Core Model Proposal #362: Implementation of multiple consumers in residential floorspace and energy demand

Product: Global Change Analysis Model (GCAM)

Institution: Joint Global Change Research Institute (JGCRI)

Authors: Jon Sampedro, Stephanie Waldhoff, Pralit Patel

Reviewers: Marshall Wise, Kanishka Narayan

Date committed: 20 May 2024

IR document number: PNNL-36034

Related sector: energy, buildings

Type of development: code, data, queries

Purpose: The purpose of this proposal is to implement demand for residential building energy services by income deciles within each GCAM region. This will represent the socioeconomic heterogeneity in demand across consumer groups within each region. For the moment, the subregional income distribution projections will be read exogenously as an input csv), which has been pre-computed using the pridr package (Narayan et al, in prep). In an upcoming CMP, the exogenous “income_shares.csv” file will be substituted by a new file generated within the gcamdata (LXX.Income.Shares) from a new R chunk.

Description of Changes and Validation Figures

The following subsections explain in detail the changes developed to the model for the implementation of multiple consumers in the residential floorspace and energy sectors. The model is structured in flexible "nodes" that automatically allow the implementation of different consumers ("gcam.consumer"), which can be easily nested as desired. Nevertheless, the implementation of multiple consumers requires several adjustments to the existing demand functions (including changing existing functional forms) to ensure that future model projections show an adequate behavior, so that consumers within a region consume the same quantities at the same income levels, all else equal. In addition, the existing representation of the buildings sector generates some issues with future service and energy demands when multiple consumers are implemented, which are not visible with a single representative consumer. Therefore, we have developed some additional changes associated with the structure of the residential sector (e.g., distinction between "modern" and "traditional" fuels) that will improve the final demand representation with multiple consumers.

This CMP is part of a larger series of changes to incorporate multiple-consumer representation in all the end-use demands within the model. In this CMP we focus on the residential floorspace and energy service demands. For the moment, the subregional income distribution projections will be read exogenously as an input csv file (extdata/socioeconomics/income_shares.csv), which has been pre-computed using the *pridr* package (Narayan et al, in prep). In an upcoming CMP, the exogenous "*income_shares.csv"* file will be substituted by a new file generated within the model DS (*LXX.Income.Shares*) from a new developed chunk. Moreover, two additional developments are planned for the mid-term, namely the implementation of the food demand and the transportation demand with multiple consumers.

In this CMP, changes have been classified in three different blocks for clarity: floorspace, residential energy service demand, and "other" changes (i.e., emissions and other adjustments). We have also included a "table of changes" that specifies which code files and R chunks have been modified for the implementation of this work.

In adapting the methods for computing both floorspace and residential and energy services (including services provided by both modern and traditional fuels) for multiple income levels (here, deciles), we have generally followed these principles and strategies for handling lack of global harmonized subnational data (described in more detail in the text but highlighted here).

- Design principle – within a region – consumers in a region should demand the same amount of either floorspace or per capita energy service when they reach the same income level.
    - This implies applying equal calibration parameters (e.g., bias adders) to each income decile.

- For both floorspace and energy services, there isn't data on service demand by income decile, so we address by -
  - First, we use the functional forms (for floorspace and energy service) to compute an estimate of demand for each decile.
  - Then we sum these decile level demands to get an estimate of total demand (for either floorspace or energy services)
  - Then, when there is data for regional total demand, we compare the estimate that was computed by summing the deciles to the data for the regional totals.
  - Then we compute a bias adder as the difference between the actual data for regional demand and the estimate computed from summing the decile demands.

Although this proposal is focused on residential buildings energy consumption – it implements a new calculation of satiation impedance by region based on base year data rather than a fixed value being read in. Because the Commercial Floorspace Demand uses the same satiation demand functional form – it now also uses this new calculation for its satiation impedance – which does result in small changes in the Commercial Buildings results as shown.

FLOORSPACE

RESIDENTIAL

The Gompertz function (shown below and updated in CMP 346) used for the estimation of future per capita residential floorspace, is by definition flexible to a potential implementation of multiple consumers (e.g., income groups) into the model.

$$f_{t,r} = (UnadjSat_r - a * \log(PD_{t,r})) * exp(-b * exp(-c * \log(GDPpc_{t,r}))) - k_r$$

Where,

UnadjSat is the unadjusted saturation value is the maximum per capita floorspace value a consumer demands at their maximum income level,

PD is the Population density in period t and region r.

k is a bias correction adder.

a, b, and c are parameters applied to the population density and per capita income.

However, some adjustments need to be developed in order to ensure the appropriate behavior of subregional (decile-level) floorspace demands. These are related to **calibration** and to the evolution of **per capita GDP** for different consumers.

Calibration

By default, the latest observed residential floorspace data in all GCAM regions is 2005, due to lack of more updated global data. Before the implementation of the Gompertz demand function (CMP 346), the approach was to take the observed 2005 per capita floorspace and keep it constant until the final calibration year (2015). Therefore, the 2005-2015 change on absolute residential floorspace was exclusively driven by population growth. In that CMP (346), we changed this approach and we implemented that the residential per capita floorspace on post-2005 historical years (2010 and 2015) is the value that results from the application of the Gompertz function in those years. However, we made this structure flexible so that if more updated residential floorspace data becomes available for any GCAM region, the user would just need to add it and indicate it in the "*regions_with_obs_data*" vector in the LA144.building_det_flsp chunk (Line 295).

At this moment we only have post-2005 observed data for USA and China, so these are the only two GCAM regions where residential floorspace in the final calibration year is actually observed and not estimated with the Gompertz function. With the implementation of multiple consumers, we needed to decide how to calculate this regional 2010-2015 calibration data for all regions (except China and USA, which have observations): either using national averages (average GDPpc) or estimating the floorspace values for each consumer (decile) and adding them up. Now that multiple consumers are implemented, the estimation can be completed at decile-level (and added up for regional values), we have tested these approaches by comparing the estimated per capita floorspace to observed values in those regions where we have real observations (China and USA):

PC FLSP (m2/pers)				
2015	Obsv	Est_avg	Est_quintiles	Est_deciles
China	38.02	18.46	19.51	20.37
USA	56.74	69.89	66.07	63.85

The comparison shows that the estimation using the deciles provides the better fit for the national estimation of residential per capita floorspace. Note that, while there are some differences between the observed and estimated values, these will be corrected with the bias adder that captures the "unobservable" effects that cannot be captured with the existing model variables. This bias adder parameter is calculated at national level and split equally across all consumers, to ensure all the consumers are in the "same path" and have exactly the same functional form for the estimation of future residential floorspace. In regions with no data, the bias adder is 0.
To summarize the steps

- Gompertz function is used to compute each decile's floorspace
  - Average floorspace computed for region from these decile floorspaces
  - Bias adder computed from difference between estimated average and observed
    - This bias adder is then applied to each decile Gompertz estimate of floorspace

Decile level income effect

Considering that per capita GDP is the main driver for per capita floorspace demand, and that there is no building vintaging in the model, a reduction on per capita GDP will directly imply a reduction on per capita floorspace. This means that, if population is kept fixed, a decrease in per capita GDP will reduce absolute residential/commercial floorspace, which would be interpreted as "destruction" of existing building stock.

In the core version of the model, with a single representative consumer, this effect is rarely observed, because most GCAM regions assume a positive growth rate for per capita GDP. However, it is actually happening in the few regions with decreasing per capita income (e.g., South America_Northern). With the implementation of alternative consumers, the evolution of subregional income distribution is a new factor that can generate that one consumer group gets relatively poorer from one period to another, despite average GDP increases in the region. In order to avoid this "destruction" of buildings, which will be more observable now due to consumer heterogeneity, we have assumed that residential floorspace in any region, period, and for each consumer group, will be calculated as the maximum between the value estimated for that period and the value in final calibration year (2015):

$$Flsp_{r,t,i} = \max\left(\widehat{Flsp_{r,t,i}}, Flsp_{r,0,i}\right)$$

- Floorspace is constrained to not go below base year values.

Note that this approach could also be applied to the commercial sector. While there are not multiple consumers implemented in the commercial sector, we can still see that commercial floorspace decreases from one period to another if there is a decrease in national average per capita GDP in a certain region. For example, in the core model we can see that commercial floorspace in SAN decreases around 31% from 2015 to 2020, which represents around -174M km2. Applying a similar approach would avoid these dynamics. However, it may make more sense with commercial buildings, since they may be closed if/as GDP decreases, so it is not modified within this CMP.

COMMERCIAL

In this CMP, there is no implementation of multiple consumers in the commercial sector. However, the adjustments to the satiation demand function, which is applied by both the commercial floorspace and residential modern energy service demands (see next subsection) also has a direct impact on future commercial floorspace projections. In the core version, satiation impedance is read in as 10.5 for all regions, and the regional differences (estimated - observed) are captured by a regional satiation adder as part of calibration. The adjusted satiation function developed for the demand of residential modern services changes this approach, and calibrates a region-specific satiation impedance that will be equal across income consumers, while the bias adder will be calculated as the difference between observed and estimated values, and equally distributed across consumer groups (see next subsection). The new, adjusted satiation function replaces the "satiation adders" used in the core with the "bias adders" used here.

With the application of this adjusted satiation function for commercial floorspace demand, which means with an updated (region-specific) satiation impedance, we will have some impacts for commercial floorspace which will differ around 15% at a global level compared to the core, with differences across regions (larger divergences in developing economies, see Figure 4). While this is a non-minor difference compared to the core, keeping this updated structure will simplify future implementations of multiple consumers in the commercial sector. In any case, this can be rolled-back if required by the review team.

RESIDENTIAL ENERGY/SERVICE DEMAND

Observed data shows how difficult it is to fit a single functional form for all the services and regions over the world. However, we have validated that the satiation principle could adequately represent the long-term residential service demand. In the model, the building service demand per unit of floorspace is represented by the following functional form (for service "s"):

$$f_{r,t,s} = k_{r,s} * (DD_{r,t} * \eta_{r,t} * R_{r,t} - \lambda_r * IG_{r,t}) * [1 - \exp(\frac{-ln2}{\mu_r} \frac{GDPpc_{r,t}}{P_{r,t,s}})])$$

The equation has three components. First (blue), a calibration parameter (*k*), that is calculated for each region. The second part of equation (orange) represents the *thermal load*, and it includes a range of parameters. Heating/cooling degree days (HDD or CDD) represent how much (in degrees), and for how long (in days) the temperature was below (above) a determined level. The model includes the option to read in HDDs and CDDs from different earth-system models, and with and without incorporating the future effects of climate change. η represents the building shell conductivity. We note this parameter combines indices for cooling and heating conductivity in a single value, so it indicates the "overall" building efficiency. The parameter can be set per region and sector (residential/commercial), and by default, it assumes a gradual efficiency improvement up to 2040 based on projections from the Annual Energy Outlook (U.S. EIA, 2016), with no further improvement for the longer term. *R* is the ratio of floorspace to building area. By default, it is assumed to be 5.5 for all regions without varying over time. It could also be set per region and sector, if desired. *IG* represents the internal gains per unit of floorspace associated with the heat released by "other" energy services (i.e., appliances) dividing the final energy of other services by the efficiency of each technology. λ is the internal-gains-scalar that adjusts the internal gains for the different regions by multiplying the regional internal gains with the ratio of degree days to USA degree days.

Finally, the third term (purple) represents the economic effect. It indicates the fraction of the satiated demand that is achieved, which increases with the affordability of the services. The affordability is measured as per capita income divided by the price of the service. Therefore, affordability and, consequently, service demand increase if per capita income increases or if the price of the service decreases. The μ parameter is the satiation impedance, and it represents the level of affordability required to achieve half of its satiation. This parameter is calculated based on the per capita service demand and the affordability (per capita income and service price) in the final calibration year. Parameters are listed below for convenience.

k is a calibration parameter,

HDD or CDD are heating and cooling degree days,

η is the building shell conductivity,

R is the ratio of floorspace to building area, set by default to 5.5,

λ is the internal-gains-scalar,

IG represents the internal gains per unit of floorspace associated with the heat released by "other" energy services,

μ parameter is the satiation impedance.

This proposal also implements a regionally-specific computation (or calibration) of satiation impedance. This is done both here in modern residential energy services demand and for commercial floorspace. The formula is given below.

$$\mu_{r,s} = \log(2) * \frac{\frac{pcGDP_{2015,r}}{P_{r,s}}}{\log(\frac{SatLevel_{r,s}}{SatLevel_{r,s} - D_{2015,r,s}})}$$

where

PcGDP is the base (2015) per capita GDP in $1975

P is the price of service s

SatLevel is the satiation level exogenously set for service s

D is the base year (2015) service demand per unit of floorspace in region r, and service s.

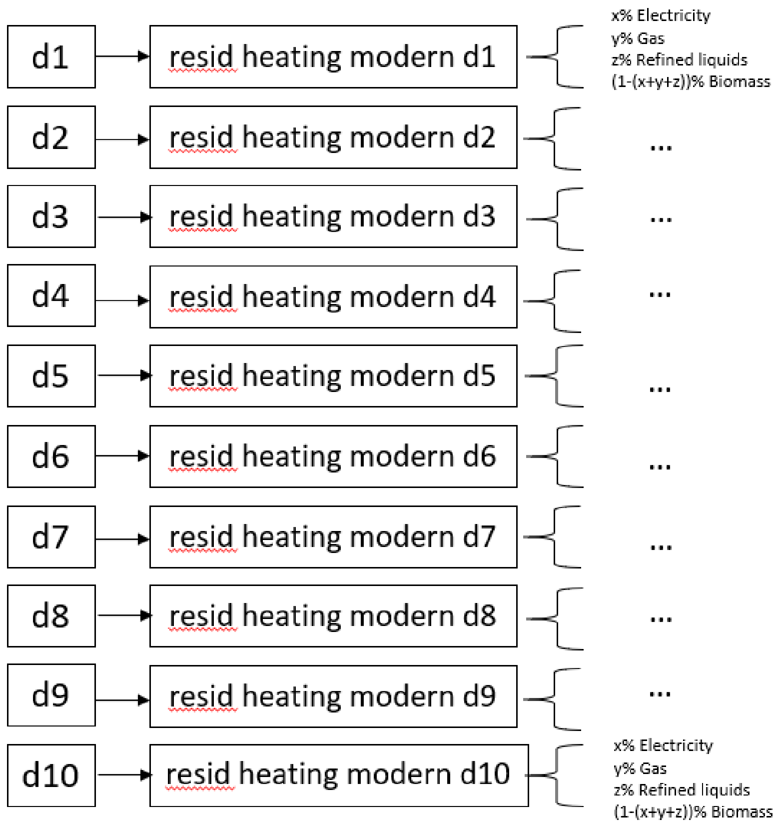
- Satiation level for residential floorspace is explained in previous CMP 346 when we set the Gompertz-type function and is not changed here.
- Satiation level for commercial floorspace varies per region class, and it comes from a workbook by Jiyong Eom
- Satiation for residential energy services come also from that workbook.

The table below shows the direction of the effects in total residential energy service demand

Driving force	**Direction of the effect**
Floorspace	+
Satiation level	+
GDP/cap	+
Average service price	-
HDD/CDD	+
Building shell efficiency	-
Floorspace-to-surface ratio	+
Internal gains	-

While a developed analysis suggests that the existing functional form would be an adequate representation for building energy service demand, the current structure will produce some undesired results if a user implements multiple consumer groups, that are not visible with a single representative consumer. In order to have a more accurate representation, we have first divided the sectors/fuels, distinguishing between "**modern**" versus "**traditional**" fuels.

Traditional fuels include coal and traditional biomass, and are represented individually. The following figure shows a representative example of the new structure.



NOTE: Technologies for each decile have equal shareweights, logits, costs, and efficiencies, so the technology shares within each sector are going to be equal across consumer groups. (x =x , y=y…). Emission factors are also the same across technologies. These are probably simplified assumptions, but there is no global data to capture the differences across the consumer groups in a non-arbitrary manner. Anyway, the flexibility of the structure allows to individualize costs, EFs, and assumptions if data gets available.

Modern sectors/fuels are the main energy source in developed economies, and we assume that their demand will increase as income rises, until a satiation level is achieved. Therefore, the already used satiation demand function is an adequate form for representing the demand of these modern services, but with some adjustments as shown below. On the other hand, we need to fit alternative functional forms for the demand of coal and traditional biomass, as their consumption has been observed to decrease as income rises.

Sector/Fuel	**Function**	**Description**
Modern services	Adjusted satiation demand	Increases with affordability (I/P) until a satiation level is achieved
Coal	Inverse	Decreases as affordability rises
Traditional biomass	Inverse	Decreases as affordability rises

The following subsections explain these functions and the estimation of parameters in more detail. In addition, the implementation of this new structure (modern vs traditional fuels) along with the implementation of multiple consumers, requires some changes to the assumptions for the distribution of the building energy across the different energy services (i.e., cooling, heating, and others in global GCAM). This allocation is based on different assumptions, which are translated into the "A44.share_serv_fuel" input file to be read by the data system. In order to ensure a more realistic behavior of the different services across different consumers, we have applied some modifications to the assumptions to allocate residential energy across the different services within each region. We acknowledge that some of these assumptions are arbitrary and/or convenient for calibration purposes. Future work ought to focus on these allocations, particularly for the development of studies with a regional focus.

For calibration, the model uses World Energy Balances by the International Energy Agency (IEA)[9], which provide residential energy consumption by fuel (not by service) from 1971 to 2015 (final calibration year). To allocate residential energy across the different services modelled by GCAM (cooling, heating, and others), we use alternative superregional data sources, summarized in the following table.

Table S2: Summary of data sources for allocation of residential energy across different services

Region	**Source**
USA	Dataset Building ENergy Demand, BEND[10] (detailed US building model)
Canada	IEA "30 years of energy use" report for residential space heating fuel breakdown
Western Europe	Odyssee Database
Japan	IEA Dataset
Korea	IEA Dataset + Korea Energy Consumption Survey of 2007
Southeast Asia	Expert consultation
India	Chaturvedi et al, 2014[11]
Australia_NZ	IEA dataset + Australian Bureau of Agricultural and Resource Economics (ABARE)
Africa	Expert consultation
Middle East	Expert consultation
Latin America	Expert consultation
Former Soviet Union	World Bank- Energy Efficiency in Russia: Untapped Reserves

China	China Energy Databook
Eastern Europe	Odyssee Database

Regarding the implementation of the multiple consumers, considering the lack of subnational, income-decile data for residential energy consumption, we use the functional forms, which were calibrated to average income data, to estimate the income decile-level consumption within regions. However, the total of the estimated income decile consumption levels may not exactly equal the national (regional) observed consumption. Thus, we calculate a bias-correction adder, equal to the difference between the estimated values and the observed values. In some cases, the bias correction may induce negative demand for some income deciles, so we need some additional adjustments to this bias-correction parameter for decile-level calibration and forward-looking analysis.

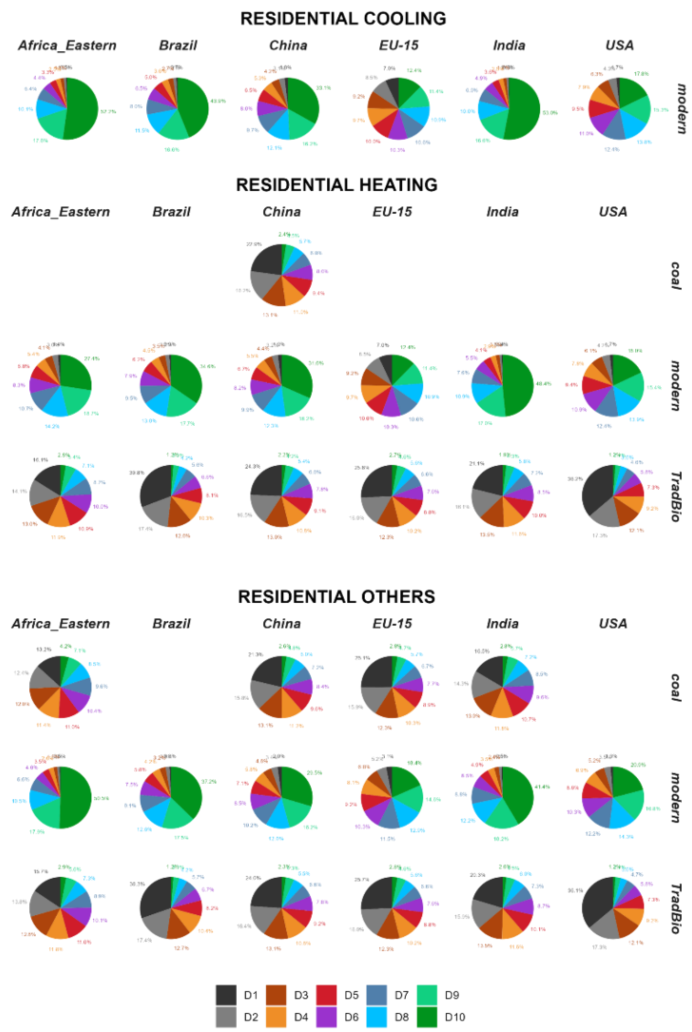
The approach, developed for each service, has 6 steps, which are described in more detail below:

1. Calculate e*stimated* decile-level consumption
2. Calculate total *estimated* regional consumption (sum of all deciles)
3. Calculate *estimated* income decile energy shares of total regional consumption
4. Compute difference between *observed* and *estimated* regional consumption (bias correction)
5. Apply bias correction to each decile
6. Adjust bias correction to avoid negative demands in some low-income deciles: phase bias correction out over three model periods

For steps 1-3, we first calculate the share of the service demand (*share*) for each decile (*i*) within each region (*r*) and service (*s*), by using the corresponding energy demand function (which is different for modern and traditional fuels). We calculate an "estimated" service level for the different deciles (ss,r,i), and divide it by the sum of the total estimated service to get the shares, as described in the following equation:

shares,r,i= ss,r,ii=110ss,r,i

These estimated shares are summarized in the following figure:



For steps 4-5, we calculate the difference between the observed $(s_{s,r})$ and the estimated value $(s_{s,r})$, which we define as the bias-correction parameter (bias-adder; *a*). This adder is necessary to avoid a sudden change in consumption levels between the calibration and first modeled periods. Note that the estimated value is the sum of all deciles:

$$s_{s,r} = \sum_{i=1}^{10} s_{s,r,i}$$

$$a_{s,r} = s_{s,r} - s_{s,r}$$

This bias correction parameter is calculated for each region and service. In order to ensure that all consumers are in the same path, this bias correction adder needs to be equally split across the different consumer groups:

$$a_{s,r,i} = \frac{a_{s,r}}{10}$$

However, using these "equally-split" bias-adder parameters ($a_{s,r,i}$) creates some (minor) negative demand values in the initial modeled periods (2020 and 2025) in some specific regions, so some additional adjustments are required. For that purpose, we first calculate the decile-level "observed" value, by allocating the region-level energy observed consumption across different deciles using the computed shares:

$$s_{s,r,i}=s_{s,r} * share_{s,r,i}$$

Then, we calculate a "decile-level" adder, which ensures that there are no negative demand values.

$$a'_{s,r,i}=s_{s,r,i}- s_{s,r,i}$$

While this bias correction parameter ($a'_{s,r,i}$), does not generate any negativities in the energy service demands, we need to ensure that all the consumers (deciles) are in the same path and all calibration parameters are equal, so that all consumers would demand the same energy amount at a similar income level. Therefore, we transition from decile-specific adders ($a'_{s,r,i}$) to the equally-split adders ($a_{s,r,i}$) in three model periods (which eliminate problems with negative demands), so that $a'_{s,r,i}$ is linearly interpolated to $a_{s,r,i}$ by 2030.

Given the differences between modern and traditional services, the following subsections provide more detailed information about the assumptions and functional forms associated with each type of services, which are used for estimating future residential energy demand.

Modern services: Adjusted satiation demand function

For the estimation of modern energy services, we make use of the satiation demand function (by default in GCAM), but with some changes. The main modification to make the satiation demand function flexible for multiple GCAM consumers, is to move the **estimation of the calibration parameters from the consumer level to the region level**, to ensure that all the consumers within each region will have the same functional form and will just "move through the x-axis" depending on their income level, with no other tuning parameters affecting the shape of their demand function. Taking residential heating as a representative example, estimated heating demand in region *r*, period *t*, and consumer *i* with the adjusted service function per unit of floorspace would have the following representation:

$$h_{t,r,i}=k_{h,r}*HDD_{t,r}* \eta_{r,t,i}*R_r+ \lambda_{h,r}*IG_{t,r,i}*1-exp-\ln 2 \mu_{h,r}* GDPpc_{t,r,i} P_{h t,r}+a_{h,r} \#$$

The function shows that the two calibration parameters (*k* and *μ*) are estimated at region level (they don't have the i sub-index). Note that the prices used in the estimation of the calibration parameters (P) are equal for all consumers due to lack of subnational data. Note that service price in the final base year (used for the calibration of parameters like the satiation impedance) is exogenously read in by the model DS (based on an "observed value" from a previous run).

Future model updates may imply that this exogenous value (slightly) differs from the "final calibration price" calculated within the model cpp files. In order to solve this issue, we include a "price-adjustment" parameter (PriceAdjustParam), which is computed in the cpp code and kept invariant during the whole-time horizon (see *building_service_function.cpp*). The price adjustment is just the difference between the read in price and the solved final calibration price.

This parameter will avoid the potential undesirable 2015-2020 jumps in modern service demand. Ideally, in the future, we should calculate the bias directly in the cpp code so that we do not read in any calibrated prices.

The bias-adder parameter (a) represents the difference between the observed and estimated values and is set to be equal across all the consumer groups as explained in the previous subsection.

Building shell efficiency (η) is also adjusted for the different regions and consumers to capture that richer regions and groups will demand more efficient households over time:

$$Eff_{r,t,i} = Eff_r * \left(\frac{GDPpc_{r,t,i}}{GDPpc_{r,t}}\right)^{PrElast}$$

Note that where Prelast represent price elasticity and is assumed to linearly decrease (become less elastic) from -0.1 to -0.07 over the century, which is not a change from the core.

Traditional fuels: Coal and traditional biomass (TradBio)

One important aspect to consider before analyzing the demand of traditional fuels, is the mapping of traditional biomass. In the core version of the model, there is an assumption that exogenously assigns traditional biomass to some regions but not to others (TradBio =1 in A_regions.csv). This is inconsistent with actual data (e.g., fuelwood for heating in the US is not currently counted as traditional biomass). We have modified this assumption and implemented that all the Primary Solid Biofuels reported by the IEA balances will be mapped into traditional biomass in GCAM in all the regions. Therefore, non-traditional biomass in GCAM ("commercial biomass") will be represented as the sum of the following IEA categories: charcoal, industrial waste, municipal waste, other primary biofuels, and other liquid biofuels. This implies that some (developed) regions (USA) will have no (commercial) biomass, but only traditional biomass in the building sector, which would represent firewood (as primary solid biofuel). This is summarized in the following table:

IEA Category	**Core**	**New version**
Primary solid biofuels	•Traditional biomass if TradBio = 1 in _A_regions.csv_ •Biomass if TradBio = 0 in _A_regions.csv_	Traditional Biomass
Charcoal	Biomass	Biomass

Industrial waste	Biomass	Biomass
Municipal waste	Biomass	Biomass
Other primary biofuels	Biomass	Biomass
Other liquid biofuels	Biomass	Biomass

In terms of implementation, future demand for traditional fuels does not follow the same path as modern sources. Historical data clearly shows that the consumption of these traditional fuels decreases as income rises. Therefore, we fit a functional form that represents the inverse of service affordability, measured as the income divided by the fuel/service prices. Using affordability instead of income allows us to capture some price dynamics that may have an impact on specific scenarios. For example, in a scenario with a higher price of coal, the phase-out of residential coal would be even faster, as would be expected in the real world. The functional form per unit of floorspace is described by the following equation:

$$TradDmnd_{r,t,s,f,i} = \frac{X_{s,f}}{\frac{GDPpc_{r,t,i}}{P_{r,t,s,f}} + Y_{s,f}} + a_{r,t,s,f,i}$$

Where

s is service,

f is fuel (coal, traditional biomass),

P is average service price,

X and *Y* are calibration parameters estimated from historical data, and

a is the bias-adder, which is calculated at region level and split equally across consumers.

So, demand for a traditional service (*s, heating or others),* fuel (*f, coal or TradBio*) in region r, period t, and for consumer group I, will be estimated based on per capita GDP, service price (P), two calibration parameters (X and Y), and a bias adder (a) that will be calculated at region level and split equally across consumers. In addition, as we do with floorspace, if there is a decrease in per capita GDP, the structure of the function implies an increase on traditional fuels, and switching back to these traditional fuels can be a bit odd as regions would be already using modern fuels. Therefore, we assume that the consumption of traditional biomass will not increase more than the value consumed in final calibration year:

$$TradDmnd_{r,t,f,s,i} = \min\left(\widehat{TradDmnd}_{r,t,f,s,i}, TradDmnd_{r,0,f,s,i}\right)$$

As explained, apart from the estimation of future traditional fuel demand, the functional form function is also used for the allocation of the observed regional base service across deciles (per each region and service). In summary, we need to generate two datasets:

- "Shares" to allocate observed service quantities across deciles (in the DS, L244.GenericShares and L244.ThermalShares)
- X and Y coefficients to be read by the cpp files to estimate future traditional fuel demand by region, period, service, and decile. Note that these coefficients will be equal for all deciles (to ensure all are in the same path) and will be constant over time.

Specifically, we follow the following steps:

1. The GCAM data-system generates a data frame with region-level service output by service for a range of historical years (1971-2015) (*L144.base_service_EJ_serv_fuel* in the GCAM data-system). This dataset is generated combining building energy use from IEA balances with the assumptions/workbook to allocate that energy across end-use services (adjusted from initial assumptions).
2. We add regional-average GDP_pc, and regional service prices to calculate affordability (GDP_pc/price) and fit:

   $$TradDmnd_{r,t,f,s} = \frac{X_f}{Afford_{r,t,f,s} + Y_f}$$

   The fit is developed at <u>REGIONAL LEVEL.</u> These estimated regional parameters produce values that are far from the observed amounts. This is because they don't adequately represent the behavior of the very low-income levels. Therefore, the estimated parameters need to be improved and re-fitted as explained in the following steps. However, these estimated (non-refitted) parameters are going to be used to calculate the shares that allocate observed regional service quantities across different deciles. Note that we don't fit the parameters "by service", because the estimation of heating creates some problems (singular gradient error when applying nls estimation).

3. We create L244.GenericShares and L244.ThermalShares and then adjust the base-service data frames (L244.GenericBaseService and L244.ThermalBaseService), so that they now include the subregional (decile-level) calibration data in historical years for each region and service.
4. As explained, we need the X and Y parameters to estimate future traditional fuel demand within the cpp files for future periods. Instead of reading the already fitted parameters, which create overestimation problems at low-income levels, we re-estimate the parameters in order to get a better fit using the following approach:
   a. Take the recently generated L244.GenericBaseService and L244.ThermalBaseService datasets, which now include historical service quantities for each region, service, <u>and decile</u>.
   b. Join the subregional GDP_pc data (I), generated combining region-level GDP_pc and subregional income shares (latest from Task 1.3). Also add the regional service prices to calculate decile-level affordability as:

      $$Afford_{r,t,f,s,i} = \frac{I_{r,t,i}}{P_{r,t,f,s}}$$

c. With decile-level affordability and service amounts, we can re-fit the parameters as:

$$TradDmnd_{r,t,f,s,i} = \frac{X_{f,s}}{Afford_{r,t,f,s,i} + Y_{f,s}}$$

The parameters estimated at income decile level provide much better estimates that are closer to the observed data. In addition, it does not create a problem if the estimation is performed at service level (s), so we do that and get fuel-and-service-specific X and Y parameters (x_heating, x_others, y_heating and y_others).

d. These new parameters are written to specific data frames (e.g., L244.GenericTradBioCoef or L244.ThermalTradBioCoef) that will be read by the model (cpp) for future projections.

e. Note that these parameters are going to be used to calculate the BiasAdder, which is calculated as the difference between the observed and the estimated value. Calculating the estimated value using these new re-fitted parameters makes sense because they are the ones that will be used for future model simulation.

OTHER CHANGES

EMISSIONS

The changes in the residential sector, particularly the distinction between modern and traditional fuels make that the emissions associated with the residential sector need to be changed accordingly to maintain consistency between the systems. Specifically, the following emissions need to be adjusted:

- Primary particulate matter (BC/OC) emissions: zchunk_L114.bcoc_en_R_S_T_Y.R
- Hydrofluorocarbons (HFCs) and Perfluorocarbon (PFC): zchunk_L142.pfc_R_S_T_Y.R
- Other nonCO2 emissions: zchunk_L201.en_nonco2.R
- Emission controls that graduate future emission factor for the residential technologies: zchunk_L151.ctrl_R_en_S_T.R
- MACC curves need to be adjusted to incorporate the updated residential sector with multiple consumers: zchunk_L252.MACC

Apart from these changes, all the input mapping files need to be adjusted to represent the added modern and traditional subsectors.

At this moment, it is assumed that all the technologies across all the consumer groups have equal emission factors. This may be a simplistic assumption, because it may be argued that lower deciles will use "less efficient" technologies that will emit more per unit of energy. Therefore, future work should focus on improving this representation, particularly for sector-specific or region-specific studies.

ADJUSTMENTS FOR GCAM-USA

The implementation of the changes in the global version of the model has some (minor) implications for GCAM-USA in terms of floorspace demand:

- The new calibration of the satiation impedance in the DS has some impact for future commercial floorspace demand in all states. In the core version of GCAM-USA, satiation impedance is assumed to be equal to 10.5 for all the states, and then an adder is used for "correcting" the differences across states (observed - estimated). The new calibration of the impedance in the DS allows the use of exactly the same procedure in GCAM-USA, which would improve the consistency across the two model versions. In any case, the change in commercial floorspace is relatively small (up to ~10%).

The changes in floorspace demand will have some subsequent effects in service output and building energy demand, but they are relatively small. However, while everything works adequately in GCAM-USA, the changes or adjustments applied in the global version have not been applied to GCAM-USA and have been postponed for a future CMP that may include the multiple consumers in GCAM-USA. At this point (after this CMP), in GCAM-USA:

- There is no distinction between modern and traditional services (and the corresponding demand functions)
- Biomass consumed in the residential sector is "non-traditional" or "commercial" biomass. It has not been re-mapped to traditional biomass as in the global version.

Prior to the implementation of multiple consumers in the residential sector in GCAM-USA, these changes should be completed for a more adequate representation.

The distinction between modern and traditional fuels requires several changes in different files.

First, the service-sector mapping needs to be redefined in the corresponding mapping files, such as *EIA_AEO_services* or *calibrated_techs_bld_usa*. The new names of the modern/traditional services should also be included in the "extada/gcam-usa/A44*" files, which represent the sector/subsector/technology structure and different parameters such as shareweights or logit exponents for the building sector within GCAM-USA. Likewise, the implementation would very likely require some modifications in the main chunks associated with the building sector in GCAM-USA, which include *zchunk_LA144.Residential*, *zchunk_LA144.Commercial* (if the commercial sector is revisited), *zchunk_LA142.Building_USA*, *zchunk_L244.building_USA*, and *zchunk_batch_building_USA_xml*. Finally, some additional modifications would be required in the input csv files, and chunks to represent nonCO2 emissions in the building sector (e.g., *zchunk_L273.en_ghg_emissions_USA*).

The re-mapping of the biomass categories would also be an improvement for GCAM-USA and will help to keep consistency between the two versions of the model. The EIA reports two biomass fuels that are consumed within the buildings sector: wood (WD) and waste (WS). In order to be consistent with global GCAM, wood (WD) should be mapped to traditional biomass and waste (WS) to ("non-traditional") biomass. Indeed, the EIA energy information shows that all the biomass consumed by the residential sector in all the US-states is wood (WD) while waste (WS) is only used in the commercial sector. This is directly consistent with our assumption in global GCAM for the USA, where all biomass in the residential sector is mapped to "traditional biomass" (representing "fuelwood"). In order to develop this remapping, the first step is to adjust the corresponding chunks to change the mapping defined in "extdata/energy/mappings/enduse_fuel_aggregation" (also read in by GCAM-USA chunks). In addition, there are two adjustments that would need to be rolled-back or adjusted:

- In *zchunk_LA131.enduse_noEFW_USA*, there is an adjustment to keep the existing mapping between biomass and traditional biomass that would not be needed anymore once the re-mapping is implemented (from line 109)
- In *zchunk_L244.building_USA*, there is a data table (L210.DeleteRsrcTradBio_gcamusa) that was created to avoid a warning message the model printed. This file will not be needed after the mapping, and could create some problems, so needs to be deleted.

Note that the EIA provides a detailed time series with residential (traditional) biomass consumption: {+}https://www.eia.gov/totalenergy/data/browser/index.php?tbl=T02.02#/?f=A&start=1949&end=2021&charted=9+ . If desired, the demand for traditional biomass in the US could differ from the global demand function. The new fitted/calibrated demand form should be written in the cpp files (+added to the function manager), the use of the new function for the US-states should be indicated in .input/gcamdata/inst/extdata/gcam-usa/A44.demandFn_serv.csv

ADJUSTMENTS FOR SSP NARRATIVES

In the context of the simulation of the alternative Shared Socioeconomic Pathway scenarios (SSPs), the core model includes two building-specific assumptions that differ across the narratives to capture some inter-scenario dynamics. With the implementation of multiple consumer groups, these assumptions need to be revisited to check if/how they can still be implemented.

First, already in the core, it is assumed that the floorspace satiation level will be higher in the SSP narratives with higher economic growth. With the changes to residential floorspace demand (CMP 346), only commercial floorspace has an exogenously defined satiation value (residential uses the Gompertz function). Considering that the implementation of multiple consumers does not affect the commercial sector, this assumption is not modified.

In addition, the core model assumes that the multiplier that represents the "assumed increase in demand of generic services in the USA" slightly varies across SSPs (only from -10% to +10%). The effects of this variation are quite small and will be completely dominated by other demand drivers that more largely differ across the SSP narratives (i.e., population or GDP). Considering the additional complexity of keeping this assumption in a model version in which the calibration of the building-related parameters is developed in the data-system, we have decided to not include the SSP-specific multipliers in this model version.

ADJUSTMENT FOR ELECTRIC HEATING IN CHINA (+additional regions)

In China, the share of electric heating within the heating sector is really low in the final calibration year (<1%). In the core version of the model, in a decarbonization scenario, the solution to mitigate emissions in the building sector in China was to "go back" to traditional biomass, even though this may not be a realistic response due to some factors or dynamics that cannot be captured with the model (e.g., household air pollution). After this implementation, given that traditional biomass has been modeled independently, the solution the model finds in a decarbonization scenario in China is to significantly reduce the total service demand, because there are no clean alternatives. Therefore, to correct this unrealistic dynamic, we have adjusted the shareweight of electric heating in China, so that in 2050, the share of electricity within the heating sector achieves the global average (~25%).

This approach has been extended to other regions beyond China to ensure consistent behavior. The adjustment has been expanded to regions where the electric heating production per unit of floorspace is lower than 1e-03 GJ/m2.

Adjustments of Emission factors of residential traditional biomass

After disaggregating traditional biomass in this CMP, we observe very high traditional biomass emission factors (for heating and others/cooking) throughout the century resulting in continued high air pollutant emissions. However, we understand that regions will be using more efficient (i.e., less polluting) technologies over time, so the model should have emission factors that decrease over time.

The SSP-specific files (e.g., *ssp2_emissions_factors.xml*) included in GCAM have decreasing emission factors for different techs/fuels for different pollutants. We have extracted the decreasing emission factors for traditional biomass of the ssp2 narrative and create an “add-on” file (but generated by the DS for consistency) that will be loaded by default in the configuration files (“*ssp2_tradBio_emissions_factors.xml”*). Note that if the user runs other SSP scenarios, the EFs will be overwritten by the SSP-specific files (*ssp15_emissions_factors.xml* and *ssp34_emissions_factors*).

Regarding the **source**, these decreasing emission factors come from GAINS (Rao et al 2017), which includes different EFs for different regions for different scenarios with different ambition, such as the “current legislation” (CLE) or the “maximum feasible technical reduction” (MFTR). Then, there is an input file in GCAM (A61_emfact_rules) that sets the “rules by SSP and regional groupings on how to apply emissions factors from GAINS scenarios”.

While it is beyond the scope of this CMP, future work should focus on distinguishing between heating and cooking/others and, in the longer term, have decile-specific emission factors (not just EFs, but tech efficiencies or costs).

TABLE OF CHANGES

File	**Change**	**Area (Flsp/Energy/Other /All)**
Changes in the CPP files		
building_node_input.h	Add population density in base year (TotDens)	Floorspace
building_service_input.h	Read in parameters for the residential energy demand functions (generic services)	Residential energy
thermal_building_service_input.h	Read in parameters for the residential energy demand functions (thermal services)	Residential energy
satiation_demand_function.h	Correct parsing of satiation impedance	Floorspace (commercial) and Residential energy
building_service_function.h	Add structure that works with multiple consumers Revisions to define functions for traditional services	Residential energy
building_node_input.cpp	Add population density (TotDens)	Floorspace
building_service_input.cpp	Add input coefficients (generic services)	Residential energy
thermal_building_service_input.cpp	Add input coefficients (thermal services)	Residential energy
satiation_demand_function.cpp	Correct parsing of satiation impedance	Floorspace (commercial) and Residential energy

building_gompertz_function.cpp	Add population density (TotDens) Add income effect to avoid "building destruction"	Floorspace
building_service_function.cpp	Adjust estimations for multiple consumers Add demand functions for traditional services	Residential energy
Changes in the Data System chunks		
generate_package_data.R	Add new level2 data names	Floorspace and Residential energy
GCAM_DATA_MAP.rda	Modifications to system data	All
PREBUILT_DATA.rda		
sysdata.rda		
zchunk_LA144.building_det_flsp.R	Add structure/calibration for multiple consumers	Floorspace
zchunk_LA144.building_det_en.R	Add prices + adjustments for calibration	Residential energy
zchunk_L244.building_det.R	Major changes for multiple consumers	Floorspace and Residential energy
zchunk_batch_building_det_xml.R	Write parameters into xml (to be read by cpp files)	Floorspace and Residential energy
zchunk_batch_building_SSP_xml.R	Delete satiation multipliers (just keep SSP-specific satiation values for comm floorspace)	Floorspace (comm)
zchunk_L114.bcoc_en_R_S_T_Y.R	Adjust BC/OC for multiple consumers	Other: Emissions
zchunk_L141.hfc_R_S_T_Y.R	Adjust HFCs for multiple consumers	Other: Emissions
zchunk_L142.pfc_R_S_T_Y.R	Adjust PFC for multiple consumers	Other: Emissions
zchunk_L151.ctrl_R_en_S_T.R	Adjust emission controls for multiple consumers	Other: Emissions
zchunk_L201.en_nonco2.R	Adjust nonCO2 calculations for multiple consumers	Other: Emissions

zchunk_L252.MACC	Adjust MACC curves (HFCs from cooling) for multiple consumers	Other: Emissions
zchunk_LA131.enduse_noEFW_USA.R	Adjustments for GCAM-USA: Adjustments associated with TradBio remapping in the global version Calibrate parameters in the DS to be consistent with the adjusted demand functions in the cpp files	Other: GCAM-USA
zchunk_LA142.Building_USA.R		
zchunk_LA144.building_det_en.R		
zchunk_L244.building_USA.R		
module_energy_L244.building_det.Rd	Changes in doc files	Residential energy
module_energy_LA144.building_det_en.Rd		Residential energy
input/gcamdata/R/zchunk_L251.en_ssp_nonco2.R	adjusted emissions factors because of the changed definition of traditional biomass	Emissions factors
input/gcamdata/R/zchunk_batch_ssp2_emissions_factors_xml.R	adjusted emissions factors because of the changed definition of traditional biomass	Emissions factors
Changes in the extdata csv files		
income_shares.csv	Add income distribution trajectories for deciles	All (socioeconomics)
income_shares_quintiles.csv	Add income distribution trajectories for quintiles (not used unless specified)	All (socioeconomics)
ModelInterface_headers.txt	New Model Interface headers	Floorspace and Residential energy
A44.CalPrice_bld	Add prices for calibration in the DS	Residential energy
A44.CalPrice_service_gcamusa	Add prices for calibration in the DS – GCAM-USA	Other: GCAM-USA

A44.share_serv_fuel.csv	Add modern/trad fuels and adjust energy-to-service allocation	Residential energy
A44.cost_efficiency.csv	Adjustments to distinguish between modern and traditional fuels (energy)	Residential energy
A44.demand_satiation_mult.csv		
A44.demand_satiation_mult_SSPs.csv (check to delete)		
A44.fuelprefElasticity.csv		
A44.globaltech_shrwt.csv		
A44.internal_gains.csv		
A44.sector.csv		
A44.subsector_interp.csv		
A44.subsector_logit.csv		
A44.subsector_shrwt.csv		
A44.USA_TechChange.csv		
calibrated_techs_bld_det.csv		
A44.fuelprefElasticity_SSP15.csv	Delete SSP-specific FuelPrefElasticity	Residential energy
A44.fuelprefElasticity_SSP3.csv		
A44.fuelprefElasticity_SSP4.csv		
CEDS_sector_tech_combustion.csv	Adjustments to distinguish between modern and traditional fuels (emissions)	Other: Emissions
CEDS_sector_tech_combustion_revised.csv		
CEDS_sector_tech_proc.csv		
CEDS_sector_tech_proc_revised.csv		
GCAM_sector_tech.csv		
GCAM_sector_tech_Revised.csv		
A51.max_reduction.csv		

A51.min_coeff.csv		
A51.steepness.csv		
GCAM_EPA_CH4N2O_energy_map.csv		
gcam_fgas_tech.csv		
Changes in queries		
subregional population	Population for consumer group i	Socioeconomics
subregional income	Per capita GDP for consumer group i	Socioeconomics
building commercial floorspace satiation impedance	Satiation impedance, which is the calibration parameter of the satiation demand function used for commercial floorspace demand	Floorspace
building residential floorspace regional bias adder	Regional bias adder, calculated as the difference between observed and estimated floorspace values in final calibration year. Note that it only takes non-zero values in USA and China (regions with observed floorspace post 2005)	Floorspace
building satiation impedance by energy service	Satiation impedance, used in commercial energy demand, and residential modern service demand (satiation demand function)	Building energy
building residential floorspace regional bias adder	Regional bias adder, calculated as the difference between observed and estimated values in final calibration year.	Building energy

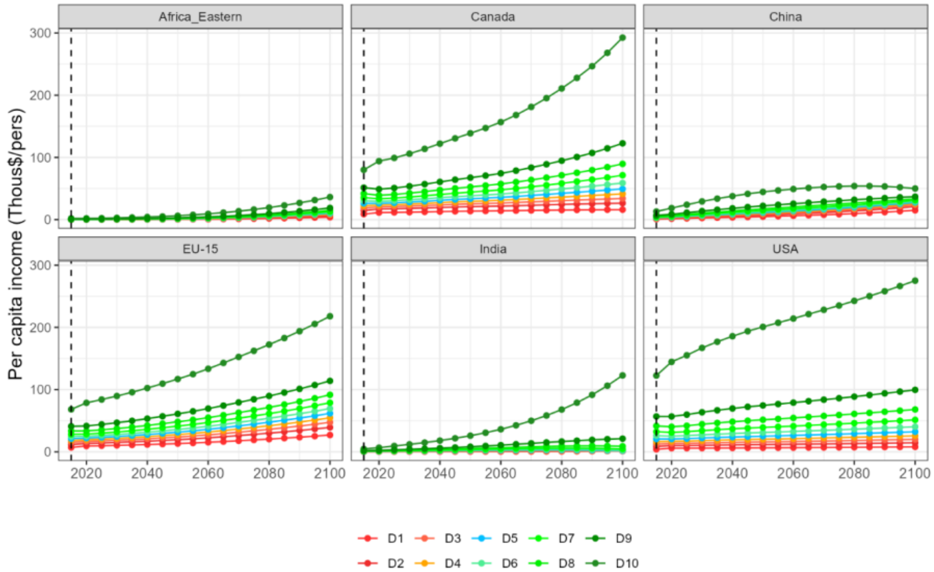
Validation

Income distribution pathways

As explained, income distribution trajectories are exogenously read by the model DS as an input csv file, which will be substituted by a new file generated within the model DS in a future CMP.

However, these income distribution trajectories are essential for understanding the floorspace and residential energy projections developed during this CMP.

Figure 1: Per capita income per period, region, and consumer group (Thous$/pers) (Income shares: SSP2).



Floorspace

RESIDENTIAL

Figure 2: Per capita residential floorspace per period, region, and consumer group (m2/pers) (Income shares: SSP2). The dashed black line represents per capita floorspace in the current Core.

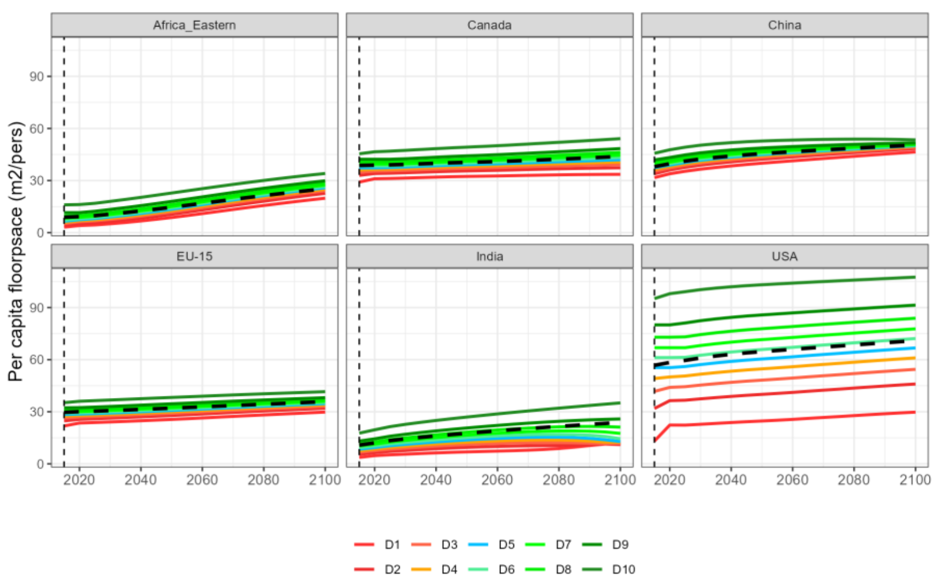
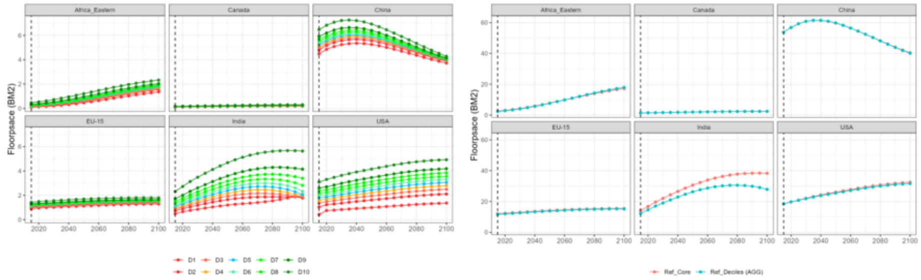
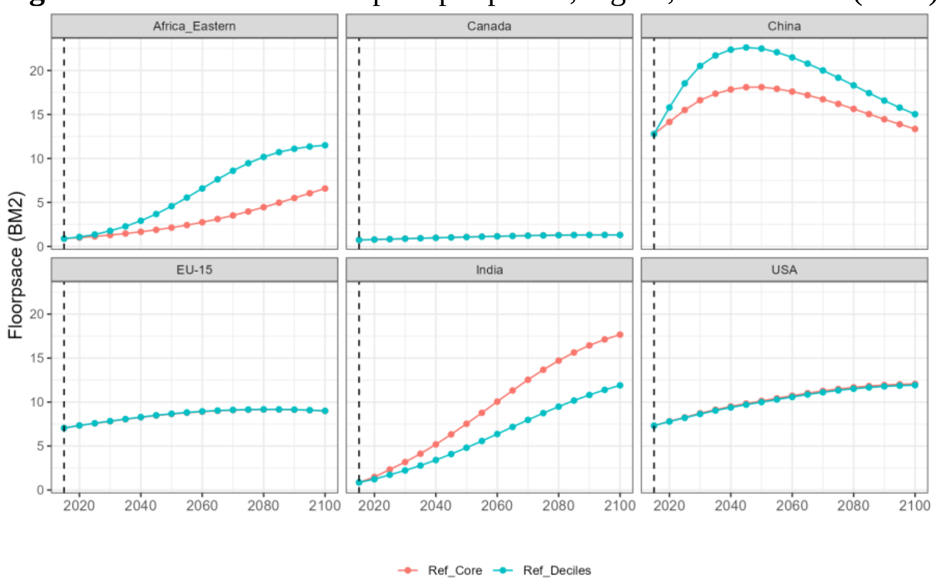


Figure 3: Residential floorspace per period, region, and consumer group (left panel), and comparison to the current core (right panel) (BM2) (Income shares: SSP2).



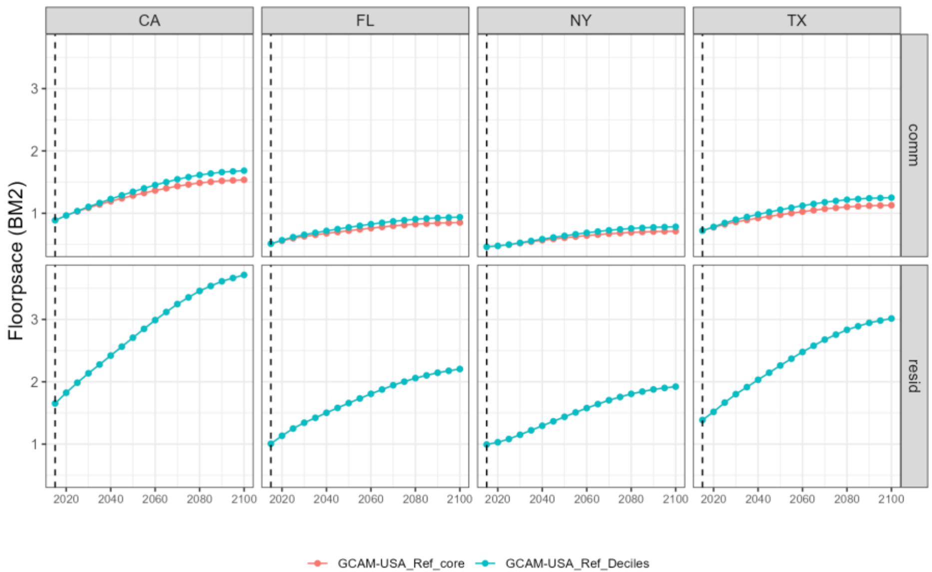
COMMERCIAL

Figure 4: Commercial floorspace per period, region, and scenario (BM2).



GCAM-USA

Figure 5: Residential and commercial floorspace per period, state, and scenario (BM2)



Service Output

Figure 6: Residential service output per period, region, service, and consumer group (EJ) (Income shares: SSP2)

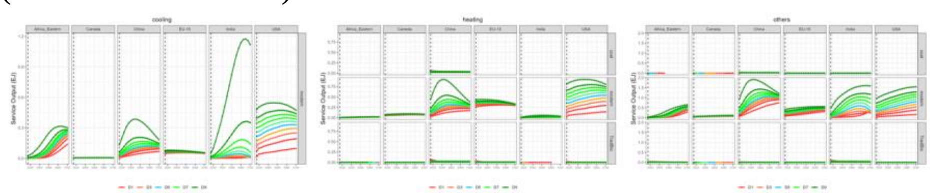
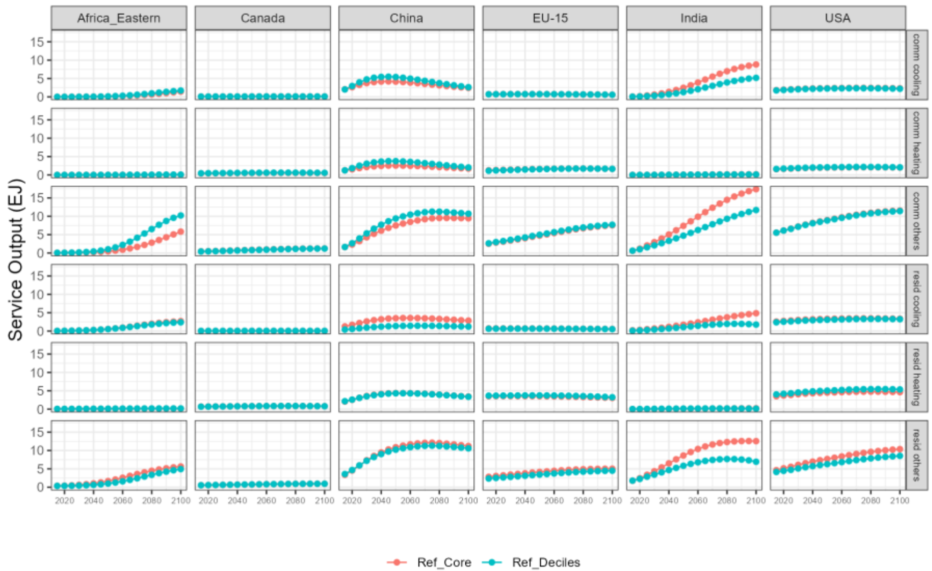
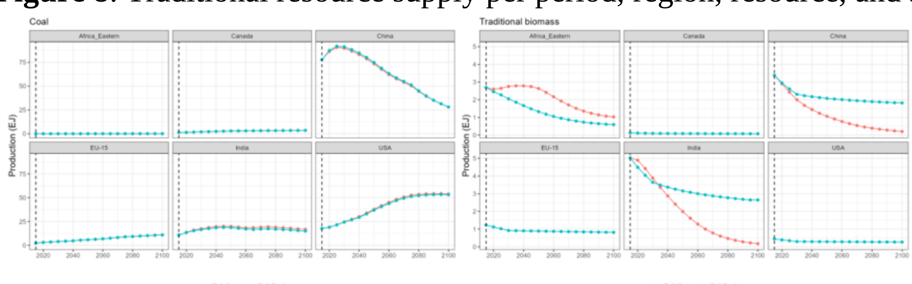


Figure 7: Residential and commercial service output per period, region, service, and scenario (EJ)



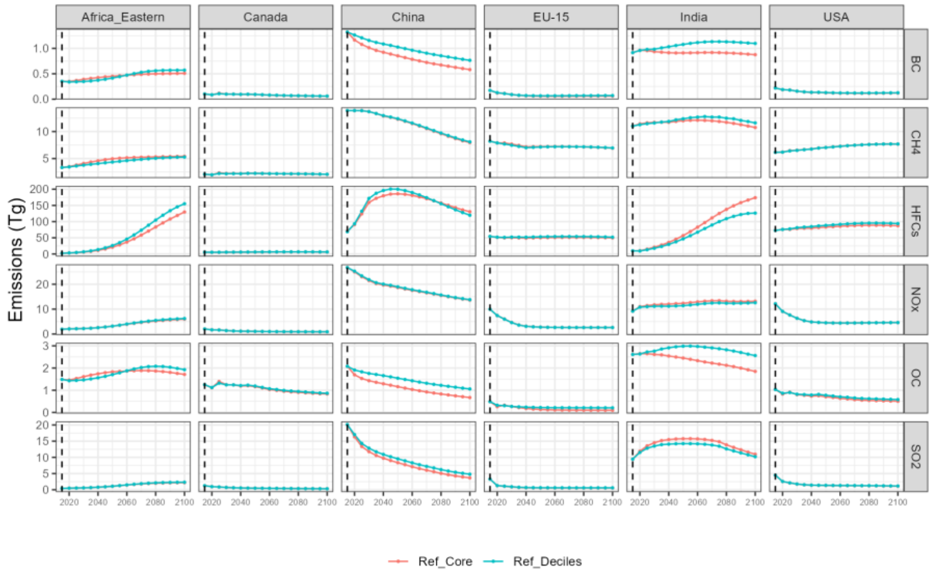
Traditional resource supply

Figure 8: Traditional resource supply per period, region, resource, and scenario (EJ)



Emissions

Figure 9: NonCO2 emissions per period, region, gas, and scenario (Tg)



EXTRA: Service output demand for regions heavily reliant on traditional biomass

In some regions, such as Africa_Eastern or Africa_Western, according to observed data, traditional biomass is now the main energy source for residential energy services. In our results, we see that, as the average GDPpc in these regions rises over time, their transition from traditional to modern services creates some unique results, particularly for "non-thermal" (other) services such as cooking or lighting, that are analyzed in this subsection.

As indicated above, region-level traditional and modern service output is allocated across the different deciles based on the application of the corresponding demand functions in the base year (using pre-computed subregional income shares). Considering the large use of traditional biomass in the aforementioned regions, especially by lower deciles, if we aggregate the decile-level non-thermal service demand from traditional biomass and modern sources, and estimate which percentage of the demand is allocated to each decile, we obtain the following results (example for Africa Eastern).

Table 1: Demand for non-thermal services in Africa Eastern in final calibration period (2015)

	Non-thermal service demand (EJ)			Percentages per decile over total (%)		
decile	TradBio	Modern	Total	TradBio	Modern	Total

d1	0.0408	0.0004	0.0413	14.85%	0.55%	**11.61%**
d2	0.0368	0.0010	0.0379	13.40%	1.28%	**10.65%**
d3	0.0346	0.0015	0.0361	12.58%	1.87%	**10.16%**
d4	0.0323	0.0021	0.0345	11.76%	2.62%	**9.69%**
d5	0.0303	0.0028	0.0331	11.02%	3.47%	**9.31%**
d6	0.0281	0.0037	0.0318	10.23%	4.59%	**8.95%**
d7	0.0251	0.0053	0.0305	9.14%	6.63%	**8.57%**
d8	0.0212	0.0084	0.0297	7.72%	10.49%	**8.35%**
d9	0.0167	0.0144	0.0311	6.06%	17.95%	**8.75%**
d10	0.0089	0.0407	0.0496	3.25%	50.54%	**13.95%**
TOT	0.2751	0.0805	0.3556	100.00%	100.00%	100.00%

These results, which have been calculated using the functions calibrated to historical data, show that non-thermal service demand may not directly increase with income when traditional biomass accounts for a large share of the service demand. The data in the final calibration year shows that the share of demand allocated to lower deciles will be larger than the share for the mid-upper deciles. While modern energy service demand is largely driven by income, existing literature and energy balances show that households rapidly get rid of traditional fuels for cooking and/or lighting as their income slightly rises, so total service demand may temporarily decrease when the households transition from traditional to modern fuels. This is also observed in the Food Demand model, when people transition from staples to non-staple calories (CMP 303).

An important point is that we are using a single efficiency value for all traditional biomass technologies across deciles. Service output in calibrated years is calculated as residential energy from the IEA balances multiplied by the efficiency. Having decile-specific efficiency parameters, particularly for biomass-related technologies, would improve this dynamic. While there is no global harmonized data on technology-level efficiencies to be implemented at this point, this may be an important aspect to revise in regional residential-focused studies. We also note that there are other externalities not captured by the model, that also can also justify the rapid decrease of traditional fuels as income increases. For example, collection of traditional biomass implies a physical effort and substantial time that is not used for labor (opportunity cost). Also, indoor solid fuel burning is a major risk for human health, particularly in the Global South, and several institutions are incentivizing households to reduce their use. In summary, all these factors can make that households rapidly decrease their traditional fuel consumption, and substitute it with modern energy sources, even though the total service demand temporarily decreases in the short-term.

Finally, we note that the income effect we see across deciles in the final calibration year is also observed in the near-term for lower deciles, which temporarily reduce their overall service demand until it can be completely substituted with modern sources (example for Africa Eastern).

Figure 10: Service output per period, and decile for Africa Eastern. Note the D5 has been included for comparison purposes.