# Core Model Proposal #376: GCAM-USA fossil resource curves

**Product:** Global Change Analysis Model (GCAM)

**Institution:** Joint Global Change Research Institute (JGCRI)

**Authors:** Yang Ou, Gokul Iyer, Ying Zhang, Jae Edmonds

**Reviewers:** Marshall Wise, Pralit Patel, Matthew Binsted

**Date committed:** 25 May 2026

**IR document number:** PNNL-39384

**Related sector:** energy

**Type of development:** data, code

**Purpose:** The purpose of this core model proposal is:
1) To incorporate state-level crude oil, coal, and natural gas resources. This will provide a more detailed representation of United States fossil fuel resource production by creating inter-state competition within the national supply sector and intra-state competition among different types of fossil fuel resources.
2) To ensure that the sum of state-level fossil fuel resource availability and production aligns with USA region fossil fuel resources from the GCAM core.

# Background

Quantitative assessment of regional energy resource availability is essential for developing future regional energy demand and supply scenarios. However, a US-focused analysis based on highly aggregated parameterization of resource production is insufficient to capture state-level details and hinders the capability to assess state-level decarbonization actions related to resource productions.

## Data source

Table 1 below shows a summary of main input data used in this CMP. Many individuals and institutions have attempted to compile and publish estimates of resources. However, the statutory responsibility for collecting and publishing authoritative statistical information on the various types of energy sources in the United States has been given to specific federal agencies. The **Energy Information Administration (EIA)** was originally created as the Federal Energy Administration (FEA) and is charged with the responsibility of monitoring and reporting U.S. energy reserves and production. Likewise, the responsibility for assessing onshore undiscovered technically recoverable oil and gas resources in the United States resides with the **U.S. Geological Survey (USGS)**, in the Department of the Interior. The USGS has conducted a few national assessments of undiscovered technically recoverable oil and natural gas resources over several decades. The most recent complete comprehensive national assessment for onshore oil and gas was completed in 1995, but USGS updates that assessment on an ongoing basis as new data become available for individual basins. Responsibility for assessment of offshore undiscovered technically recoverable oil and natural gas resources belongs to the **Bureau of Ocean Energy Management (BOEM)**, Regulation and Enforcement, also within the Department of the Interior. EIA and USGS have responsibilities for evaluating the nation's endowment of coal that are similar to their responsibilities for oil and natural gas. The resource costs are from IEA's Energy Technology Systems Analysis Program (ETSAP).

**Table 1 Summary of source data (links to data are included in the corresponding csv input file in gcamdata)**

| Type | Natural gas | Crude Oil |
|---|---|---|
| Onshore resource availability (resource basin level) | USGS National Assessment of Oil and Gas Resources Update (2013) | |
| Offshore resource availability (by PADD) | BOEM Assessment of Undiscovered Technically Recoverable Oil and Gas Resources of the Nation's Outer Continental Shelf, 2011(Includes 2014 Atlantic Update) | |
| Onshore resource cost (national) | ETSAP brief | |
| Offshore resource cost (by PADD) | BOEM supply curve | |
| Historical production (by state) | EIA, onshore and offshore | EIA, onshore by state; Federal offshore; Alaska Oil and Gas Conservation Commission Production Data |
| Price (by state) | EIA dry gas wellhead price by state | EIA crude oil first purchase price by state |

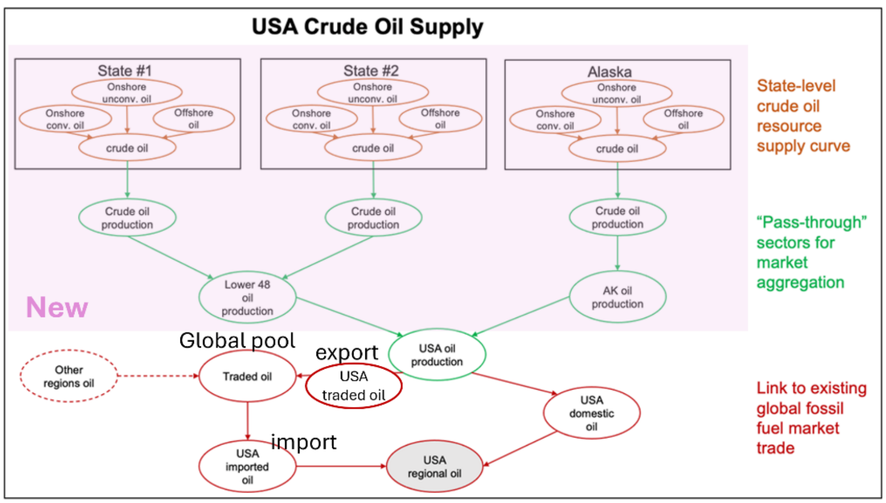
# GCAM-USA

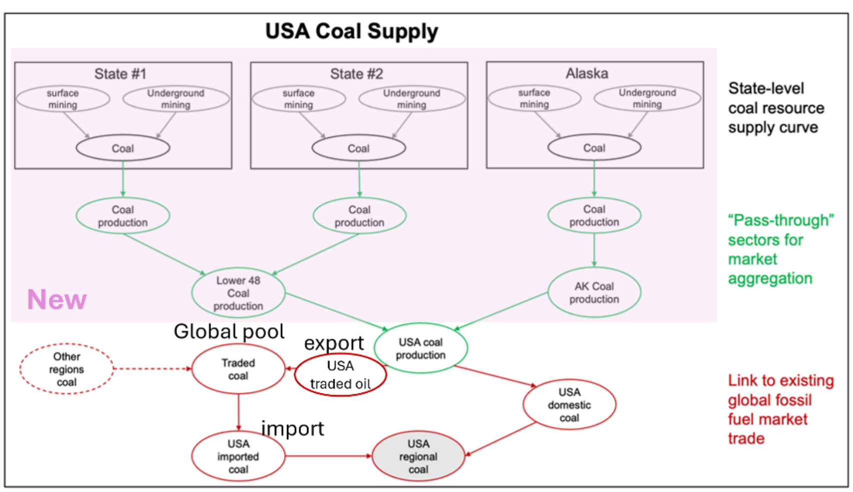
GCAM-USA has 51 regions that include the 50 states and Washington D.C. For simplicity, these regions will be referred to as "states" or, when being referred to in aggregate, as "the sum of states". The states currently contain features that have been downscaled from the aggregate USA region, including the electricity sector and consumer end-use demands. GCAM-USA also contains a separate USA region with features that have not yet been downscaled and exist only at the aggregate level, including some primary energy resources such as coal, crude oil, and natural gas.

Considering the wide-ranging impacts of fossil energy supply on the U.S. energy sector, more current and robust information on fossil resource types, extraction costs, and geographical availability is desirable. This proposal describes the incorporation of state-level coal, natural gas, and crude oil supply curves into GCAM-USA.

# Description of Changes

## Proposed model structure



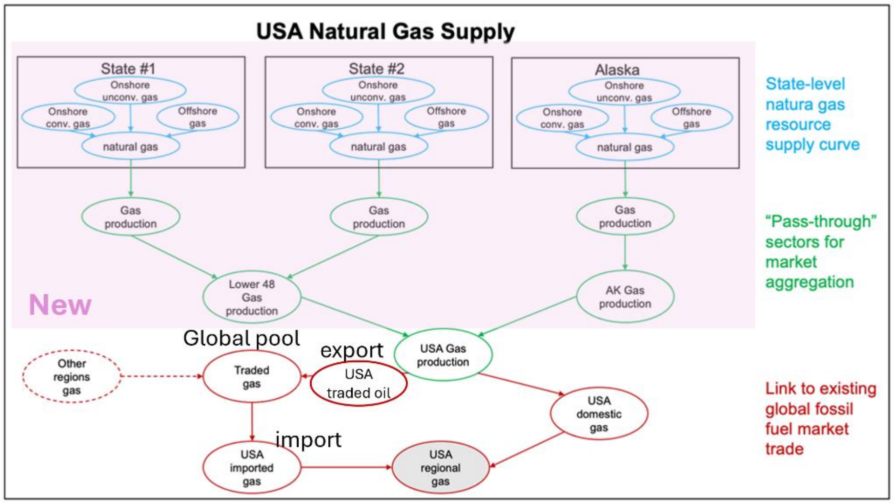



**Fig.1 Proposed state-level fossil fuel resource supply structure for coal, natural gas, and crude oil supply (top to bottom) at US state levels.** ***Note that for Crude oil supply, there is one additional state-level sub-resources of "Onshore unconv. heavy oil" besides the ones shown in the diagram. Compared to "Onshore unconv. oil" which refers to resources such as shale oil/tight oil, the "Onshore unconv. heavy oil" refers to resources such as tar sands and oil shale (Note that shale oil and oil shale are two different types of crude oil resource). The "Onshore unconv. heavy oil" requires additional energy input to extract the resource and thus use more energy and emission intensive. The counterpart of "Onshore unconv. heavy oil" in GCAM-32 is the "unconventional oil" sub-resource.***

# Input Files

**Table 2: new input files in gcamdata**

| File name | Description | Resource Type |
|---|---|---|
| GCAM assumptions (input/gcam-usa/) ALL NEW FILES | | |
| A10.fossil_sector_vertical.csv | Provides input/price units and logit information for fossil energy production and region supplysectors | All |
| A10.fossil_tech_associations.csv | Define minicam.energy.input names for US domestic fossil energy supply, with coefficients as 1 | All |
| A10.subsector_interp.csv | Define fossil energy production subsector shareweight interpolation rules (fixed) | All |
| A21.globalrsrctech_coef_USA.csv | Technology coefficients for onshore unconventional heavy oil, which requires wholesale gas and unconv oil upscaling inputs | Gas |
| Input data (input/gcam-usa) ALL NEW FILES | | |
| Alaska_offshore_gas_oil_2022.csv | Alaska Oil and Gas Conservation Commission Production Data | Oil/Gas |
| BOEM_gas_cost.csv | Offshore natural gas supply curves from BOEM 2021 report | Gas |
| BOEM_gas_supply_EJ.csv | Federal offshore gas supply from BOEM | Gas |
| BOEM_oil_cost.csv | Offshore oil supply curves from BOEM 2021 report | Oil |

| | | |
|---|---|---|
| BOEM_oil_supply_EJ.csv | Federal offshore oil supply from BOEM | Oil |
| EIA_coal_open_market_price_states.csv | US coal average market price by state | Coal |
| EIA_coal_open_market_price_states_mapping.csv | Mapping EIA coal market states into regions to fill in missing values | Coal |
| EIA_coal_prod_state_ton_surface_2001_2021.csv | US Aggregate coal mine production by state; surface | Coal |
| EIA_coal_prod_state_ton_underground_2001_2021.csv | US Aggregate coal mine production by state; underground | Coal |
| EIA_coal_reserve_2021_Mton.csv | Recoverable coal reserves at producing mines and estimated recoverable reserves and demonstrated reserve base by mining method | Coal |
| EIA_gas_market_prod_state_Bcf_coalbed.csv | US natural gas coalbed methane production by state | Gas |
| EIA_gas_market_prod_state_Bcf_shalegas.csv | US natural gas shale gas production by state | Gas |
| EIA_gas_market_prod_state_MMcf_total.csv | US natural gas onshore and offshore production by state | Gas |
| EIA_gas_market_wellhead_price_states.csv | US Natural Gas Wellhead Price by state | Gas |
| EIA_NG_prod_mapping_coalbed.csv | Mapping EIA coalbed methane production region into US states | Gas |
| EIA_NG_prod_mapping_shalegas.csv | Mapping EIA shale gas production region into US states | Gas |
| EIA_NG_prod_mapping_total.csv | Mapping EIA natural gas production region into US states | Gas |
| EIA_NG_prod_mapping_wellhead_price.csv | Mapping EIA wellhead price region into US states | Gas |
| EIA_oil_first_purchase_price_states.csv | US Domestic Crude Oil First Purchase Prices by Area | Oil |
| EIA_oil_first_purchase_price_states_mapping.csv | Mapping EIA crude oil first purchase price area into US states | Oil |
| EIA_oil_GOM_refineries_input_2022.csv | Golf Coast gross inputs to refineries by refining district (1985-2022) | Oil |
| EIA_oil_market_prod_state_thousBBL_total.csv | US crude oil onshore and offshore production by state | Oil |
| EIA_tight_oil_play_state_mapping.csv | Mapping EIA tight oil production plays into states with assumed shares | Oil |
| EIA_tight_oil_production_mbbl_per_day.csv | EIA monthly tight oil production by play | Oil |
| ETSAP_coal_cost_range.csv | Estimates for coal production costs by grade | Coal |
| ETSAP_gas_cost_range.csv | Estimates for natural gas production costs by grade | Gas |

| | | |
|---|---|---|
| ETSAP_oil_cost_range.csv | Estimates for oil production costs by grade | Oil |
| USGS_basin_state_mapping.csv | Map USGS TPS basins to states | Oil/Gas |
| USGS_gas_supply_Quad.csv | USGS natural gas resource estimates | Gas |
| USGS_oil_supply_Quad.csv | USGS oil resource estimates | Oil |
| USGS_unconv_heavy_oil.csv | estimate of two types of unconventional heavy oil potential - oil shale and tar sands - in US | Oil |
| Input data (input/gcam-usa/emissions) ALL NEW FILES | | |
| EPA_fossil_rsrc_GHG_mapping.csv | Map of EPA fossil resource categories and subcategories to GCAM resources and reserve-subresources | All |
| EPA_state_res_ghg_2021.csv | EPA state level GHG emissions from fossil resource production | All |

## Data processing

*In this section, [fossil] represents coal, natural gas, and crude oil.*

There are two general parts of processing: 1) state-level fossil supply and 2) the corresponding emissions

**1) State-level fossil supply**

A total of six scripts (one level 1, two level 2, and three level batch) transform the input files into three xml files for the model.

*Level 1:* zgcamusa_L111.rsrc_fos_Prod.R

*Level 2:* zgcamusa_L211.resources_fossil.R and zgcamusa_L2262.regional_fossil.R

*Level batch:* zenergy_xml_en_supply.R, zenergy_xml_gas_trade.R, and zgcamusa_xml_resources_fossil

*XMLs:* resources_fossil_USA.xml, en_supply_USA.xml, gas_trade_USA.xml

*resources_fossil_USA.xml* contains information on state-level [fossil] resource curves and production structures in the USA (Fig.1).

*en_supply_USA.xml* modifies the sectors originally created in *en_supply.xml*, updating the relevant minicam-energy-input names from USA [fossil] into USA [fossil] *production* in regional [fossil] and traded [fossil] sectors.

*gas_trade_USA.xml* makes similar same changes as *en_supply_USA.xml* for sectors originally created in *gas_trade.xml*.

**A brief description of the data processing is below:**

***zgcamusa_L111.rsrc_fos_Prod.R*** prepares information on state-level [fossil] resources cost, availability, and historical production. The script reads in most of the input assumption and source data files listed above. It also reads in two energy data system level1 files – *L111.Prod_EJ_R_F_Yh*, which contains historical annual fossil fuel production by fuel at the core GCAM 32-region level; and *L111.RsrcCurves_EJ_R_Ffos*, which contains graded fossil fuel resource curves for each GCAM region. Table 3 shows sub-resource mappings between this CMP and GCAM-32 for harmonization of the production and resource supply curves at the sub-resource level.

**Table 3 Sub-resource mappings between GCAM-32 and GCAM-USA in this CMP.**

| Resource | Sub-resource in GCAM-32 | Sub-resource in GCAM-USA CMP | Includes |
|---|---|---|---|
| crude oil | crude oil | onshore conventional oil | |
| | | onshore unconventional oil | tight oil (shale oil) |
| | | offshore oil | |
| | unconventional oil | onshore unconventional heavy oil | oil sands, oil shale |
| natural gas | natural gas | onshore conventional gas | |
| | | onshore unconventional gas | coalbed methane, shale gas |
| | | offshore gas | |
| | | speculative natural gas | additional resource from GCAM-32 |
| coal | coal | surface mining | |
| | | underground mining | |
| | | speculative coal | additional resource from GCAM-32 |

Historical production data for [fossil] energy is obtained from EIA, and the share of USA-level [fossil] energy production is calculated for each state/ sub-resource, and then applied to the IEA national production estimates in *L111.Prod_EJ_R_F_Yh.csv* to ensure that aggregate state-level production will equal the USA-region production from the energy data system for each historical period.

1. For natural gas, we further obtained separated historical production of coalbed methane and shale gas from EIA, the sum of which are onshore unconventional gas production. Using total onshore gas production minus the onshore unconventional gas production, the remaining part is onshore conventional production. Federal offshore production is

allocated into the neighboring states. For Federal Gulf of Mexico region, these data are available between 1997-2021, so we just allocate them into the nearest three Federal offshore states of AL, LA, and TX (data only available between 1992-1998) based on their average production share during 1992-1998.

2. For crude oil, there is no historical production of unconventional (heavy) oil (e.g., tar sands) in the US, so crude oil production is split into onshore conv., onshore unconv. (shale oil/tight oil), and offshore. Like natural gas, federal offshore oil production is allocated into the nearest states. For PADD5 (Pacific), federal offshore productions are all allocated to California, as Alaska has no offshore oil wells in federal water (all offshore oil wells are in state water). Alaska offshore oil production from state water is based on Alaska Oil and Gas Conservation Commission Offshore data. For PADD3 (Gulf of Mexico, or GOM), federal GOM production is allocated into neighbor states based on the year-specific coastal refinery input shares, obtained from EIA (*EIA_oil_GOM_refineries_input_2022.csv*).
3. For coal, EIA coal data browser includes historical coal production since 2001, separated by surface and underground mining. For GCAM historical years earlier than EIA's earliest year, the earliest available shares are used.

For Oil and Gas, resource availability by grades is processed from USGS (onshore) and BOEM (offshore), and for coal, resource availability is from EIA. For oil and gas, resource types are mapped from USGS basins to the states (*gcam-usa/USGS_basin_state_mapping.csv*), and resource grades are defined. To be consistent with the core 32-region model, resources availability of the "speculative" sub-resource category in this CMP are calculated by:

1. summing [fossil] resource data to a national total in the sub-resource categories as in GCAM-32 (see mapping of sub-resource from this GCAM-USA CMP to GCAM-32 in Table 3)
2. subtracting this total from the energy data system's [fossil] resource curve (*L111.RsrcCurves_EJ_R_Ffos.csv*) to find remaining [fossil] resources
3. putting this remaining resource in a new "speculative" sub-resource category (Table 3)
4. allocating this "speculative" sub-resource remainder to states based on their share of existing [fossil] resources

Note that crude oil doesn't have additional "speculative" subresource from the GCAM-32 resource supply curves.

It's important to note the definition of "conventional oil", which can be different given different context. In GCAM-32, the definition of 'unconventional oil' refers to the 'unconventional (heavy) oil', which is only produced historically in Canada and Venezuela. In the USA region, there's no unconventional (heavy) oil production in historical periods, although its production emerges in future periods in GCAM-32. Therefore, we add 'onshore unconventional heavy oil' in GCAM-USA with zero historical production; this resource has the potential to be produced in the future. This is important because the 'unconventional (heavy) oil' needs additional processing to become light oil, which is then delivered to refineries. That is, the 'unconventional (heavy) oil' is associated with additional energy and non-energy input, including an 'unconventional oil upscaling' input as in the GCAM-32. This input mostly serves to reflect the higher carbon content of unconventional vs. conventional oil and the fact that excess carbon is often emitted when bitumen and other heavy oil is upscaled. The production of 'unconventional (heavy) oil' is therefore associated with higher emissions than conventional oil. In sum, the 'unconventional oil' represents the unconventional heavy oil in GCAM-32, and the tight oil/shale oil in the USA region is classified as 'conventional oil' in GCAM-32. Given that there may be updates in GCAM-core to reclassify the tight oil /shale oil in the USA as unconventional (light) oil, we will keep the tight oil/shale oil as the 'onshore unconventional oil' subresource in GCAM-USA as it is now and add another subresource category of 'onshore unconventional heavy oil' in GCAM-USA, which is equivalent to the 'unconventional oil' in GCAM-32.

The state-level resource potential of unconv. heavy oil in the USA is provided in *gcam-usa/USGS_unconv_heavy_oil.csv*. The largest known deposits of oil shale are in a 16,000-square mile area of the Green River Formation in Colorado, Utah, and Wyoming with a total in-place oil shale resources estimated at 4.3 trillion barrels. (Note that "shale oil" and "oil shale" are two different types of crude oil resource.) Estimated total in-place resources are about 1.53 trillion barrels of oil for the Piceance Basin, about 1.32 trillion barrels of oil for the Uinta Basin and 1.44 trillion barrels of oil in the Greater Green River Basin. In the USA, tar sands (another type of unconv. heavy oil than oil shale) resources are primarily concentrated in eastern Utah, mostly on public lands. The in-place tar sands oil resources in Utah are estimated at 12 billion to 19 billion barrels.

The supply cost for each grade is obtained from ETSAP (for onshore resources) and BOEM (for offshore oil and gas). Due to the lack of data, here we took a relatively simplified approach by extrapolating ETSAP's high and low cost range into four different cost levels (grade.hist, grade.1, grade.2, and grade.3), then harmonized the costs of the "speculative" sub-resource with the core 32-region resource curves. Since the "speculative" sub-resource category represents additional resource availability in GCAM-32 USA region curves, resource cost of this "speculative" sub-resource **starts** (i.e., grade.1) from the maximum cost of other *onshore* sub-resource categories in GCAM-USA CMP ("benchmark") plus "delta cost" (Fig.2). The "delta cost" is based on linearly interpolated cost gradient between the two grades on GCAM-32 curve, where GCAM-USA aggregate availability ends up (in-between the two grades). The "delta cost" equals the cost difference between where GCAM-USA aggregate availability ends up on the linearly interpolated cost gradient and the lower grade in GCAM-32. Using the highest cost of *onshore* sub-resources avoids the overly high offshore costs. This also means that there could be production of "speculative" subresource before the offshore resource is depleted. Resource cost of the next grade (i.e., grade.2) of this "speculative" sub-resource would be the "benchmark" plus the cost difference between the two grades on GCAM-32 curve. See illustration of the method in Fig.2.

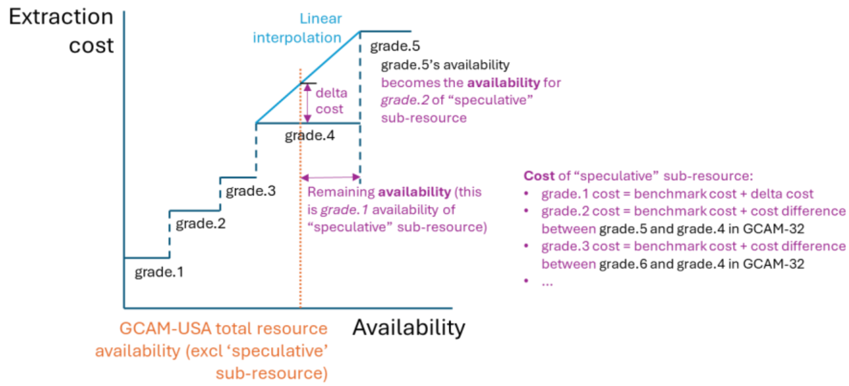



**Fig.2 Harmonization resource supply curves in this GCAM-USA CMP with GCAM-32 by creating a sub-resource category named "speculative natural gas" or "speculative coal"; there's no "speculative oil" (Table 3) as the total resource availability of crude oil in GCAM-USA CMP covers all the resource availability in GCAM-32 (no remaining availability to be counted as "speculative")**

Resource cost by grades for offshore oil and gas are obtained from BOEM, using a customized function called "compute_offshore_costs". For coal, ETSAP reported resource extraction cost by coal type. As a simplification, we assumed "lignite" represents the cost range of surface mining as lignite coal are often formed near surface and assumed "hard coal" represents the cost range of underground mining. Overall, the cost estimation part can be further improved in future if more granular cost data by resource grade are available. So far ETSAP is the "best available source".

To construct the resource curve, resource availability and costs are matched by grade. Two additional steps have been taken. First, we need to add resource to cover historical production since it is not included in the supply curves. Here we back calculated reserve additions to be exactly enough given our historical production and assumed production lifetime. Note production lifetimes may not cover the entire historical period making the calculation a bit trickier. We used the same "lag_prod_helper" created in the zchunk_L210.resources.R chunk to help project forward production by each historical vintage so we can take this into account, and this amount is defined as the "grade.hist". Second, a higher grade (n+1, where n is the current highest grade for each sub-resource) is added by the end of the resource curve, with a slightly higher cost (e.g., 10% higher) but zero availability. This will make sure GCAM's internal extrapolation function works properly when approaching the end of resource curves.

This chunk will return two tables: the state-level resource curves by sub-resources, (*L111.ResCurves_EJ_R_Ffos_USA*) and historical production by sub-resources (*L111.Prod_EJ_R_F_Yh_USA*).

***zgcamusa_L211.resources_fossil.R*** first defines the structure of the USA-level [fossil] energy production supply sector, including its logit competition (Fig.3). To account for geographic constraint, AK is set to feed in a separated "AK [fossil] production" subsector, whereas other continental states are grouped into a "Lower-48 [fossil] production". In other words, Alaska is geographically isolated from the continental U.S. and has limited transportation infrastructure connecting it to the lower 48, so we treat AK as a separate subsector to better reflects how production competes in the system to meet the demand. Based on assumptions in the *A10.fossil_sector_vertical.csv*, competition between AK and lower48 states to supply [fossil] to the USA-level [fossil] production supply sector is solved using an *absolute* logit exponent of -3, with subsector share weight for AK and lower48 fixed at their calibration values through 2100 (*A10.subsector_interp.csv*). State-level [fossil] production is defined as a technology within the [fossil] production subsector (AK [fossil] production and Lower-48 [fossil] production) so that state production can be aggregated nationally. At the technology level, the lower 48 states compete with an *absolute* logit = -6 and technology shareweight fixed as base-year calibrated shareweight (these parameters are currently defined in this R chunk and can be moved out to a .csv input file if needed). Calibration production values are defined for each stub technology (resource type) in each state for each model historical period (*L211.TNGTechProduction_USA*).

Note that we use absolute logit instead of relative logit here based on model testing and diagnosis. More importantly, we think this is a better option because, although absolute and relative logits have similar sensitivity near base-year prices, they behave differently when costs diverge substantially. Under a relative logit, an option (e.g., Texas crude oil) that becomes much more expensive (e.g., twice the cost of California crude oil oil) still retains some market share, producing a “long tail” of phasing out. In this application, however, all options feed into a common pool and function as "perfect substitutes", so an option that is significantly more costly should drop to zero share more quickly and the absolute logit reflects this behavior.

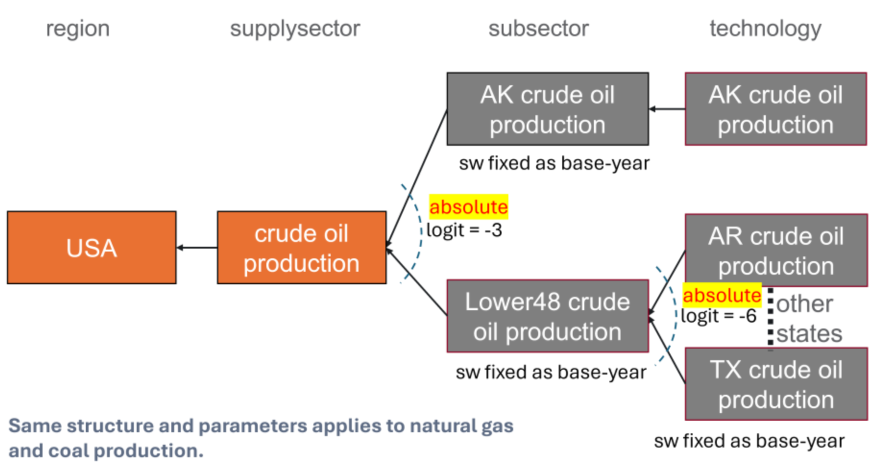



**Fig.3 USA regional oil production structure.**

State-level [fossil] production supply sectors are created to aggregate production of different [fossil] resource types (Fig.4).

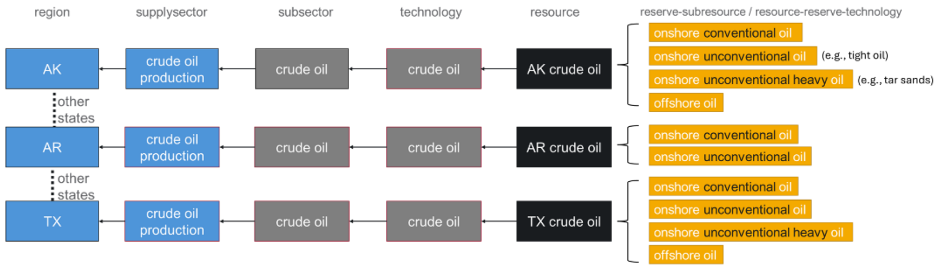



**Fig.4 state-level oil production structure. Similar structure applied to natural gas and coal production, except the sub-resource categories are slightly different.**

We constructed the entire resource-reserve technology hierarchy at state level. Historical resource production prices (at sub-resource level) are calibrated based on state-level resource extraction costs, derived from EIA (wellhead price for gas, first purchase price for crude oil, and open market price for coal). Based on historical extraction costs, we calculate the ratios between each state to the USA average price and then use these ratios as scalars to scale the historical calibration price for region USA (*L210.RsrcPrice*). Other subresource level assumptions, including reserve subresource lifetime and price adders, follow the core model assumptions (*L210.ResSubresourceProdLifetime* and *L210.SubresourcePriceAdder*). Historical calculated reserves follow the same method in *zenergy_L210.resources.R*, using the “lag_prod_helper” function again. Other resource reserve technology level assumptions, including technology lifetime, decline.phase.percent, profit.shutdown, and technology shareweight, follow the core model assumptions (*L210.ResReserveTechLifetime*, *L210.ResReserveTechDeclinePhase*, *L210.ResReserveTechProfitShutdown*, and *L210.ResTechShrwt*, correspondingly).

***zgcamusa_L2262.regional_fossil*** restructures the USA regional [fossil] sector to take in both domestic [fossil] (input = USA [fossil] production) and traded [fossil] (input = global traded [fossil] market, which also takes USA [fossil] production as an input), and sets the logit competition structure for this sector (Fig.5). Per the assumption in *A10.fossil_sector_vertical.csv*, competition between domestic [fossil] and imported [fossil] at the USA level uses a logit exponent of -3, same as GCAM-core. The share-weight assumptions are also consistent with GCAM-core (see detail in Fig.5). The original USA [fossil] resources and the USA regional [fossil] supplysectors are deleted to prevent double-counting USA [fossil] supplies/production. Calibration values for the domestic [fossil] and imported [fossil] subsectors are restructured from *L239.Production_reg_dom* and *L239.Production_reg_imp*.

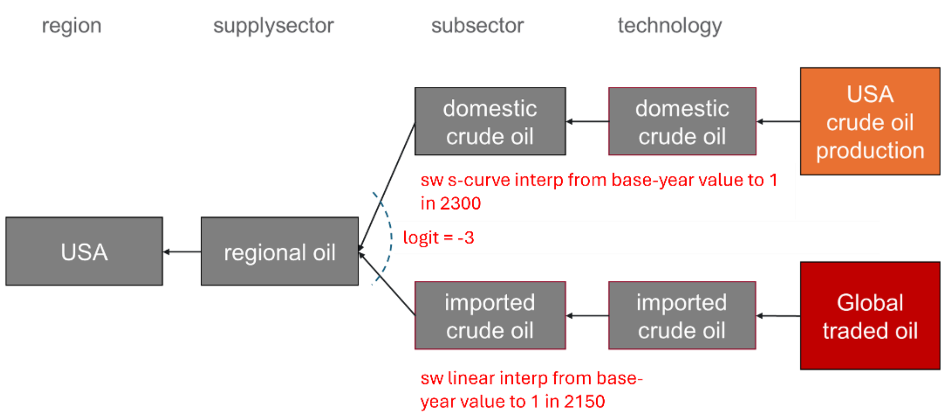



**Fig.5 USA regional oil supply structure. Same structure applies to natural gas and coal.** ***Note that regional natural gas structure defined here (in resources_fossil_USA.xml) will be replaced if the gas_trade_US_reg.xml, which has the new pipeline vs LNG structure, is included in the configuration and overwrites "regional natural gas" and "traded natural gas" in resources_fossil_USA.xml. By default, gas_trade_US_reg.xml is included. For coal, the imported coal subsector share-weight is linearly interpolated from base-year value to 1 in 2100 (different from 2150 for crude oil). All these assumptions (logit and share-weight) for crude oil, coal, natural gas are consistent with GCAM-core.***

***zgcamusa_xml_resources_fossil*** is the main batch xml chunk from this proposal, which consolidate all outcomes from *zgcamusa_L211.resources_fossil.R* (L211 tables) and *zgcamusa_L2262.regional_fossil.R* (L2262 tables) to create the final *resources_fossil_USA.xml*.

***zgcamusa_xml_en_supply_USA.R*** creates a new xml file, *en_supply_USA.xml*, which deletes and re-specifies the minicam.energy.input for USA traded [fossil] and USA domestic [fossil] to the new commodity name (USA [fossil] *production*). The traded fossil fuel supplysectors have the same structure and parameters as in the GCAM-core.

***zenergy_xml_gas_trade*** similarly creates a new xml file, *gas_trade_USA.xml*, which deletes and re-specifies the minicam.energy.input for the USA "traded LNG", "traded N.Amer pipeline gas", "gas trade statistical differences", and "regional natural gas" (originally defined in *gas_trade.xml*). The logit and share-weight assumptions are kept the same as in the GCAM-core.

## 2) Emissions of state-level fossil production

A total of six chunks (two level 1 chunks, four level 2 chunks, and two xml batch chunks) transform the input files into two xml files for the model.

*Level 1:* zgcamusa_L102.ghg_en_S_T_Y.R

*Level 2:* zgcamusa_L273.en_ghg_emissions.R, zemissions_L270.nonghg_NEI_to_GCAM.R, and zgcamusa_L273.nonghg_fossil_res_USA.R

*Level batch:* zgcamusa_xml_ghg_emissions.R and zgcamusa_xml_fossil_res_nonghg_emissions.R

*XMLs:* ghg_emissions_USA.xml and fossil_res_emissions_USA.xml

**Summary of changes**

**GHG** emissions factors are derived from the 2023 EPA state-level GHG inventory which covers GHG ($CO_2$, $CH_4$, and $N_2O$) emissions from coal, natural gas, and petroleum production among other sectors. The inventory includes production and exploration emissions for natural gas systems and petroleum systems as well as processing emissions for natural gas systems. Natural gas systems emissions are mapped to GCAM-USA's natural gas resource and petroleum systems emissions are mapped to GCAM-USA's crude oil resource. The EPA inventory also includes a more detailed breakdown of GHG emissions from coal mining, splitting emissions between underground and surface mining; these disaggregated emissions are mapped to GCAM-USA's underground and surface mining coal subresources. Emissions factors are calculated by dividing the EPA emissions by GCAM-USA's calibrated resource production data aggregated to the resource level (for crude oil and natural gas) or subresource level (for coal). The aggregated natural gas and crude oil emissions factors are assigned to all subresources (offshore, onshore conventional, and onshore unconventional oil/gas). The disaggregated underground and surface mining emissions factors are assigned directly to their respective coal subresources.

Resource-gas-state-year combinations with missing emissions factors (which occur when resource production and/or emissions data are missing or zero) are assigned the national median emissions factor for the corresponding resource-gas-year. Emissions factors above an outlier threshold of 20 times the national median are also replaced with the national median.

**Non-GHG** emissions factors are derived from the NEI inventory, which covers emissions of various pollutants from petroleum and natural gas systems, among other sectors; it does not, however, include any pollutant emissions from coal production. GCAM-USA implementation of non-GHG emissions is consistent with GCAM-32 (except for outlier replacement, which is done state-by-state rather than comparing the total US to other GCAM regions). In the GCAM-32, "petroleum production" and "natural gas production and distribution" categories of the NEI data are aggregated into combined oil and gas emissions. Oil and gas emissions are split based on energy shares and the corresponding emissions factors are computed. The emission factors are compared with other GCAM regions to flag and replace outliers. Outliers are replaced by the median of emissions factors in the regions cumulatively producing 99.75% of the resource. This avoids skewing the model with extreme values from low-production regions. For the USA, the following emissions factors are replaced:

- NH3 oil and NH3 gas (1975 and 1990)
- NOx oil and NOx gas (2005, 2010, and 2015)

**A brief description of the data processing is below:**

**1) GHGs:**

***zgcamusa_L102.ghg_en_S_T_Y.R*** calculates emissions factors for state-level resource production GHGs and produces a new output of *L102.res_ghg_tgej_USA*

- Use mapping file to map EPA categories to GCAM resources and subresources (gas and oil, which aren't mapped to the subresource level, get NAs in subresources)
- Separate into resource-level and subresource-level emissions
- Join each with resource production (*L111.Prod_EJ_R_F_Yh_USA*) aggregated to the corresponding level (resource or subresource) and calculate emissions factors
- For each subresource, add up subresource- and resource-level emissions factors
- For EPA categories reported at the national level rather than the state level, calculate a single national emissions factor and add to the corresponding subresources for each state
- For subresource- GHG- state combinations with no production and/or emissions data (or zero for either), assign the national median emissions factor; also replace emissions factors above 20 times the national median with the national median.

***zgcamusa_L273.en_ghg_emissions.R*** produces a new output of *L273.res_ghg_emfact_USA*

- In *L102.res_ghg_tgej_USA*, rename CO2 to CO2_FUG and convert from Tg $CO_2$ to MTC; add technology level (same as subresource)

***zgcamusa_xml_ghg_emissions.R*** creates *ghg_emissions_USA.xml*

- Rename columns of *L273.res_ghg_emfact_USA* to meet header requirements and add the resource GHG emissions factors

**2) Non-GHGs:**

***zemissions_L270.nonghg_NEI_to_GCAM.R*** and ***zgcamusa_L273.nonghg_fossil_res_USA.R*** produces state-level emissions factors for non-GHG emissions from fossil resource production and a new output of *L273.nonghg_state_fossil_res_EF_USA*

- Inputs: *L270.nonghg_tg_state_fossil_res_F_Yb* (from modified chunk above), *L211.RsrcCalProd_USA*, *L222.StubTech_en_USA*, *gcam-usa/states_subregions, gcam-usa/emissions/BC_OC_assumptions, gcam-usa/emissions/BCOC_PM25_ratios*
- Use the BC_OC_assumptions and BCOC_PM25_ratios input files to split out BC and OC emissions from PM2.5 emissions
- Aggregate emissions from NEI sectors "petroleum_production" and "NG_production_distribution"
- Aggregate energy production for crude oil and natural gas; calculate emissions factors and assign them to each crude oil and natural gas subresource
- For both crude oil and natural gas, assign the national median emissions factor to region-non-GHG combinations with missing resource production and/or emissions; also replace emissions factors above a threshold (20 times the median)

***zgcamusa_xml_fossil_res_nonghg_emissions.R*** creates a new xml - *fossil_res_emissions_USA.xml*

- Add fossil resource production non-GHG emissions factors from *L273.nonghg_state_fossil_res_EF_USA*

# Validation

Key GCAM-USA outputs from this CMP (GCAM-USA-CMP) are provided below. In summary, we show the results of fossil resource production (aggregated national level and state level, and by sub-resource), U.S. fossil resource consumption and trade, and emissions

## Fossil resource production, consumption, and trade

Fossil resource production results from GCAM-USA-CMP are shown below. We first show the aggregated national results by resource and compare them with GCAM-core results and other external sources (Fig.6). We then dive deeper into the aggregated national results by resource and sub-resource in GCAM-USA-CMP (Fig.7). After that, we show the state-level results, particularly the top producing states (Fig.8). We also show the results of fossil resource trade (Fig.9). We provide brief interpretation of the results as well.

This CMP results in smaller coal production in the near-term but larger coal production in the later years towards 2100 than the current core. This is mainly due to difference between this CMP and core in their sub-resource supply curves. In this CMP, the resource supply curves are at the state-level with more sub-resource categories, compared to the aggregate single sub-resource category curve in the core. GCAM shows generally higher coal production than AEO. In GCAM, the large production volume of coal mostly goes to export, instead of domestic consumption (Fig.9); this reflects the relatively low resource costs in and large resource availability in the US. and we do not impose any export constraint of coal in the GCAM core or the CMP.

USA crude oil production is lower than other sources shown in Fig.6 starting around 2030. Note that AEO 2025 Reference case projects a lower crude oil production in 2050 than AEO 2023 Reference case, and the GCAM-core projections align better with AEO 2025 than AEO 2023. Crude oil production is lower starting around 2030 in this CMP compared to other sources for two reasons. First, this CMP assigns a short lifetime (10 years) for tight oil wells and has limited tight oil resources available beyond what is currently in production. This is evident in Figures 7 and 8b where onshore unconventional crude oil production declines sharply in 2030. Recent reports (e.g., USGS 2026 https://pubs.usgs.gov/fs/2026/3059/fs20263059.pdf) contain more recent data on undiscovered tight oil and gas resources; an additional core model proposal is planned to incorporate this newer data soon. Second, once tight oil resources are exhausted, the other remaining unconventional heavy oil sub-resource (i.e., the tar sands and oil shale) have relatively higher resource costs compared to the costs implemented in GCAM core; most of the production after 2050 are the production of these heavy unconventional oil in both the CMP (Fig.7) and the current GCAM-32 core (figure not shown). In this CMP, the production of onshore unconventional heavy oil is mainly in CO and WY (Fig.8b).

For natural gas, the natural gas production in GCAM-USA-CMP generally aligns with GCAM-USA-core. However, AEO projections show much higher projections through 2050. One reason is that in GCAM-USA-CMP, the U.S. is projected to become an increasingly larger net importer of natural gas over time (Fig.9). Therefore, the domestic production of natural gas is projected to decrease while imports increase. However, the overall regional demand of natural gas in USA in

2050 in GCAM-USA-CMP is 37 EJ (Fig.9), still lower than the AEO projection of 46 EJ (Fig.6).

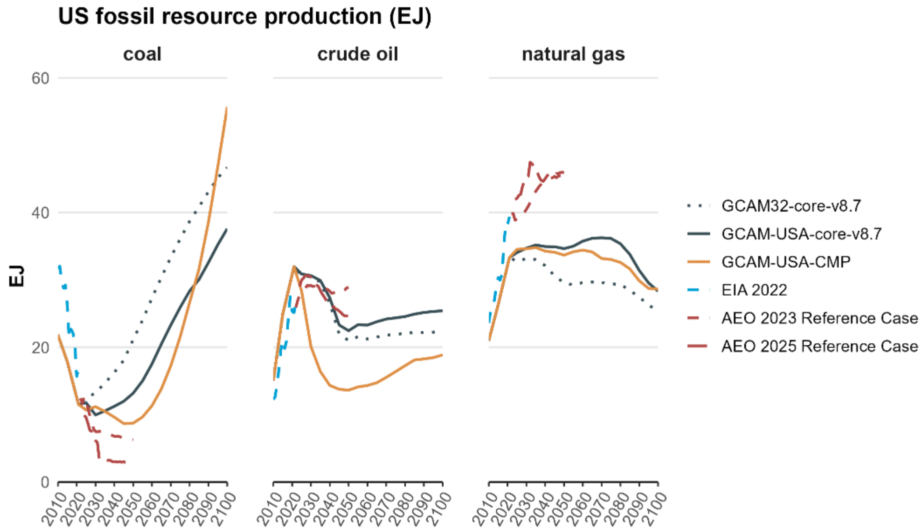



**Fig.6 GCAM-USA outputs of fossil resource production based on this CMP (GCAM-USA-CMP) aggregated from state-level production to national-level production, compared with the current master, i.e., GCAM-core-v8.7 (both GCAM and GCAM-USA) and other external sources**

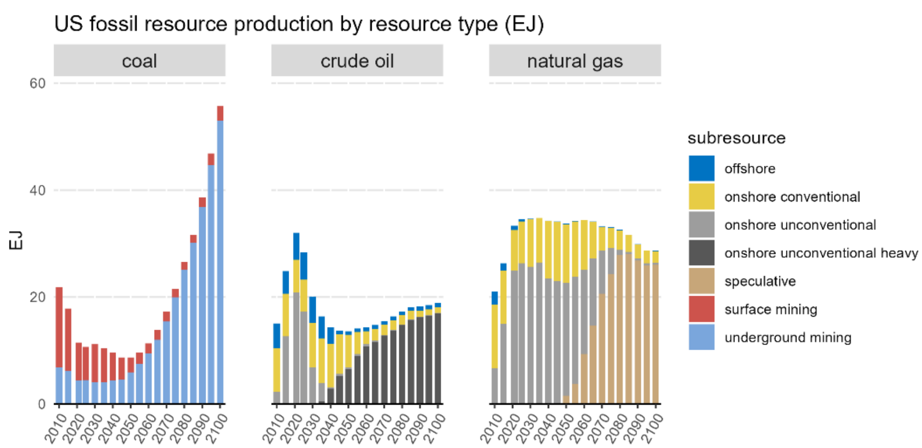
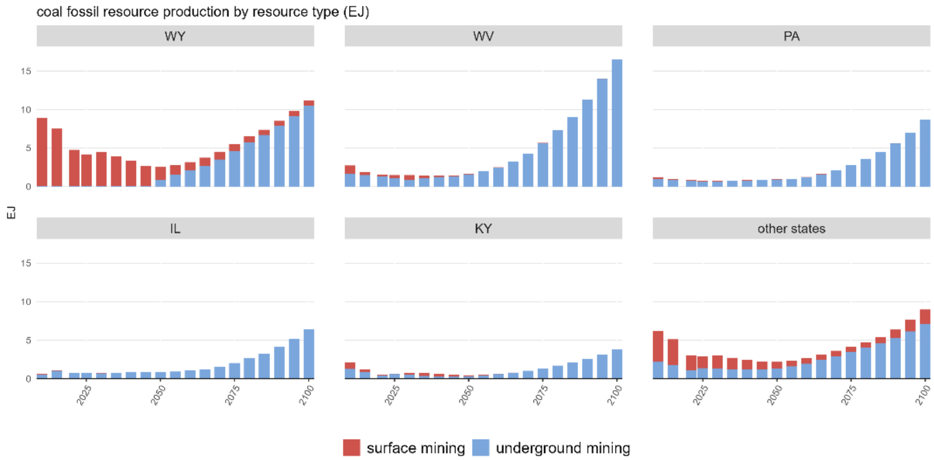
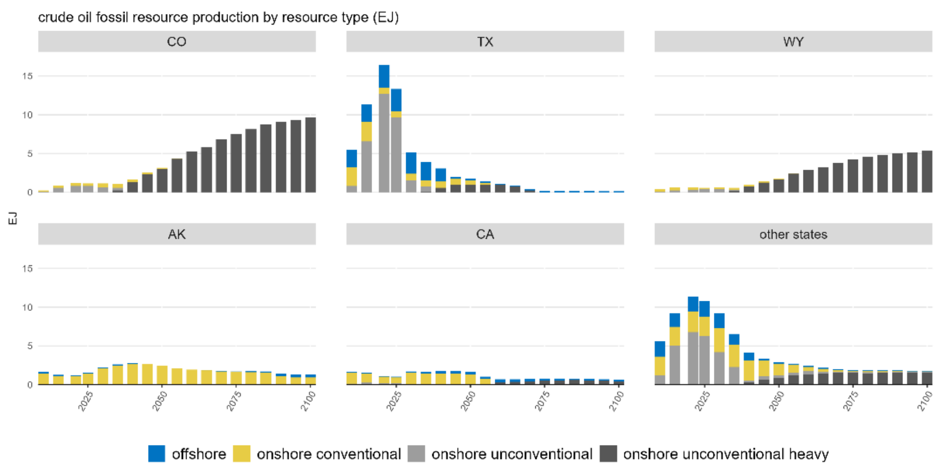


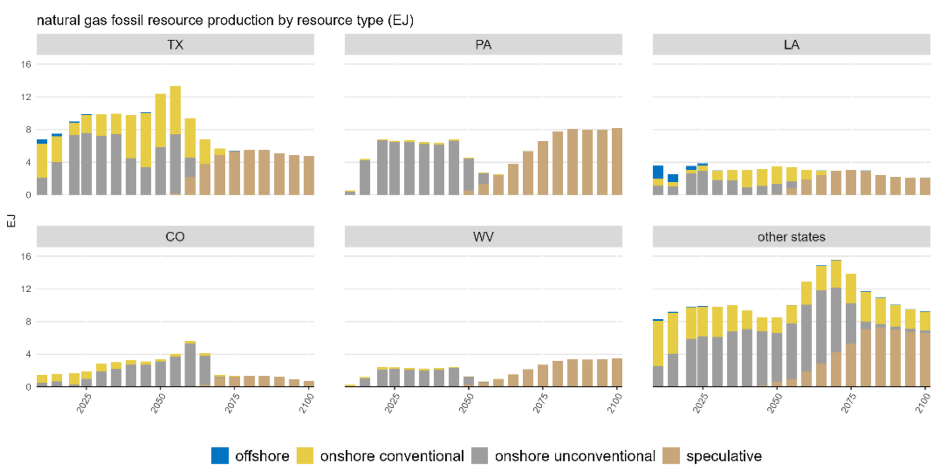

**Fig.7 U.S. fossil resource production by resource and sub-resource based on this CMP aggregated from state-level production**




(a)




(b)




(c)

**Fig.8 Resource production of (a) coal, (b) crude oil, and (c) natural gas by sub-resource in top five producing states based on cumulative production over 1990-2100, results from GCAM-USA-CMP**

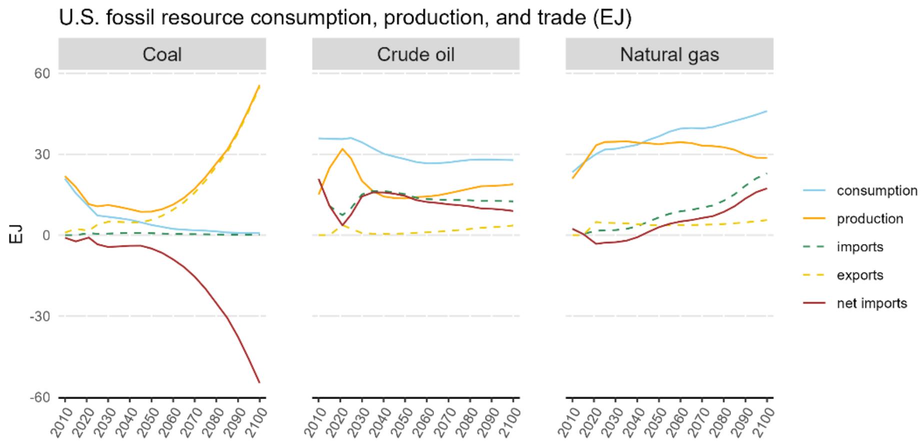



**Fig.9 U.S. fossil resource consumption, production, and trade, results from GCAM-USA-CMP**

## Emissions

The CO2 emissions from resource production, comparing GCAM-USA-CMP with GCAM-core (v8.7), are very different for coal and crude oil, while similar for natural gas (Fig.10). The difference in $CO_2$ emissions for coal production is expected due to the implementation of new emission factors for surface and underground mining, respectively, based on the splits provided in the EPA inventory. This is different from GCAM-core where only one sub-resource category "coal" is included. The significant higher $CO_2$ emissions in GCAM-core from crude oil production (Fig.11 and 12) come from the much higher production of unconventional (heavy) oil in the core than GCAM-USA-CMP as discussed above. Note that the "unconventional oil" in the core represents the heavy unconventional oil type such as oil sands, which is associated with much higher $CO_2$ emissions than conventional oil per unit of production.

U.S. GHG emissions are comparable among this CMP, GCAM-core, and GCAM-USA-core (Fig.13 and 14). The majority of GHG emissions from coal and natural gas production are $CH_4$, while $CO_2$ is the main source of GHG emissions for crude oil production starting around 2050 (Fig.15). This is further shown in detail by fossil sub-resource category (Fig.16).

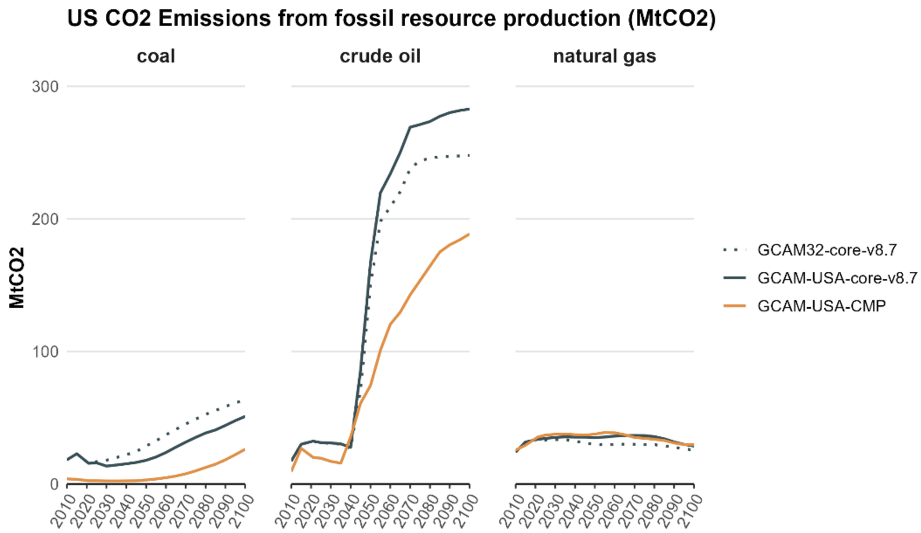



**Fig.10 U.S. $CO_2$ emissions from fossil resource production, comparing this CMP and GCAM-core**

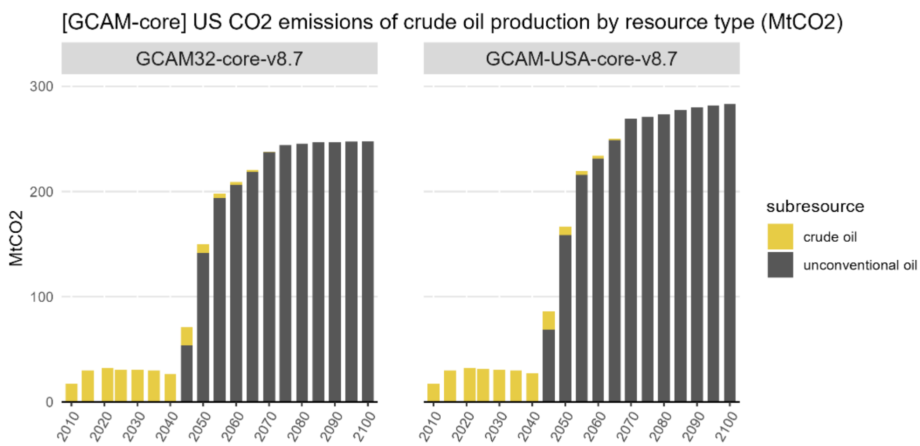



**Fig.11 U.S. $CO_2$ emissions from crude oil production by sub-resource in the GCAM-core (v8.7). Note that the "unconventional oil" in the core represents the heavy unconventional oil type such as oil sands, which is associated with much higher CO2 emissions than conventional oil per unit of production.**

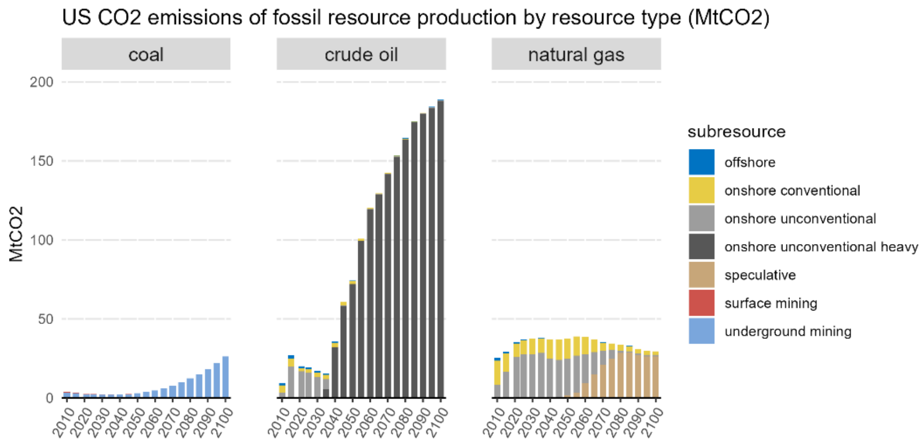



**Fig.12 U.S. $CO_2$ emissions from fossil resource production by sub-resource in the GCAM-USA-CMP**

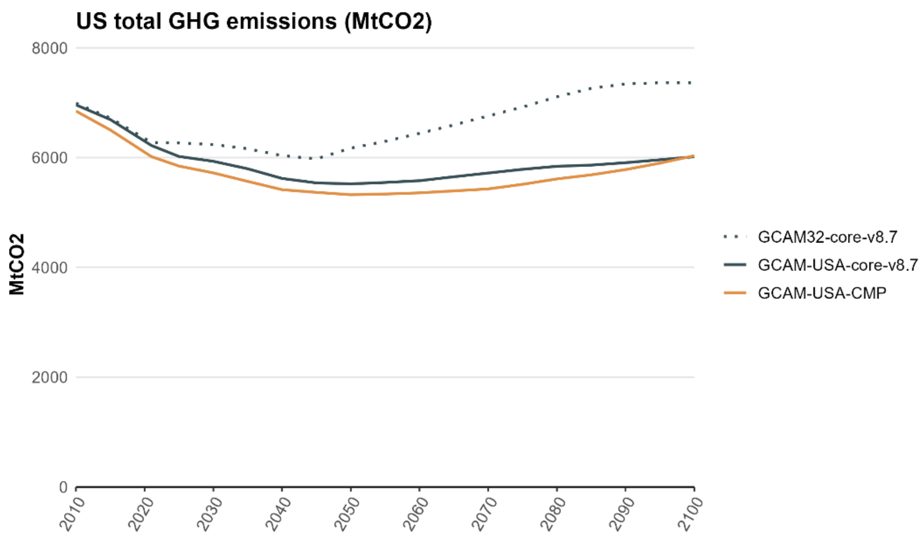



**Fig.13 U.S. total GHG emissions comparison between results based on GCAM-core and this CMP**

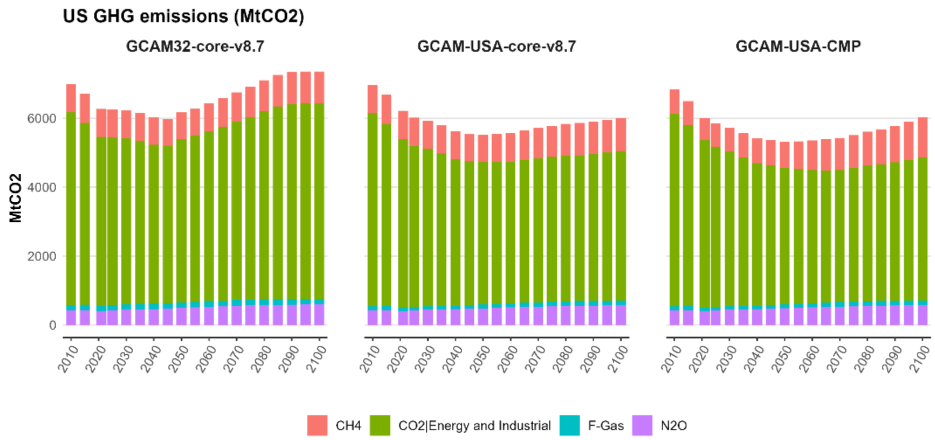



**Fig.14 U.S. GHG emissions by gas type - comparison between results based on GCAM-core and this CMP**

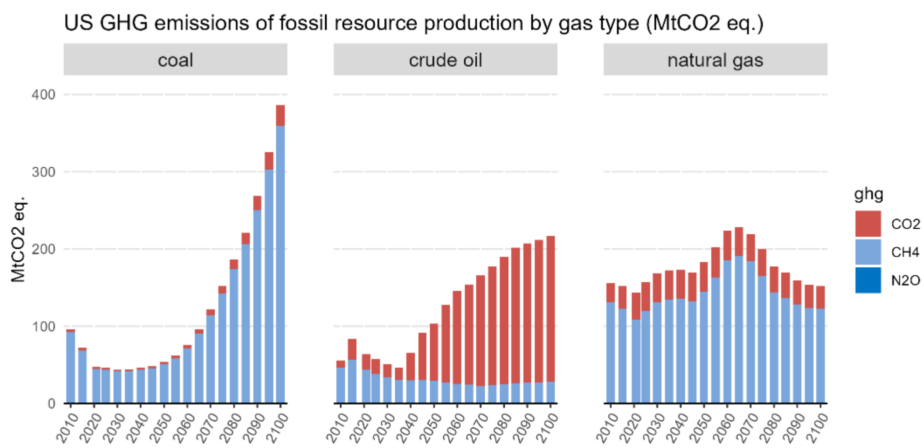
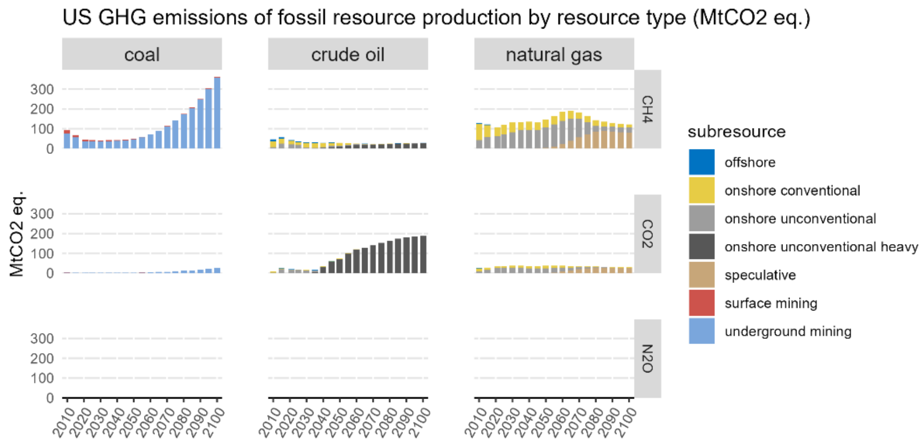
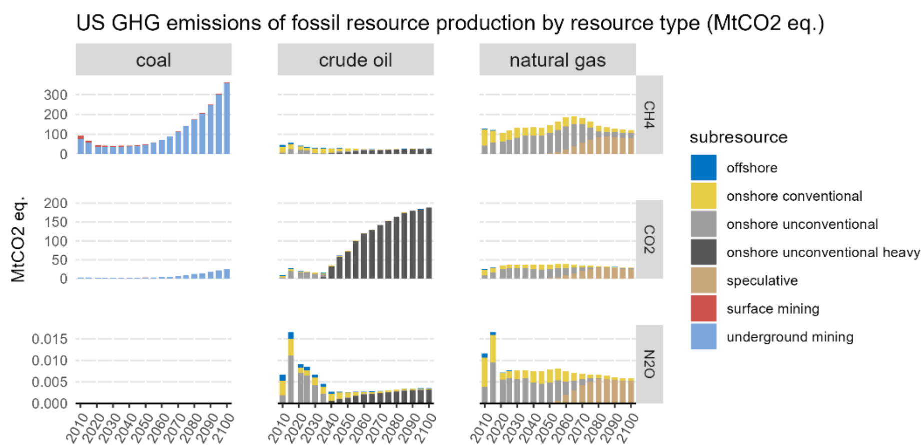



**Fig.15 U.S. GHG emissions from fossil resource production by gas type in the GCAM-USA-CMP**




(a)




(b)

**Fig.16 U.S. GHG emissions from fossil resource production by sub-resource in the GCAM-USA-CMP. (a) shows the results in the same y-axis scale; (b) shows different scales of y-axis.**