The International Journal of Life Cycle Assessment (2022) 28:156–171
https://doi.org/10.1007/s11367-022-02113-1

DATA AVAILABILITY, DATA QUALITY

Derivation and assessment of regional electricity generation emission factors in the USA

Tapajyoti Ghosh[1,4] · Wesley W. Ingwersen[2] · Matthew Jamieson[3] · Troy R. Hawkins[4] · Sarah Cashman[4] · Troy Hottle[4,5] · Alberta Carpenter[1] · Kirti Richa[1]

Received: 29 April 2022 / Accepted: 15 November 2022 / Published online: 19 December 2022

Abstract

Purpose Electricity production is one of the largest sources of environmental emissions—especially greenhouse gases (GHGs)—in the USA. Emission factors (EFs) vary from region to region, which requires the use of spatially relevant EF data for electricity production while performing life cycle assessments (LCAs). Uncertainty information, which is sought by LCA practitioners, is rarely supplied with available life cycle inventories (LCIs).

Methods To address these challenges, we present a method for collecting data from different sources for electricity generation and environmental emissions; discuss the challenges involved in agglomerating such data; provide relevant suggestions and solutions to merge the information; and calculate EFs for electricity generation processes from various fuel sources for different spatial regions and spatial resolutions. The EFs from the US 2016 Electricity Life Cycle Inventory (eLCI) are analyzed and explored in this study. We also explore the method of uncertainty information derivation for the EFs.

Results and discussion We explore the EFs from different technologies across Emissions & Generation Resource Integrated Database (eGRID) regions in the USA. We find that for certain eGRID regions, the same electricity production technology may have worse emissions. This may be a result of the age of the plants in the region, the quality of fuel used, or other underlying factors. Region-wise life cycle impact assessment (LCIA) ISO 14040 impacts for total generation mix activities provide an overview of the total sustainability profile of electricity production in a particular region, rather than only global warming potential (GWP). We also find that, for different LCIA impacts, several eGRID regions are consistently worse than the US average LCIA impact for every unit of electricity generated.

Conclusion This work describes the development of an electricity production LCI at different spatial resolutions by combining and harmonizing information from several databases. The inventory consists of emissions, fuel inputs, and electricity and steam outputs from different electricity production technologies located across various regions of the USA. This LCI for electricity production in the USA will prove to be an enormous resource for all LCA researchers—considering the detailed sources of the information and the breadth of emissions covered by it.

Keywords Life cycle inventory · Database · Electricity · Life cycle assessment · Uncertainty · Sustainability

Communicated by Martin Baitz

✉ Wesley W. Ingwersen
ingwersen.wesley@epa.gov

1 National Renewable Energy Laboratory, Golden, CO, USA

2 U.S. Environmental Protection Agency, Office of Research and Development, Atlanta, GA, USA

3 KeyLogic, Morgantown, USA

4 Eastern Research Group, Inc, Concord, MA, USA

5 Oak Ridge Institute for Science and Education (ORISE), Oak Ridge, TN , USA

1 Introduction

The generation, distribution, and transmission of electricity contributed to nearly 30% of US anthropogenic greenhouse gas (GHG) emissions in 2016 (EPA 2018a) and criteria air pollutants (CAPs), such as sulfur dioxide (64% of total SO2 emissions in the USA in 2016) (Tiseo 2021). Because electricity generation technologies have evolved over time and are varied across different regions of the USA, the potential life cycle environmental impacts of generation within the country vary by region; as a result, the potential life cycle environmental impacts of electricity generation within the USA differ across regions (Mutel et al. 2012; Tamayao et al. 2015). The environmental releases and wastes associated with electricity production, as well as the generation amount and other production-related statistics, are often reported to relevant US federal government agencies as part of regulatory compliance. These agencies, including the Energy Information Administration (EIA) and the US Environmental Protection Agency (EPA), publish these data in inventory databases, such as the National Emissions Inventory (NEI) and Toxic Release Inventory (TRI). The data are released to the public on annual, biennial, or triennial bases. Systematic reviews of their contents and regular releases also make these datasets valuable for benchmarking. The datasets encompass different electricity production technologies across multiple years and are regularly updated. Thus, using these datasets for creating life cycle inventories (LCIs) ensures regular updates of the emission inventories. However, such data sources present many challenges to integration into LCIs (Sengupta et al. 2015; Cashman et al. 2016), including overlap of emissions reported by the same facility to different databases as well as unreported or missing emissions.

A large number of studies have been aimed at building inventories for emissions from the electricity production sector in the USA (IEA 2016; Gately and Hutrya 2018; Argonne National Laboratory 2019; Gagnon et al. 2021). Jaramillo and Muller (2016) analyzed the air pollution emissions and damages from electricity production and focused on PM2.5, SO2, NOx, NH3, and volatile organic compound (VOC) emissions. Hutchins et al. (2017) compared five fossil fuel–based GHG inventories for electricity production in the USA. de Chalendar et al. (2019) tracked emissions in the US electricity system and used multi-regional input output (MRIO) models to quantify pollutant flows. These studies included CO2, SO2, and NOx emissions. The Emissions & Generation Resource Integrated Database (eGRID) (EPA 2020) is one of the comprehensive databases for CAPs from the electricity production sector. Many individual states have developed their own GHG emission inventories for energy production (CARB 2016; Gately and Hutrya 2018; Oregon DEQ 2019; DVRPC 2018). However, most of these inventories focus mainly on GHG emissions and a few other air pollutants.

The Greenhouse Gases, Regulated Emissions, and Energy use in Technologies (GREET) Model (Argonne National Laboratory 2019) is one of the most widely used emission inventories for life cycle assessments (LCAs). However, GREET includes only 9 or 10 emissions in total. The US Life Cycle Inventory Database from the National Renewable Energy Laboratory (NREL) is the only true LCA inventory that models many pollutants—not only to air, but to water and soil—from electricity production and other industrial and commercial activities in the USA. However, the current version of the database lists this activity information—dated from 2010—which results in significant temporal misalignment if used for studying current technologies (NREL 2022). Koffler et al. (2019) published a paper on the relevance of Scope 3 emissions and power trade for regional LCIs of electricity consumption in the USA, which refers to LCI data for eGRID subgrids, using eGRID as a basis for emission factors. However, it only includes emissions from the eGRID database.

For addressing the challenges of creating a life cycle inventory for electricity generation that includes many pollutants and can be easily updated with time, we constructed a Python framework for combining information from several data sources and created a new LCI for electricity generation technologies in the USA. The compiled inventory is known as Electricity Life Cycle Inventory (eLCI), and it is available within the Federal LCA Commons repository (EPA 2022). The modeling platform for recreating the inventory is available on GitHub.

This study aims to explore the derivation of emission factors (EFs) in eLCI for electric power production facilities by combining information from different emission databases, solving challenges such as merging data coherently from multiple databases and handling data issues and using relevant assumptions to derive scientifically sound EFs. Along with EF derivations, we also explore the derivation of log-normal distribution parameters for the inventory that users can apply for a Monte Carlo analysis. These uncertainty data are made available to users of the LCI. Log-normal distribution, reliability, or quality scores based on pedigree matrices for the EFs are also supplied, providing a quick overview of the data quality. The final database, generated by collecting, refining, and modifying the raw information, is transformed for import and use in LCA software, such as openLCA.

Section 2 presents the calculation steps to determine the EFs, from the raw data to the final aggregated values. A case study in Section 3 is presented to show a generic built version of the EF database, and we explore the effects of various assumptions on the results. EF analyses comparing different regions, technologies, and pollutants are performed using the database.

2 Methodology

To create the electricity generation EF database, the two components needed are electricity production data and the respective emissions of electricity production by facility and technology. Databases maintained by the EPA and EIA are used, as shown in Fig. 1. The Emissions & Generation Resource Integrated Database (eGRID) (EPA 2020) is a primary source for obtaining electricity generation and GHG emission information. Several other emissions are

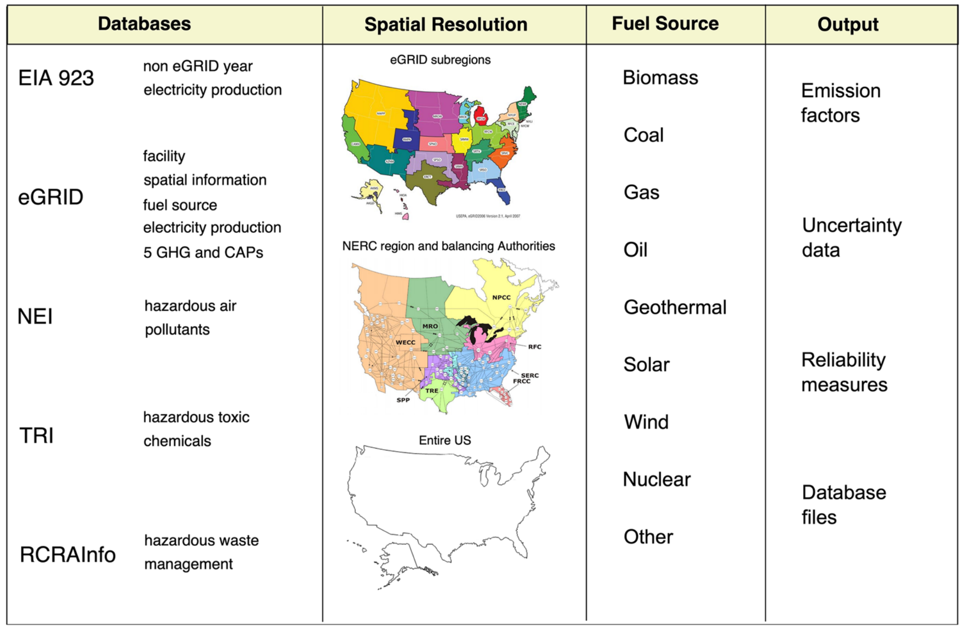

Fig. 1 Overview of data sources and determination of fuel source disaggregated generation-level EFs for different spatial levels. The columns list the range of options included or used in the framework for each column-header characteristic

derived from EPA-maintained national emission and waste inventories, including the Toxics Release Inventory (TRI) (EPA 2016), NEI (EPA 2015a), and the Resource Conservation and Recovery Act (RCRA) Biennial Report (EPA 2015b), stored in the RCRAInfo system. Brief descriptions of these databases are provided in Section 2.1 of this paper.

Important assumptions and their effects at every step of the eLCI generation method are described in Section 2.1. The EFs discussed and highlighted in this paper are in situ (i.e., at the plant) and do not include upstream activities. The EFs also exclude downstream activities of waste treatment. However, the actual eLCI framework builds the database with upstream fuel use and power generation plant construction. The framework also includes processing and extraction of the fuel within upstream activities. Our intention for this manuscript was to describe in detail how the direct emissions for generation were estimated from an array of sources that covers nearly all generators in the USA and to present the method along with a thorough evaluation of the resulting EFs.

2.1 Data sources used for EF calculation

Several emissions databases were combined and assessed to produce the eLCI database. These include eGRID (EPA 2020), NEI (EPA 2015a), TRI (EPA 2016), RCRAInfo (EPA 2015b), and EIA-923 (EIA 2015). Detailed information about these databases is provided in Section 1 of the "Supplementary information" (SI).

2.2 Description of spatial disaggregation levels

Three distinct levels of spatial aggregation were identified for the creation of the eLCI database, as shown in Fig. 1. The smallest spatial resolution is in the balancing authority (BA) regions, followed by eGRID, the North American Electric Reliability Corporation (NERC), and the entire US electricity grid, respectively, based on size. EFs vary in connection with the spatial resolution chosen for constructing the LCI. The appropriate spatial level and corresponding EFs are selected depending on the user's choice and requirements. Using a single EF, which is averaged over the entire country, introduces greater uncertainty in LCA studies. Alternatively, for systems with large spatial variations or for studying the utilization of electricity by nonpoint sources, a broader spatial scope is required. For those purposes, using eGRID or NERC EFs may be better. Trade of electricity may affect these results and is outside the scope of this paper. Trade methodologies and their effects on EFs have been investigated in an extension of this work by Hottle and Ghosh (2021). Descriptions of the different spatial regions are provided in Section 2 of the SI.

2.3 Modeling inputs, electricity generation, and fuel use

For building the spatially relevant emissions LCI database for electricity generation, the major inputs required are electricity generation, facility locations, fuel inputs, spatial locations, and emissions. eGRID supplies all of these input data but only five emissions; therefore, all other emissions are obtained from other databases, as listed in Section 2.1.

2.3.1 Data from eGRID

Electricity generation processes are based on facility-reported data from the eGRID. As shown in Fig. 1, various information is obtained from the eGRID database. Plant names defined in the eGRID "PLNT" sheet are the electricity generating facilities used as the data pool for creating the LCI. A screenshot of the PLNT sheet is provided in the SI, Fig. 3. The net generation quantity reported in the "Plant Annual Net Generation (MWh)" column is used for electricity production data. The US Department of Energy/EIA Office of Regulatory Information Systems (ORIS) plant or facility code is used as a relational bridge column to match with facility emissions from other databases. Balancing authority, eGRID subregion, and NERC region names are extracted from their respective columns to preserve spatial information, which provides the ability to perform EF calculations at different spatial levels.

The fuel source for electricity production is a common aggregation level for EFs. As shown from the spatial resolution column in Fig. 1, emissions for different spatial levels are obtained for different fuel sources separately. The primary fuel sources are listed in Fig. 1. The fuel types used for the aggregation are the eGRID fuel categories: coal, oil, natural gas, synthetic gas, nuclear fuel, hydroelectric, biomass, wind, solar, geothermal, and other fuel. Fuel source information is obtained from the "Plant Primary Fuel" and the "Plant Primary Coal/Oil/Gas/Other Fossil Fuel Category" columns in the eGRID database PLNT sheet. Combined heat and power plant useful thermal output (metric million British thermal units [MMBtu]) information is used to estimate the steam byproduct of combined heat and power production plants. The only input considered in the eLCI databases for electricity production is the fuel. From the eGRID PLNT sheet, plant total annual heat input (MMBtu) provides the total heat inflow to any facility. For fuel sources of coal, oil, natural gas, and biomass, rather than providing the heat input in energy units, the values are converted to physical flows of the fuel being used by the production facility. For the conversion, information about the fuel sources, along with the EIA data for the heat content of those respective fuel sources, is used. For renewable fuel sources of production, this conversion is not possible, and inputs are kept in units of megajoules (energy). For every facility reported in eGRID, the data collected includes the facility name, facility ID, primary fuel type, eGRID subregion, BA, NERC region, electricity generation, fuel input or energy input, and criteria air pollutants. These eGRID data are reported yearly. Thus, the temporal boundary of the database being generated is the eGRID publishing year used for building the database.

2.3.2 Determination of electricity production for non-eGRID years

eGRID publishes data every alternate even year, which matches the reporting years for the NEI and TRI databases. However, RCRAInfo data are reported every alternate odd year. Thus, when these databases are combined, a temporal mismatch arises. To address this situation, for calculating the EFs of waste in RCRAInfo, eGRID electricity production data are not used. Instead, production data are obtained from the EIA-923 (EIA 2015) database. For this modification, facilities are matched between the eGRID PLNT sheet and the EIA-923 sheet from the same year as RCRAInfo, and corresponding electricity production data in eGRID are replaced with electricity production data from EIA-923 only for RCRAInfo EF calculations.

2.3.3 Determination of the primary fuel source of the plant

Many facilities, as observed from eGRID, use a combination of different fuels as energy/fuel sources for generating electricity. However, the EF database for electricity generation needs to be differentiated by fuel, which creates the challenge of determining the primary fuel source of each plant. For this purpose, the fuel used for more than 90% of total electricity production is chosen as the primary fuel category for a facility.

2.4 Emission sources

As described previously, the different databases supply emissions information for EF calculations. While eGRID supplies information about the five CAPs, the NEI supplies facility-level emissions of air pollutants, and TRI supplies emissions to air, water, and soil. RCRAInfo primarily provides information on waste going to land. Utilization of the national release databases provides a common year of data and an adequate data collection period.

2.4.1 Matching emissions from databases

Connecting facilities between the different databases is a major challenge. The facility ID reported in eGRID was different than the ID reported for the same facility in NEI or TRI for all facilities. Thus, to match the same facilities between the databases, the Facility Registry Service (FRS) (EPA 2018b) is used as a relational bridge. FRS contains eGRID IDs as well as facility ID codes from NEI, TRI, and RCRAInfo. Thus, FRS serves as the easiest method to match the facilities between different databases and obtain emissions information from NEI, TRI, and RCRAInfo for generating plants in eGRID. For facilities without FRS ID matches, longitude-latitude matching up to three and four

eGRID2016 Plant file sequence number	Plant state abbreviation	Plant name	DOE/EIA ORIS plant or facility code	eGRID subregion acronym	Plant total annual heat input (MMBtu)	Plant annual net generation (MWh)	Plant annual NOx emissions (tons)	Plant ozone season NOx emissions (tons)	Plant annual SO2 emissions (tons)	Plant annual CO2 emissions (tons)	Plant annual CH4 emissions (lbs)
SEQPLT16	PSTATABB	PNAME	ORISPL	SUBRGN	PLHTIANT	PLNGENAN	PLNOXAN	PLNOXOZ	PLSO2AN	PLCO2AN	PLCH4AN
363	AZ	IRC Generator Facility	56899	NWPP	192	-21	0	0	0	16	1
6224	NE	Don Henry	2243	MROW	1,998	-16	0	0	0	117	4
146	AK	Viking	56147	AKMS	0	-10	0	0	0	0	0
1367	CA	Pala Energy Storage Yard	60566	CAMX	0	-4				0	0
3453	IN	Evonik Industries AG	54835	RFCW	120	-4	0	0	0	10	1
1349	CA	Ortega Highway Energy Storage	60567	CAMX	0	-2				0	0
131	AK	Thorne Bay Plant	7414	AKMS	653	-1	1	0	0	53	4
291	AR	Warren Lumber Mill	50640	SRMV	0	0			0	0	0
754	CA	Covanta Delano Energy	10840	CAMX	0	0			0	0	0
755	CA	Covanta Mendota	10837	CAMX	0	0			0	0	0
811	CA	Dinuba Energy	60100	CAMX	0	0			0	0	0
3440	IN	Crawfordsville Power Plant	1024	RFCW	0	0	0	0	0	0	0
3498	IN	Jasper 2	6225	RFCW	0	0			0	0	0
3719	KS	Synata Hugoton	58613	SPNO	0	0			0	0	0
4415	ME	Mead Rumford Cogen	10491	NEWE	0	0	0	0	0	0	0
4440	ME	Red Shield Envir Old Town Facility	10700	NEWE	0	0	0		0	0	0
5187	MO	Missouri City	2171	SRMW	0	0	0	0	0	0	0
9390	WI	Appleton Coated, LLC	55558	RFCW	0	0	0		0	0	0
36	AK	ESS Battery	58405	AKMS		0					
41	AK	Flywheel Energy Storage System	60563	AKMS		0					
70	AK	Kodiak	6281	AKMS		0					
4523	MI	Dafter	1868	RFCW	3,571	1	5	3	1	292	24
7464	OH	ST-1/1A Engine No 1	56422	RFCW	90	2	0	0	0	7	1
6431	NJ	Deutsche Bank- Piscataway Solar	60803	RFCE	18	2				0	0
6589	NJ	Sabert Solar	60810	RFCE	18	2				0	0
711	CA	Coca Cola American Canyon	57807	CAMX	20,035	3					46
7443	OH	PS ST-8 Engine No 1	56421	RFCW	66	4	0	0	0	5	0
7351	OH	Dodge Park Engine No 1	56423	RFCW	96	4	0	0	0	8	1
6916	NY	Central Hudson High Falls	579	NYUP	1,357	147				0	0
2239	CT	Torrington Terminal	565	NEWE	2,229	148	1	1	1	185	15
8280	TN	IKEA Memphis 508	61211	SRTV	1,385	150				0	0
2826	IA	Eastern Iowa Solar	60876	MROW	1,394	151				0	0
5181	MO	Marshall Solar Farm (MO)	60707	SRMW	1,394	151				0	0
9597	WV	Chemours Belle Plant	10788								

Fig. 2 PLNT Excel sheet of eGRID database. Data problems are marked using colored boxes and explained in Section 2.5

digits, as well as fuzzy logic-based name matching, was performed to link facilities between the different data sources. Details about creating a standardized inventory, Standardized Emission and Waste Inventories (StEWI) from different databases are explained in Young et al. (2022).

2.4.2 Emissions database hierarchy

While sourcing air emissions data from different databases, overlap occurs if the same facility has reported emissions data for the same pollutant to different databases. For example, carbon dioxide emissions overlap between the eGRID air emissions database and the NEI. Similar overlap occurs between the NEI and TRI. This problem was resolved by establishing a hierarchy between the databases. The databases are arranged as eGRID > NEI > TRI, with relative importance decreasing from left to right, i.e., data sources are prioritized in order of eGRID, then NEI, then TRI. Therefore, if any pollutant emission information from a facility is reported in a particular database, that data source supersedes all other data sources for the same pollutant from the same facility from all databases to the right. For example, if a certain power generation facility reports its carbon dioxide emissions to both the eGRID emissions database and the NEI, the NEI data are ignored and the eGRID data are used for developing the inventory.

eGRID-reported emissions were maintained above other sources because these emissions are reported along with the electricity generation values used in EFs. The selection of NEI emissions over TRI emissions when overlap occurs is maintained from other models where these sources are used together to estimate EFs (Cashman et al. 2016; Yang et al. 2017). The NEI contains unit-level information that allows for better harmonization across data sources; TRI is reported at the facility level rather than by unit, so it has more limitations. TRI is preferred over RCRAInfo because it reports the actual chemical quantity to the environment, represented as an elementary flow (RCRAInfo will report the total quantity of waste, which may contain many chemicals).

StEWI, which was developed by the EPA, is the tool for combining emissions from different databases for electricity generating facilities and creating a standardized, all-inclusive source of total emissions for electricity generation. StEWI v0.9.1 (Ingwersen et al. 2019; Young et al. 2022) was used in this study.

2.5 Data cleaning and processing

In direct visual verification of reported data, as shown in Fig. 2, many inconsistencies, errors, and data issues can be observed that must be rectified. Missing, erroneous, and irretrievable data points must be pruned from the databases. These problems and the adopted solutions are explained as follows:

- Box A (purple outline): For several plants, plant annual net generations are negative. Such facilities are dropped from inclusion in the database. The net generations may be negative because of the following:

  The plant was not operational.
  The plant is used as backup for a larger power production plant.

A mistake was made in the data during eGRID generation.

- Box B (red outline): Plant annual heat input and plant annual net generation are the bases of all EF calculations. If these data are missing, the facility is dropped from being included in the database.
- Box C (gray outline): Plant total annual heat input (MMBtu) and plant annual net generation (MWh) are used to calculate EFs. Some emissions reported in eGRID are calculated from the annual heat input values using EFs reported in eGRID. Other emissions, however, are measured using Continuous Emission Monitoring Systems (CEMS), which provide real-world data. During the calculation of EFs from these emissions, the total emission rates are divided by the annual net generation of the plants. A problem is thus introduced when the value for net generation is too low, an example of which can be seen in Box C, where the heat input is high but the net generation is extremely low. This can occur if the emissions or annual net generation are reported incorrectly. For these plants, the calculated EFs will be too high because of the denominator being small in value. Essentially, the mismatch between the total heat input and net generation causes EFs to be too high and skews the results for the entire database for a certain region or fuel source. Such problems are negated with the help of filters, as explained in Section 2.6, that calculate plant efficiencies based on their total electricity generation and total annual heat input and then retain only those facilities that are within the upper and lower ranges of electric generation efficiencies.
- Box D (blue outline): Various emissions reported in the databases do not have any values and occur as blanks. To handle these issues, we first duplicated the entire dataset, creating two copies: copy 1 and copy 2. In copy 1, we removed all facilities with blank emissions. In copy 2, blank data points were replaced with 0 on the assumption that these facilities are reporting actually zero or no emissions. EFs are calculated for copy 1 and copy 2 separately. The two EFs are then combined to calculate a final EF using weights. The weights are based on the electric generation values of the two data frames. The electricity generations of copy 1 and copy 2 will be different because in copy 1, facilities with blank emissions have been removed.
- Box E (green outline): Plants that report their emission values as 0 are assumed to be actually zero or no emissions.

2.6 Filters

As shown in Fig. 3, four data filters are used to remove outliers and misreported values from the data sources and calculate EFs within ranges of thermodynamic and practical possibilities. The filters are essential for not allowing outliers or erroneous data to affect the EF derivation. These filters were derived after in-depth exploration of the outlying data, calculation of initial EFs, comparison of the values with general scientific and engineering knowledge and other validated LCIs and searching for the erroneous outliers as well as determining proper methods to eliminate such data. The filters listed in Fig. 3 are applied in parallel (i.e., not in series or one after the other). Once individual filtered datasets are obtained from each filtering operation, they are merged with only common facilities being preserved. The following filters are used:

- Filter 1: This filter ensures that facilities chosen for calculating eLCI inventory have positive electricity generation values. If a facility was not producing a viable amount of electricity, including these facilities was unnecessary, as it resulted in infinite or negative EFs.
- Filter 2: This filter ensures that facilities selected for EF calculations have efficiencies between chosen upper and lower ranges. The efficiency is calculated by dividing the electricity generated by the heat input to the plant (after conversion to the same units). As explained previously, plants with low efficiencies will generally have overly high EFs because of the electricity generation being too small compared to heat energy input and emissions.
- Filter 3: This filter ensures that for a certain plant, if a certain fuel contributes to more than a certain percentage of the electricity generation of that facility, only then is that facility chosen for EF calculations and marked as a facility of that respective fuel type. A higher number should generally be chosen for this range (preferably greater than 90%). For example, if a plant uses both natural gas and oil for producing electricity, and 92% of electricity production can be attributed to energy from natural gas with the remainder from oil, the plant is categorized as a natural gas plant and preserved for calculating EFs of natural gas-type power plants.
- Filter 4: This filter ensures that a plant producing electricity is a primary power production facility and not a facility producing electricity as a byproduct. For example, facilities such as paper plants are removed. This filter uses North American Industry Classification System (NAICS) codes to match and remove such facilities.

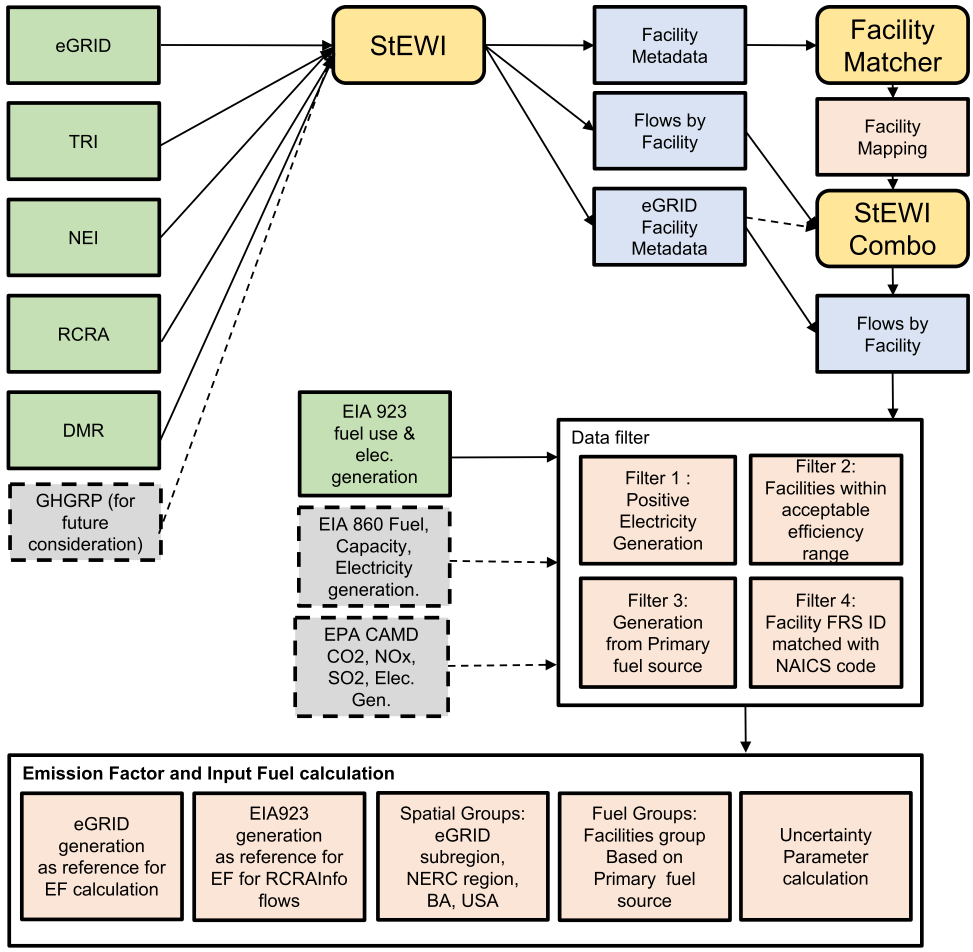

Fig. 3 Flowchart showing appropriate steps and applicable databases for calculation of EFs. StEWI is the supplier of the standardized total emissions for electricity generation facilities

2.7 Calculation of emission and waste factors

Emission and waste factors are developed using the information from the different databases. If f denotes total electricity production of a facility and e is the flow of the emission species x from the facility, then the EF for the x emission compound is calculated as:

$$y_x = \frac{\sum_j e_x^j}{\sum_j f^j} \tag{1}$$

where j denotes the facility.

The EFs are calculated for all different emissions, as obtained from the different databases grouped by a particular fuel source and belonging to a particular spatial region. Both the numerator and denominator change with variation in spatial location (different regions), spatial resolution (e.g., BA, eGRID, NERC), and fuel sources.

2.8 Generation of uncertainty information

Prediction intervals of 90% for each EF were calculated for the different releases and grouped by the spatial region and facility fuel type. The prediction intervals were evaluated as a function of the standard error of the prediction around the expected mean of EFs using Eq. 2. The parameters of the distribution are calculated such that the expected value of the log-normal distribution is set equal to the release factor, and the 95th percentile of the distribution is set to approximate the high-end range of the 90% prediction interval. Prediction intervals P, representing 90% confidence values, are developed for each EF for each emission x. Prediction intervals for a given flow are expressed as a percentage of the expected EF:

$$P_{y_x} = \frac{s_{y_x}(\sqrt{1 + \frac{1}{n}})t^*}{y} \tag{2}$$

where is the prediction interval for the EF for the emission. The standard error is, n is the number of data points, is the mean of EF, and is the critical value. This approach assumes that uncertainties of a given flow from multiple sources are independent. Negative release factors can occur when prediction intervals exceed 100%. To ensure non-negative releases, aggregate uncertainties for each flow are approximated using a log-normal distribution. The parameters of the distribution μ and σ are calculated such that the expected value of the distribution is set equal to the release factor, and the 95 percentile of the distribution is set to approximate the high-end range of the 90% prediction interval:

$$E[y_x] = e^{\mu+0.5\sigma^2} \tag{3}$$

$$CDF[y_x(1 + P_{y_x})] = 0.95 \tag{4}$$

$$CDF[y_x(1 + P_{y_x})] = 0.5 + 0.5erf\left[\frac{ln(y_x(1 + P_{y_x}))-\mu}{\acute{o}\sqrt{2}}\right] \tag{5}$$

where erf refers to the error function.

3 Case study: development of 2016 US electricity generation eLCI

In this case study, an electricity emissions inventory database is generated for 2016 at the spatial eGRID resolution. The eGRID regions are listed in Table 1. The fuel input, electricity generation quantities, emissions, and waste amounts are obtained from 2016 versions of eGRID, NEI, TRI and 2015 versions of RCRAInfo and EIA-923. These data are collected separately via queries to StEWI. A total of 7966 eGRID facilities are identified to have possible matches across the other EPA inventory systems. Of these facilities, the number of matching facilities with emissions reports for the years of interest found in NEI, RCRAInfo, and TRI are 2514, 407, and 799, respectively. The number of matches was low with eGRID, because eGRID includes the full set of electricity generation units, including all renewable electricity generation facilities. From these facilities, the StEWI flow-by-facility data are used as the primary data input for emissions and waste. The complete flow of information is visualized in Fig. 4.

3.1 Filtering

The reduction in the number of eGRID facilities covered in these filtering steps is insightful and is depicted in Fig. 4. Before EFs are calculated, facilities and their associated flows are filtered in several steps, corresponding to the filters described in Section 2.6.

Table 1 eGRID region list and geographic location

AKGD	Alaska Grid/Alaska Power Grid
AKMS	Miscellaneous/Alaska Power Grid
AZNM	Southwest/Western Power Grid
CAMX	California/Western Power Grid
ERCT	ERCOT all/ERCOT Power Grid
FRCC	FRCC All/Eastern Power Grid
HIMS	Miscellaneous/Hawaii Power Grid
HIOA	Oahu/Hawaii Power Grid
MROE	East/Eastern Power Grid
MROW	West/Eastern Power Grid
NEWE	New England/Eastern Power Grid
NWPP	Northwest/Western Power Grid
NYCW	NYC/Westchester / Eastern Power Grid
NYLI	Long Island/Eastern Power Grid
NYUP	Upstate NY/Eastern Power Grid
RFCE	East/Eastern Power Grid
RFCM	Michigan/Eastern Power Grid
RFCW	West/Eastern Power Grid
RMPA	Rockies/Western Power Grid
SPNO	North/Eastern Power Grid
SPSO	South/Eastern Power Grid
SRMV	Mississippi Valley/Eastern Power Grid
SRMW	Midwest/Eastern Power Grid
SRSO	South/Eastern Power Grid
SRTV	Tennessee Valley/Eastern Power Grid
SRVC	Virginia/Carolina/Eastern Power Grid

3.1.1 Removal of facilities with negative and zero (kWh) electricity generation

From an original total of 9709 facilities in the eGRID 2016 data, approximately 2171 facilities (22%) are removed that reported negative net generation (in filter 1). Electricity production values are the denominators in EF calculation; therefore, having such values affects the determination of EFs adversely. Hence, considering that either these facilities were not operative, or data were entered incorrectly, the facilities are excluded from consideration.

3.1.2 Plants with very low or high efficiency for electricity production

Of all the reported 9709 facilities, 2302 facilities (23.7%) are removed by filter 2 for efficiencies of less than 10% or greater than 100%. Most economically and technologically viable power production facilities lie between these efficiency ranges. These low and high ranges were obtained through discussions with experts on power production

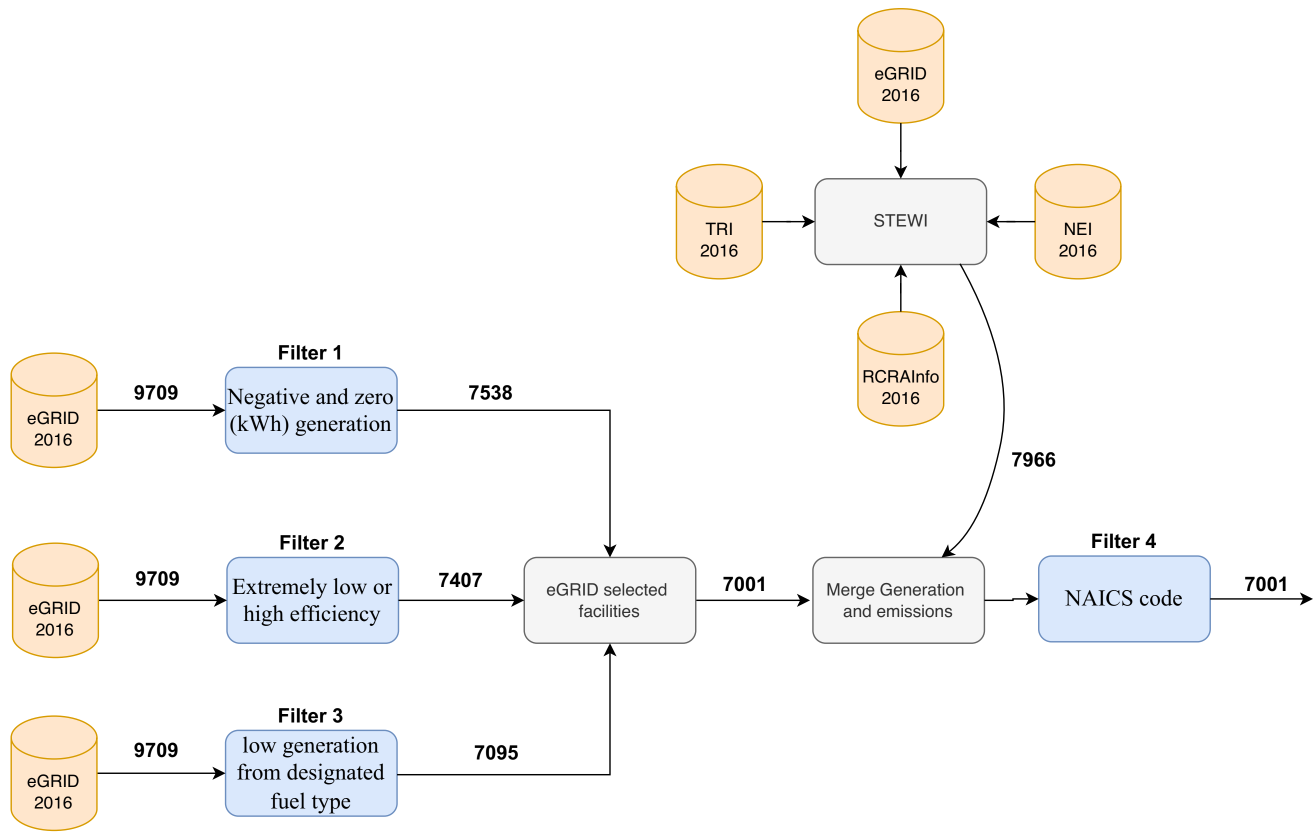

Fig. 4 Filtering operation on the eGRID facilities to screen out data issues within the inventory. The numbers on the arrows denote the number of facilities being passed on to the next step in the framework

facilities and a survey of various power plants. Choosing these efficiency numbers is subjective; for this study, we chose these ranges for exploring our EFs, including both nonrenewable and renewable energy generation technologies. Some of the facilities had high-efficiency numbers (combined heat and power), and for that reason, the decision was made to include efficiencies up to 100%. The advantage of the eLCI framework is that the database can be recreated by running the model (coded in Python) with the user's subjective choices of parameters (e.g., these efficiencies).

3.1.3 Plants with low generation from designated fuel type

A 90% cutoff was chosen for filter 3. The filter removes 2614 facilities (27% of all reported facilities) because the percentage of generation from their designated fuel type is < 90%. This filter is used for determining plant fuel types. A plant might burn more than one type of fuel; however, for generating the database, every facility needed to be assigned a particular fuel type. Therefore, the fuel that contributed more than 90% of the total fuel burned was chosen as the fuel type for that facility. The 90% amount was based on expert judgment and can be varied.

3.1.4 NAICS code not matching electricity generation code

Facilities whose NAICS code does not match the code for power generation are required to be removed if such facilities exist in the final dataset. Industries such as paper mills are occasionally reported as power generators as they product electricity as byproduct but are not "power-production" facilities. The NAICS filter does not remove any facilities as these facilities were already removed through StEWI.

Figure 4 demonstrates the filtering option. From StEWI, the emissions inventory obtained after combining the different emission databases from the EPA (TRI, NEI, eGRID, and RCRAInfo) contains 7966 unique eGRID facilities. Seven thousand one facilities are obtained after applying filters 1, 2 and 3. These two datasets are merged to obtain a final dataset of 7001 facilities. Initially, before any filters were used, the number of emission data points obtained was 120,219. Because of the filters, the number of emission data points was reduced; after the operation of the last filter, 34,415 data points remained from eGRID, 51,000 from NEI, 5699 from TRI, and 1873 from RCRAInfo. The filters can be considered one of the major assumptions in calculating EFs, with the expectation that

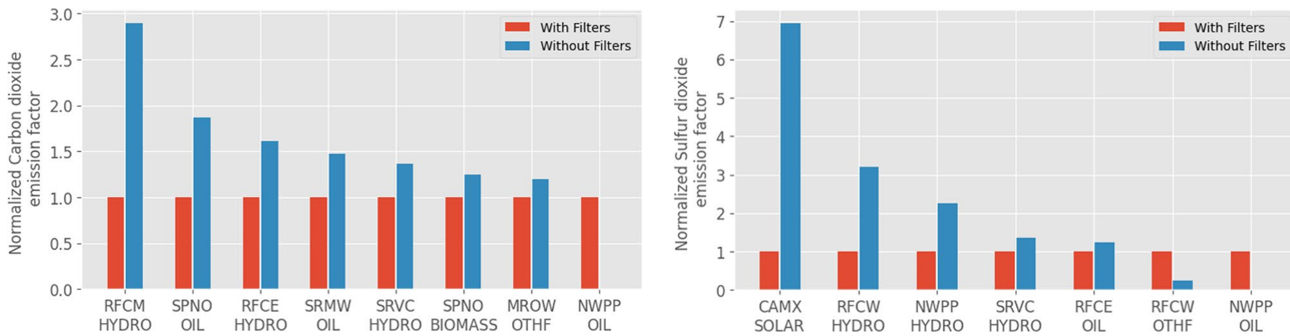

Fig. 5 Comparison of normalized EFs with and without filter application for carbon dioxide (left) and sulfur dioxide (right) for different eGRID subregions and fuel use. The filters help in removing some large outliers that significantly affect EF values

changing these assumptions will have a large effect on the value of EFs. To explore such effects, we calculated EFs for some examples to observe how EFs are affected when no filter is applied versus when all filters are applied. The graph in Fig. 5 shows these results. On the y-axis, the normalized EFs for different pollutants are shown. The larger the value of the y-axis, the greater the difference between EFs derived from filtering and without filtering. The x-axis corresponds to different regions and fuel sources for various pollutants. For certain cases, such as for RFCM Hydro, CAMX Solar, and SPNO Oil (the acronyms represent eGRID regions as shown in Fig. 1 as well as in Table 1), the difference in the normalized factors is very large between EF values with and without filters. The in situ emissions from renewable technologies of solar and hydro are probably a result of diesel generators or transportation equipment at the facility. Such differences, if not removed, lead to significant changes in the final derived spatial EF results. However, less than 3% of the facilities cause a large difference in the EFs of certain fuel sources. Few EFs of certain regions or certain fuel sources are affected significantly by these filters. These filters also help in removing outliers in the data. As shown in Fig. 6, outliers are removed when these filters are applied, which results in correct EF calculations for the combined US electricity grid. Without these calculations, the EF factor values would be well outside physically possible ranges.

3.2 Handling unreported emissions

As explained in Section 2, unreported emissions are handled by creating two databases: one with unreported emissions dropped and one with unreported emissions considered as a zero-emission rate. After building two separate databases, we compared the EF values and explored their differences. Some of the highest differences are shown in Fig. 7. The methods described to handle the unreported emissions

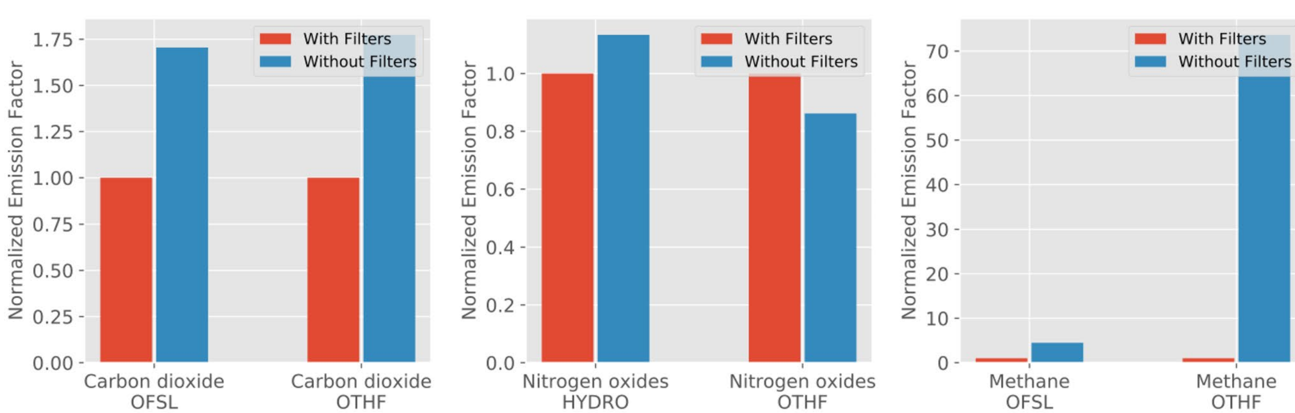

Fig. 6 Comparison of normalized EFs with and without filter application for different pollutants for complete US electricity grid and fuel use. OFSL (other fossil); OTHF (other fuel)

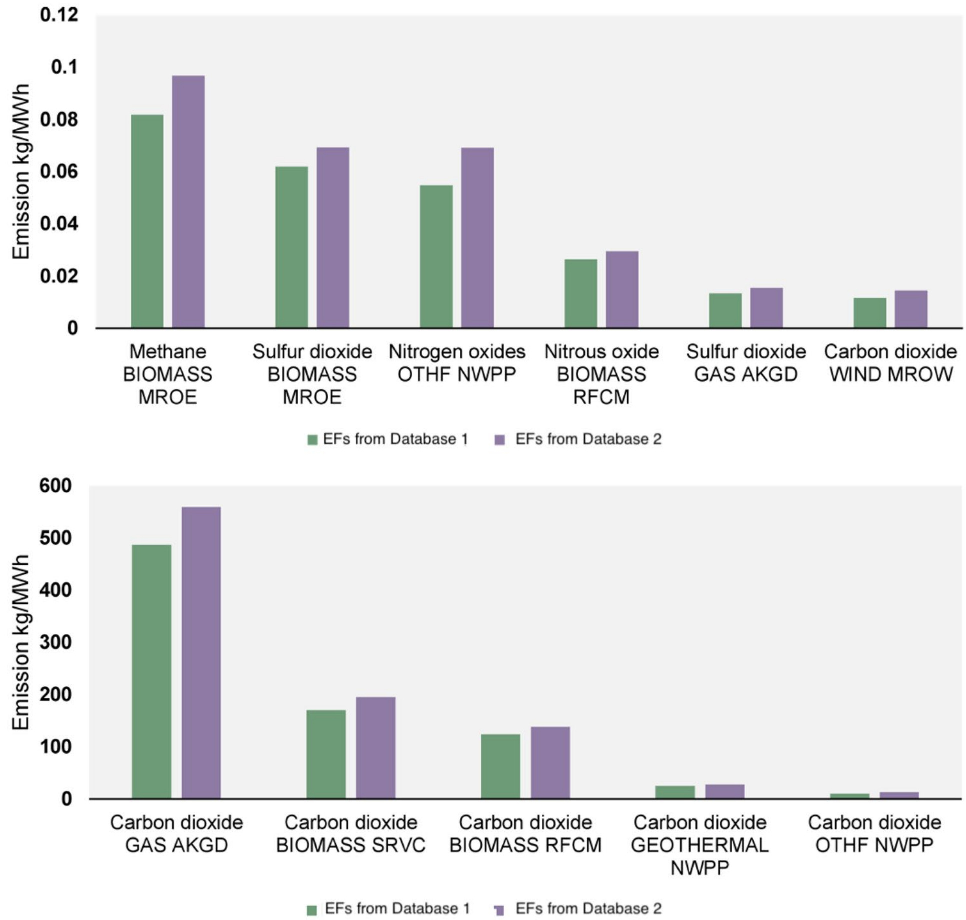

Fig. 7 Comparison of EFs when blank emissions are dropped from the calculations (Dataframe 1) with the EFs when blank emissions are considered 0 (Dataframe 2) according to the Methodology described in Section 2.5

within data cleaning and processing in Section 2.5 generally overestimate the EFs. Some cases for which the EFs are underestimated, although their number is lower than 3%.

4 Results

The final EF database contains EFs for 516 unique chemical compounds spread across compartments of air, water, and soil. Contributing to the calculation of these EFs were 7001 different facilities and approximately 65,000 different emission data points. The EFs are reported along with the geometric mean and geometric standard deviation, calculated using the assumption of log-normal distribution, as explained in the generation of uncertainty data in the Section 2.

4.1 Comparing the EFs between regions for particular fuel sources

This exploration compares how the same energy (technology) sources for producing electricity among different eGRID regions fare against one another in terms of EFs. Figure 8 shows the results for randomly chosen pollutant emissions for different energy sources. ERCT has a high EF for the following pollutants: carbon disulfide, PM2.5, 2-chloroacetophenone, and volatile organic compounds (VOCs) for coal-based electricity production. Alternatively, RMPA has EFs on the lower side for these pollutants for coal electricity production. Note that the EF variation should not be related to total emissions; for example, a certain eGRID subregion may have just one highly emitting coal power plant, which would highlight the region as a bad emitter for that fuel technology. However, fuel technologies that need to be upgraded or improved can be explored using spatial analysis of the EF generation in eLCI. Maps (3224 maps in total) plotting the combination of various pollutants and production technologies distributed across the USA. eGRID regions are available as a ZIP file with the SI.

4.2 Comparison of LCIA impact indicators for total generation mix from eGRID regions

The generation mix of electricity for individual eGRID regions combines the EFs from different energy sources based on their respective electricity production ratios. This

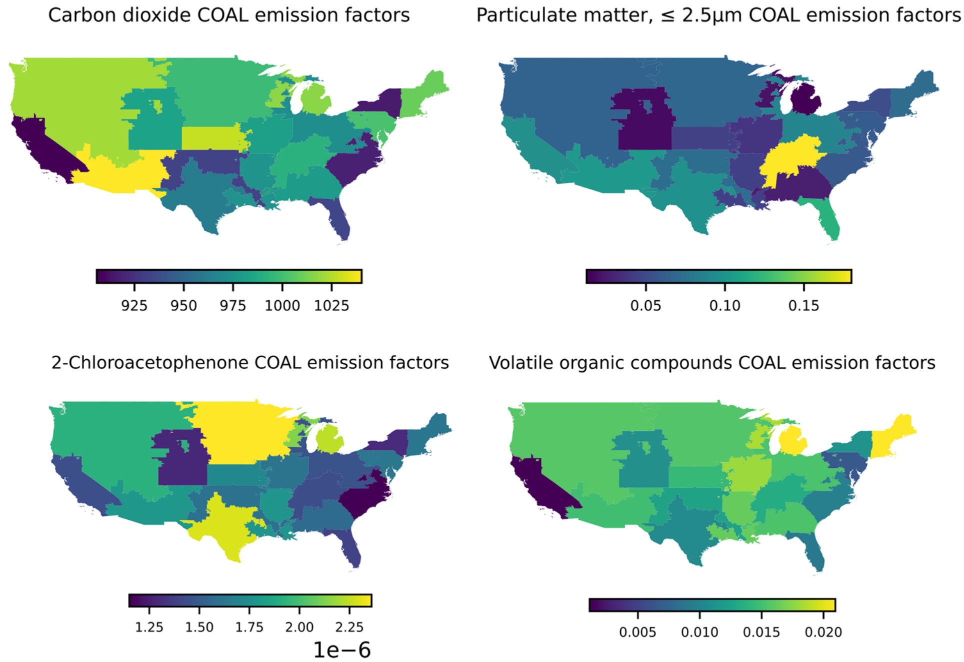

Fig. 8 EFs (kg/MWh) of different pollutants from different electricity production technologies for different eGRID regions across the contiguous USA

exploration is useful to understanding the final EFs, spatially disaggregated for performing LCA, because the generation-mix EFs are highly representative of the electricity that is actually consumed in the industrial and residential sectors; however, note that trading has significant effects on the final EFs (Hottle and Ghosh 2021). This analysis is extended from emission flow quantities to LCIA impact categories. Combining electricity from different energy sources and plotting EFs provides an overview of the actual average emissions resulting from production of every MWh of electricity, as shown in Fig. 9. NYUP demonstrates the best emissions profile for different impact categories in the USA, while SRMW shows the worst profile. The SRMW profile may be caused by the large concentration of coal power plants in that eGRID region. Figure 9 shows that CAMX, NEWE, and NYUP are regions with low GWP impact rates for every unit of electricity generated. SRMW and MROE are regions with very high EFs for GHG emissions. For particulate matter emissions, NEWE and NYUP have low emission rates. Analysis of the environmental impact of electricity power production thus needs to account for other emissions in addition to GHG, to aim for holistic sustainability.

4.3 Comparison with U.S. average EFs and LCIA

A major contribution of this work is that it shows spatially disaggregated EF information is extremely important, because using an average EF for the entire USA will result in an unreliable impact analysis. This study reinforces the findings of Mutel et al. (2012). The eLCI database sources emissions information from several EPA databases and agglomerates that information into EFs for different eGRID regions and fuel sources. Subsequently, EFs are combined in each region to create generation mixes, and finally, generation mixes of all regions combine to produce the US average generation mix.

The plots in Fig. 10 demonstrate how varied the EFs and subsequent impact category values are between different eGRID regions and the US average. Rather than showing individual emissions in the figure, different pollutants are combined in each emission indicator category, and LCIA results are calculated for Tool for Reduction and Assessment of Chemicals and Other Environmental Impacts (TRACI 2.1) impact indicator categories. Many regions have significantly high impacts for GWP, and acidification compared to the US average. The values have been normalized in Fig. 10, with the US average chosen as the baseline (value of 1).

4.4 Validation of eLCI database by comparison against eGRID database

The calculations explained and performed in this study, such as derivation of EFs for eGRID region generation mixes and the US generation mix, are also performed by the original eGRID database, albeit using only five of the emissions available in that database. In addition, deriving the US eLCI database entails using filters and statistical analysis to filter out potential errors in reported emissions data, as explained in Section 3.1. Thus, understanding the difference of the EFs of the five eGRID-reported pollutants between these two databases is useful. To harmonize and compare

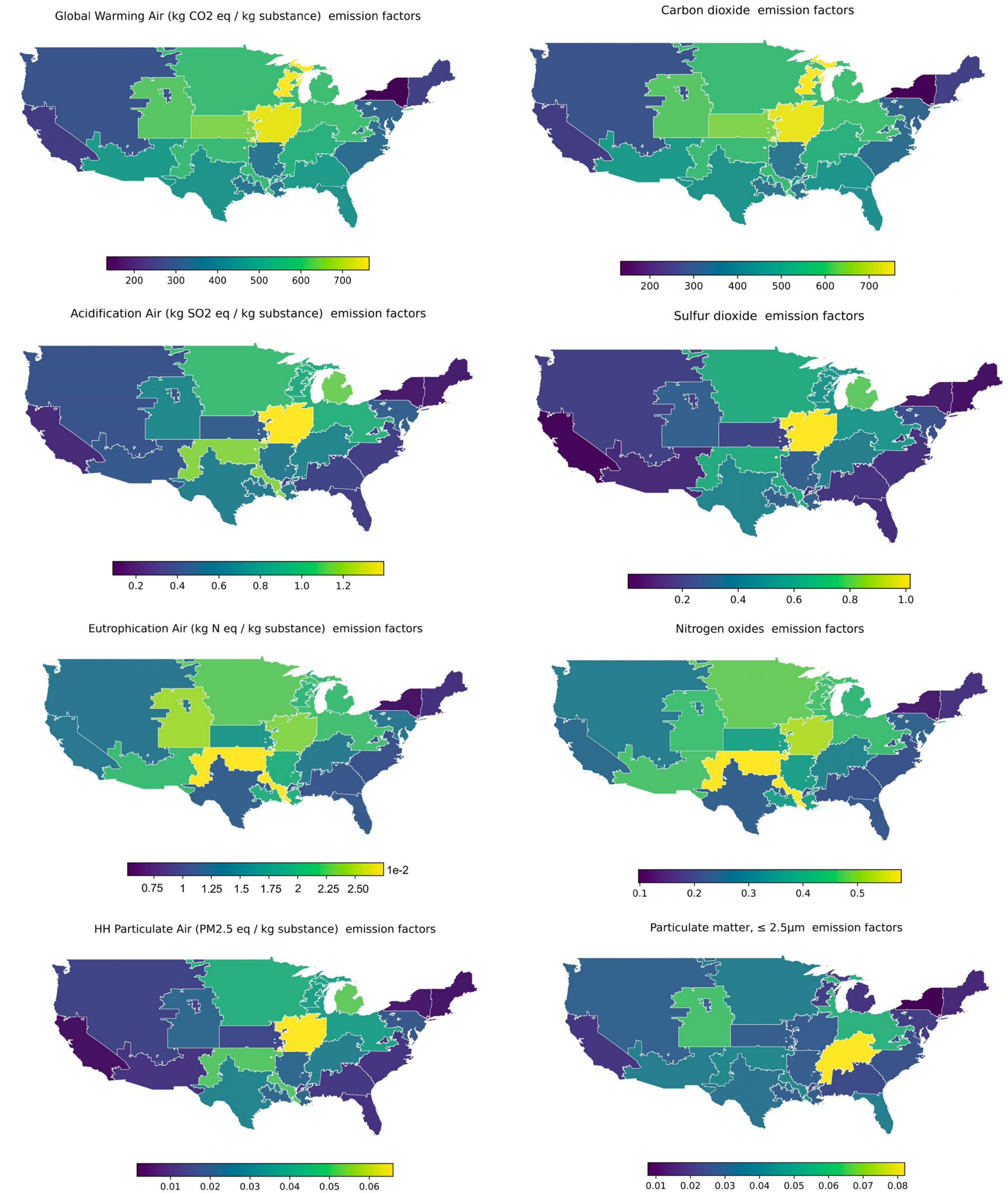

Fig. 9 EFs in kg/MWh (left) and associated LCIA impacts (right) varied across the different eGRID regions. The EFs shown here are derived from the generation mix of electricity that combines electricity production from different energy sources

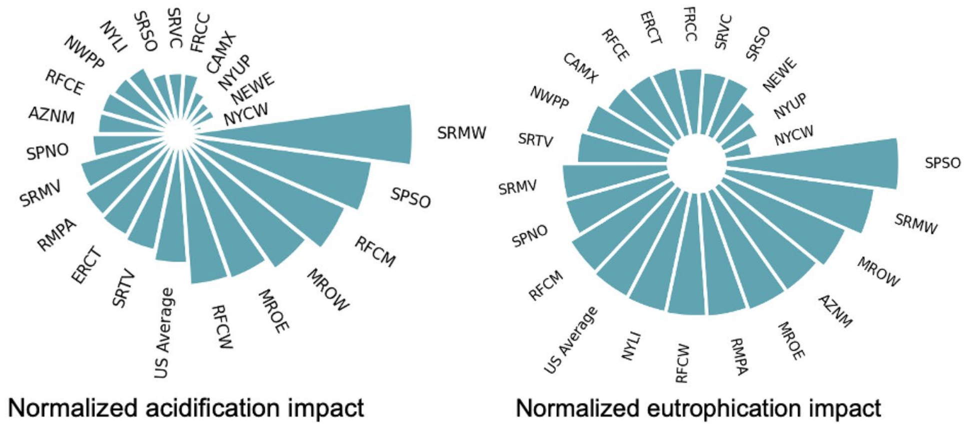

Fig. 10 Comparison of impact category values per MWh for different eGRID regions with the US average impacts

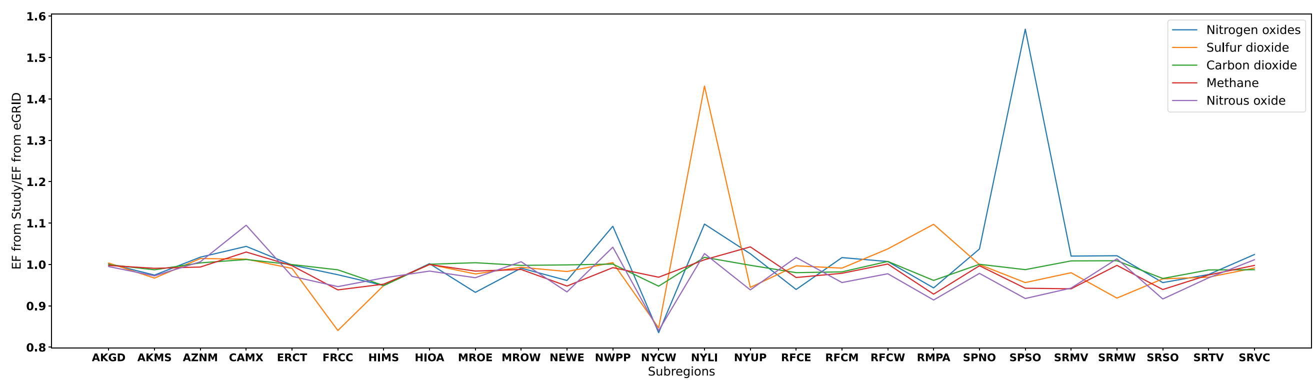

Fig. 11 Comparison of criteria air pollutant EFs for different eGRID regions with the US average EFs

the pollutants together, rather than plotting absolute values, we normalize the values by plotting the EF from the eLCI study divided by the EF from the eGRID database, as shown in Fig. 11. Values equal to 1 show that the EFs are the same; values lower than 1 indicate that EFs from the eLCI database are lower, and vice versa. In Fig. 11, we find large differences for the regions of NYLI and SPSO for sulfur dioxide and nitrogen oxide emissions. The probable cause of these large overestimations for the study might be low reporting of emissions when plants had extremely high efficiency, resulting in low EFs. Such plants were removed from the analysis by using filters, as described in Section 3.1.

4.5 Emission from renewable power generation facilities

Direct emissions from renewable power plants generally come from operations and maintenance activities (e.g., vehicle operation). Some power plants have an on-site office or other building(s) that could plausibly have a generator for backup power. Emissions may come from chemicals used for cleaning or lubricant replacement. Hydropower plants will have some emissions from biomass decay in the reservoir. Explanations of why databases like eGRID and TRI report such emissions are provided in their respective technical documents. As the eLCI framework directly uses these reported emissions for deriving EFs, renewable power generation technologies do appear to have emissions in this study.

5 Discussion

5.1 Conclusion

In this study, we described the generation steps of an electricity production LCI at different spatial resolutions by combining and harmonizing information from several databases. Numerous challenges involved with this process, as well as strategies for resolving those issues, were explained in detail. The inventory consists of emissions, fuel inputs and electricity, and steam outputs from different electricity production technologies located across various regions of the USA. The spatial resolution can be varied according to the user's preferences. Many emissions in air, water, and soil were compiled for the activities in this inventory. Finally, generation mix activities were produced, which combined electricity from different technologies in the ratio of their respective generation quantities.

We explored EFs from different technologies across eGRID regions in the USA. We found that for the same electricity production technology, a certain eGRID region may have worse emissions, which may be a result of the age of the plants in that region, the quality of coal used, or other underlying factors. Multiple regional LCIA impacts for total generation mix activities provided a better sustainability profile of electricity consumption in particular regions, rather than only considering GWP. We also found that for certain LCIA impacts, several eGRID regions are consistently worse than the US average LCIA impact for every unit of electricity generated. The variation in the LCIA impacts can help estimate improvement potential and guide research and development expenditure.

5.2 Limitations of the approach

All electric generation facilities reporting to the EPA as a large CAP source are incorporated in eGRID (EPA 2020). The data collection coverage for generation, therefore, likely represents over 80% of the current market, with small CAP emitters potentially excluded. However, for some inventory sources, quantity-based and categorical exclusions apply to facilities generating electricity for the grid. Therefore, further research is needed to determine what gaps exist in reported data for emissions and waste for facilities, as well as how those gaps can be filled with data or estimates from other sources.

The authors also recognize the subjective nature of the filter numbers chosen for the case study and database generation. These decisions were made with the best possible knowledge during the construction of the framework, and a decision was made to provide this choice to the user. Thus, the framework can easily be recreated with the user's chosen parameters. Additionally, allocation was not performed during the derivation of the EFs. This database was primarily meant for application using OpenLCA. Once this database is imported, when performing LCA study, OpenLCA can perform the allocation when requested, as steam generation from combined heat and power technologies is listed as a byproduct within the inventory.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11367-022-02113-1.

Acknowledgements The authors wish to thank Emily Newes and Dan Bilello for their feedback on this manuscript. This research was conducted as part of the EPA Air, Climate, and Energy National Research Program. Support was provided by the National Renewable Energy Laboratory through Interagency Agreement DW-89-92448301-1, and by Eastern Research Group (ERG) through EPA contract EP-C-16-015, Task Order 008. We acknowledge the valuable contributions of Timothy Skone, Gregory Cooney, George Zaimes, and Ambica Pegallapati at the National Energy Technology Laboratory and KeyLogic.

Author contribution Conceptualization, W. W. I.; software, W. W. I., T. G. M. J., T. H.; data curation, W. W. I, T. G., M. J., T. H, K. R.; investigation, W.W.I, T.G. M.J., T.H.; methodology, W.W.I, T.G. M.J., T.H.; validation, W.W.I, T.G., M. J., T. H., A. C., T. R. H., S. C., A. C.; visualization, W. W. I, T. G.; writing, W. W. I, T. G., M. J.,T. R. H., S. C.

Funding Funding was provided by the US EPA Office of Research and Development's Chemical Safety and Sustainability National Research Program and the former National Risk Management Research Laboratory. The US Environmental Protection Agency, through its Office of Research and Development, funded and conducted the research described herein under an approved Quality Assurance Project Plan.

Data availability The datasets generated and/or analyzed during this study are available in the electricity baseline life cycle inventory repository. Source code is available on a GitHub repository.

https://www.lcacommons.gov/lca-collaboration/Federal_LCA_Commons/US_electricity_baseline/datasets

https://github.com/USEPA/ElectricityLCI

Declarations

Conflict of interest The authors declare no competing interests.

Disclaimer The work has been subjected to review by the EPA Office of Research and Development and approved for publication. Approval does not signify that the contents reflect the views of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

References

Argonne National Laboratory (2019) *GREET vehicle-cycle model*. Argonne National Laboratory: Argonne, IL, United States. Available at: https://greet.es.anl.gov. Accessed 1 Nov 2020

CARB (2016) California greenhouse gas emission inventory program | California Air Resources Board. https://ww2.arb.ca.gov/our-work/programs/ghg-inventory-program. Accessed 10 Apr 2022

Cashman SA, Meyer DE, Edelen AN, Ingwersen WW, Abraham JP, Barrett WM, Gonzalez MA, Randall PM, Ruiz-Mercado G, Smith RL (2016) Mining available data from the United States Environmental Protection Agency to support rapid life cycle inventory modeling of chemical manufacturing. Environ Sci Technol 50(17):9013–9025. https://doi.org/10.1021/acs.est.6b02160

de Chalendar JA, Taggart J, Benson SM (2019) Tracking emissions in the US electricity system. Proc Natl Acad Sci 116(51):25497–25502. https://doi.org/10.1073/pnas.1912950116

Department of Environmental Quality (2019) Oregon greenhouse gas sector-based inventory data : air quality programs : State of Oregon. https://www.oregon.gov/deq/aq/programs/Pages/GHG-Inventory.aspx. Accessed 10 Apr 2022

DVRPC (2018) Delaware valley regional planning commission regional greenhouse gas emissions inventory. https://www.dvrpc.org/energyclimate/inventory/. Accessed 10 Apr 2022

EIA (2015) Form EIA-923 (EIA-906/920). https://www.eia.gov/electricity/data/eia923/. Accessed 11 Nov 2022

Gagnon P, Frazier W, Cole W, Hale E (2021) Cambium documentation: *Version 2021*; NREL/TP-6A40–81611; National Renewable Energy

Lab. (NREL), Golden, CO (United States). https://doi.org/10.2172/1835075
Gately C, Hutrya LR (2018) CMS: CO2 emissions from fossil fuels combustion, ACES Inventory for Northeastern USA. 43202.655434 MB. https://doi.org/10.3334/ORNLDAAC/1501
Hottle T, Ghosh T (2021) Regional Electricity Consumption Mixes Using Trade Data for Representative Inventories. Int J Life Cycle Assess 26(6):1211–1222. https://doi.org/10.1007/s11367-021-01876-3
Hutchins MG, Colby JD, Marland G, Marland E (2017) A Comparison of Five High-Resolution Spatially-Explicit, Fossil-Fuel, Carbon Dioxide Emission Inventories for the United States. Mitig Adapt Strateg Glob Change 22(6):947–972. https://doi.org/10.1007/s11027-016-9709-9
International Energy Agency (2016). https://www.iea.org/data-and-statistics/data-products. Accessed 10 Apr 2022
Ingwersen W, Bergmann M, Liadov M, Ghosh T, Li M, Cashman SA (2019) Standardized emission and waste inventories (StEWI), v0.9.1
Jaramillo P, Muller NZ (2016) Air pollution emissions and damages from energy production in the U.S.: 2002–2011. Energy Policy 90, 202–211. https://doi.org/10.1016/j.enpol.2015.12.035
Koffler C, Hengstler J, Thellier L, Stoffregen A (2019) On the relevance of scope 3 emissions and power trade for regional life cycle inventories of electricity consumption in the USA. Int J Life Cycle Assess 24(8):1360–1375. https://doi.org/10.1007/s11367-018-1566-1
Mutel CL, Pfister S, Hellweg S (2012) GIS-Based Regionalized Life Cycle Assessment: How Big Is Small Enough? Methodology and Case Study of Electricity Generation. Environ Sci Technol 46(2):1096–1103. https://doi.org/10.1021/es203117z
NREL (2022) U.S. Life Cycle Inventory Database. https://www.nrel.gov/lci/index.html. Accessed 10 Apr 2022
Sengupta D, Hawkins TR, Smith RL (2015) Using national inventories for estimating environmental impacts of products from industrial sectors: a case study of ethanol and gasoline. Int J Life Cycle Assess 20(5):597–607. https://doi.org/10.1007/s11367-015-0859-x
Tamayao M-AM, Michalek JJ, Hendrickson C, Azevedo IML (2015) Regional variability and uncertainty of electric vehicle life cycle CO2 emissions across the United States. Environ Sci Technol 49(14):8844–8855. https://doi.org/10.1021/acs.est.5b00815
Tiseo I (2021) U.S. electric utilities SO2 emissions 1970–2020. Statista. https://www.statista.com/statistics/504390/volume-of-sulfur-dioxide-emissions-from-fuel-combustion-in-electric-utilities-us/. Accessed 29 Apr 2022
US EPA (2015a) Environmental Protection Agency- National Emissions Inventory (NEI). https://www.epa.gov/air-emissions-inventories/national-emissions-inventory-nei. Accessed 11 Nov 2021
US EPA (2015b) Environmental Protection Agency- RCRAInfo Overview. https://www.epa.gov/enviro/rcrainfo-overview. Accessed 11 Nov 2021
US EPA (2016) Environmental Protection Agency- Toxics Release Inventory (TRI) Program. https://www.epa.gov/toxics-release-inventory-tri-program. Accessed 11 Nov 2021
US EPA(2018a) Environmental Protection Agency- Our Nation's Air. https://gispub.epa.gov/air/trendsreport/2018/#naaqs. Accessed 11 Nov 2021
US EPA (2018b) Environmental Protection Agency- Facility Registry Service (FRS). https://www.epa.gov/frs. Accessed 12 Nov 2021
US EPA (2020) Environmental Protection Agency- Emissions & generation resource integrated database (eGRID). https://www.epa.gov/egrid. Accessed 11 Nov 2021
US EPA (2022) Environmental Protection Agency- Federal_LCA_Commons/US_electricity_baseline | LCA Collaboration Server. https://www.lcacommons.gov/lca-collaboration/Federal_LCA_Commons/US_electricity_baseline/datasets. Accessed 28 Apr 2022
Yang Y, Ingwersen WW, Hawkins TR, Srocka M, Meyer DE (2017) USEEIO: A New and Transparent United States Environmentally-Extended Input-Output Model. J of Clean Prod 158:308–318. https://doi.org/10.1016/j.jclepro.2017.04.150
Young B, Ingwersen WW, Bergmann M, Hernandez-Betancur JD, Ghosh T, Bell E, Cashman SA (2022) System for Standardizing and Combining U.S. Environmental Protection Agency Emissions and Waste Inventory Data. Appl Sci 12 (7), 3447. https://doi.org/10.3390/app12073447