IM3

INTEGRATED MULTISECTOR MULTISCALE MODELING

Reproducible Integrated Multisectoral, Multi-model Frameworks – Lessons Learned from IM3
(It Takes a Village…and a Mayor)

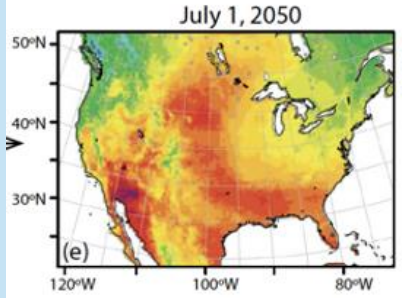
Jennie S. Rice, Casey D. Burleyson, Casey McGrath, Kendall Mongird, Em Rexer, Travis Thurber, Chris R. Vernon (all PNNL)

This research is supported by the U.S. Department of Energy, Office of Science, as part of research in MultiSector Dynamics, Earth and Environmental System Modeling Program

Pacific Northwest National Laboratory · Berkeley Lab · Oak Ridge National Laboratory · Sandia National Laboratories · Baylor University · Boston University · Cornell University · NC State University · The University of North Carolina at Chapel Hill · UC Davis University of California · PennState

Integrated Multisector, Multiscale Modeling Example

How will the price of electricity, power outages, and the infrastructure landscape be affected by future climate, energy system transitions, multisectoral energy-water-land-economy interactions, and socioeconomics?

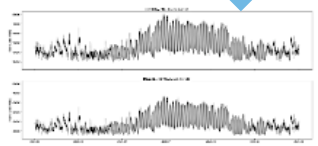


Thermodynamic Global Warming Scenarios

IM3 HYPERFACETS

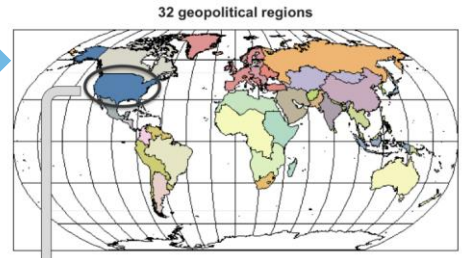
Socioeconomic Projections (SSP3 and SSP5)

IM3



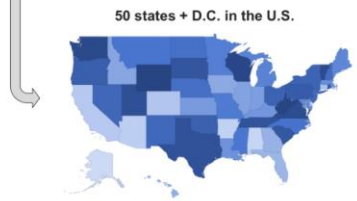
TELL: Total ELectricity Loads Model

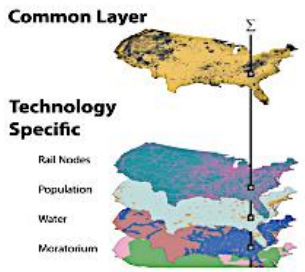
IM3



GCAM-USA

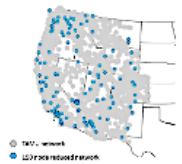
GCIMS





CERF: Infrastructure Siting

IM3



GOWEST: Grid Operations West

IM3

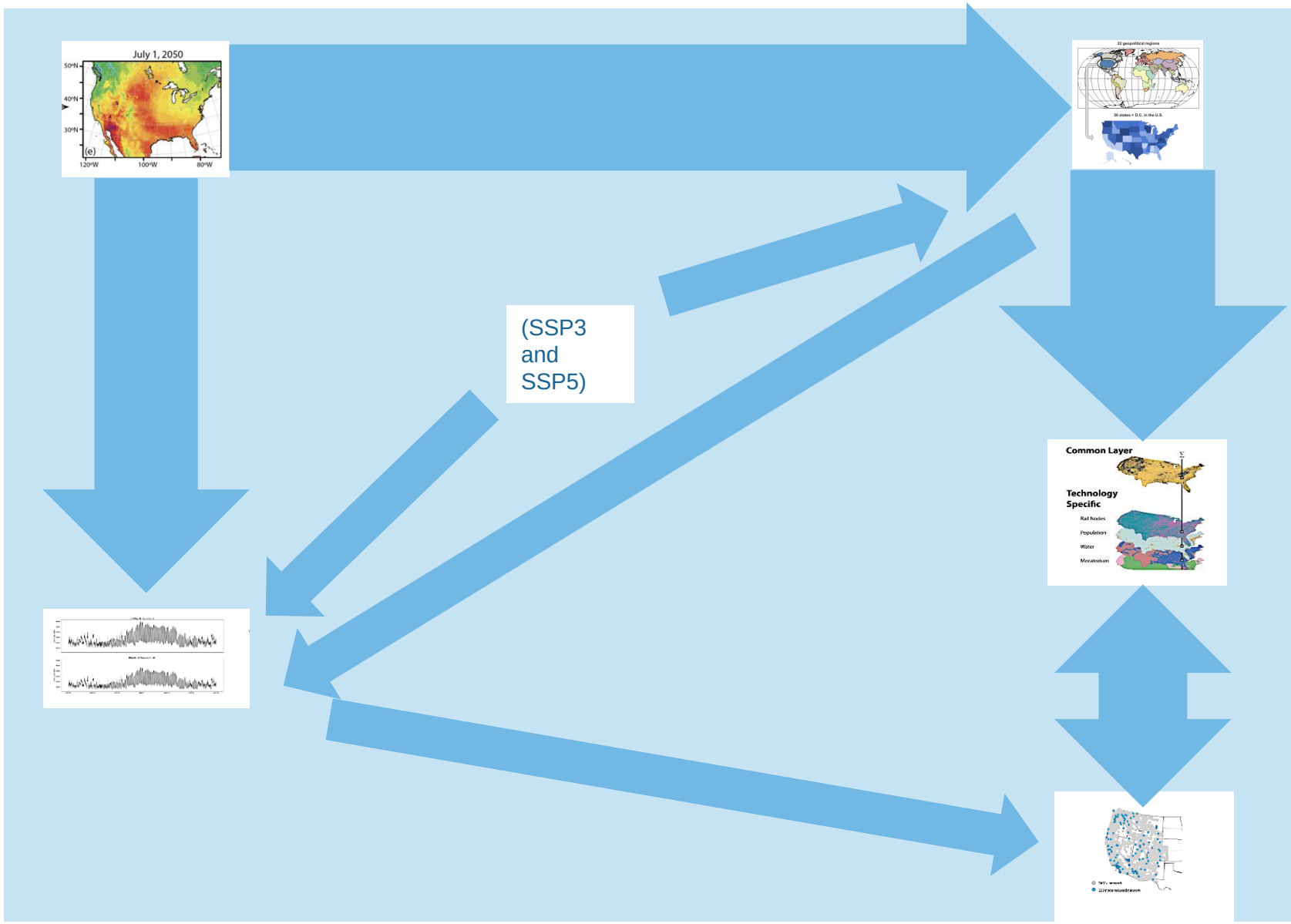
Global/ National

Local

Mongird et al. in prep; Akdemir et al. in prep; Wild et al. in prep

The Real Work of Integrated Modeling is in the Arrows

How will the price of electricity, power outages, and the infrastructure landscape be affected by future climate, energy system transitions, multisectoral energy-water-land-economy interactions, and socioeconomics?



Key Challenges For Multi-Model Integration and Reproducibility

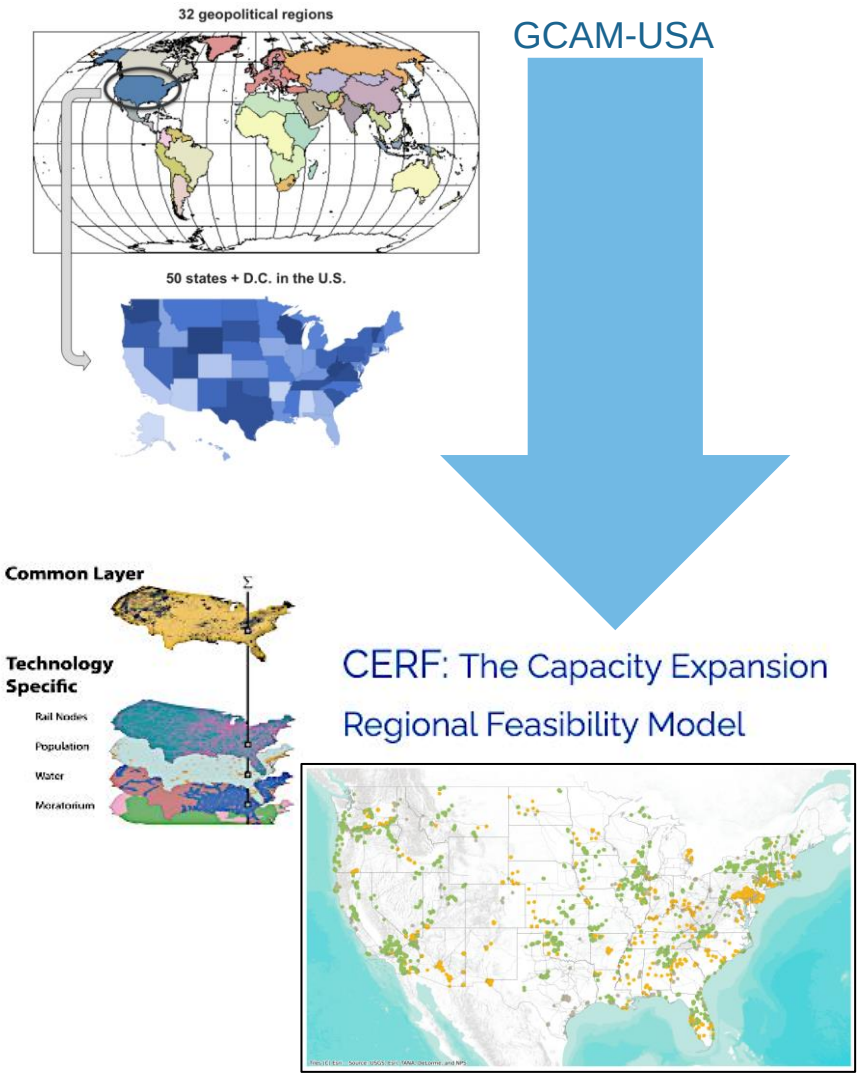
- Many best-in-class models used in the project were designed originally for standalone use by a specific community of users
- Differences in spatial, temporal, and process resolution between models
- A large team across eleven collaborating institutions
- Varying levels of software engineering expertise across collaborators
- Terminology/jargon differences across subject matter areas; You don't know what you don't know

- Disparate experiences with open science
- Transitioning from desktop to HPC environment
- Time and resources to manage QA/QC and integration
- Documenting the complexity

Structural: management, funding, staffing, resources

"But I always use MATLAB!"

Data Extraction/Transformation Example: GCAM-USA to CERF (State-level Capacity Expansion Plan to On-the-ground Power Plant Sitings)



- **Convert**
  - New Vintage Annual Energy Generation to Capacity by tech and state
  - New Vintage Capacity to # of power plants by technology and state
    - Account for solar and wind resource availability
  - Retirement Schedule to individual power plant retirements by state
- **Extract**
  - Variable costs by tech
  - Plant lifetime
  - Fuel costs and fuel cost escalation
  - Carbon prices (if RCP4.5)
  - Emissions content of fuels
  - CCS efficiency
- **Key Challenges**
  - GCAM-USA extraction tools and data structures not designed for these needs
  - Terminology and variable name misunderstandings between teams
  - Harmonizing assumptions implicit in GCAM input data (e.g., unit size)

Lesson Learned #1: It Takes a Village

- Multisectoral “translators” needed to converge the jargon
- Scientists from each model willing to work together and co-develop the transformation scripts
- All: Defining the variables needed in sufficient detail so that data transformation scripts can be written
- Software engineers to support the data transformations and sometimes optimize/refactor code
- Experts to design QA/QC for model outputs and for outputs of data transformation scripts—especially for first time coupling
- Multisectoral review panels to bring different perspectives in QA/QC
- Data management experts to mint the datasets and code versions; detailed metadata; MSD-LIVE
- Collaboration experts (e.g., Slack channels devoted to each arrow

Lesson Learned #2: …and a “Mayor”

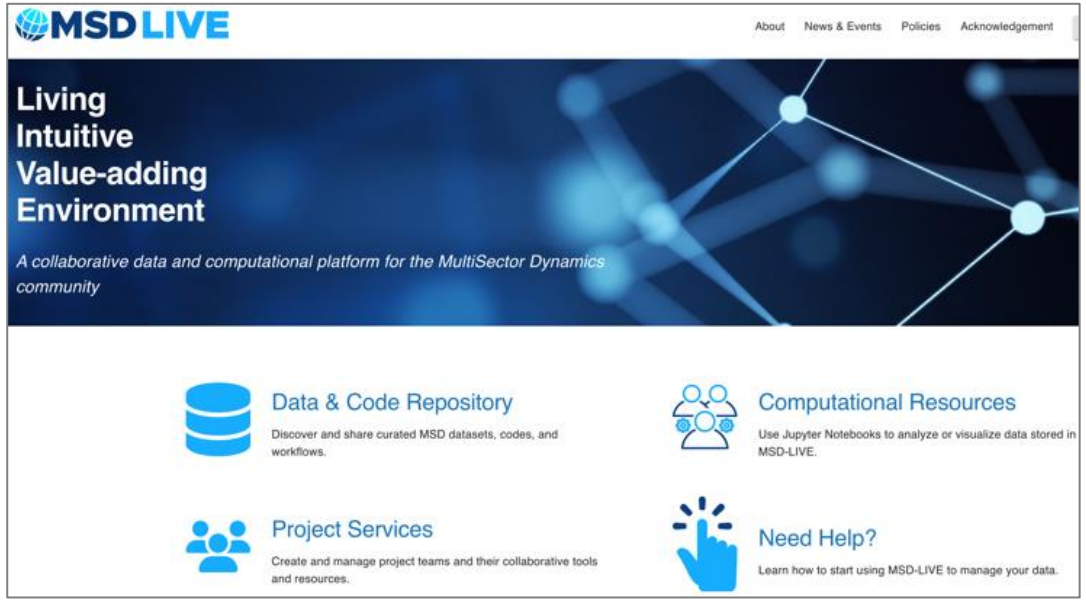
PI/Experiment Lead Role:

- Require open source – all data, models, tools, solvers, etc.
- Require workflow documentation – the meta-repository
- Work with the team to translate the science question into the necessary multi-model workflow

- Be accountable for the entire modeling chain – or delegate this
- Be hands-on: go into “the weeds” to help team identify and resolve issues while keeping an eye on the larger picture

Leading Open Source and Fair Data Approaches for MSD Community

- IM3 requires all open-source code, data, and tools – no exceptions
- IM3 ideas and data management objectives led to MSD-LIVE project led by Casey Burleyson



- IM3 developed the concept of a "meta-repository" to accompany each submitted manuscript (Vernon, C.R. 2023) **https://immm-sfa.github.io/metarepo/**

Input data

Dataset	Repository Link	DOI
GCAM-USA Output	https://data.msdlive.org/records/43sy2-n8y47	https://doi.org/10.57931/1989373
TGW Weather Forcing	https://data.msdlive.org/records/cnsy6-0y610	https://doi.org/10.57931/1960530

Output data

The output of the TELL model is stored in the data repository linked below. The post-processed files (resulting from the analysis scripts itemized below) are stored in the /data directory in this meta-repository.

Dataset	Repository Link	DOI
TELL Output	https://data.msdlive.org/records/r0rvc-kjw89	https://doi.org/10.57931/2228460
Post-Processed Data	https://github.com/IMMM-SFA/burleyson-etal_2023_applied_energy/tree/main/data	https://doi.org/10.5281/zenodo.10278502

Contributing modeling software

Model	Version	Repository Link	DOI
GCAM-	v5.3	https://data.msdlive.org/records/r52tb-	https://doi.org/10.57931/1960381

Reviewer feedback: *"I agree that this paper should be deemed fully reproducible and given the* ***highest rating for reproducibility.****"*

IM3

INTEGRATED MULTISECTOR MULTISCALE MODELING

Thank you

jennie.rice@pnnl.gov

This research is supported by the U.S. Department of Energy, Office of Science, as part of research in MultiSector Dynamics, Earth and Environmental System Modeling Program

Pacific Northwest National Laboratory · Berkeley Lab · Oak Ridge National Laboratory · Sandia National Laboratories · Baylor University · Boston University · Cornell University · NC State University · The University of North Carolina at Chapel Hill · UC Davis University of California · PennState