



Editing Tools for Reconciling Housing Characteristics and Energy Billing Data

FCSM 2023

L. Kaili Diamond and S. Grace Deng

October 25, 2023 | College Park, MD

The analysis and conclusions contained in this presentation are those of the authors and do not represent the official position of the U.S. Energy Information Administration or the U.S. Department of Energy.

Residential Energy Consumption Survey (RECS) overview

- RECS is a nationally-representative study of household energy use and energy-related characteristics
- First conducted in 1978
- The most recent iteration collected information on energy-related characteristics for the housing unit, usage patterns, and household demographics for the 2020 calendar year
- RECS consists of two parts
 - The Household Survey (HS) is a sample survey that collects energy-related characteristics and household demographics for primary residences
 - The Energy Supplier Survey (ESS) collects energy consumption and expenditure data from energy suppliers for the respondents to the HS. Fuels reported include electricity, natural gas, fuel oil, and propane

RECS editing overview, 2020

- 2020 RECS Data:
 - Household Survey – 18,525 completed surveys between September 2020 and April 2021
 - Energy Supplier Survey – 27,729 completed surveys between July 2021 and February 2022 (16,894 electricity, 9,087 natural gas, 674 fuel oil, and 1,074 propane)
- Each survey was edited independently for internal data errors and inconsistencies.
 - Household Survey had 18,496 respondents after removing out of scope cases
 - ESS had 16,706 electricity, 9,000 natural gas, 666 fuel oil, and 1,063 propane bills after individual editing
- Final data editing phase is Reconciliation of the two data sources
 - We compare the expected household energy consumption based on household survey responses to the actual energy consumption provided by suppliers
 - Seemingly accurate housing characteristics can be incompatible with seemingly accurate supplier data.

Reconciliation editing overview, 2020

- 3 types of edits:
 - ESS edits: Identify and remove invalid or incorrect ESS consumption data
 - Batch edits: updates made programmatically to cases where the failure of an edit had known outcome
 - Example: reported NG water heating but ESS NG consumption is too low to support water heating
 - Household characteristic edits: Manually review and edit housing characteristics based on inconsistencies between bills and preliminary engineering models based on HS responses
 - Example: HS respondent claimed one fuel as primary heating fuel but consumption for that fuel obtained during ESS does not indicate seasonality associated with space heating
- 39 total reconciliation edits including ESS, HS, and Batch edits
- Editing guidelines were created to reduce editing bias and enhance repeatability

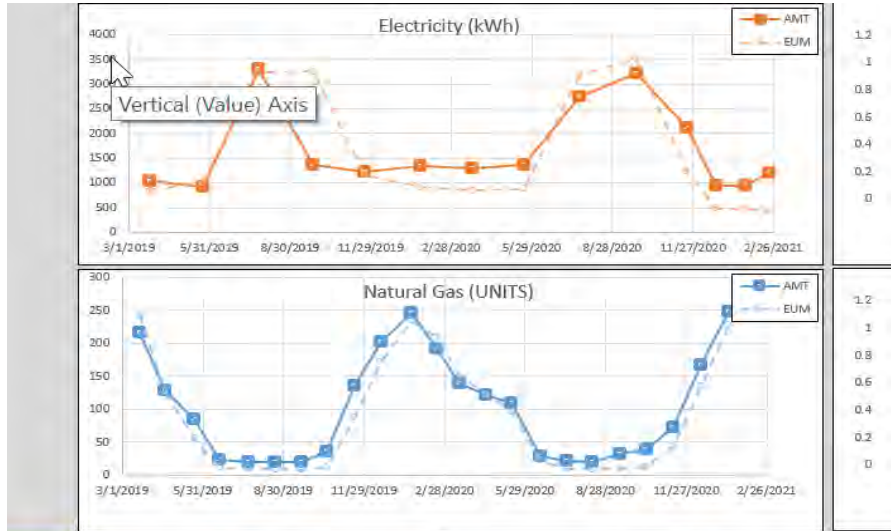
Reconciliation Editing Methodology

- Each edit was listed on a separate spreadsheet tab. Columns were created on each tab for all related variable changes
- Edits were performed in a hierarchical order. Editing programs were re-run periodically to reduce edit failure redundancy
- Analysts checked out cases and determined if they should be changed. Cases without a clearly defined resolution were discussed as a group
- A dashboard was created that showed all supplier bills for all fuels and the preliminary engineering end-use modeled estimates
- Supplier data were deleted in cases where it seemed inaccurate. For example, consumption appeared to be for a whole apartment building rather than the sampled unit

Engineering end-use models (EUM) and dashboard overview

- Expected consumption was derived from engineering based end-use models using household survey responses
 - End-use models rely on responses to the household survey and attached location and weather data
 - End-use models predict consumption for many devices in the home
- For example, we can estimate the electricity consumption for a home's main refrigerator based on characteristics such as the size, age, and style (French door, top freezer, existence of in-door water or ice, etc.) of the refrigerator
- A dashboard was created to compare end-use modeled consumption to actual energy consumption provided by the energy suppliers
- Only obvious inconsistencies between the two data sources led to changes

Reconciliation dashboard tool, consistent case

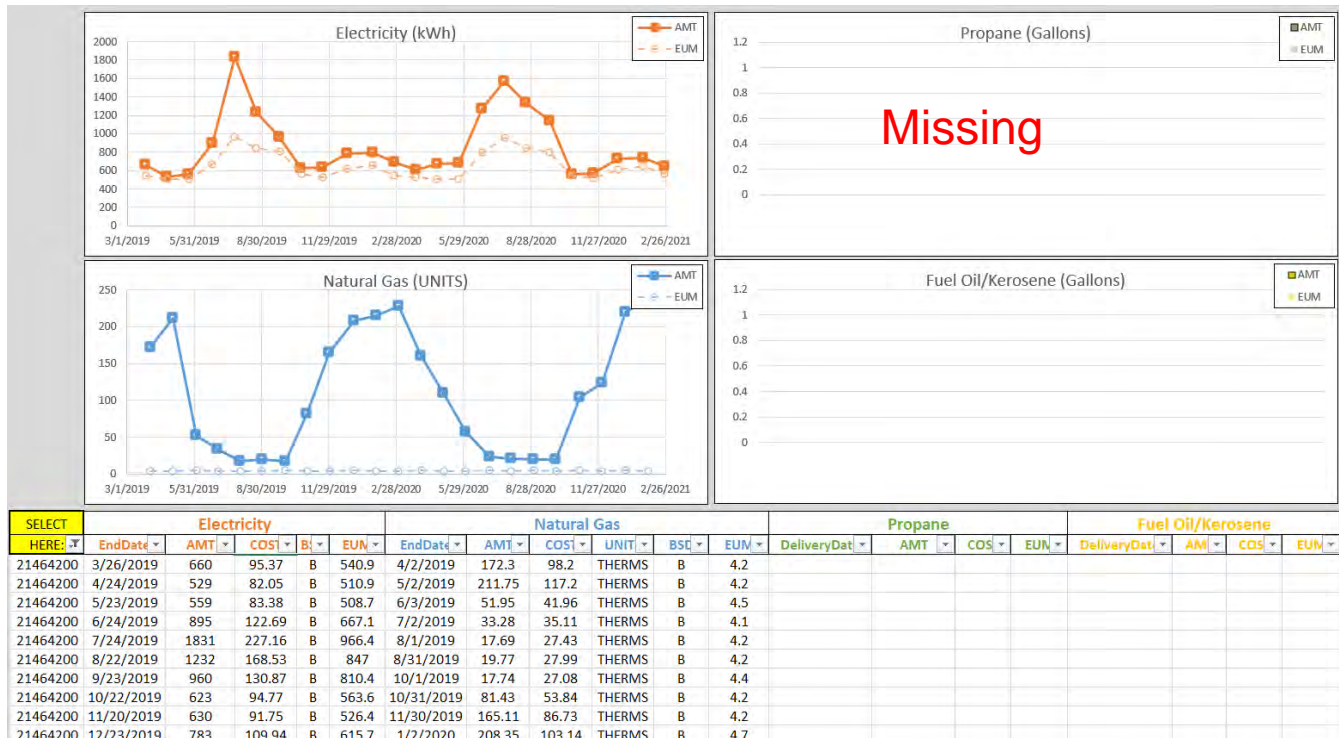


- Expected consumption (dashed line) closely matches reported energy consumption from suppliers (solid line)
- Expected seasonality from space heating shown in Natural Gas (NG)
- No changes to fuel or equipment for the case

Note: In this case, EL consumption data from the supplier was reported on a bimonthly basis.

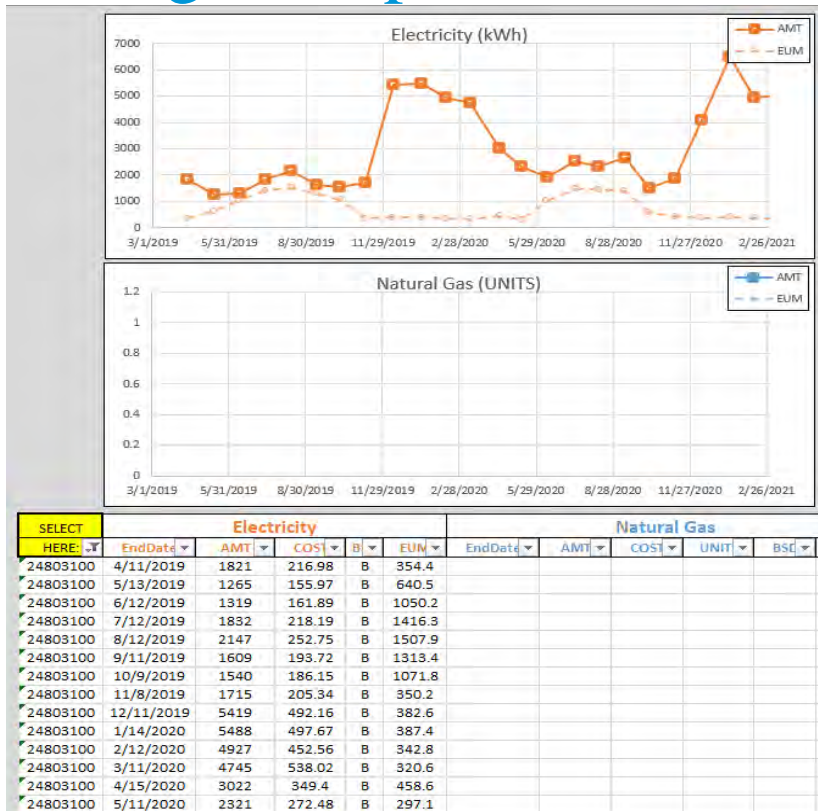
SELECT	Electricity					Natural Gas					
HERE:	EndDat	AMT	CO2	B	EUM	EndDat	AMT	CO2	UNIT	BS	EUM
13108600	3/25/2019	1037	212.75	B	848.2	3/20/2019	216	216.13	THERMS	B	242.4
13108600	5/24/2019	921	186.69	B	1014.2	4/18/2019	128	158.54	THERMS	B	127.1
13108600	7/25/2019	3308	674.14	B	3238.8	5/21/2019	85	132.13	THERMS	B	55.8
13108600	9/25/2019	1360	291.74	B	3258.7	6/19/2019	24	62.74	THERMS	B	11.7
13108600	11/23/2019	1219	254.66	B	1165.1	7/22/2019	20	54.47	THERMS	B	10.2
13108600	1/24/2020	1340	279.93	B	916.4	8/21/2019	20	52.85	THERMS	B	8.8
13108600	3/23/2020	1290	259.65	B	843.3	9/19/2019	20	52.39	THERMS	B	8.4
13108600	5/21/2020	1364	278.55	B	883.5	10/18/2019	36	83.85	THERMS	B	12.6
13108600	7/22/2020	2741	599.27	B	3184	11/18/2019	135	161.26	THERMS	B	87
13108600	9/24/2020	3220	693.95	B	3501.8	12/18/2019	203	211.83	THERMS	B	172.2
13108600	11/20/2020	2121	452.04	B	1236.5	1/21/2020	247	260.3	THERMS	B	234.5

Dashboard tool, potential natural gas vs propane reporting issue example



- Propane reported for main heating fuel
- Modeled use shows seemingly accurate EL data but shows unaccounted for NG use with seasonality consistent with NG space heating
- Change main heating fuel to NG

Dashboard tool, potential natural gas vs electric space heating example



- Reported main space heating fuel was natural gas with wood for back-up, but no natural gas use data was available
- ESS data shows seasonality consistent with electric space heating
- Main heating fuel changed to electricity

Reconciliation editing outcomes

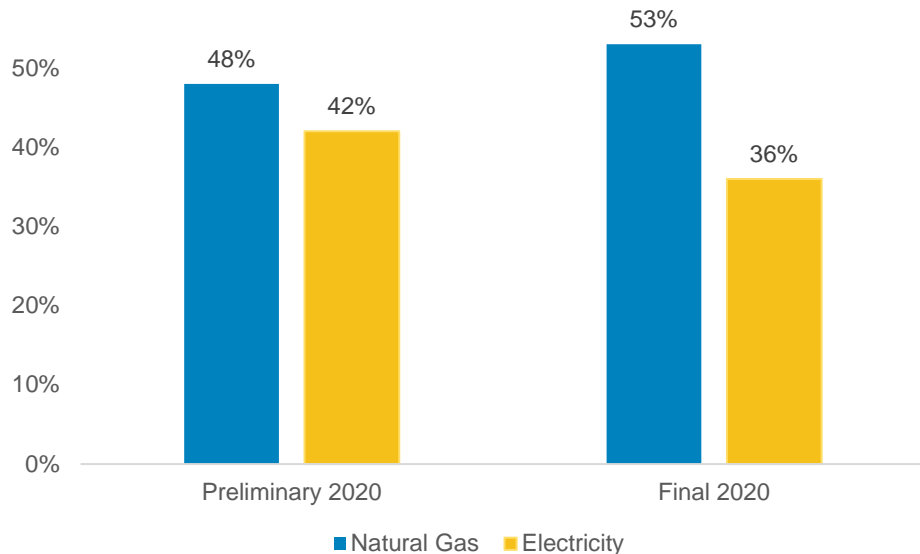
- All reconciliation edits:
 - 4631 (25%) cases failed at least one reconciliation edit
 - 2917 (16%) cases required at least one data change
- HS reconciliation edits:
 - 2958 (16%) cases failed at least one HS edit
 - 1765 (10%) cases required at least one household characteristic change
- Changes to main heating fuel tended to be from electricity to natural gas
- 10% of single family and 6% of required at least one analyst edit to housing characteristics

Results of reconciliation, data changes

- Most common variables changes:
 - 1,203 (7%) Main space heating fuel
 - 931 (5%) Main water heating fuel
 - 453 (2%) Main space heating equipment
 - 230 (1%) Backup fuel for central heat pump
 - 186 (1%) Secondary space heating fuel
 - 181 (1%) Secondary space heating equipment
- Space heating equipment only changed when related fuel was changed and the reported equipment was not compatible with identified fuel
 - For example, if main space heating equipment was reported as a built-in electric heater but reconciliation found that space heating must have been natural gas

Impact of main space heating fuel reconciliation edits

Percent of homes using NG or EL for main space heating fuel, weighted estimates for homes that heat

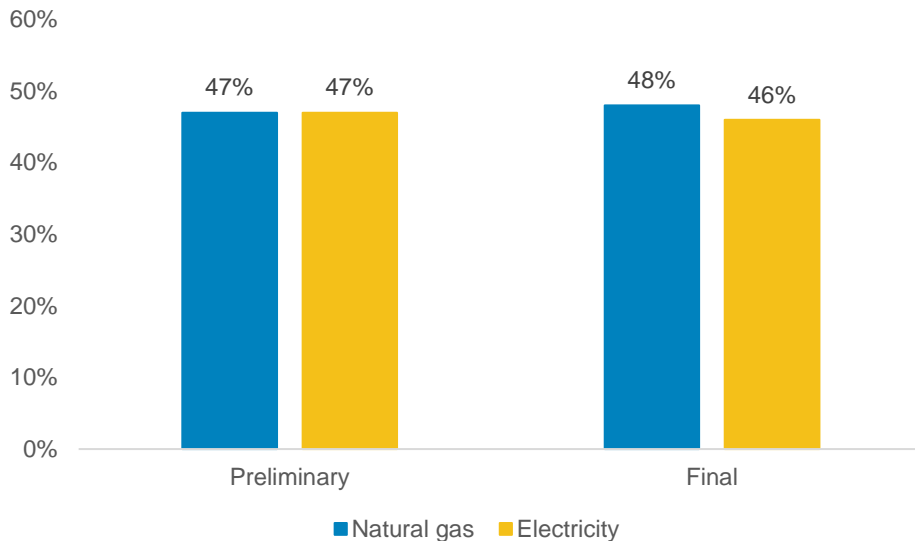


- Changes tended to be from electricity to natural gas
- Changing main space heating fuel required checking that the reported heating equipment is compatible with the new fuel
- 1,203 changes in fuel led to 453 equipment changes

Data source: U.S. Energy Information Administration, 2020 Residential Energy Consumption Survey

Impact of main water heating fuel edits

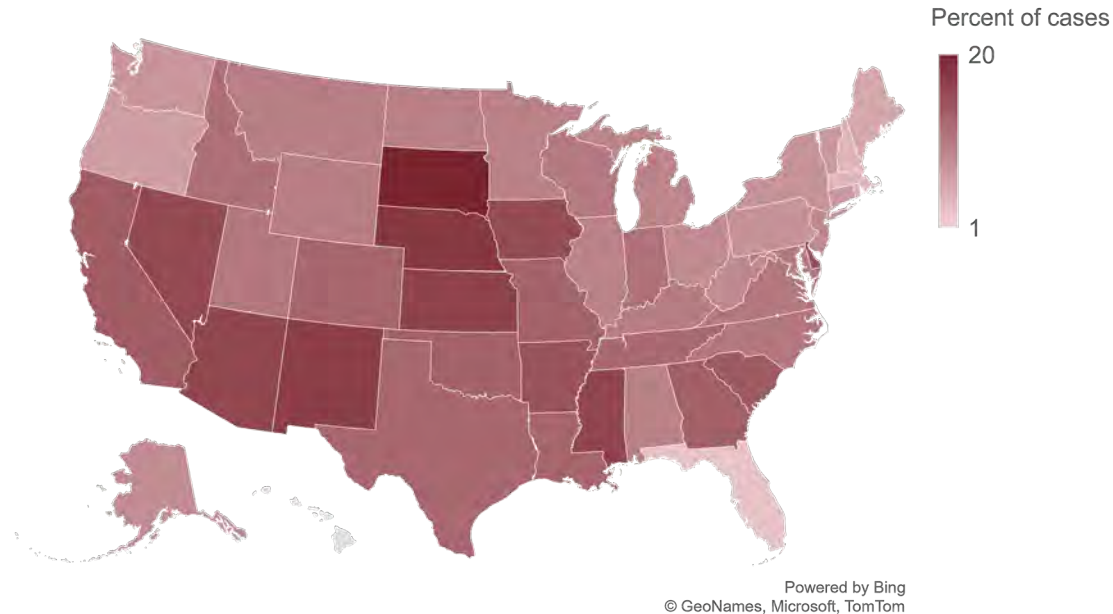
Preliminary versus final distribution of main water heating fuel, weighted estimates for most changed fuel types types



- Most main water heating fuel edits were from electricity to natural gas
- No change in equipment was necessary for these edits
- Energy consumption for water heating has seasonality, though less than space heating

Data source: U.S. Energy Information Administration, 2020 Residential Energy Consumption Survey

Percent of cases requiring a data change during reconciliation, by state



Note: Percents calculated based on reconciliation changes to main space heating fuel, main space heating equipment, secondary space heating fuel, secondary space heating equipment, or main water heating equipment

The Reconciliation process improves our estimates of household energy characteristics

- Impacts of weighted distributions to secondary space heating fuel and equipment were negligible at the national and regional levels
- Some data changes from reconciliation forced data inconsistencies with other variables that also needed to be updated
 - Adding unreported residential on-site solar generation opened solar capacity and solar installation date. These variables were imputed using modal imputation
 - Adding heat to a home that previously reported not using space heating required changes to HEATHOME, EQUIPM, and FUELHEAT
- Although changes had minimal impacts on distributions, these changes are necessary as EIA makes micro data files available to stakeholders

eia Residential Energy Consumption Survey (RECS) Dashboard

Reset dashboard

3. Select a measure

Space heating consumption per household using space heating (MMBtu)

Air-conditioning (AC) consumption per household using AC (MMBtu)

Other end use consumption per household (MMBtu)

Refrigerator consumption per household using a refrigerator (MMBtu)

Space heating consumption per household using space heating (MMBtu)

Water heating consumption per household using water heating (MMBtu)

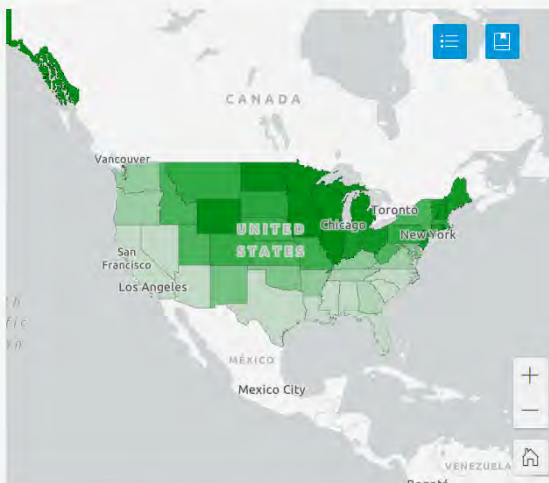
4. Select a state

South Dakota

About this dashboard

This dashboard includes selected estimates from the U.S. Energy Information Administration's *2020 Residential Energy Consumption Survey*. You can find more information on these data, including relative standard errors and withheld estimates, on our [metadata](#) page.

RECS Main Page | Metadata | Contact Us



South Dakota

47.9

Space heating consumption per household using space heating (MMBtu)

Household characteristics

0.35

Number of housing units (million)

2,056

Total square footage per housing unit

87.1

Total site consumption per household (MMBtu)

58%

Natural gas as main space heating fuel

United States

34.4

Space heating consumption per household using space heating (MMBtu)

Household characteristics

123.53

Number of housing units (million)

1,818

Total square footage per housing unit

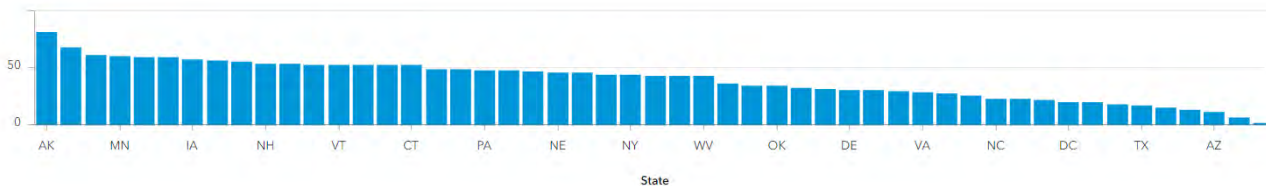
76.8

Total site consumption per household (MMBtu)

51%

Natural gas as main space heating fuel

Space heating consumption per household using space heating (MMBtu)



Impact of implementing a hierarchal editing system

- The 39 reconciliation edit specifications were run in 11 batches
 - A single household characteristic reporting error could trigger multiple related edit failures
 - Re-running edit criteria with previously updated housing characteristics should remove these cases from related edit failures
- Iterative and hierarchal editing lowered the number of cases requiring review
 - Initial run of the reconciliation edits yielded 10,683 failures (*Note: a single case can appear in more than one edit failure*)
 - Re-running reconciliation edit criteria in 11 batches lowered the failure count to 5,680, saving analysts from reviewing 5,003 failures, or 47%

Conclusions from the Reconciliation process

- The use of the editing dashboard that combined modeled consumption estimates based on housing characteristics with reported consumption from suppliers greatly aided the reconciliation process
- Applying updates and re-running the edit criteria between related edits helped reduce analyst burden and led to a timelier reconciliation process and data release
- 12% of paper respondents and 11% of web respondents required at least one change during reconciliation, showing minimal mode effect on editing
- Most of the changes to housing characteristics were to non-imputed values
- Making edits, even when the impact on overall estimates is minimal, limits the error associated with the entire vector of end-use level consumption estimates

Future research and possible changes to the reconciliation process

- Make consumption dashboard interactive to see how changes to household energy characteristics would impact the expected consumption in real-time
- Continue to evaluate editing thresholds and edit order for further reduction of analyst burden while maintaining data quality
- Build a more comprehensive editing process to include household survey edits, ESS data edits, and reconciliation edits
- Utilize outcomes of reconciliation editing to improve imputation methods

For more information

U.S. Energy Information Administration home page | www.eia.gov

Residential Energy Consumption Survey (RECS) home page | www.eia.gov/recs

L. Kaili Diamond | Kaili.Diamond@eia.gov

S. Grace Deng | Shaofen.Deng@eia.gov