I1H1	I1H1|Less CA36HiDenResid|More CL03RetailandFood	0.067612
I1H1	I1H1|Less CA36HiDenResid|More CL04Services	0.044128
I1H1	I1H1|Less CA36HiDenResid|More CL05WhiteCollar	0.084201
I1H1	I1H1|Less CA36HiDenResid|More Not working	0.047194
I1H1	I1H1|Less CA35DetResid|More CL01BlueCollar	0.03948
I1H1	I1H1|Less CA35DetResid|More CL02Health	0.006644
I1H1	I1H1|Less CA35DetResid|More CL03RetailandFood	0.025706
I1H1	I1H1|Less CA35DetResid|More CL04Services	0.027615
I1H1	I1H1|Less CA35DetResid|More CL05WhiteCollar	0.052541
I1H1	I1H1|Less CA35DetResid|More Not working	0.038024
I1H1	I1H1|More CA36HiDenResid|More CL01BlueCollar	0.052796
I1H1	I1H1|More CA36HiDenResid|More CL02Health	0.014058
I1H1	I1H1|More CA36HiDenResid|More CL03RetailandFood	0.067612
I1H1	I1H1|More CA36HiDenResid|More CL04Services	0.044128
I1H1	I1H1|More CA36HiDenResid|More CL05WhiteCollar	0.084201
I1H1	I1H1|More CA36HiDenResid|More Not working	0.047194
I1H1	I1H1|More CA35DetResid|More CL01BlueCollar	0.03948
I1H1	I1H1|More CA35DetResid|More CL02Health	0.006644
I1H1	I1H1|More CA35DetResid|More CL03RetailandFood	0.025706
I1H1	I1H1|More CA35DetResid|More CL04Services	0.027615
I1H1	I1H1|More CA35DetResid|More CL05WhiteCollar	0.052541
I1H1	I1H1|More CA35DetResid|More Not working	0.038024

Figure 4. Target versus calibrated Model for Household categories.



When clustercalib.py runs, it creates calibration output files called "ClusterCalib_n.csv" where "n" is the iteration number, and these files report how close or far is each household option to achieve the targets. The option weight parameter is reported in the file called TechnologyOptionsI_n.csv, where the n is the iteration number. The file called "AllClusterCalib.csv" appends all the files from that calibration session (e.g. all the iterations specified in the settings when running the calibration). This file allows the analyst to track the error coming down in each iteration and is created in the scenario folder.

For the Import Export calibration, the script file name is iecalib.py, while the settings should be specified in the file called iecalib.yml. To run the imports and exports calibration one target files with the amounts of imports and exports is required. Import and Export target amounts are presented in Table 6. The file is called by default ImportExportCalibI.csv, but the analyst can choose a different file name and adjust the settings accordingly.

Table 6. Target amounts for Import Export calibration

Activity	Commodity	MorU	Target
Import Providers_CG11AgFor	CG11AgFor	M	686226111.2
Import Providers_CG21Mining	CG21Mining	M	9324843987
Import Providers_CG313233Manu	CG313233Manu	M	76010003186
Import Providers_CS42Whlsale	CS42Whlsale	M	5267640561
Import Providers_CS4445Retail	CS4445Retail	M	81451704.3
Import Providers_CS4849Trans	CS4849Trans	M	6814066111
Import Providers_CS51Info	CS51Info	M	5258358024
Import Providers_CS52Finance	CS52Finance	M	9640186875
Import Providers_CS53RealEst	CS53RealEst	M	11499728707
Import Providers_CS54ProfTech	CS54ProfTech	M	3201489519
Import Providers_CS56AI56Admin	CS56AI56Admin	M	1660925157
Import Providers_CS61EduServ	CS61EduServ	M	1692406352
Import Providers_CS62Health	CS62Health	M	9422078834
Import Providers_CS71Arts	CS71Arts	M	2712958266
Import Providers_CS72Accom	CS72Accom	M	1312649975
Import Providers_CS81Other	CS81Other	M	2957018312
Export Consumers_CG11AgFor	CG11AgFor	U	-47684496
Export Consumers_CG21Mining	CG21Mining	U	-122401414.7
Export Consumers_CG22Util	CG22Util	U	-1277605505
Export Consumers_CG23Constr	CG23Constr	U	-898603816.9
Export Consumers_CG313233Manu	CG313233Manu	U	-41608009218
Export Consumers_CS42Whlsale	CS42Whlsale	U	-24141761793
Export Consumers_CS4445Retail	CS4445Retail	U	-1618968740
Export Consumers_CS4849Trans	CS4849Trans	U	-24241629486
Export Consumers_CS51Info	CS51Info	U	-32661966365
Export Consumers_CS52Finance	CS52Finance	U	-19065584534
Export Consumers_CS53RealEst	CS53RealEst	U	-29630687842
Export Consumers_CS54ProfTech	CS54ProfTech	U	-12577559921
Export Consumers_CS55Manag	CS55Manag	U	-3791883143
Export Consumers_CS56AI56Admin	CS56AI56Admin	U	-6567854702
Export Consumers_CS61EduServ	CS61EduServ	U	-4487651088
Export Consumers_CS62Health	CS62Health	U	-10508381617
Export Consumers_CS71Arts	CS71Arts	U	-2854654349

Activity	Commodity	MorU	Target
Export Consumers_CS72Accom	CS72Accom	U	-3037371584
Export Consumers_CS81Other	CS81Other	U	-2635904869

When iecalib.py runs, it creates calibration output files called "ImportExportCheck_n.csv" where "n" is the iteration number, and these files report how close or far is each import or export option to achieve the target amounts. The option weight parameter is reported in the file called TechnologyOptionsI_n.csv, where the n is the iteration number. The file called "AllImportExportCheck.csv" appends all the files from that calibration session (e.g. all the iterations specified in the settings when running the calibration). This file allows the analyst to track the error coming down in each iteration and is created in the scenario folder.

5. Floorspace calibration

In general, the floorspace calibration can be understood as the process of adjusting the quantity of space to achieve prices target that better match observations; normally, adding space of a given type will decrease the price of that space, and vice versa.

Adjusting the floorspace quantities to exactly match the observed prices used to be the recommended modelling approach, however based on previous experiences with several PECAS models, it was identified that this could distort the distribution of different types of space in cases of substitutable space types. Given that both quantity and price data contain errors and interpretation mismatches, and that the model itself is a representation of the reality but is not the reality, floorspace calibration attempts to balance between matching the observed space quantity and space price, while respecting the spatial distributions of activity (households and employment) and the floorspace consumption functions. To accomplish this, a "tolerance" value is assigned to both the quantity targets and the price targets for each LUZ/space type combination, representing the uncertainty of that target. If, for example, the tolerance for the price target is high while that for the quantity target is low, then after calibration, the quantity will match its target more closely than the price matches its target. In other words, the tolerances can be set by the analyst based on the reliability of the collected dataset.

The details about the mathematical approach of the Floorspace Calibration are contained in a technical documentation called "Theoretical and Mathematical approach of the Floorspace Calibration" (Hill 2012).

Space quantity targets are normally derived from the model's floorspace inventory by parcel, which is the estimated space by type and TAZ from the most updated parcel database. While price targets are developed from the updated space rent estimation by land use zone (LUZ) based on a sample of rent observations from the Costar data for 2020. There is a reasonable sample of observations for the rent estimation by zone, between 10,000 and 30, 000 depending on the space type.

For the Floorspace Calibration, the script is called weightfscalib.py, and the settings are specified in the weightfscalib.yml file. To run the floorspace calibration two target files are needed, one containing space quantities by type by TAZ and the second one with average zonal price per unit of space per space type and TAZ. Excerpts of these target files are shown in Table 7and Table 8 respectively. The target file containing the space quantities by TAZ is called FloorspaceTargets.csv and the one with the price targets per space and zone is called ExchangeResultsTargets.csv. In these two tables a column with the tolerance for each combination of space type and zone has been indicated. In this case, a smaller tolerance has been set up for the price targets.

Table 7:. Excerpt of quantity targets for the floorspace calibration

TAZ	Commodity	Quantity	Tolerance	LUZ
1	CA29AgMin	0	1000	551
1	CA30Indust	0	1000	551
1	CA31Retail	1476	1000	551
1	CA32Office	0	1000	551
1	CA33Instit	0	1000	551
1	CA35DetResid	578209	231283.6	551
1	CA36HiDenResid	0	1000	551
2	CA29AgMin	0	1000	551
2	CA30Indust	0	1000	551
2	CA31Retail	0	1000	551
2	CA32Office	89668	35867.2	551
2	CA33Instit	16540	6616	551
2	CA35DetResid	1597314	638925.6	551
2	CA36HiDenResid	0	1000	551
3	CA29AgMin	0	1000	551
3	CA30Indust	0	1000	551
3	CA31Retail	341275	136510	551
3	CA32Office	4562	1824.8	551
3	CA33Instit	122584	49033.6	551
3	CA35DetResid	4703497	1881398.8	551
3	CA36HiDenResid	0	1000	551
4	CA29AgMin	0	1000	485
4	CA30Indust	0	1000	485
4	CA31Retail	0	1000	485
4	CA32Office	0	1000	485
4	CA33Instit	90751	36300.4	485
4	CA35DetResid	2497911	999164.4	485
4	CA36HiDenResid	0	1000	485
5	CA29AgMin	0	1000	485
5	CA30Indust	12330	4932	485
5	CA31Retail	58985	23594	485
5	CA32Office	0	1000	485
5	CA33Instit	3862	1544.8	485
5	CA35DetResid	1706742	682696.8	485
5	CA36HiDenResid	150750	60300	485

Table 8. Excerpt of price targets for the floorspace calibration

Commodity	ZoneNumber	TargetPrice	Tolerance
CA30Indust	1	5	0.5
CA31Retail	1	13.58354805	1.358355
CA32Office	1	15.35139811	1.53514
CA33Instit	1	15.35139811	1.53514
CA35DetResid	1	15.35139811	1.53514
CA36HiDenResid	1	11.29106972	1.129107
CA30Indust	2	5.136099266	0.51361
CA31Retail	2	14.19169732	1.41917
CA32Office	2	16.23501853	1.623502
CA33Instit	2	16.23501853	1.623502
CA35DetResid	2	16.23501853	1.623502
CA36HiDenResid	2	11.40595484	1.140595
CA30Indust	3	17	1.7
CA31Retail	3	15.584015	1.558402
CA32Office	3	18.51578619	1.851579
CA33Instit	3	18.51578619	1.851579
CA35DetResid	3	18.51578619	1.851579
CA36HiDenResid	3	7.927753945	0.792775
CA30Indust	4	5	0.5
CA31Retail	4	15.14981414	1.514981
CA32Office	4	16.23842123	1.623842
CA33Instit	4	16.23842123	1.623842
CA35DetResid	4	16.23842123	1.623842
CA36HiDenResid	4	10.04090467	1.00409
CA30Indust	5	4.555516904	0.455552
CA31Retail	5	15.40789412	1.540789
CA32Office	5	20.61274192	2.061274
CA33Instit	5	20.61274192	2.061274
CA35DetResid	5	20.61274192	2.061274
CA36HiDenResid	5	10.26877081	1.026877

When weightfscalib.py runs, it creates calibration output files called "FloorspaceCalib_n.csv" where "n" is the iteration number, and these files report how close or far is each LUZ to achieve the targets in terms of quantities and prices per space type, it also reports an Error for each combination of space by zone (The formula of this error is sown in Equation 1). Another calibration output file is created when the iterations finish, it is called "AllFloorspaceCalib.csv", this one appends all the files from that calibration session (e.g. all the iterations specified in the settings when running the calibration). This file allows the analyst to track the error coming down in each iteration and is created in the scenario folder.

Since floorspace calibration is a balance between matching the quantity and price, its success is measured using the *total error*. For a given type of space in a given zone, the error for that combination is given by the following equation:

Equation 1

$$Error = \frac{(ModelledPrice - TargetPrice)^2}{PriceTolerance^2} + \frac{(ModelledSpace - TargetSpace)^2}{SpaceTolerance^2}$$

These error values can be aggregated to produce various total error measurements. The total error across all zones for each space type, both before and after calibration, is shown in Table 9.

Table 9: Total error by space type before and after floorspace calibration

Space type	Total error before	Total error after
Agriculture & Mining	0.0	0.0
Industrial	0.0	0.0
Retail	3.6	0.8
Office	1.4	0.2
Institutional	0.1	0.1
Detached Residential	40.3	17.5
High-Density Residential	85.4	17.0

More detail about how the individual zones changed because of floorspace calibration is shown in Figures 5 and 6. In each graph, the target price (in dollars per square foot per year) appears on the horizontal axis, while the modelled price appears on the vertical axis. Each circle represents the price of a certain space type in one zone. The black diagonal line represents a perfect match; if the model was perfect, or if price match was given an extremely low tolerance as compared to the floorspace quantity targets, all circles would lie exactly on the 45 degrees line.

The non-residential space types appear in Figure 5. Even before the calibration, all of the circles were very close to the ideal line, indicating that the prices were already reasonable. After calibration, they were even closer.

Figure 5: Fit to price targets – non-residential space

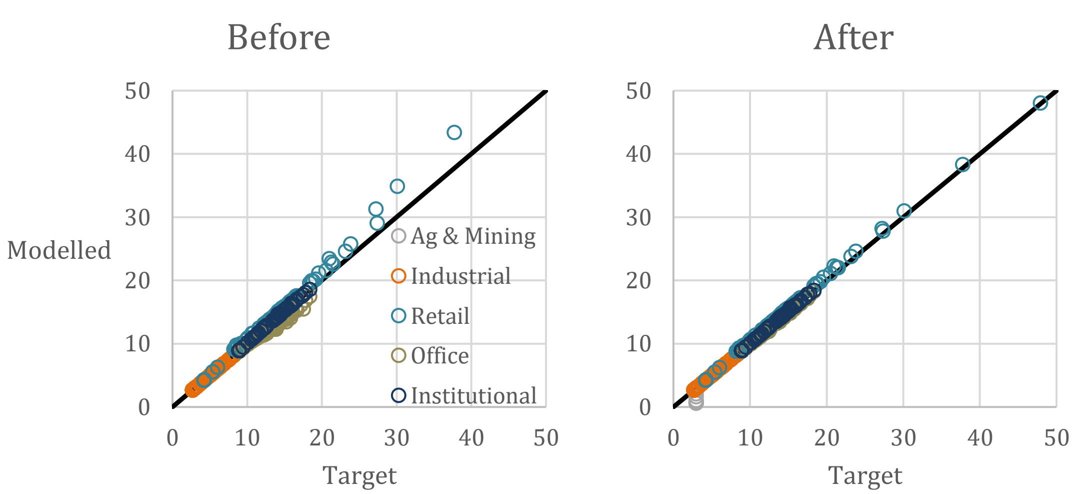
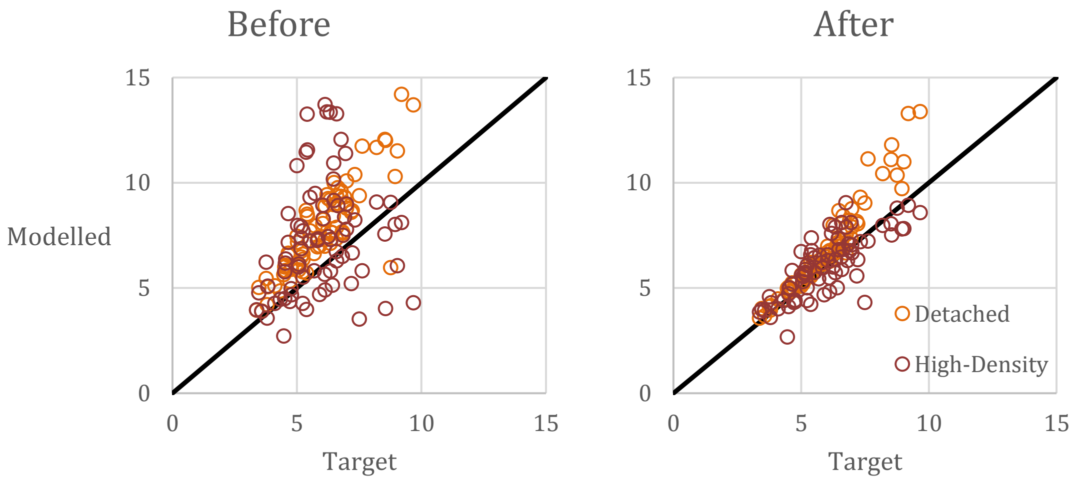


Figure 6 shows the price of residential space. Both detached and high-density residential space showed significant improvements in matching the observed prices. Before calibration, the price of detached residential space was consistently high, while high-density residential had a weak correlation between the observed and modelled price in a zone. Both of these shortfalls were addressed by floorspace calibration, though the final fit was not perfect because of the competing need to avoid large changes in space quantity.

Figure 6: Fit to price targets – residential space



6. The new calibration PECAS system

The PECAS Calibration Software has been updated to be consistent with the PECAS run system. Within a terminal window (e.g. Git Bash under Windows), the latest calibration software is installed with a series of commands, executed from the scenario directory.

- make _install (installs the proper version of PECAS for the scenario)
- make _install_latest_from_repository (PECAS uses GitHub repositories to perform version control of the calibration code developed for the AA and SD modules)
- make _upgrade_calib (to to get the latest software for calibration)
- source ve-win/Scripts/Activate for a windows computer, but if the scenario is located in a Unix machine then the command to run is: source ve/bin/activate

Once the calibration software is installed and upgraded, the calibration scripts can be run.

- tlcalib (to run Trip Length Calibration software) and settings should be specified in the tlcalib.yml file
- clustercalib (to run Household Calibration software) and settings should be specified in the clustercalib.yml file
- iecalib (to run Imports and Exports Calibration software) and settings should be specified in the iecalib.yml file
- weightfscalib (to run Floorspace and Rents Calibration software) and settings should be specified in the weightfscalib.yml file
- Model conditions (e.g. constrained or not, zero iteration or to convergence) this should be specified in the pecas.yml file

7. References

Abraham, J. E. 2014. *Calibration Sequence for the Sacramento PECAS Model*. Calgary, AB. http://www.hbaspecto.com/pecas/pecas_videos/.

Abraham, John E., Kevin J. Stefan, and J. D. Hunt. 2012. "Population Synthesis Using Combinatorial Optimization at Multiple Levels." In . http://trid.trb.org/view.aspx?id=1130260.

Fuenmayor Molero, Geraldine J. "Building a Spatial Economic Model for Caracas Using PECAS." Ph.D., University of Calgary, 2016. https://prism.ucalgary.ca/handle/11023/3061?show=full.

Fuenmayor, Geraldine, and John E. Abraham. 2013. "Generating PECAS Base Year Built Form for Clayton County in Atlanta." HBA Specto Incorporated.

Hill, Graham. 2012. "Theoretical and Mathematical Approach of the Floorspace Calibration." HBA Specto Incorporated.

———. 2014. "Bayesian Calibration Software for the SD Module." HBA Specto Incorporated.