United States
Environmental Protection
Agency

EPA 600/R-23/228 | January 2024 | www.epa.gov/research

USEEIO State Models v1.0

Environmentally-Extended Input-Output Models for U.S. States

Office of Research and Development
Center For Environmental Solutions and Emergency Response

This page is intentionally left blank

EPA 600/R-23/228 I January 2024 I www.epa.gov/research

USEEIO State Models v1.0

Environmentally-Extended Input-Output Models for U.S. States

by
Wesley W. Ingwersen
Catherine Birney
Center for Environmental Solutions and Emergency Response
Cincinnati, OH 45208

Ben Young
Jorge Vendries
Eastern Research Group, Inc.
Concord, MA 01742

This report was produced with the assistance of Eastern Research Group under Contract 68HERC21D0003, Task Order 68HERC21F0133.

Disclaimer

The U.S. Environmental Protection Agency, through its Office of Research and Development, funded and conducted the research described herein under an approved Quality Assurance Project Plan (K-LRTD-0030017-QP-1-3). The Contractor's role did not include establishing Agency policy. It has been subjected to the Agency's peer and administrative review and has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Table of Contents

List of Tables

List of Figures

Acknowledgments

The research was conducted with support of the USEPA Sustainable and Health Communities National Research Program and the Office of Resource Conservation and Recovery.

This report is the result of the support of numerous individuals whose contribution is briefly described and acknowledged below.

Peer Review
Gretchen Newman, Washington Department of Ecology
Nathan Robbins, Maine Department of Environmental Protection
David Meyer, Center for Environmental Solutions and Emergency Response, ORD, USEPA
Priscilla Halloran, Office of Resource Conservation and Recovery, Office of Land and Emergency Management, USEPA

Peer Review Coordination
Brian Dyson, Center for Environmental Solutions and Emergency Response, ORD, USEPA

Task Management
Sarah Cashman, Eastern Research Group, Inc.

Abbreviations

BEA	Bureau of Economic Analysis of the U.S. Department of Commerce
BLS	Bureau of Labor Statistics of the U.S. Department of Labor
CBECS	Commercial Business Energy Consumption Survey
CBEI	Consumption-Based Emission Inventory
EEIO	Environmentally-Extended Input-Output model
EIA	Energy Information Administration of the U.S. Department of Energy
FBS	Flow-By-Sector format for environmental totals by sector
G/S	Good or Service
ICF	Interregional Commodity Flows
LCA	Life Cycle Assessment
LCIA	Life Cycle Impact Assessment
MECS	Manufacturing Energy Consumption Survey
MRIO	Multi-Regional Input-Output
NAICS	North American Industry Classification System
NEI	National Emission Inventory
RoUS	Rest of U.S. (excluding SoI)
RoW	Rest of the World (excluding U.S.)
SoI	State of Interest (U.S. State)
StateIO	State-level economic Input-Output models (EPA developed)
TRI	Toxic Release Inventory
USEEIO	U.S. EEIO models for the U.S.

1 Executive Summary

Many U.S. states and communities have established policies and plans to steer themselves toward both cleaner and more prosperous economies. Progress towards achieving such goals cannot be easily measured due to the complexities of economies and related environmental impacts occurring outside of their territories but that are due to local consumption. Quantitative models that capture the relationship between the economy and environment can provide information to inform these policies. For states, these models need to account for the interrelationships between different economic sectors, across the regions they trade with beyond their borders as well as their own, and that potentially drive different types of environmental and social impacts. If states are to be confident in their assessments of technologies or industries, models must capture the relationship between consumption of goods or services (G/S), the state's own industries, dependencies upon imports, and the associated potential environmental, human health, and economic impacts occurring within and outside their borders.

EPA has developed a first set of models, called the USEEIO State models, that are intrinsically built on these relationships and can quantify these potential impacts for existing G&S produced or consumed in a U.S. state. The new models are extensions of the existing national level U.S. Environmentally-Extended Input-Output (USEEIO) models. They use established methods to combine regularly reported economic statistics with environmental data and impact models to characterize the production of goods and services in a given state along with the rest of the U.S. The models can capture impacts from the supply chains as well as the final use of those G/S, whether they are used as inputs to another industry, used by households, investors, or government, both within or out-of-state, or exported internationally. Resources used, emissions, hazardous waste generation, and jobs are modeled for each G/S category by state of interest (SoI) and the rest of the U.S. (RoUS). Total production of all G/S and final consumption is estimated by state and can be used to estimate total associated impacts and compare environmental performance among the G/S classes, within and across states. Evaluating G/S consumed in a state can reveal whether inputs come from within or outside of the state and where the associated impacts may occur.

An annual time-series from 2012-2020 of USEEIO State models for each of the 50 U.S. states are available with a resolution of 73 G/S categories, two-regions, SoI and RoUS, with accounting for impacts from imports, and can report 16 environmental, human health, resource use, waste, and economic indicators. The models cover for each G/S the "cradle-to-gate" life cycle stages, including from resource mining through manufacturing or service provision. The models have known applications in sustainable material management, sustainable procurement, economic sector analysis, hotspot analysis, life cycle assessment, and consumption-based emissions inventories, among others. The models have been built using a tested and validated modeling framework and have been through internal and external review. The models are distributed via a number of outlets and formats. The full model in a native format requires advanced understanding to use as-is, but the models could also be incorporated into more easy-to-use interfaces for broader audiences and more targeted uses.

The strengths of these models include that they are comprehensive of all economic sectors and their relationships, that they are based on public data combined through established algorithms, that they are freely available and that the source code to build the models is available to the public to inspect and improve upon with minimal restrictions.

The models lack the capacity to characterize and differentiate between specific products within a category, and due to some data source limitations and associated model assumptions, they may assume common economic structure and related environmental profiles across states.

This report provides an overview of the new USEEIO State models. Example results from the models are provided to show and compare emissions intensities of G/S, ranking sectors by contribution to potential impacts, and exposing the source of potential impacts by source across the supply chains. Some initial findings of embodied impacts per dollar are explored by commodity and state for 2020. Example results for a SoI are provided to show how users might use a USEEIO State model for a state to rank sectors by contribution to potential impacts, exposing the region where impacts come from, and using models for multiple years for a SoI to evaluate how its total or sector-specific impacts have changed in recent years. In the near-future, EPA will use these models to conduct consumption-based emissions inventories for some U.S. states. Links to access models, results and source code are provided. An appendix provided further details on data sources, methodologies and software used.

2 Purpose and Overview

The purpose of this report is to describe the first version (v1.0) of USEEIO State models. We describe the background and current context of the problems that the models are addressing. It includes an overview of the model characteristics, general functions, availability, usage, strengths, and limitations. We validate the models to assure they will function to compute results as intended. We provide examples of results that the models can generate and a discussion around the insights these examples provide. We included a brief note on immediate future plans. We provide additional details in an appendix including theoretical background on EEIO, reasoning for data selection and methodologies employed, and a description of software used to create the models.

The intended audience for this report, broadly, are potential USEEIO State models users. While a potential user base might include individuals and organizations with many purposes, EPA is providing these resources with representatives from organizations in mind that desire state-specific information on environmental and economic aspects related to the production, procurement or use of goods and services in their state of interest, and that might use this information to support or provide insight into their programs and policies. The authors do not assume readers of this report have a technical background in the formal construction of EEIO (environmentally-extended input-output models). However, the text does assume basic knowledge of economic statistics, environmental and human health threats, and concepts of quantitative models. Representatives from state government agencies of Georgia, Minnesota, Oregon, and Washington showed interest and provided early feedback on our work. When policies, examples, or results are given where not all states can be covered, these four states are highlighted.

3 Background

The U.S. government as well as many U.S. state and local governments are interested in making their regions cleaner or “greener” while improving welfare and opportunities for residents and visitors. In many states and communities this interest is strongly motivated by threats related to global climate change including those that can endanger U.S. ecosystems and human well-being (Program, 2023), and how actions might be able to reduce these pressures through more sustainable production and consumption. This is often captured in the concept of a clean or “green” transition. There are some fundamental facts of the relationship between the economy, environment, and human health that are essential to consider in order to guide such transitions.

The four fundamental facts (4Fs) are:

1. Sectors are interdependent. Evaluating opportunities in one sector affect others. For example, new transportation technologies will not just change tailpipe emissions or vehicle fuel efficiency; they may impact the manufacturing, utilities, and agricultural sectors.
2. There are multiple environmental, health, and resource concerns. While one environmental or health issue may be of primary concern, such as climate change, other concerns cannot be ignored, such as children’s health, environmental justice, and safe drinking water.
3. Consumption as well as production is relevant. While production has typically been addressed more directly in environmental regulations that impose pollution limits in target sectors, new technologies often imply increased consumption of certain raw or manufactured goods and also rely on changes in final consumer or government consumption to be successful. Furthermore, consumption-oriented policies may be the only alternative for reducing impacts that are occurring outside of their political boundaries.
4. Regions and nations are deeply interconnected through trade and cross-boundary environmental and health issues. The U.S. relies heavily on imports to satisfy industrial and consumer demand for raw materials and goods. Goods and services flow freely across U.S. state borders. Local political boundaries within the U.S. provide no clear demarcation of economic boundaries. Environmental problems like climate change are global and water and air pollution are cross-boundary.

U.S., state and local agencies and lawmakers are leading their regions towards these cleaner production or consumption transitions in many ways. These strategies have appeared through the lenses of sustainable material management and circular economy policies, sustainable procurement guidelines, climate action plans, and programs or policies that promote green innovation and clean technology. The laws, policies, and programs they enact may be driven by more targeted concerns, such as waste generation, pollution reduction, or climate-related threats, but what is common across these strategies is that they must innately acknowledge the 4Fs to be successful.

Some examples of these policies, programs, or initiatives in U.S. states are:

- Oregon has a 2050 Material Management Vision (OR DEQ, 2012) that describes a future state with more efficient and clean production.

- The state of Washington has made sustainable materials management the core of its

Solid and Hazardous Waste Plan (WA Department of Ecology, 2021). Core aspects of this plan included the abilities to evaluate all life cycle phases of products and to be able to evaluate the impacts driven by consumption of goods and services.

- States and cities have also approached this more recently with Climate Action Plans. Minnesota's climate plan as well as plans in other states and cities address strategies in many economic sectors, like transportation, energy, agriculture, and the built environment (State of MN, 2022).

- Georgia is addressing these issues of sustainability through an innovation lens, looking to develop and attract new industries and technologies that will generate returns in energy and material efficiency that can cascade across the economy (Georgia Center of Innovation, 2023).

Evaluating the potential impacts of public policy to support any of the aforementioned goals requires analytical tools that embed the 4Fs through quantitative relationships. At the national level through the sustainable materials management program, EPA has used such an analytical framework for high-level evaluation of sustainable management strategies (EPA, 2009). Since that time the EPA has developed and maintained the United States Environmentally-Extended Input-Output (USEEIO) family of models as the core of this analytical framework (Ingwersen et al., 2022; Yang et al., 2017).

USEEIO embeds the 4Fs as attributes or capabilities within its quantitative framework. This family of models has been used for estimating the carbon and other environmental impacts associated with U.S. economic activity. USEEIO models are based on an internationally-standardized and regularly-issued statistical data product – the U.S. Supply (also known as Make) and Use tables – that embed the economic activity of all industries and their interrelationships, trade, and consumption, addressing 4F #1 and #3. These Supply and Use tables are paired with datasets of resource use, emissions, wastes which enable the reporting of environmental impact, human health, and economic indicators, addressing 4F #2. At the national level, they have only to a limited extent addressed environmental issues driven by imports, by assuming they are produced like in the U.S. and not providing trade partner-level accuracy.

At the U.S. state level, USEEIO has not been capable of addressing these needs in a way that considers the economies in a manner that is specific to the states. EPA piloted efforts with the State of Georgia to apply USEEIO models to provide insight into state and community opportunities for sustainable materials management and cleaner economic development (EPA, 2020; Ingwersen et al., 2020; Ingwersen and Meyer, 2016). Through one of these pilot efforts, a prototype GA-specific model was developed for internal evaluation, but not for external use and not extendable to other states.

A major challenge to creating more robust, state-specific models is the lack of state level economic and environmental data that are available in national level statistics. However, recent work to develop a model that provide Supply (aka Make) and Use tables for each of the 50 states provides a resource to fill in the gap of needed economic data for state-level models (Li et al., 2022), partially addressing 4F #4. This report further describes the work to provide the associated environmental data, building off recent advances in the USEEIO family of models captured in v2.0 (Ingwersen et al., 2022) along with additional environmental modeling improvements, and combining those with the new state-level economic model to create state-specific USEEIO models for all 50 U.S. states.

4 Model Overview

USEEIO State models are formally environmentally-extended input-output (EEIO) models, using a multi-regional framework, also known as multi-regional input-output (MRIO) models with environmental extensions. The models are also a form of a Life Cycle Assessment (LCA) model and can be provided in formats that are ISO 14048-compliant as well as compliant with the Federal LCA Commons data conventions. The models are two-region models, where a model for a state includes a region of the state of interest (SoI) and another region represents the rest of the US (RoUS). Table 1 summarizes the coverage the models across various aspects. The following sections explain those components in more depth. The details of the Methodology used to derive all the model components is provided in the Appendix and reference publications.

Table 1. Model Coverage

Model Aspect	**Coverage**
States	All 50 States
Years	2012-2020
Sectors (Commodities)	73
Sector Classification	BEA Summary Level Commodities (2012 classification)
Demand Types	Total production, total regional consumption, state only production, consumption of state-only goods
Demand Unit	Current U.S. Dollar
Indicators - Human Health	Respiratory Effects, Cancer from Toxic Releases, Noncancer from Toxic Releases, Impacts of Toxic Releases (combined)
Indicators - Waste	Commercial Hazardous Waste Generation
Indicators - Resource Use	Freshwater withdrawals, Land use
Indicators - Environmental	Greenhouse Gases, Acidification, Eutrophication, Freshwater Ecotoxicity Potential, Ozone Depletion, Smog Formation Potential, Hazardous Air Pollutants
Indicators - Economic/Social	Value Added, Jobs Supported
Commodity life cycle scope	Cradle-to-gate

4.1 Economic Components

The original economic components, namely the input-output tables in a Make and Use format, come from the EPA StateIO models (Li et al., 2022). Models of this form explicitly represent all the interactions between industries, commodities (where a commodity is a good or service), and final users within and between these two regions. They are balanced such that the value of all commodities produced by industries is used by another industry as an input into production, or by a final user, or exported. Thus, production and consumption of all goods and services in both model regions are accounted for. International exports and imports are exogenous in the model; imports are included but they are assumed to be produced and have the same environmental profiles as regional commodities and assumed to be used the same as regional commodities once they cross the border. International exports are included but the destination is unknown.

StateIO two-region models are built for every state in the U.S., with a model for each state and each year from 2012 to 2020, for a total of 450 unique models (50 states x 9 years). Each model has 73 distinct commodities, with the production and use of that commodity uniquely represented in each of the two regions, for a total of 146 commodities. Table 2 lists and categorizes all commodities in a StateIO model present for each region. These commodities are defined according to BEA's Summary level industry data (U.S. Bureau of Economic Analysis, 2019). These BEA codes are either official North American Industry Classification (NAICS) codes at the 2, 3, or 4 digit level from the 2012 code set (U.S. Census Bureau, 2019). For example, '22' is a 2-digit NAICS for 'Utilities' and BEA '111CA' is an aggregation of NAICS codes '111' and '112'). Codes starting with 'G' are government sectors; codes starting with 'F' are final uses, a 'Used' code to represent used goods (of any type), and 'Other' for goods that do not fit within existing classification because they are generally imported and no equivalent goods are made in the U.S. Final uses are by type and include household consumers, private investment, change in inventories, international imports, international exports, federal government, and state and local government uses. Government is further broken down into defense and non-defense. Uses are broken down into consumption expenditures vs. investments for each user. Investments are further broken out into equipment, structures, and intellectual property. Table 3 lists the final use categories present in a model for each region. A complete crosswalk that matches BEA summary sectors to NAICS codes is available as a part of each USEEIO State model excel file (see Model Availability).

Table 2. Model Commodities

Code	Name	Category
111CA	Oilseeds, grains, vegetables, fruits, animal farms and aquaculture	11: Agriculture, Forestry, Fishing and Hunting/
113FF	Raw forest products, wild-caught fish and game, agriculture and forestry support	11: Agriculture, Forestry, Fishing and Hunting/
211	Unrefined oil and gas	21: Mining, Quarrying, and Oil and Gas Extraction/
212	Metal ores, dimensional stone, nonmetallic minerals	21: Mining, Quarrying, and Oil and Gas Extraction/
213	Well drilling and support activities for	21: Mining, Quarrying, and Oil and Gas

Code	Name	Category
	mining	Extraction/
22	Electricity, natural gas, drinking water, and wastewater treatment	22: Utilities/
23	Construction	23: Construction/
321	Wood products (e.g., plywood, veneer)	31-33: Manufacturing/
327	Clay, glass, cement, concrete, and other nonmetallic mineral products	31-33: Manufacturing/
331	Primary and secondary ferrous and nonferrous metals	31-33: Manufacturing/
332	Fabricated metal products (e.g., architectural and structural metal products)	31-33: Manufacturing/
333	Machinery (except computers)	31-33: Manufacturing/
334	Computers and relevant parts, conductors, measuring devices, communication devices	31-33: Manufacturing/
335	Lights and light fixtures, switch boards, transformers, and home appliances	31-33: Manufacturing/
3361MV	On-road vehicles (excluding motorcycles) and accompanying parts	31-33: Manufacturing/
3364OT	Other vehicles (e.g., aircraft, water vessels), missiles, and accompanying parts	31-33: Manufacturing/
337	Furniture and shelving	31-33: Manufacturing/
339	Medical supplies, entertainment and sporting goods, fashion goods, advertising products	31-33: Manufacturing/
311FT	Food and beverage and tobacco products	31-33: Manufacturing/
313TT	Textiles and textile-derived products (except clothes)	31-33: Manufacturing/
315AL	Clothing and leather	31-33: Manufacturing/
322	Paper products and paper production facilities	31-33: Manufacturing/

Code	Name	Category
323	Print media and printing support	31-33: Manufacturing/
324	Petroleum fuels, asphalt, and other petroleum and coal products	31-33: Manufacturing/
325	Agricultural, pharmaceutical, industrial, and commercial chemicals	31-33: Manufacturing/
326	Plastics and rubber products	31-33: Manufacturing/
42	Wholesale trade	42: Wholesale Trade/
441	Vehicles and parts sales	44-45: Retail Trade/
445	Food and beverage stores	44-45: Retail Trade/
452	General merchandise stores	44-45: Retail Trade/
4A0	Other retail	44-45: Retail Trade/
481	Air transport	48-49: Transportation and Warehousing/
482	Rail transport	48-49: Transportation and Warehousing/
483	Water transport (boats, ships, ferries)	48-49: Transportation and Warehousing/
484	Truck transport	48-49: Transportation and Warehousing/
485	Passenger ground transport	48-49: Transportation and Warehousing/
486	Pipeline transport	48-49: Transportation and Warehousing/
487OS	Couriers, messengers, transportation for leisure activities	48-49: Transportation and Warehousing/
493	Warehouses	48-49: Transportation and Warehousing/
511	Media, literature, and software	51: Information/
512	Film and sound-based entertainment	51: Information/
513	Radio, TV, telecommunication	51: Information/
514	Data processing, internet publishing, and other information services	51: Information/
521CI	Monetary authorities, depository and nondepository credit intermediation and related activities	52: Finance and Insurance/
523	Financial investments, exchanges, and advising	52: Finance and Insurance/

Code	Name	Category
524	Insurance agencies, carriers, and brokerages	52: Finance and Insurance/
525	Funds, trusts, and financial vehicles	52: Finance and Insurance/
HS	Housing	53: Real Estate and Rental and Leasing/
ORE	Other real estate	53: Real Estate and Rental and Leasing/
532RL	Renting and leasing of goods, equipment, vehicles, and nonfinancial intangible assets	53: Real Estate and Rental and Leasing/
5411	Legal services	54: Professional, Scientific, and Technical Services/
5415	Computer programming and systems design	54: Professional, Scientific, and Technical Services/
5412OP	Miscellaneous professional, scientific, and technical services	54: Professional, Scientific, and Technical Services/
55	Company and enterprise management	55: Management of Companies and Enterprises/
561	Administrative and support services	56: Administrative and Support and Waste Management and Remediation Services/
562	Waste management and remediation services	56: Administrative and Support and Waste Management and Remediation Services/
61	Educational institutions and services	61: Educational Services/
621	Healthcare professions, laboratories, and ambulances	62: Health Care and Social Assistance/
622	Hospitals	62: Health Care and Social Assistance/
623	Nursing, community, mental health, and substance abuse facilities	62: Health Care and Social Assistance/
624	Child day care, community food services, housing services, and other relief services	62: Health Care and Social Assistance/
711AS	Performing arts, spectator sports, museums, and related activities	71: Arts, Entertainment, and Recreation/
713	Amusement facilities, gambling facilities, resort and recreation facilities	71: Arts, Entertainment, and Recreation/

Code	Name	Category
721	Hotels and campgrounds	72: Accommodation and Food Services/
722	Food and beverage establishments	72: Accommodation and Food Services/
81	Other services, except government	81: Other Services (except Public Administration)/
GFGD	Federal general government (defense)	Other Activities/Government
GFGN	Federal general government (nondefense)	Other Activities/Government
GFE	Federal electric utilities and postal service	Other Activities/Government
GSLG	State and local general government	Other Activities/Government
GSLE	Other state and local government enterprises including transit and utilities	Other Activities/Government
Used	Scrap, used and secondhand goods	Other Activities/
Other	Noncomparable imports and rest-of-the-world adjustment	Other Activities/

Table 3. Model Final Uses

Code	Name
F010	Personal consumption expenditures
F02S	Nonresidential private fixed investment in structures
F02E	Nonresidential private fixed investment in equipment
F02N	Nonresidential private fixed investment in intellectual property products
F02R	Residential private fixed investment
F030	Change in private inventories
F040	Exports of goods and services
F050	Imports of goods and services
F06C	Federal national defense: Consumption expenditures
F06S	Federal national defense: Gross investment in structures
F06E	Federal national defense: Gross investment in equipment

Code	Name
F06N	Federal national defense: Gross investment in intellectual property products
F07C	Federal national nondefense: Consumption expenditures
F07S	Federal national nondefense: Gross investment in structures
F07E	Federal national nondefense: Gross investment in equipment
F07N	Federal national nondefense: Gross investment in intellectual property products
F10C	State and local: Consumption expenditures
F10S	State and local: Gross investment in structures
F10E	State and local: Gross investment in equipment
F10N	State and local: Gross investment in intellectual property products

An example may be useful for understand the kind of information available in the economic components of State models. The model Make and Use tables can reveal, respectively, which industries produce a given commodity, and the uses of a commodity produced in a state.

The OR 2020 Make (V) table shows that the Computer and electronics manufacturing industry (NAICS code 334) produced $11.6 billion in computers and electronics. 13 other industries also produced computers and electronics but their combined output was very small in comparison (< $9 million).

The OR 2020 Use (V) table shows the quantities of these OR produced computers/electronics were purchased and used in production by 30 industries in OR and 30 industries in the RoUS (68 of 71). About $390 million of the OR computers/electronics are used by OR industries and nearly $2 billion are used by RoUS industries. Another $243 million are consumed by households, $5.2 billion as private investment. The largest share of OR computers and electronics, $6.4 billion, are exported internationally. OR also imports $7 billion of computers and electronics from other countries, and OR industries consume $1 billion in electronics from the RoUS and OR final consumers consume $2.4 billion in RoUS computers and electronics.

Data in the Use table also include monetary quantities of what commodities that industries consume and how much value added they provide. The Computer and electronics industry in OR purchases $1.2 billion in OR commodities and about $1.2 billion in RoUS commodities to make computers, and provides $8.35 billion in wages to employees, $252 million in taxes minus subsidies, and generates $6.95 billion in surplus.

4.2 Industry Production Requirements

In the State EEIO model, the industry transaction data from StateIO on (purchases of commodities by industries) data are normalized by industry output to create direct requirements in the form of purchases of each commodity per dollar industry output. Direct requirements are analogous to recipes to make a product. These direct requirements per industry are transformed into direct requirements by commodity using the market share data that is derived from the industry production data in the Make table to get a representative model of what is used to make each commodity.

After the normalization step, OR computer and electronics require about 7 cents in OR commodities and 7 cents in RoUS per dollar output; and another 2-3 cents of imported commodities. The requirements of each commodity from each region plus imports are available in the model. For each requirement from a commodity, the model includes the inputs required to make that commodity – whether it is made in the SoI or the RoUS – in this same table. Those requirements could be considered secondary requirements, as they are the inputs to production to make the inputs for the commodity of interest. From each of those, tertiary requirements can be determined, and so on, ad infinitum. The requirements are all combined into a total requirements table, using a standard input-output analysis technique. The State EEIO Models include versions of the direct and total requirements that include the inputs of internationally imported commodities as well as the SoI and RoUS, and other versions that include only inputs from SoI and RoUS.

4.3 Final Demand Vectors

The Use table also has final uses of commodities by households, inventors, or for export. The quantity of commodities purchased by final users is also known as final demand. The final uses are not included in the direct and total requirements because those are reserved for industries. The final uses are kept as totals that can be used for analysis of the entire economy or segments of it, because the final uses along with a total requirements matrix scaled to provide the final demand is how total use (total = intermediate + final) of all commodities associated with the use of a commodity estimated.;

IO models are final demand-driven models. There are collections of final demand for commodities (also called demand vectors) that represent totals amounts of all commodities consumed under different scenarios. These include:

1. **complete consumption** – consumption of a region (SoI or RoUS) includes international imports but removes international exports;
2. **complete production** –production for a region that excludes international imports but includes international exports;
3. **domestic consumption** – consumption of a region of just domestic commodities (from SoI and RoUS), and;
4. **domestic production** – production in a region only using domestic inputs.

For each of the four cases, the region can be SoI or RoUS, and therefore there are eight collections. The complete consumption vectors can be explained with Equation 1.

$$\mathrm{CompleteConsumption} = \mathrm{C} + \mathrm{I} + \mathrm{G} \qquad (1)$$

where *C* = Consumers (households), *I* = Investment (private business), and *G* = Government.

The *C*, *I*, and *G* terms already include international imports in the complete model.

In the case of just domestic consumption vectors, they remove the international imports (Equation 2).

$$\mathrm{DomesticConsumption} = \mathrm{C} + \mathrm{I} + \mathrm{G} - \mathrm{M} \qquad (2)$$

The production vectors can be explained with Equation 3.

$$\text{Production} = \text{C} + \text{I} + \text{G} + \text{X} - \text{M} + \delta\text{S} \quad (3)$$

where *M* is International imports, and *X* is International exports, and δS is change in stocks or inventories. Note that exports are included, while imports are removed, and the change in stock is included. Both the complete and the domestic production vector are calculated using the same basic approach.

In the model result calculations, the domestic demand vectors should only be used along with the domestic total requirements, and the complete demand vectors should only be used with the complete demand vectors.

4.4 Environmental and Employment Data

The environmental, employment and value added component (compensation, taxes, and gross operating surplus) data are represented by 1,581 unique environmental, resource, waste, job and monetary flows. These data are represented as flows, which are a combination of a substance, an optional environmental compartment/media and a physical unit (e.g., Carbon dioxide/air/kg; Freshwater/ground/kg'). These are primarily chemicals that are released to the environment but also include water use, jobs, and specific hazardous waste types. The flow quantity is given for each industry for each relevant flow (if the industry produces or consumes that flow) and for each model region that emits/releases or consumes that flow. For example, a kg of freshwater withdrawn from the ground is a flow in the model represented as the data object Water, fresh (substance)/resource/water/subterranean/freshwater body (compartment)/kg (unit). The withdrawal of 4.02E11kg of freshwater from the ground by farms in MN in the MN 2020 model is an exchange using that flow. Table 4 lists the environmental data and sources used to make State EEIO models.

Table 4. Primary Environmental Data Sources

Name	Sources	Creator	Years
Greenhouse gas emissions	EPA Disaggregated State Inventories	EPA	2012-2020
Water withdrawals	Water Use in the US	USGS	2015
Criteria and Hazardous Air Emissions	National Emissions Inventory; Toxic Release Inventory	EPA	2014, 2017, 2020;2012-2020
Point source industrial releases to ground	Toxic Release Inventory	EPA	2012-2020
Point source releases to water	Toxic Release Inventory; Discharge Monitoring Report	EPA	2012-2020
Land use	Public Land Statistics; Commercial Building Energy Consumption Survey; Manufacturing Energy Consumption Survey; Major Uses of	BLM; EIA; EIA;	2012; 2012; 2014; 2012

Name	Sources	Creator	Years
	Land in the United States	USDA	
Commercial RCRA-defined hazardous waste	National Biennial RCRA Hazardous Waste Report	USEPA	2013, 2015, 2017, 2019
Employment	Quarterly Census of Employment and Wages	BLS	2012-2020
Value Added	Annual Gross Domestic Product By State	BEA	2012-2020

Table 5 documents the year of each environmental data source for a given model year, which is the year that the economic input-output data represents. Environmental data years generally match the year of the model. However, in some cases, data are not available for the model year. For example, non-point source, off-road mobile equipment, and on-road air emissions data are only available every three years, so a single environmental data year is used for multiple model years. Similarly, primary data for land use and water use are only available for a single year (2012 and 2015 respectively), so all models use the same environmental data.

Table 5. Environmental Data Inputs by Model year

Name	2012	2013	2014	2015	2016	2017	2018	2019	2020
Greenhouse gas emissions	2012	2013	2014	2015	2016	2017	2018	2019	2020
Water withdrawals	2015	2015	2015	2015	2015	2015	2015	2015	2015
Criteria and Hazardous Air Emissions	2014	2014	2014	2017	2017	2017	2020	2020	2020
Point source industrial releases to ground	2012	2013	2014	2015	2016	2017	2018	2019	2020
Point source releases to water	2012	2013	2014	2015	2016	2017	2018	2019	2020
Land use	2012	2012	2012	2012	2012	2012	2012	2012	2012
Commercial RCRA-defined hazardous waste	2013	2013	2015	2015	2017	2017	2019	2019	2019
Employment	2012	2013	2014	2015	2016	2017	2018	2019	2020
Value Added	2012	2013	2014	2015	2016	2017	2018	2019	2020

The following environmental data included in the USEEIO national v1 and v2.0 models are not included, due to the absence of readily-available data at the state level to prepare sector attribution models for these environmental flows:

1. Pesticide losses to the environment

2. Energy use
3. Mineral use

To link the environmental data to the core economic components of the model (the Make and Use tables), the following steps are taken. For all commodities in the model, the totals by sector are normalized by the economic output of that commodity in USD. Since the flow totals data is for industries, first the values are normalized by industry output to form a coefficient. We refer to this coefficient as an exchange coefficient. Then the market shares, or how much of the total output of a commodity that an industry produces, are used to calculate a weighted average of the industry output-normalized flow totals to base the normalized value on commodity output. Expanding on the previous example with freshwater withdrawal in MN, for freshwater withdrawal from the ground by farms in MN, the normalized exchange coefficient is 2.8 kg/USD. This is also referred to a direct intensity, as it represents the direct withdrawal of water from the ground per $ farm commodity produced. The MN model for 2020 has 26,018 exchange coefficients, with 7,385 associated with MN industries or final users, and 18,633 associated with RoUS industries or final users.

4.5 Indicator Data

Each model can provide results for 16 indicators covering potential environmental, human health, resource use, waste, and economic/social impacts. Table 6 includes lists and sources of indicators. These indicators are a subset of those used in the USEEIO v2.0 model. Not all the USEEIO v2.0 indicators are used because the environmental data are not all available to compute results for those indicators.

Table 6. Indicator Data Inputs

Name	Code	Creator	Sources
Greenhouse Gases	GHG	USEPA	TRACI 2.1 (Young et al., 2021a)
Acidification Potential	ACID	USEPA	TRACI 2.1
Eutrophication Potential	EUTR	USEPA	TRACI 2.1
Freshwater Ecotoxicity Potential	ETOX	USEPA	TRACI 2.1
Human Health - Cancer	HCAN	USEPA	TRACI 2.1
Human Health - Noncancer	HNCN	USEPA	TRACI 2.1
Human Health Toxicity	HTOX	NA	Aggregation of HNCN and HCAN
Human Health - Respiratory Effects	HRSP	USEPA	TRACI 2.1
Ozone Depletion	OZON	USEPA	TRACI 2.1

Name	Code	Creator	Sources
Smog Formation Potential	SMOG	USEPA	TRACI 2.1
Freshwater withdrawals	WATR	USEPA	FEDEFL Inventory Methods v1.0.0 (Young et al., 2021b)
Land use	LAND	USEPA	FEDEFL Inventory Methods v1.0.0
Hazardous Air Pollutants	HAPS	USEPA	FEDEFL Inventory Methods v1.0.0
Value Added	VADD	USEPA	USEEIOv1.1 - Elementary Flows and Life Cycle Impact Assessment (LCIA) Characterization Factors (Ingwersen et al., 2017)
Jobs Supported	JOBS	USEPA	USEEIOv1.1 - Elementary Flows and Life Cycle Impact Assessment (LCIA) Characterization Factors
Commercial RCRA Hazardous Waste	CRHW	USEPA	Commercial Waste National Totals by NAICS and US Satellite Tables for USEEIO (Ingwersen et al., 2019)

The indicators consist of characterization factors that relate the given quantities of relevant exchanges to a potential environmental impact based on their potential impact they might have. The exchanges are multiplied in the model with characterization factors, and then the characterization factor values are added together across all the exchanges that have a non-zero value using a common denominator of a dollar. The result is an impact coefficient, given in a value of an indicator per dollar (e.g., kg CO2e/$ for Greenhouse Gases (GHG) or kg freshwater/$ for freshwater withdrawal). For the environmental impact indicators, these characterization factors come from the EPA Tool for Reduction and Assessment Chemical Impacts (TRACI) life cycle impact assessment methodology (Bare, 2015), a collection of characterization factors for different environmental impact categories. Resources, waste, jobs, and value added characterization factor values are all just 1 and used to sum different freshwater types, different waste types, etc. The environmental flow and indicator total intensities (direct + indirect) are estimated by combining the total economic requirements with the flows or the indicators. This results in flow or indicator amounts for each commodity per dollar commodity output that reflect the direct + indirect flows per dollar or indicator value per dollar, also referred to as direct + indirect flow or impact coefficients.

4.6 Adjustment Factors for Price Type and Year

All coefficient values have dollar denominators. These denominators are in the currency value of the given model year and in producer price. Producer price is equivalent to the cost of production. It differs from purchaser prices, which includes the costs plus the margins which are the additional costs associated with transportation, wholesale and retail, and the price type that final consumer generally sees on items. Coefficients in the model can be converted into other dollar years using commodity-specific adjustment values (also known as deflators), and similarly, all coefficients can be converted from producer price to purchaser price using commodity- and year-specific adjustment values. These adjustment values are available for all commodities from years 2002 to 2021.

4.7 Result Calculations

Results calculated with the provided demand vectors are not provided, but the calculation functionality where results can be provided with these or custom demand vectors is present in selected model formats.

The basic approach of a result calculation is to scale up a coefficient matrix of direct + indirect impacts by a final demand vector (Equation 4).

$$Result = Directandindirect(total)impactorflowcoefficients * demand \quad (4)$$

Depending on ordering of the components and operations (detailed in Ingwersen et al. (2022)), results can be provided that associate the impacts with the commodities where the direct emissions/resource use/waste generation occurred, or results can be associated with the final products consumed. The former is considered the direct perspective and the latter the final perspective. The direct perspective is useful to identify what industries (or more properly the commodities they produce) create the pollution associated with what is finally consumed, in other words it shows the source of impacts within the supply chain. The final perspective is useful to associate the impacts with the final products and their uses that are driving the impacts.

4.8 Components as Matrices and Vectors

A USEEIO State model is a collection of these components in the form of tables of data and metadata as well as matrices that are available with other USEEIO models. The most commonly used matrices and vectors (a single column matrix) are identified in Table 7. Detail on all components can be found in the *useeior* Model format specs. The availability of all model components and information depends upon the format through which the model is accessed, which is described in the next section.

Table 7. Model Matrices and Vectors. Phy = a physical flow unit (e.g., kg).

Symbol	Name	Indices	Units	Dimensions
V	Make matrix	industries x commodities	$	142 x 146
U	Use matrix	commodities + value added x industries + final uses	$	152 x 182
U_d	Domestic use matrix	commodities + value added x industries + final uses	$	152 x 182
A	Direct requirements matrix	commodities x commodities	$/$	146 x 146
A_d	Direct domestic requirements matrix	commodities x commodities	$/$	146 x 146

Symbol	Name	Indices	Units	Dimensions
B	Exchange coefficient matrix (Satellite) matrix	flows x commodities	Phy/$	1581 x 146
C	Characterization factor matrix	equivalencies x flows	Equivalency/Phy	16 x 1581
D	Direct impact coefficients	equivalencies x commodities	Equivalency/$	16 x 146
L	Total requirements matrix	commodities x commodities	$/$	146 x 146
L_d	Total domestic requirements matrix	commodities x commodities	$/$	146 x 146
M	Direct + indirect flow coefficients	flows x commodities	Phy/$	1581 x 146
M_d	Direct + indirect domestic flow coefficients	flows x commodities	Phy/$	1581 x 146
N	Direct + indirect impact coefficients	equivalencies x commodities	Equivalency/$	16 x 146
N_d	Direct + indirect domestic impact coefficients	equivalencies x commodities	Equivalency/$	16 x 146
&Phi	price type adjustment matrix	commodity x year	$/$	146 x 20
&Rho	currency year adjustment matrix	commodity x year	$/$	146 x 20
q	Commodity output	commodities	$	146
x	Industry output	industries	$	142
y	a final demand vector (8 total)	commodities	$	142

The shape of the two-region models is different from the single region national USEEIO models in that wherever a matrix index (rows or columns) has commodities, value added, or final demand components, they are listed twice, one for each region. Therefore, all matrices with such indices that are regionally-specific are at least twice as large as the national model matrices, and for the economic matrices where both rows and columns are indexed with regionally-specific data (e.g., V, U, A, L), they are four times as large. In all such cases, the SoI always appears first in the index, followed by RoUS.

4.9 Life Cycle Stages Covered

For users familiar with LCA or more general life cycle thinking associated with goods and services, the USEEIO State models provide life cycle results for goods and services produced and consumed in U.S. states. However, the model results only cover what is commonly considered "cradle-to-gate" data. The models do incorporate data on pre-consumer transport, wholesale, and retail activities, as well as what may be considered the use phase, which is equivalent to the final users' activities, and the end-of-life stage, which may be a waste management or recycling activity, but those phases are not directly associated with the commodities in the models. The USEEIO-based Margin Emission Factors which are part of the Supply Chain Factors (Ingwersen and Li, 2020a) are based on a methodology that uses model data to incorporate the transportation to consumer, wholesale and retail stages, and the same methodology could be used with the state models to incorporate data on these phases, but those additions are not available with the v1.0 models.

5 Model Use

The use of USEEIO State models is like other USEEIO and EEIO models more generally. This section provides some ways the model data can be used. Generally, the models can be used to evaluate the quantities of impacts associated with categories of goods and services produced or consumed within a U.S. state, or the large collections such as the entire production or consumption of a state. Users can extract or trace economic data on commodities and industries through the output vectors (q and x) or Make (V), Use (U, U_d), Direct Requirements (A, A_d) and Total Requirements (L, L_d) matrices. These data can inform the make and uses of commodities within the SoI and RoUS regions as well as the flow of commodities between regions and internationally. Advanced users perform advanced analysis, such as regional economic impact analysis, using these matrices. Users interested in environmental and economic impacts results can inspect and use values in the coefficient matrices (B, D, M, M_d, N, N_d), which are composed of environmental intensities per dollar that are either direct or indirect. Users can perform calculations to estimate potential impacts of the final consumption of goods and services either for individual commodities, or by collections of commodity amounts that represent scenarios or regional totals. The collections might be the provided final demand vectors or using custom collections of commodity amounts consumed (custom final demand vectors). Calculations can be performed using the models through *useeior*, the USEEIO API, useeio.js, or in the users tool of choice. See a note on EPA planned future additions to and uses of USEEIO State models in the Future Work section.

6 Model Validation

To confirm that the two-region state models are built correctly, a set of validation procedures are implemented to assess the economic and environmental values present in the models. These validation procedures were performed for all state models for all years reflected in this report (2012-2020). The full validation results are available are available upon request or users can build the models and run the validation functions using the source code (see Model Availability).

The main economic validation consists of checking that the sector (commodity) totals in each state model can be recreated using the Total Requirements matrices (L, L_d) and appropriate final demand vectors. This is similar to how the national models are validated (Ingwersen et al., 2022), with the main difference being that the final demand vectors need to be adjusted to account for both in-state and out-of-state demand, to comply with the two-region nature of the models. For the checks to pass, the validation calculation result must be within 1% of the original commodity totals. Out of over 87,000 economic validations performed (146 sectors for 50 states over six years for both domestic and non-domestic models), only 16 checks failed. All of the economic failures were related to sector 211, Unrefined oil and gas, for specific states (e.g., 211/US-MD), with no failures related to RoUS sectors. These failures are caused by the relatively low 211 commodity totals when compared to the totals of other sectors. The low 211 values result in differences between the original commodity totals and model calculations being larger than 1% of the total for that sector, when comparisons for other sectors which result in a similar absolute difference between the original commodity totals and calculated values being smaller than the 1% threshold. More generally, the small commodity totals for the 211 sectors for some states can be understood as a proportionally small amount of oil and gas extraction for specific states when compared to national and other states' totals.

For the environmental results, similar validation checks are performed. Specifically, we verify that the environmental flow totals by sector can be recreated based on the model satellite matrix (B), the Total Requirements matrices (L, L_d), and the appropriate final demand vectors. These validation result for flows must also be within 1% of the total flows for the checks to pass. Out of all the checks performed, about 6,000 individual flow-sector combinations failed, which is less than a 0.0001% failure rate. Most of these failures are related to the failure of the 211 sector in various states. In summary, since the commodity totals are used as denominators for creating environmental emissions factors, the low state commodity totals for this sector results in an inflated emissions value which ultimately results in a total flow amount that surpasses the 1% tolerance value for the validation check. Failures for flows related to other sectors can also be due to similar mismatches between the original commodity totals and the calculated commodity totals passed the economic validation but failed the flow validation due to the addition of the B matrix coefficients.

7 Initial Findings

In this section we present a selection of results featuring models from all states to reveal initial findings from the USEEIO State models. For figures featuring results for many commodities, we use the following colors to the group them into broader categories to facilitate interpretation Table 2:

- GREEN = 11, Agriculture, Forestry, Fishing and Hunting
- BLACK = 21, Mining, Quarrying, and Oil and Gas
- GRAY = 22, Utilities
- LIGHT GRAY = 23, Construction
- RED = 31-33, Manufacturing
- BLUE = 42, Wholesale
- LIGHT BLUE = 44-45, Retail
- ORANGE = 48-49, Transportation and Warehousing
- MUSTARD = 51, Information
- PURPLE = 52-53, Finance and Insurance and Real Estate
- LIGHT PURPLE = 54-56, Professional, Scientific, Technical, Management and Administrative Services
- BROWN = 61-62, Educational and Healthcare Services
- GOLD = 71-72, Arts, Recreational, Accommodation and Food Services
- PINK = 81, Other Services
- LIGHT GREEN = Government Services, Used and Other Goods

Figure 1 shows the cradle-to-gate impacts per dollar commodity produced in each sector across all 50 states for three selected impact categories: acidification potential, global warming potential, and human health respiratory impacts. These results are drawn from the 2020 models for each state.

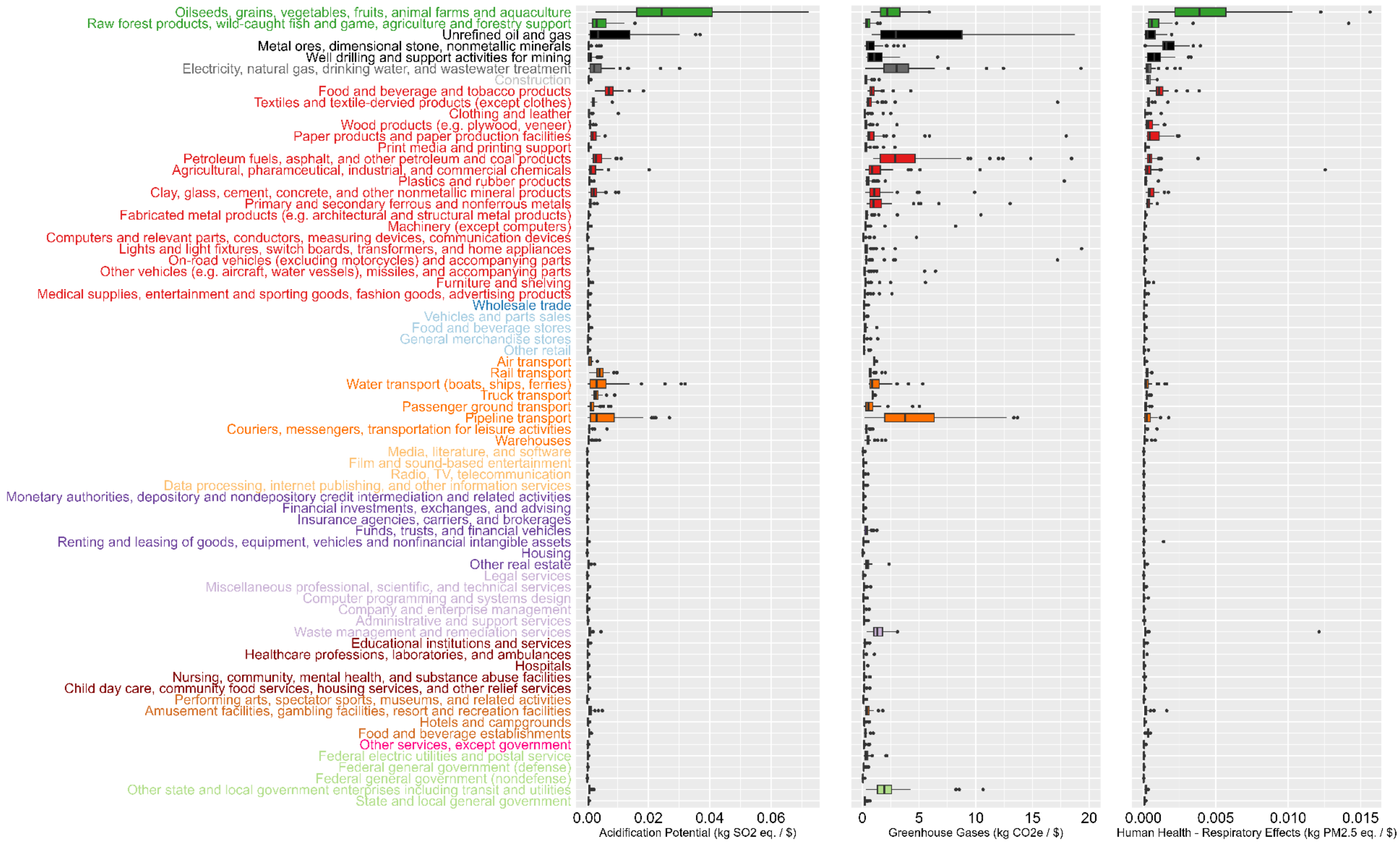

Figure 1. Range of exchange coefficients in indicator units per dollar for Acidification Potential (ACID), Global Warming Potential (GHG), and Human Heath - Respiratory (HHRP) Impacts by State for year 2020. The ranges are wider for agricultural, oil and gas, utilities and water and pipeline transport than for other commodities. The ranges are wider for some indictors like GHG, than for others, like HHRP.

Some sectors show significant variability in intensities across states. For example, in the Oilseeds, grains, vegetables, fruits, animal farms, and aquaculture sector (111CA), the highest acidification potential by a state is 0.07 kg of SO2 equivalent (eq.) per dollar (AL), while the lowest is 0.003 (HI). This broad range is partly due to the variable livestock management practices and differences in crops planted by state, resulting in varying ratios of emissions with acidification potential to the total economic output by state. While AL has a total economic output for the 111CA commodity of 2.4 billion dollars and total ammonia emissions to air of 80 million kg, HI has an economic output of 770 million dollars and ammonia emissions of 235 thousand kg, respectively. In AL, these emissions result from activities related to (in decreasing order) poultry, fertilizer application, and beef cattle, while in HI, the primary driver of ammonia emissions to air is goat farming. These values result in a ratio of ammonia emissions to air to total economic output of approximately 0.03 kg/USD for AL, while the same ratio in HI is approximately 0.0003 kg/USD, causing the high variability shown for this sector. Similarly, greenhouse gas emissions per dollar from Electricity, natural gas, drinking water, and wastewater treatment (22) show high variability likely resulting from the distinct electricity fuel mixes across the US, while PM emissions per dollar also vary considerably by state and commodity.

We render the same cradle-to-gate GHGs per dollar commodity produced result (as in middle figure above) in a map format for two commodities to show the spatial variation of CO2e per dollar commodity across the U.S. in Figure 2. We highlight the GHG emission intensities of agricultural products like raw grains, fruits, vegetables, and animal products, which as mentioned in the previous figure, showed a high variation. In this figure the higher intensities are found in south central and upper mountain region states that are more associated with animal production or crops with potentially higher GHG impacts, like rice (AR). The emissions of associated with Nonmetallic mineral products which include cement, concrete, ceramics, and glass are also presented. Here the variation is less uniform across regions as higher intensity states are scattered. These figures highlight that emissions intensities can vary considerably both between across states and sectors.

Figures with these maps for each indicator and each commodity for 2020 can be found on figshare - see Model Availability.

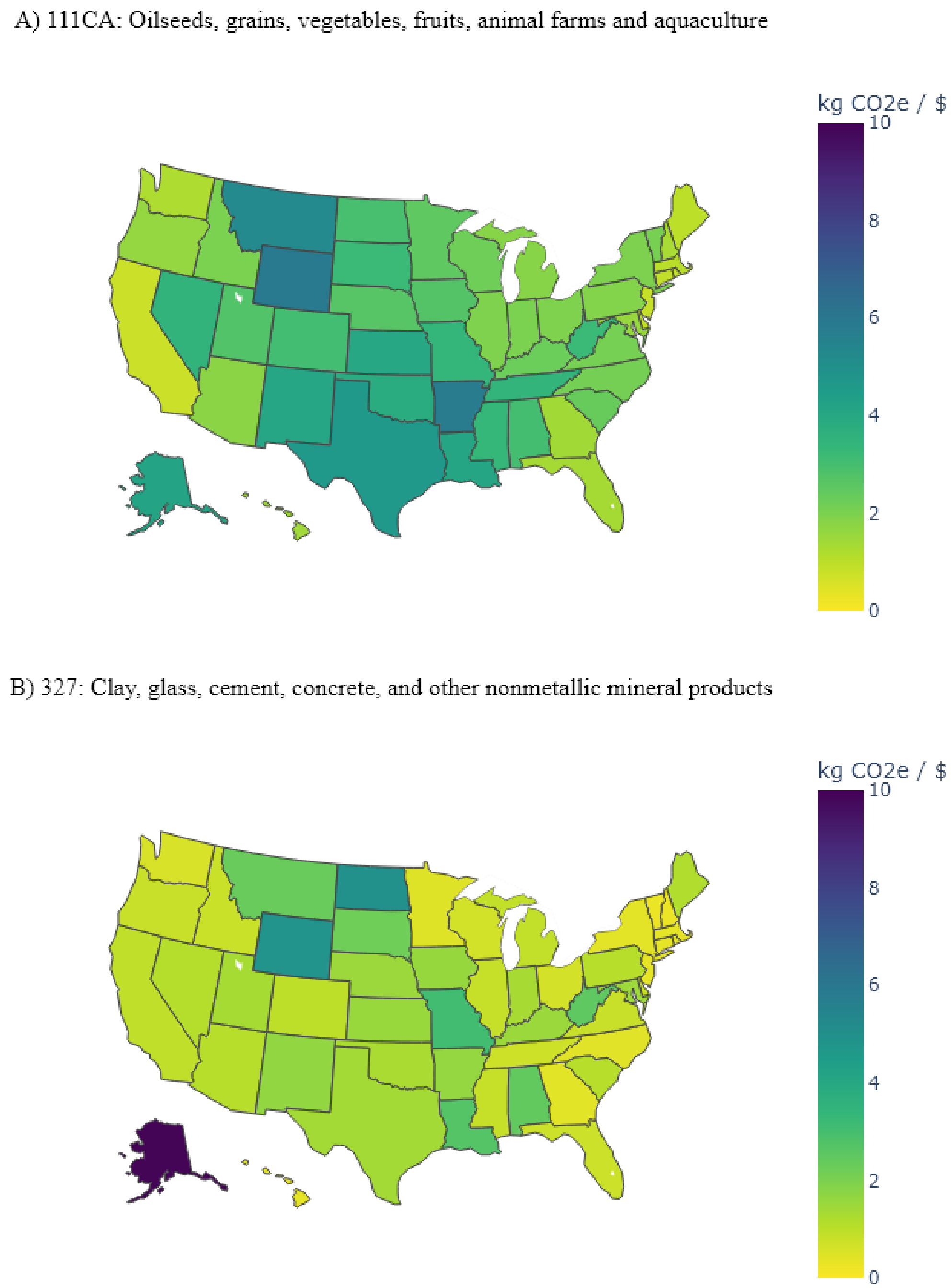

Figure 2. Intensity of cradle-to-gate GHG emissions per dollar produced by state for (a) Oilseeds, grains, vegetables, fruits, animal farms and aquaculture (111CA) and (b) Nonmetallic mineral products (327)

8 Example Results for a State of Interest

In this section, we provide examples of several types of results that users might want to explore for their SoI. Each result can be derived from a single state model for a given year (or multiple years for the Time Series). For each result, we use one or more of the partner states (see Purpose and Overview) as the SoI(s) for demonstrative purposes, and for results using a single indicator, we use GHGs. These results can be calculated for any of the 50 states and for any of the indicators present in Table 6. Separately we share figures made with all indicators and for all states. We also provide links to source code to enable advanced users to generate similar results for their SoI and year(s) of interest, which will also enable users to modify the format or style of the figures or access data underlying the figures. See the Model Availability section for links to figures and source code.

8.1 Sector rankings by state

The ranking of sectors by impacts for the four sample states based using two model calculation variants are shown in Figure 3. The first calculation uses the consumption demand vector and the final perspective, and the second calculation uses the production demand vector and the direct perspective. The ranking procedure used the same algorithms for ranking sectors across indicators that is described in the USEEIO v2.0 paper (Ingwersen et al., 2022) equations 34 & 35. For each indicator, the impacts of purchases by each state from each sector are ordered and ranked, helping identify sectors that have exceptionally high or low impacts relative to other indicators. Sectors are ordered for the figure by highest overall rank across all indicators, with all indicators weighted equally. For example, purchases of *Food and beverage and tobacco products* have the highest overall impacts for all states; however, the contributions to total impacts differ by indicator for each state. For all states, the Acidification Potential indicator is the indicator that contribute the most to the impacts caused by *Food and beverage and tobacco products* purchases. However, for OR and GA, the Water Withdrawals and Land Use indicators are the next two top contributors; for WA, it is the Water Withdrawals and Human Health - Respiratory Effects; and for MN, it is Land Use and Human Health - Respiratory Effects. In addition, the sector with the second most overall impacts is not the same for all sample states. For OR, WA, and MN, the *Construction* sector has the second highest overall impacts, while for GA the *Federal General Government (defense)* sector has the second highest impacts. Additionally, rankings also change when considering the production demand vector, using the Direct perspective. With this perspective, it is the *Farms* sector that has the highest overall impacts for all four sample states followed by *Chemical products*. These results by state can also be compared to national level results calculated from a USEEIO v2.0 summary sector model presented in Figure 2 of Li et al. (2022).

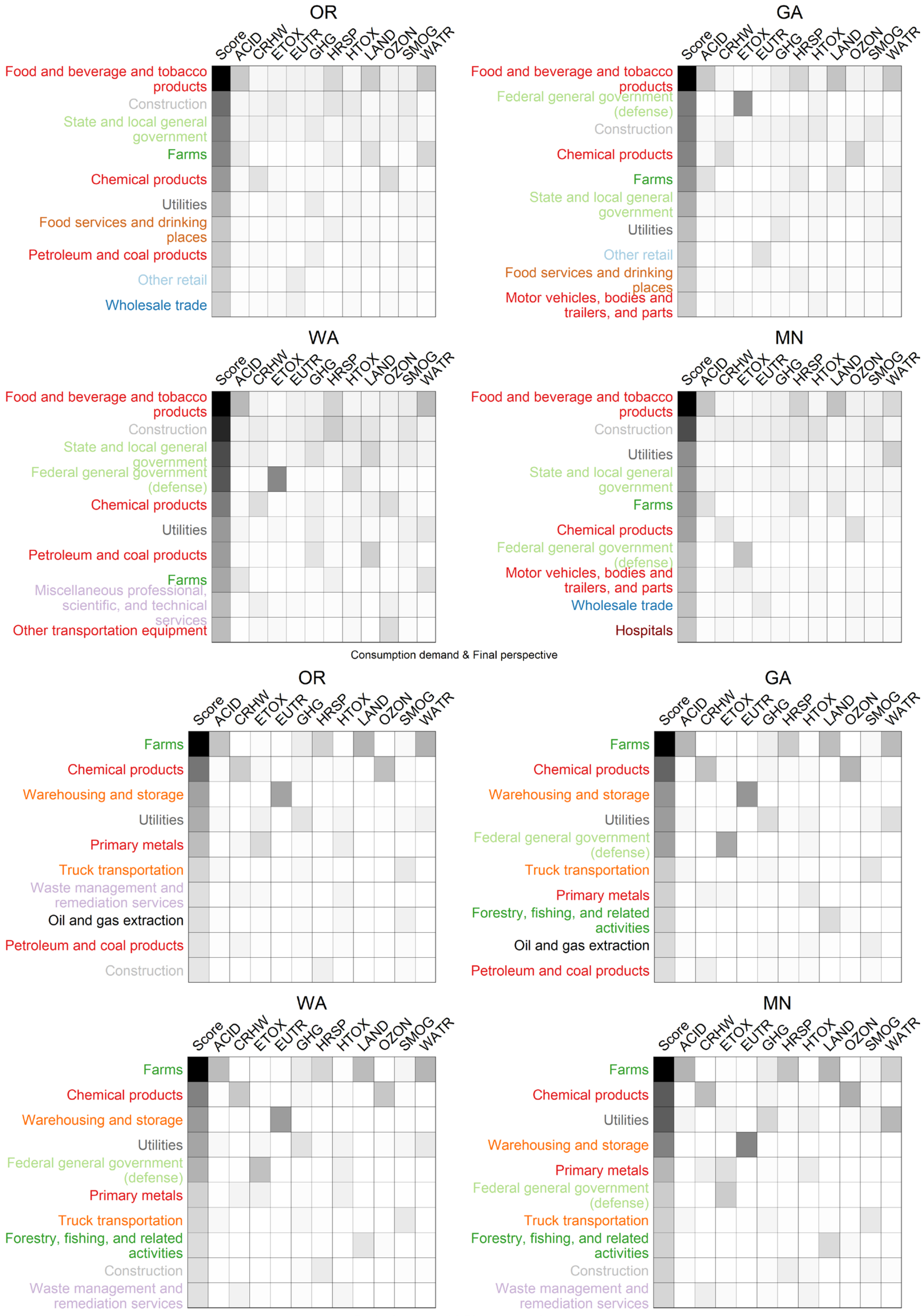

Figure 3. Sector rankings, from higher (top) to lower (bottom) impact, based on composite scores across environmental indicators for select states, for (a) Consumption demand & Final perspective and (b) Production demand & Direct perspective. The darker

the square, the higher the impact for that sector relative to other sectors for the indicator in the column. All sectors not shown (63 sectors) have less impact. Indicator acronym definitions can be found in Table 6. Indicator Data Inputs.

8.2 Impacts By Region

USEEIO State models are capable of tracking locations of impacts or associated emissions/resource use by the two-model regions (SoI, RoUS) as well as the Rest of the World (RoW). Figure 4 reflects the emissions by sector for the consumption demand vector in Minnesota from a Direct Perspective. The Direct perspective describes the location of the environmental impact resulting from purchases of all goods by the SoI. Based on this perspective, 62.9% of GHG emissions in Minnesota are generated domestically. Alternatively, emissions from *Oilseeds, grains, vegetables, fruits, animal farms and aquaculture ("111CA")* from RoUS and RoW are significantly greater than those from Minnesota; 68.1% of embedded GHG emissions are from outside the state.

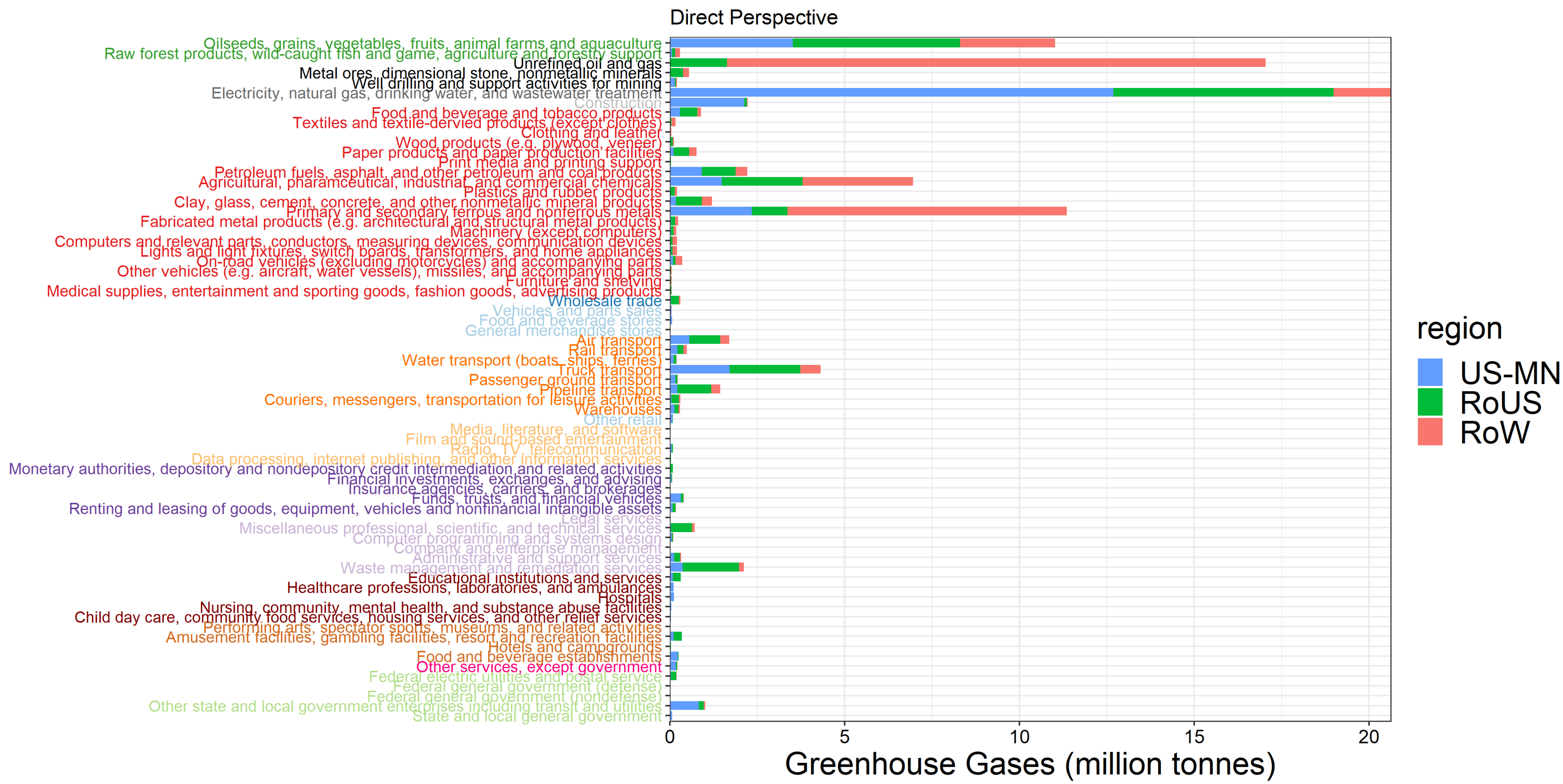

Figure 4. GHG emissions by commodity consumed in Minnesota showing region in which emission occurred. Note the RoW portions are approximated here based on domestic intensity for the commodity scaled based on quantity imported. See the Model Limitations and Next Steps sections for related plans for updates to improve RoW impact estimation.

8.3 Time series

Another potential application for the USEEIO State models is to track the evolution of results for individual indicators across several years. This type of analysis can be used to see rapid changes in emissions due to changes in consumption patterns or to evaluate the improvements in emissions controls for specific sectors across time. Figure 5 shows the GHG emissions for the state of Washington between 2012 and 2020, aggregated to the BEA Sector level (as indicated in the figure legend) from a direct perspective using the production demand vector. In this figure, Summary level sectors resulting in net negative emissions were dropped. This was only the case for the *Unrefined Oil and Gas (211)* sector, for which WA had a net negative level of final demand due to imports.

The values in this figure represent emissions resulting from production by in-state industries and governments of both in-state and out-of-state commodities. The figure shows varying level of GHG emissions by year, from approximately 96 million metric tons (MMT) CO2eq in the year with lowest emissions (2020) to over 118 MMT CO2eq in 2014. A main driver for the variability in emissions in WA is the emissions related to Utilities. Manufacturing has seen a steady decrease since 2014, while Professional and business services have seen an upward trend.

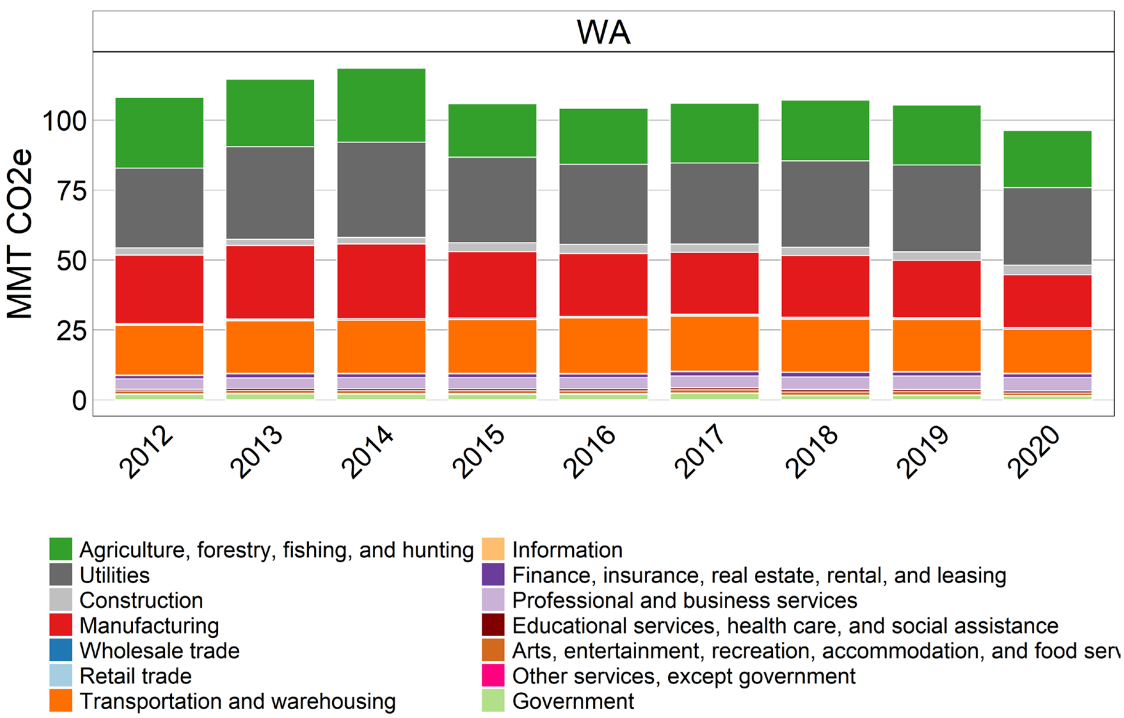

Figure 5. Time series for GHG emissions for Washington. Annual emissions are broken out by commodity category.

9 Comparison of State v1.0 models to the National-level USEEIO v2.0 models

USEEIO State models are built similarly to the national USEEIO v2 models. The models have the same components and can produce similar results as described in Ingwersen et al. (2022) and are available in the same formats. However, there are a number of notable differences.

Representation of commodities and final demand categories is more aggregated (less resolved) in the State models. State models have 73 commodities per region. The coverage of the economy is complete, but these commodities are aggregations of one or more commodity from the national models.

v1.0 of the State models represents a full time-series where economic data are unique to the model year and go up to 2020, whereas for national v2.0 model, only one year of economic data was present (2012) and it is mixed with more recent environmental data (e.g., 2016 for GHGs).

Environmental data for pesticides, energy use, mineral use, commercial solid waste generation are not present in the State models, and associated indicators for energy use, mineral use, and commercial solid waste are therefore not present.

.

10 Model Strengths and Associated Opportunities

The USEEIO State models are the first U.S. state EEIO models to be made widely and freely available. Purely economic input-output models have been available in the U.S. at state or sub-state level, but until very recently they have been proprietary, or former public efforts like the Bureau of Economic Analysis (BEA)'s Regional Input-Output Modeling System (RIMS) models are no longer free, and generally regional level IO models do not cover the environmental aspects that the USEEIO State models cover. Indicators related to environmental health in many areas including air and water pollution, water and land use, climate, freshwater ecosystems, human health, and hazardous waste generation, along with more traditional measures of employment and value-added, are all available in USEEIO State models. Therefore, they can be used to show tradeoffs across various environmental and economic results that are not just one-dimensional.

The USEEIO models are built primarily using methodologies that are well-established in the scholarly literature. Additional nuances in USEEIO models along with the full version 1.0 and 2.0 national models, most of which are incorporated into the USEEIO State models, have been peer-reviewed (Ingwersen et al., 2022; Yang et al., 2017). The StateIO models that provide the underlying economic data have been peer-reviewed in an international scholarly journal (Li et al., 2022). The models are furthermore developed with open-source software that has been through a peer-review process and that incorporate testing to verify proper functionality (Birney et al., 2022; Li, Mo et al., 2022; Young et al., 2021c).

EEIO-based models that include the full economy have some advantages over models traditionally used in LCA to evaluate or compares goods and services. Namely, they provide a full network of the economy to include all goods and services and are based on an approach that enforces a balance between the supply and use of commodities at a regional scale, and thus there are no arbitrary cutoffs that result in not including service inputs or inputs of goods that might be considered irrelevant or for which data are not available. Furthermore, the use with the provided final demand vectors allows for estimations of regional totals that include indirect effects but prevent double-counting which is often a risk in such models. The results of different commodities within the models are highly comparable because of the same assumptions and data are being used to construct them. For USEEIO models in particular, another advantage is that data sources used of emissions, resource use, waste generation or employment that are considered authoritative or represent collections of reports that are formally provided by states to EPA or that EPA collects as part of a regulatory program. Additionally, the USEEIO models can be validated such that all reported environmental data at regional or national levels that is used within a model can be recalculated by summing the totals that are directly or indirectly produced though final consumption. Achieving such validation is generally not possible with standard models used for life cycle assessment.

All aspects of the USEEIO State models, like with the national model, are based on public data, and the data are all obtained, further prepared, and used in models that run on publicly-available source code with a minimally-restrictive, open-source license (MIT license). Therefore, all those with expertise to understand and operate the source code have the ability to deeply inspect and evaluate the treatment of all data, the effects of model assumptions and verify any model results that depend only on the model data. The source control and collaboration platform on which the source code is maintained, GitHub, also enables code-related communication and a system for external contributions following review and testing. This source code availability, collaboration platform, and associated documentation, could provide external users with the ability to improve upon the models or create alternative version of the models using different datasets, assumptions, etc. Links to all source code repositories are provided in Appendix B - Source Code Overview and Links.

11 Model Limitations

Any computational model is limited in its ability to provide useful insight. EEIO models as well as LCA models have recognized limitations for evaluating environmental impacts associated with goods and services. Some particular limitations of results associated with the listed uses are:

- The “recipes” for making commodities in SoI or RoUS, coming from the StateIO models, are originally based on the national “recipe” included in the Use table. They are only altered to balance the commodity and industry output of an industry within a state if this balance is not identical to the national model.
- International imports are modeled as if produced in the U.S. This limitation of USEEIO national models is discussed in the USEEIO papers (Ingwersen et al., 2022) and Supply Chain Factors report (Ingwersen and Li, 2020b). See the Future Work section for a note on plans to address this limitation.
- Interstate trade data is not officially collected and reported in the U.S. Trade in goods is estimated using models based on commodity flow surveys, and trade in services is based mainly on assumptions on their local vs non-local nature and the relative size of service industries across different states.
- Environmental data are often estimated based on reports from years prior to the model year, and often not reported at the level of industry resolution used. Therefore, models that make simplistic assumptions, to provide greater level of resolution, such as using employment to assume that water use of specific manufacturing industries can be used to allocate water use estimates across manufacturing industries. The temporal and technological accuracy of environmental data varies significantly across the data types. The sector attribution model (described and referenced in the appendices) provide information on data years and methods for attribution to industries.
- Environmental data or other data represented by flows that (emissions, resource use, wastes, and employment) are linked to the model through industry economic output, to create exchange coefficients. However, the flow totals and industry output are independently estimated. This presents a risk of disconnects when environmental and economic models do not produce comparable data. For example, if the environmental dataset used in the model has reports of direct GHG emissions for the Talc mining industry in Wyoming, but the economic data show no economic output (or very low) of Talc mining in WY, this would result in an invalid (or impossibly high) coefficient of GHGs/$ talc mined for WY. Furthermore, the environmental data for different types of flows (e.g., hazardous air pollutants vs. water use), come from independent sources that are compiled separately, leaving opening the possibility of different characterizations of industry environmental performance across environmental datasets, due to differences perhaps in classification or lack of required reporting for certain sectors in certain states for an emission type.
- The characterization factors used in estimating environmental impacts are not based on models that estimate damage to final human health and environmental endpoints, but rather midpoint models that address potential stress or potency. Any chemical fate and transport models associated with characterization factors in v1.0 are not state- or regionally-specific but rather represent averages of fate, transport, potential exposure, and impact across larger areas like the U.S. or Europe.
- The results from current calculations do not include the impact of final users, including households, when calculating a total impact for a given final demand.

- The USEEIO State models currently lack some elementary and waste flows as well as indicators that are present in the national USEEIO v2.0 model (U.S. EPA, 2020). Indicators that are not present in these v1.0 USEEIO State models because the associated flows are not present include commercial construction and demolition debris (CCDD) or commercial municipal solid waste (CMSW), mineral use (MRNL), energy use (ENRG) including renewable (RNRG) and non-renewable (NNRG) and pesticides (PEST).

12 Future Work

EPA plans to continue to develop USEEIO models for various applications and the associated environmental and economic datasets using the source code that is maintained on GitHub. This will enable State USEEIO models to continue to be updated with more recent and accurate environmental and economic data. The EPA plans to use the USEEIO State models to conduct consumption-based GHG emissions inventories (CBEI) for northeastern U.S. states. For the CBEI work the EPA will disaggregate the current utilities industry and commodity into electricity, water, and natural gas commodities, which will enable greater insight into the potential impacts related to these commodities. The CBEI will include foreign import emissions modeled by region of origin using external EEIO models with global coverage by deriving import factors, which will provide more accurate estimates of these emissions and address a limitation described in this report. It will also include household and other final users direct GHG emissions in the CBEI results. The disaggregated Utilities will become available for future USEEIO State models. The import factors and the inclusion of final users' emissions in result totals will be available for use both in USEEIO State models and in national-level models in the future. Further disaggregation of other model sectors is a longer-term potential improvement given interest and data availability.

13 Model Availability

The USEEIO State models v1.0 are available in some formats/distributions that have already been prepared and shared, while for others, availability is by request. EPA may continue to make the post the models in additional formats. Interested users are encouraged to email useeio@epa.gov to request addition to the mailing list, or to follow updates on the EPA's USEEIO web pages.

13.1 Published, openly-available distributions

The models are available, at a minimum in the following formats and sources:

1. Spreadsheets for 2020 models for each of the 50 states, stored as Microsoft Excel XML-based files (xlsx) are available: USEEIO State Models v1.0 for 2020.
2. useeior Model format, stored as R Data Serialization files (RDS), with one file per year that includes models for all states. See the EPA Data Commons USEEIO-State folder.

13.2 Result Figures

The result figures in this report in addition to figures showing results for other indicators and sectors not included here are available in the USEEIO State Models v1.0 for 2020 - Supporting Figures collection on figshare.

13.3 Distributions available upon request

EPA will provide users with these additional formats or versions of the models, through requests to useeio@epa.gov, given that resources are available to fulfill the request:

1. .xlsx spreadsheet format for one or more state models within the 2012-2019 time range.
2. openLCA schema zip archive. See the USEEIO v2.0 model on the Federal LCA Commons as an example.
3. JSON objects from the USEEIO API.
4. JavaScript widgets for integration of widgets into web pages via the useeio-widgets.

13.4 Source Code to Build and Compute Examples

The *useeior* model specification files and example code to build models and produce results is available at USEEIO-State on USEPA github.

13.5 Other EPA Outlets

EPA may later make the models available through the Sustainable Materials Management Prioritization Tools Suite.

14 References

Bare, J., 2015. Tool for Reduction and Assessment of Chemicals and Other Environmental Impacts (TRACI) [WWW Document]. URL https://www.epa.gov/chemical-research/tool-reduction-and-assessment-chemicals-and-other-environmental-impacts-traci

Birney, C., Young, B., Chambers, M., Ingwersen, W., 2023. FLOWSA v2.0.0.

Birney, C., Young, B., Li, M., Conner, M., Specht, J., Ingwersen, W., 2022. FLOWSA: A python package attributing resource use, waste, emissions, and other flows to industries. Applied Sciences 12, 5742. https://doi.org/10.3390/app12115742

Bureau of Economic Analysis, 2019. BEA summary import matrices before redefinitions.

Census of agriculture 2017 (No. AC-17-A-51), 2017. U.S. Department of Agriculture.

Commercial buildings energy consumption survey 2012, 2016. U.S. Energy Information Administration.

Discharge Monitoring Report (DMR) Pollutant Loading Tool, 2018. U.S. Environmental Protection Agency.

EIA, 2021. Manufacturing energy consumption survey (MECS) of 2018.

Georgia Center of Innovation, 2023. Georgia energy technology and development [WWW Document]. URL https://www.georgia.org/center-of-innovation/areas-of-expertise/energy-tech (accessed 8.17.2023).

Ingwersen, W., Garvey, T., Bergmann, M., 2017. USEEIOv1.1 - Elementary Flows and Life Cycle Impact Assessment (LCIA) Characterization Factors. https://doi.org/10.23719/1368541

Ingwersen, W., Heyns, L., Thomas, V., 2020. Industry Comparison Tools for Sustainable Communities.

Ingwersen, W., Li, M., 2020a. Supply chain greenhouse gas emission factors for US industries and commodities. https://doi.org/10.23719/1517796

Ingwersen, W., Li, M., 2020b. Supply Chain Greenhouse Gas Emission Factors for US Industries and Commodities (No. EPA/600/R-20/001). U.S. Environmental Protection Agency.

Ingwersen, W., Li, M., Meyer, D., 2019. Commercial Waste National Totals by NAICS and US Satellite Tables for USEEIO. https://doi.org/10.23719/1503688

Ingwersen, W., Li, M., Young, B., Vendries, J., Birney, C., 2022. USEEIO v2.0, the US Environmentally-Extended Input-Output Model v2.0 (USEEIOv2.0). Scientific Data 9, 194. https://doi.org/10.1038/s41597-022-01293-7

Ingwersen, W., Meyer, D., 2016. Sustainable materials management GA model - preliminary results.

Inventory of U.S. Greenhouse gas emissions and sinks: 1990–2020 (No. EPA 430-R-22-003), 2022-04-22. U.S. Environmental Protection Agency.

Li, M., Ferreira, J.P., Court, C.D., Meyer, D., Li, M., Ingwersen, W.W., 2022. StateIO - Open Source Economic Input-Output Models for the 50 States of the United States of America. International Regional Science Review 016001762211458. https://doi.org/10.1177/01600176221145874

Li, M., Ingwersen, W., Ferreira, J.-P., Court, C., Meyer, D., 2023. Stateior. Zenodo. https://doi.org/10.5281/ZENODO.8221432

Li, Mo, Ingwersen, Wesley, Young, Ben, Vendries, Jorge, Birney, Catherine, 2022. Useeior: An Open-Source R Package for Building and Using US Environmentally-Extended Input-Output Models. Applied Sciences 12, 4469. https://doi.org/10.3390/app12094469

Manufacturing energy consumption survey (MECS) of 2014, 2017. U.S. Energy Information Administration.

National emissions inventory 2017, 2019. U.S. Environmental Protection Agency.

OR DEQ, 2012. Materials management in oregon: 2050 vision and framework for action (No. 12-LQ-036). Department of Environmental Quality.

Program, U.S.G.C.R., 2023. Fifth national climate assessment. U.S. Global Change Research Program, Washington, DC, USA. https://doi.org/10.7930/NCA5.2023

State of MN, 2022. Minnesota's climate action framework.

Toxics Release Inventory 2017, 2018. U.S. Environmental Protection Agency.

U.S. Bureau of Economic Analysis, 2019. Input-Output Accounts Data [WWW Document]. URL https://www.bea.gov/industry/input-output-accounts-data (accessed 1.28.2021).

U.S. Bureau of Labor Statistics, 2023. Quarterly census of employment and wages 2020.

U.S. Census Bureau, 2019. 2-6 digit 2012 NAICS Code File [WWW Document]. URL https://www.census.gov/naics/2012NAICS/2-digit_2012_Codes.xls (accessed 3.5.2021).

U.S. EPA, 2009. Sustainable materials management: The road ahead. US Environmental Protection Agency, Office of Land; Emergency Management.

U.S. EPA, 2023. State GHG Emissions and Removals.

U.S. EPA, 2020. EPA researchers working to improve life-cycle assessment capabilities for communities [WWW Document]. URL https://www.epa.gov/sciencematters/epa-researchers-working-improve-life-cycle-assessment-capabilities-communities (accessed 10.20.2021).

U.S. EPA, 2020. US Environmentally-Extended Input-Output (USEEIO) Technical Content [WWW Document]. URL https://www.epa.gov/land-research/us-environmentally-extended-input-output-useeio-technical-content (accessed 1.26.2021).

WA Department of Ecology, 2021. The state solid and hazardous waste plan moving washington beyond waste and toxics (No. 21-04-050). Hazardous Waste; Toxics Reduction Program Solid Waste Management Program, Olympia, WA.

Water data for the nation 2015, 2018. U.S. Geological Survey.

Yang, Y., Ingwersen, W.W., Hawkins, T.R., Srocka, M., Meyer, D.E., 2017. USEEIO: A New and Transparent United States Environmentally-Extended Input-Output Model. Journal of Cleaner Production 158, 308–318. https://doi.org/10.1016/j.jclepro.2017.04.150

Yang, Y., Ingwersen, W.W., Meyer, D.E., 2018. Exploring the relevance of spatial scale to life cycle inventory results using environmentally-extended input-output models of the United States. Environmental Modelling and Software 99, 52–57. https://doi.org/10.1016/j.envsoft.2017.09.017

Young, B., Birney, C., Ingwersen, W.W., 2024. Dataset of 2012-2020 U.S. National- and State-Level Greenhouse Gas Emissions by Sector. Data in Brief.

Young, B., Srocka, M., Ingwersen, W., Morelli, B., Cashman, S., Henderson, A., 2021a. TRACIv2.1 for FEDEFLv1. https://doi.org/10.23719/1522413

Young, B., Srocka, M., Ingwersen, W., Morelli, B., Cashman, S., Henderson, A., 2021b. FEDEFL Inventory Methods v1.0.0. https://doi.org/10.23719/1522414

Young, B., Srocka, M., Ingwersen, W., Morelli, B., Cashman, S., Henderson, A., 2021c. LCIA formatter. Journal of Open Source Software 6, 3392. https://doi.org/10.21105/joss.03392

Zhuang, X., Balassiano, K., 2021. EPA Data Commons v0.1 [WWW Document]. URL http://edap-data-commons.s3.amazonaws.com/data_commons_search.html (accessed 12.12.2021).

Appendix A - Methodology

StateIO model framework

The modeling framework used for the development of the state economic models, referred to as a the StateIO models, is described in Li et al. (2022), and is summarized below.

The StateIO models are developed using the *stateior* R package (Li et al., 2023), which is publicly available. This package constructs two-region models for the state of interest (SoI) and the rest of the US (RoUS). Each region is described in terms of industries and commodities (goods and services) produced and consumed within each region, as well as between regions, at the BEA summary level (i.e., 71 industries and 73 commodities). Models are available for each US state; thus, 50 individual two-region models are available. Each model includes distinct steps in its creation: (1) creating the state supply model, (2) creating the state demand model, (3) estimating interregional trade, and (4) balancing the final model.

The state supply (or make) model is an industry by commodity table that describes the commodity production by industries in a state. Where state-specific industry data is available, this is used to estimate the state production of commodities by specific industries. Otherwise, state information on value-added by industry is used as a proxy to regionalize US industry totals. A similar approach is used for estimating state commodity totals: state-specific commodity data is used where available, otherwise national commodity data is used to estimate regional values using SoI-RoUS commodity output ratios.

The state demand model is a commodity by industry table that estimates both the intermediate and final consumption of commodities by region (the Use table and Final Demand vectors in input-output parlance). The intermediate consumption describes the commodities used by industries for their production processes and is estimated using state-to-US industry output ratios. This is approach is used due the lack of state-specific intermediate consumption data, and the assumption that production technologies for most industries are similar across different states. The final consumption of commodities includes consumption by households, federal, state, and local governments, changes in inventories, international imports and exports, and investments. Different data sources are used to regionalize the consumption by these categories among the different states. For example, personal consumption expenditure data by state is used to estimate regional household consumption, while US Census Bureau data is used to estimate regional export data.

In addition to SoI and RoUS components, the two-region model also describes the interregional trade between the two regions. Interregional trade is important as it allows the model to estimate the regional economic and environmental impacts of commodity consumption. That is, it allows the model to differentiate the impacts of in-state consumption of in-state products vs. the impacts of in-state consumption of products imported from the RoUS. To accomplish this, for each state model, four groups of economic flows, labeled Interregional Commodity Flows (ICFs) are established:

- In-state commodity consumption of in-state production (SoI2SoI).
- Exports of state production outside the state (SoI2RoUS).
- Commodities produced and consumed outside the state of interest (RoUS2RoUS).
- Commodities produced outside the state of interest but imported into the state (RoUS2SoI).

Different, sometimes conflicting, data sources are used to estimate the ICFs for goods and services. As a result, model balancing is performed to ensure that impossible scenarios, such as states exporting more than what they produce, are not present in the model.
After this step, the Domestic two-region tables are complete. However, these tables do not include international imports. A final step is then performed to include the value of international imports to the appropriate region in the model (SoI and RoUS). The imports data is sourced from the BEA (Bureau of Economic Analysis, 2019) and assumes that the production of imported commodities uses technologies similar to those produced in the US (i.e., the domestic technology assumption). The inclusion of imports in the models results in the creation of the Total two-region tables. These tables maintained separately from the Domestic two-region tables to allow the option of domestic-only or total impact calculations. Once all the model components for a two-region model are finalized (Make, Use, interregional trade, etc.), these are exported to the US EPA data commons, where they can be accessed by other users and applications (Zhuang and Balassiano, 2021).

Two-region models in USEEIO

While the two-region StateIO models are built using *stateior*, further steps are required to turn these components into environmentally-extended input-output (EEIO) models. This is accomplished by using the *stateior* data products with the new functionality available in the useeior v1.4.0 to handle two-region Make and Use tables, industry and commodity output.

The *useeior* package is used to develop, maintain, and expand on the USEEIO family of models (Li, Mo et al., 2022). Until now, all the models created and used for environmental modeling in USEEIO were one-region, national models. One-region models have similar components to the two-region models built in the *stateior* package; indeed, the main basic components are the Use table, Make table, Final Demand vectors, etc. However, there are several important differences between the two packages. First, *useeior* uses these initial components to calculate the direct requirements and total requirements tables (A and L tables, respectively) that are used to estimate economic impacts from changes in economic activity (Ingwersen et al., 2022); *stateior* does not produce these tables as data products. Additionally, *useeior* combines these tables with environmental satellite tables, which in turn enables the models to estimate environmental impacts from changes in economic activity. Finally, while the basic components themselves are conceptually similar, the technical implementation for building the full EEIO models in USEEIO was only able to use the structure of one-region models (e.g., it does not have the capability to handle SoI2RoUS flows). Accordingly, adjustments to the *useeior* package were necessary to allow the creation and use of two-region EEIO models. Table 8 shows the differences in functionality between the *stateior* and *useeior* packages prior to the changes introduced by this work.

Table 8. Stateior vs useeior functionality.

Functionality	stateior	useeior
Generates 1R model components		✓
Generates 2R model components	✓	

Functionality	stateior	useeior
Exports model components	✓	✓
Capacity to perform economic impact calculations		✓
Capacity to perform environmental impact calculations		✓

Major changes were introduced in the *useeior* package to enable construction of two-region EEIO models. The following changes were made:

- Create new model specification formats for two-region models.
- Look for and download *stateior* model components once two-region model specs are identified in the model build process.
- Adjust the model build process to handle two-region structures (SoI2SoI, SoI2RoUS, RoUS2SoI, RoUS2RoUS).
- Adjust the calculation of A and L matrices for two-region models.
- Adjust calculations of final demand vectors to represent total consumption, household consumption and total production for two-region models.
- Create new validation functions to ensure that two-region EEIO models are built correctly.
- Adjust results calculations for two-region models.

The resulting model is a two-region SoI-RoUS model built from StateIO models that can be used to estimate both economic and environmental impacts using the useeior framework. It should be noted that while national models are available in *useeior* at the BEA detail level (~400 sectors), two-region models are only available at the BEA summary level due to data constraints at the state level; however, both national and state models are available at the summary level for the years 2012-2020. Additionally, changes to the satellite tables were necessary to estimate emissions at the state, rather than national level. While no changes were necessary to handle the satellite table data for two-region models once they are read by the package during model build, new satellite tables needed to be created specifically for the two-region models, containing data at the state level. This is described in the next section.

Finally, it should also be noted that this framework conceptually allows the creation of models with more than two regions. However, Yang et al. (2018) showed that building two-region SoI and RoUS models for each specific state allows for models to be built using far fewer data requirements (e.g., no trading data between specific states required) while maintaining the accuracy of the environmental estimates at a relatively high level.

Environmental Data Preparation

Environmental data are attributed to economic sectors using EPA's Flow Sector Attribution (FLOWSA) v2.0 Python package (Birney et al., 2023). FLOWSA attributes primary emissions or resource use data to economic sectors using secondary attribution sources (Birney et al., 2022). These flow sector attribution models generate Flow-by-Sector (FBS) datasets which are used to build State EEIO model satellite tables. For the state models, primary data are sourced from datasets that provide state, county, or facility data, enabling the differentiation of emissions and resource use between states. In some cases, where no alternative data source was identified, data sources may be national. In the cases where national data without state-specific resolution are used, we attribute emissions or resource use using a proxy measure of sector activity in that state, such as sector output. These state models are designed to attribute emissions and resource use to 3- and 4-digit NAICS to align with the BEA Summary sector codes used in the USEEIO State models. National versions of each of these methods were first described for USEEIO v2.0 (Ingwersen et al., 2022). Updates or changes to adjust data, or identify alternative data sources, for the state EEIO models are described below.

Employment

The Bureau of Labor Statistics (BLS) Quarterly Census of Employment and Wages (QCEW) is the primary data source for the state FBS datasets, as in the national models. BLS QCEW data (U.S. Bureau of Labor Statistics, 2023) are standardized into annual tables after estimating suppressed data. To estimate suppressed values, known data from a higher level of aggregation, such as at 3-digit NAICS, are equally attributed to suppressed values at more detailed sectors, such as 4-digit sectors. For example, if BLS QCEW reports 200 employees for sector 423 and publishes values for related 4-digit NAICS that sum to 150 employees while suppressing values for 4232 and 4234, the suppressed values can be approximated by equally allocating the unattributed employees among the sectors. In this scenario, the two suppressed 4-digit sectors would each be assigned an estimated 25 employees. The state method differs from the national method only in that state employment data is used rather than national data. Subsequent data that use state employment data as an attribution source relies on these employment FBS.

Chemical releases to air

Criteria, hazardous and other non-GHG chemical releases to air are sourced from the National Emissions Inventory (NEI) and Toxic Release Inventory (TRI) (*National emissions inventory 2017*, 2019; *Toxics Release Inventory 2017*, 2018). As in the national flow sector attribution model, data for Nonpoint, Nonroad, and Onroad NEI emissions are aggregated by state. Emissions are attributed to economic sectors according to activity sources, as identified by Source Classification Codes. Attribution sources include the USDA Census of Agriculture (for agricultural emissions sources) (*Census of agriculture 2017*, 2017) EIA Manufacturing Energy Consumption Survey (MECS) (for manufacturing emissions sources) (EIA, 2021), and the employment FBS described above. Additionally, we impute annual detail make and use tables based on published annual summary level tables for use as an allocation source (in an approach described in Young et al. (2024)). To better align EIA MECS energy consumption data with the primary emissions data from each state, energy consumption for each of the four census region for each economic sector is first apportioned to individual states proportional to employment in that sector. Facility point-source data from NEI and TRI are attributed to sectors based on facility reported NAICS, a method that remains unchanged from the national model. Nonpoint, Nonroad, and Onroad datasets from NEI are available every three years, so interim FBS are not available.

Chemical releases to water and ground

Nutrient and chemical releases by facility to water and to ground are sourced from the Discharge Monitoring Reports (DMR) (*DMR*, 2018) and TRI (*Toxics Release Inventory 2017*, 2018). Facility emissions are then aggregated by state after sector attribution. Other than the state aggregation, the methodology is unchanged from that used in USEEIO v2.0 (Ingwersen et al., 2022).

Greenhouse gas emissions

Greenhouse gas emissions are sourced from EPA's Inventory of U.S. Greenhouse Gas Emissions and Sinks (*Inventory of U.S. Greenhouse gas emissions and sinks*, 2022-04-22), which are attributed to states using EPA's State GHG Emissions and Removals dataset (U.S. EPA, 2023). This state level dataset aligns with the national "Inventory of U.S. Greenhouse Gas Emissions and Sinks". This flow sector attribution model is described fully in Young et al. (2024).

Land use

The state model attributing U.S. land use to economic sectors generally follows the same approach as the national model. However, as state-level commercial and manufacturing land use are unavailable, the state model incorporates an additional attribution method that is not required in the national model. The EIA CBECS dataset (*Commercial buildings energy consumption survey 2012*, 2016), which publishes commercial land use, and the EIA MECS dataset (*Manufacturing energy consumption survey (MECS) of 2014*, 2017), which publishes manufacturing land use, are proportionally attributed to states using the employment FBS described above.

Water withdrawal

State-level USGS water withdrawals (*Water data for the nation 2015*, 2018) are attributed to economic sectors using the same approach as the national model. State water withdrawals for *Industrial* and *Mining* are proportionally attributed to sectors using the state employment FBS described above instead of national employment data. The state and national FBS results differ due to differences in state and national employment datasets.

Value added

For two-region models, total value added by sector are calculated within *useeior* as is done for the national model (Li, Mo et al., 2022). State level value added data are originally sourced from the national Use table and attributed to states by industry based on state gross value added (Li et al., 2022).

Appendix B - Source Code Overview and Links

Table 9 summarizes the tools used to build the state EEIO models and their functionality.

The *useeior* R package is the primary tool used to assemble and output the USEEIO State Models. *useeior* v1.4.0 includes new functionality to build the two-region EEIO models from the State IO economic models. Only the correct version of *useeior* along with the model specification files are needed to actually build the complete models.

FLOWSA is the tool used to prepare the environmental data. The LCIA Formatter prepares the indicators. As described above, *stateior* is used to prepare the StateIO models.

Table 9. Tools used to create USEEIO State models.

Module	Language	Version	Functionality
useeior	R	v1.4	Assemble EEIO model
FLOWSA	Python	v2.0	Run environmental sector attribution models
stateior	R	v0.2.1	Produce State IO models
LCIA Formatter	Python	v1.0.4	Produce characterization factors for indicators

The FLOWSA, *stateior*, and LCIA Formatter outputs are stored on a data server and are retrieved automatically upon first running the building the models in *useeior* and stored locally.

Links to FLOWSA Method Files and Data Output

The method files that represent the flow sector attribution models used to attribute environmental data to sectors and links to the output Flow-by-Sector data are included in Table 10.

Table 10. FLOWSA data method and output files.

Name	Year	Method File	Data File
Criteria and Hazardous Air Emissions	2014	CAP_HAP_state_2014_m1 Method	CAP_HAP_state_2014_m1 Data
	2017	CAP_HAP_state_2017_m1 Method	CAP_HAP_state_2017_m1 Data
	2020	CAP_HAP_state_2020_m1 Method	CAP_HAP_state_2020_m1 Data
Commercial RCRA-defined hazardous waste	2013	CRHW_state_2013 Method	CRHW_state_2013 Data
	2015	CRHW_state_2015 Method	CRHW_state_2015 Data

Name	Year	Method File	Data File
	2017	CRHW_state_2017 Method	CRHW_state_2017 Data
	2019	CRHW_state_2019 Method	CRHW_state_2019 Data
Employment	2012	Employment_state_2012 Method	Employment_state_2012 Data
	2013	Employment_state_2013 Method	Employment_state_2013 Data
	2014	Employment_state_2014 Method	Employment_state_2014 Data
	2015	Employment_state_2015 Method	Employment_state_2015 Data
	2016	Employment_state_2016 Method	Employment_state_2016 Data
	2017	Employment_state_2017 Method	Employment_state_2017 Data
	2018	Employment_state_2018 Method	Employment_state_2018 Data
	2019	Employment_state_2019 Method	Employment_state_2019 Data
	2020	Employment_state_2020 Method	Employment_state_2020 Data
Greenhouse gas emissions	2012	GHG_state_2012_m1 Method	GHG_state_2012_m1 Data
	2013	GHG_state_2013_m1 Method	GHG_state_2013_m1 Data
	2014	GHG_state_2014_m1 Method	GHG_state_2014_m1 Data
	2015	GHG_state_2015_m1 Method	GHG_state_2015_m1 Data
	2016	GHG_state_2016_m1 Method	GHG_state_2016_m1 Data
	2017	GHG_state_2017_m1 Method	GHG_state_2017_m1 Data
	2018	GHG_state_2018_m1 Method	GHG_state_2018_m1 Data
	2019	GHG_state_2019_m1 Method	GHG_state_2019_m1 Data
	2020	GHG_state_2020_m1 Method	GHG_state_2020_m1 Data
Land	2012	Land_state_2012 Method	Land_state_2012 Data

Name	Year	Method File	Data File
Point source industrial releases to ground	2012	GRDREL_state_2012 Method	GRDREL_state_2012 Data
	2013	GRDREL_state_2013 Method	GRDREL_state_2013 Data
	2014	GRDREL_state_2014 Method	GRDREL_state_2014 Data
	2015	GRDREL_state_2015 Method	GRDREL_state_2015 Data
	2016	GRDREL_state_2016 Method	GRDREL_state_2016 Data
	2017	GRDREL_state_2017 Method	GRDREL_state_2017 Data
	2018	GRDREL_state_2018 Method	GRDREL_state_2018 Data
	2019	GRDREL_state_2019 Method	GRDREL_state_2019 Data
	2020	GRDREL_state_2020 Method	GRDREL_state_2020 Data
Point source releases to water	2012	TRI_DMR_state_2012 Method	TRI_DMR_state_2012 Data
	2013	TRI_DMR_state_2013 Method	TRI_DMR_state_2013 Data
	2014	TRI_DMR_state_2014 Method	TRI_DMR_state_2014 Data
	2015	TRI_DMR_state_2015 Method	TRI_DMR_state_2015 Data
	2016	TRI_DMR_state_2016 Method	TRI_DMR_state_2016 Data
	2017	TRI_DMR_state_2017 Method	TRI_DMR_state_2017 Data
	2018	TRI_DMR_state_2018 Method	TRI_DMR_state_2018 Data
	2019	TRI_DMR_state_2019 Method	TRI_DMR_state_2019 Data
	2020	TRI_DMR_state_2020 Method	TRI_DMR_state_2020 Data
Water	2015	Water_state_2015_m1 Method	Water_state_2015_m1 Data

USEEIO State Model Specifications

The USEEIO State model specifications files used here are available in the USEEIO-State GitHub repository. See the model specification file format to interpret the meaning of the field names.

United States
Environmental Protection
Agency

EPA 600/R-23/228 I January 2024 I

Office of Research and Development
(8101R)Washington, DC 20460