The Journal of Open Source Software

tell: a Python package to model future total electricity loads in the United States

Casey R. McGrath [1], **Casey D. Burleyson** [1], **Zarrar Khan** [2], **Aowabin Rahman** [1], **Travis Thurber** [1], **Chris R. Vernon** [1], **Nathalie Voisin** [1], and **Jennie S. Rice** [1]

1 Pacific Northwest National Laboratory, Richland, WA., USA **2** Joint Global Change Research Institute, PNNL, College Park, MD, USA

DOI: 10.21105/joss.04472

Software

- Review
- Repository
- Archive

Editor: Frauke Wiese

Reviewers:

- @euronion
- @chrklemm

Submitted: 03 June 2022
Published: 08 November 2022

Summary

The purpose of the Total ELectricity Load (`tell`) model is to generate 21st century profiles of hourly electricity load (demand) across the Conterminous United States (CONUS). `tell` loads reflect the impact of climate and socioeconomic change at a spatial and temporal resolution adequate for input to an electricity grid operations model. `tell` uses machine learning to develop profiles that are driven by projections of climate/meteorology and population. `tell` also harmonizes its results with United States (U.S.) state-level, annual projections from a national- to global-scale energy-economy model. This model accounts for a wide range of other factors affecting electricity demand, including technology change in the building sector, energy prices, and demand elasticities, which stems from model coupling with the U.S. version of the Global Change Analysis Model (GCAM-USA). `tell` was developed as part of the Integrated Multisector Multiscale Modeling (IM3) project. IM3 explores the vulnerability and resilience of interacting energy, water, land, and urban systems in response to compound stressors, such as climate trends, extreme events, population, urbanization, energy system transitions, and technology change.

Statement of Need

To understand and plan for the resilience of the electricity system in the U.S., we need load projections that are responsive to compounding anthropogenic and natural stressors and have sufficient spatial and temporal resolution for grid stress modeling (Carvallo & Goldman, 2018; Oikonomou & Voisin, 2022). `tell` fills this gap. `tell` is unique from other load projection models in several ways. It coherently blends aspects of short- and long-term load projections. Short- and medium-term load models most commonly relate meteorology and day-of-week parameters to loads (Hong T. & Fan, 2016 and references therein). Longer-term models also use meteorology/climate as explanatory variables, typically relying on climate trends, but also include "macro" variables like the decadal evolution of population or economic indicators (Al-Hamadi & Soliman, 2005; W. Hong T. & Xie, 2014; Lindberg & Korpås, 2019). There is limited research that combines short- and long-term modeling approaches (Behm & Praktiknjo, 2020; Boßmann & Staffell, 2015; Lindberg & Korpås, 2019). The most relevant existing approach is Behm et al. (Behm & Praktiknjo, 2020), which used an artificial neural network approach to simulate hourly electricity loads in a handful of European countries based on population-weighted meteorology and date parameters and then used exogenous future annual peak loads to scale their hourly projections. The scaling to future peak loads is analogous to how we use the national to global energy-economy model in `tell`.

`tell` has a spatial component that allows us to distribute projected loads where they would occur spatially within a high-resolution grid operations model. Most other models are for a specific geographic context (i.e., a single utility or Balancing Authority [BA]) and therefore do not have any spatial element to their projections (Al-Hamadi & Soliman, 2005; Carvallo & Goldman, 2018; W. Hong T. & Xie, 2014; Lindberg & Korpås, 2019). In contrast, `tell` is a meta-model made up of an aggregation of individual MLP models for 54 BAs. `tell` covers the entire CONUS so that its output can be flexibly aggregated as input to a grid operations model in any of three U.S. grid interconnections. Finally, `tell` is based entirely on publicly available, open data and is being released as an extensively documented open-source code base so that it can be freely and easily reused by others. While `tell` is designed for 54 BAs in the CONUS and relies on the GCAM-USA model, the core MLP modeling approach could be readily adapted to work in other regions with adequate historical data.

Design and Functionality

`tell` integrates aspects of both short- and long-term projections of electricity demand in a coherent and scalable way. `tell` takes time series meteorological data at one-hour resolution as input and uses the temporal variations in weather to project hourly time-series of total electricity demand. The core predictions in `tell` are based on a series of multilayer perceptron (MLP) models (Pedregosa & Duchesnay, 2011) that relate historical meteorology to coincident BA-scale hourly loads for 54 independent BAs. The BA load projections are first disaggregated to the county-level and then summed and scaled to match the annual state-level total electricity demands projected by the U.S. version of the Global Change Analysis Model (GCAM-USA) (Binsted & Wise, 2022; Iyer & Williams, 2017). GCAM-USA is designed to capture the long-term co-evolution of energy, water, land, and economic systems. This approach allows `tell` to reflect changes in the shape of the load profile due to variations in weather and climate as well as the long-term evolution of energy demand due to changes in population, technology, and economics. The projections from `tell` are quantitatively and conceptually consistent across county-, state-, and BA-scales.

The basic workflow for `tell` proceeds in six sequential steps: 1. Formulate empirical models that relate the historical observed meteorology to the hourly time-series of total electricity demand for 54 BAs that report their hourly loads in the EIA-930 dataset (Fig. 1a). 2. Use the empirical models to project future hourly loads for each BA based on IM3's future climate scenarios generated using the Weather Research and Forecasting (Skamarock & X.-Y., 2019) model. 3. Distribute the hourly loads for each BA to the counties that BA operates in and then aggregate the county-level hourly loads from all BAs into annual state-level loads (Fig. 1b-g). 4. Calculate annual state-level scaling factors that force the bottom-up annual state-level total loads from `tell` to match the annual state-level total loads from GCAM-USA. 5. Apply the state-level scaling factors to each county- and BA-level time-series of hourly total demand. 6. Output yearly 8760-hr time-series of total electricity demand at the county-, state-, and BA-scale that are conceptually and quantitatively consistent with each other.

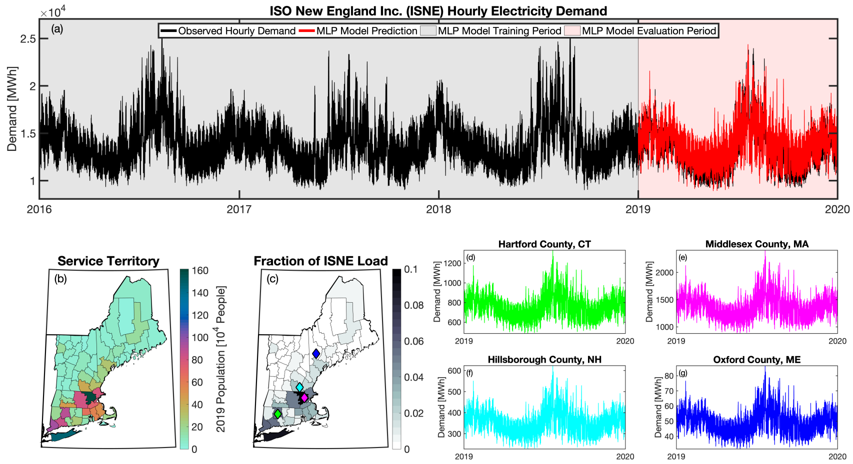

Fig. 1. a) Time-series of observed and projected hourly electricity demand within the Independent System Operator of New England (ISNE) BA during the training (2016-2019) and evaluation (2019) periods for TELL; b) County-level populations within the ISNE BA service territory in 2019; c) Fraction of the total population within the BA that lives in each county; and d-g) Time-series of projected hourly electricity demand in 2019 for select counties based on their population weights.

`tell` is an open source model that can be accessed via the GitHub Repository. The repository also includes a Jupyter Notebook that provides a walk through of the core functionality of `tell`. This notebook also provides easy access to statistical validation results between the forecasts and observed data. Finally, more details about how the model was formulated and its intended purpose can be found in the `tell` User Guide.

Acknowledgements

This research was supported in part by the U.S. Department of Energy, Office of Science, as part of research in MultiSector Dynamics, Earth and Environmental System Modeling Program. The Pacific Northwest National Laboratory is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830. The views and opinions expressed in this paper are those of the authors alone.

References

Al-Hamadi, H. M., & Soliman, S. A. (2005). Long-term/mid-term electric load forecasting based on short-term correlation and annual growth. *Electric Power Systems Research*, *74(3)*, 353–361. https://doi.org/10.1016/j.epsr.2004.10.015

Behm, N., C., & Praktiknjo, A. (2020). How to model european electricity load profiles using artificial neural networks. *Applied Energy*, *277*, 115564. https://doi.org/10.1016/j.apenergy.2020.115564

Binsted, I., M., & Wise, M. (2022). GCAM-USA v5.3_water_dispatch: Integrated modeling of subnational U.S. Energy, water, and land systems within a global framework. *Geoscientific Model Development*, *15(6)*, 2533–2559. https://doi.org/10.5194/gmd-15-2533-2022

Boßmann, T., & Staffell, I. (2015). The shape of future electricity demand: Exploring load curves in 2050s germany and britain. *Energy*, *90*, 1317–1333. https://doi.org/10.1016/j.energy.2015.06.082

Carvallo, L., J. P., & Goldman, C. A. (2018). Long term load forecasting accuracy in electric utility integrated resource planning. *Energy Policy*, *119*, 410–422. https://doi.org/10.1016/j.enpol.2018.04.060

Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, *32(3)*, 914–938. https://doi.org/10.1016/j.ijforecast.2015.11.011

Hong, W., T., & Xie, J. (2014). Long term probabilistic load forecasting and normalization with hourly information. *IEEE Transactions on Smart Grid*, *5(1)*, 456–462. https://doi.org/10.1109/TSG.2013.2274373

Iyer, L., G., & Williams, J. H. (2017). Measuring progress from nationally determined contributions to mid-century strategies. *Nature Climate Change*, *7*, 871–874. https://doi.org/10.1038/s41558-017-0005-9

Lindberg, S., K. B., & Korpås, M. (2019). Long-term electricity load forecasting: Current and future trends. *Utilities Policy*, *58*, 102–119. https://doi.org/10.1016/j.jup.2019.04.001

Oikonomou, T., K., & Voisin, N. (2022). Core process representation in power system operational models: Gaps, challenges, and opportunities for MultiSector dynamics research. *Energy*, *238*, 122049. https://doi.org/10.1016/j.energy.2021.122049

Pedregosa, V., F., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Skamarock, K., W. C., & X.-Y., H. (2019). A description of the advanced research WRF version 4. *NCAR Technical Note NCAR/TN-556+STR*, 145.