



OPEN

DATA DESCRIPTOR

# Scenarios of future Indian electricity demand accounting for space cooling and electric vehicle adoption


Marc Barbar  , Dharik S. Mallapragada , Meia Alsup & Robert Stoner

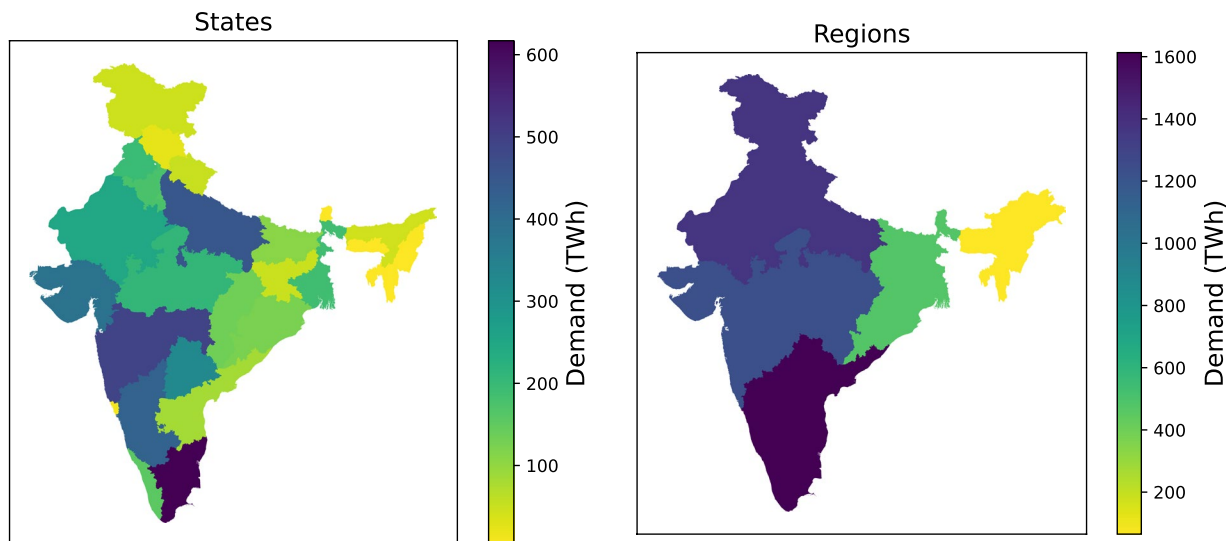
India is expected to witness rapid growth in electricity use over the next two decades. Here, we introduce a custom regression model to project electricity consumption in India over the coming decades, which includes a bottom-up estimate of electricity consumption for two major growth drivers, air conditioning, and vehicle electrification. The model projections are available at a customizable level of spatial aggregation at an hourly temporal resolution, which makes them useful as inputs to long-term electricity infrastructure planning studies. The approach is used to develop electricity consumption data sets spanning various technology adoption and growth scenarios up to the year 2050 in five-year increments. The aim of the data is to provide a range of scenarios for India's demand growth given new technology adoption. With long-term hourly demand projections serving as an essential input for electricity infrastructure modeling, this data publication enables further work on energy efficiency, generation, and transmission expansion planning for a fast-growing and increasingly important region from a global climate mitigation perspective.

## Background & Summary

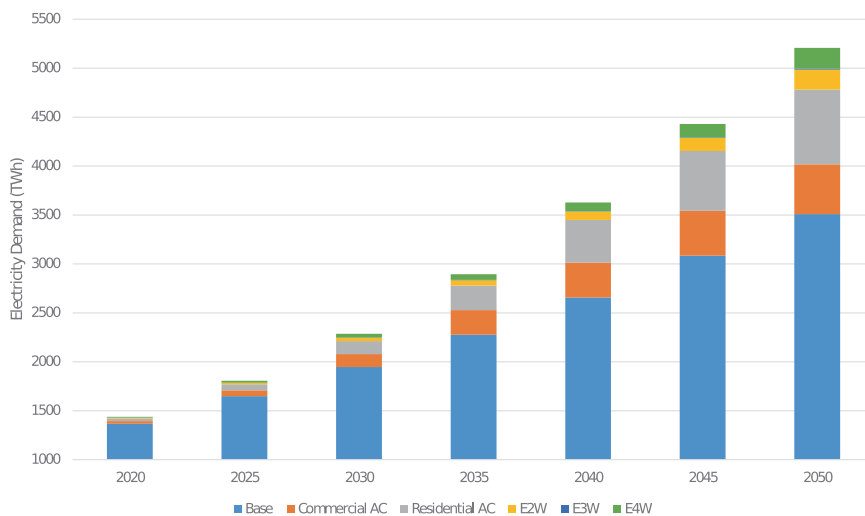
Many assessments of future electricity demand in India project large increases in electricity consumption from adoption of air conditioning technologies in the buildings sector over the next two decades<sup>1–3</sup>. This large growth is likely to make India among the top nations in terms of electricity consumption, implying that technology choices related to energy consumption and production in India are likely to play a significant impact on global climate change mitigation efforts. Additionally, the Indian government has been pushing for the transportation sector's electrification, starting with two- and three-wheel vehicles, which is further likely to increase overall electricity demand. As of 2020 in India, there are 152,000 registered electric vehicles<sup>2</sup>. Air conditioning (AC) related electricity demand accounted for 32.7 TWh, contributing to less than 2.5% of the total demand in 2019<sup>3</sup>. However, both air conditioning and transport electrification are anticipated to introduce structural changes in the temporal and spatial trends in electricity consumption patterns, that has important ramifications for long-term resource planning for the electricity sector<sup>4</sup>. This paper presents an bottom-up approach to estimate electricity consumption in India for various scenarios of technology and policy adoption with a specific focus on providing aggregated consumption estimates as well as spatio-temporally resolved consumption profiles that would be relevant for regional and national electricity system planning studies. The approach enables quantifying the impact of various growth and technology adoption scenarios on quantity and pattern in electricity consumption. The datasets detailed in this paper include annual energy consumption at India's state, regional, and national levels as visualized in Fig. 1, as well as underlying consumption profiles at an hourly time resolution. The annual energy consumption is forecasted on a five-year increment to 2050. Figure 2 shows one scenario of national electricity demand forecast. In addition to the snapshot of annual consumption, hourly load profiles are developed at the same resolution as seen in Fig. 3.

The forecasting is divided into two steps: business-as-usual and technology. Business-as-usual is a statistical model that infers data it can be trained on i.e. historical electricity demand. The technology model is a bottom-up approach that adds new loads to the total demand. Among new loads, we focus on residential and commercial cooling as well as various electric vehicles (EV). Some key insights from cooling<sup>3</sup> and EV<sup>2</sup> studies highlighting

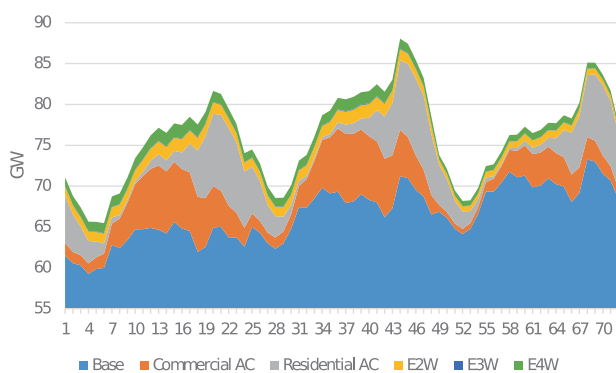
Massachusetts Institute of Technology, Cambridge, 02139, USA.  e-mail: [mbarbar@mit.edu](mailto:mbarbar@mit.edu)



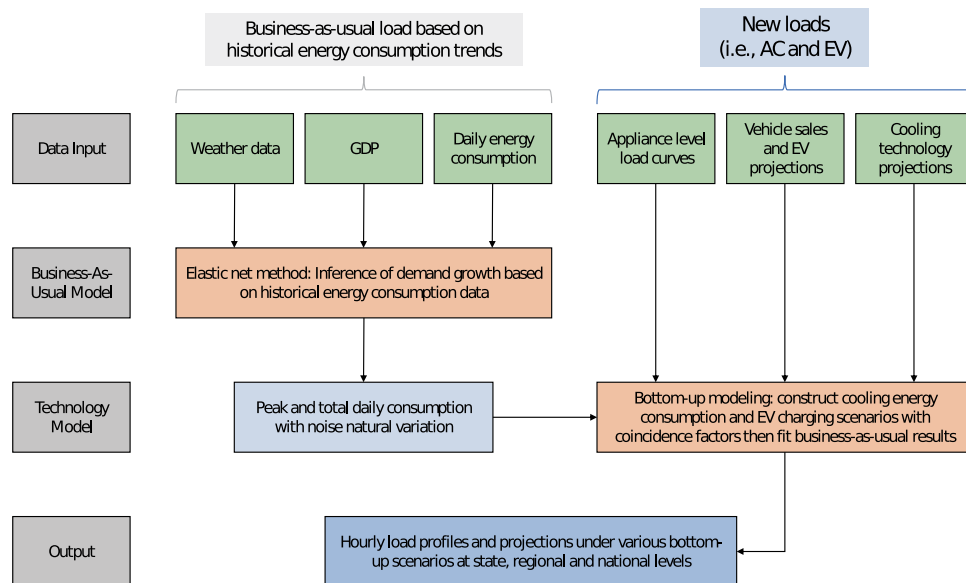
**Fig. 1** State and regional level distribution of annual electricity 2050 for stable GDP growth, baseline cooling, and home electric vehicle (EV) charging scenario.



**Fig. 2** Summary results of India's electricity demand forecasting at national level with stable GDP growth, baseline cooling, and home electric vehicle (EV) charging.



**Fig. 3** 2030 Load profile for southern Region across three days in summer. Scenario: stable GDP growth, baseline cooling, home electric vehicle (EV) charging.



**Fig. 4** Simplified schematic of methods; inputs in green, models in red, outputs in blue.

peak demand development motivate the need for demand forecasting at the hourly resolution. Cooling demand due to mainly split unit air conditioning installation in India is expected to increase the peak to mean ratio (also sometimes referred to the “peakiness”) of electricity demand in India as well as shift the timing of peak demand from evenings to midnight<sup>3</sup>. While electric vehicles do not constitute a large portion of the total demand, certain charging schemes can contribute significantly to the peak demand<sup>2</sup>. Numerous energy demand forecast for India have recently been published as decadal snapshots<sup>1,4,5</sup>, however granularity of demand at an hourly resolution has not been presented in these studies. Our approach enables quantifying the impact of different technology and structural elements, such as adopting energy efficient vs. baseline cooling technology or work-place charging vs. home charging for EVs, on the hourly electricity consumption profiles. These insights and the accompanying data sets are essential to carry out generation and transmission expansion as well as distribution network planning, and are thus essential for a sustainable energy infrastructure development in the Indian context.

Similar to other forecasting studies, we model Gross domestic product (GDP) growth<sup>6</sup> to be the main economic driver of the business-as-usual demand forecasting, and thus three scenarios are introduced: slow, stable, and rapid GDP growth. We examine two AC load scenarios: energy-efficient equipment and baseline equipment per the International Energy Agency’s Future of Cooling study<sup>3</sup>. Finally, we evaluate three EV charging mechanisms: home, work, and public charging. This totals the number of data sets spanning three input dimensions to 18 scenarios. Technology adoption growth has been correlated with economic growth under the assumption that new technologies are adopted faster when the economy is growing faster and vice versa. We present two cooling scenarios to highlight the difference in energy-efficient and regular air conditioning units and bring attention to the need for policy and programs that favor energy-efficient cooling unit sales. Furthermore, we present various EV charging mechanisms to inspect the demand impacts that electric vehicle charging can have on the electric grid at different times. The produced data can be used as input to electricity infrastructure planning both at the distribution and transmission level.

## Methods

Figure 4 illustrates the major steps of our proposed demand forecasting approach. We use two models to estimate future electricity demand in India. In the first model — business-as-usual — we use a linear regression model to project daily peak and consumption on a regional basis; this is the business-as-usual scenario. We then add natural variation to the projections by finding the error between the training data and results and scaling it to every region based on seasonality. Then we fit the projected peak and total consumption to an annual hourly load profile for 2015<sup>7</sup> featuring an evening peak<sup>8</sup>. In the second model — technology model — we take AC and EV adoption into account as an additive component on top of the business-as-usual predictions. GDP data, which is an independent variable in the model, is chosen to be the main driver of growth of the business-as-usual scenario as well as technology adoption rates. The input data used are publicly available and are referenced in Table 1.

**Input data processing.** Although GDP is widely used for forecasting energy demand, it is specifically essential in the case of India, where economic growth is expected to ramp up over the next few decades similar to the recent trends in China<sup>9</sup>. We based our demand forecast on GDP projections from a PricewaterhouseCoopers (PwC) report<sup>10</sup>, that projected India’s GDP to grow from 3.6 trillion in 2020 to reach 28 trillion USD in 2050. Considering the historical national GDP data for India starting in 1990, we fit and project an exponential curve for rapid growth and an Gompertz curve for slow growth<sup>11</sup> as detailed in Table 2. We use PwC’s projections to define the stable GDP growth scenario. Curve fitting and projection results are illustrated in Supplementary Fig. 1.

Data	Source
State-wise Historical GDP	Ministry of Statistics and Programme Implementation <sup>6</sup>
Vehicle Sales and Registration	Society of Indian Automobile Manufacturers <sup>24</sup>
Air Conditioning Stock and Capacity	International Energy Agency <sup>3</sup>
Load Profile	United States Agency for International Development <sup>20</sup>
State-level sector-wise energy consumption	Power System Operation Corporation <sup>14,28</sup>

**Table 1.** Input Data Sources.

Slow	Rapid
Gompertz Growth Curve	Exponential Growth Curve
$y = Ae^{-\frac{\mu e(B-x)}{A} + 1} + C$	$y = Ae^{Bx} + C$
$A = 7 \times 10^{-29}$	$A = 1 \times 10^{-64}$
$B = 4.64 \times 10^{-2}$	$B = 8.7 \times 10^{-2}$
$C = 0$	$C = 0$
R-squared = 0.949669	R-squared = 0.989361

**Table 2.** GDP projections curve fit results.

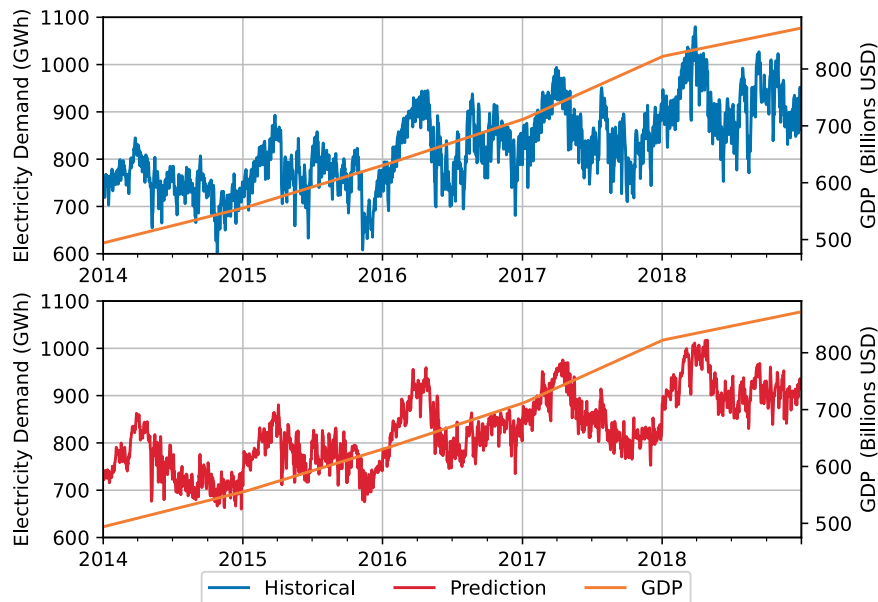
The rapid growth scenario produces an annual average growth rate of 9.5%, PwC's growth rates start at 7.8% for the first projected decade and ends at 6.2% in the final projected decade. The slow growth scenario starts at 7.2% growth rate in the first projected decade and ends at 3.9% in the final projected decade. To break down the regional energy consumption projections to state level we use the ratio of GDP per capita of the corresponding state to the GDP per capita of the region it is in. For each GDP growth scenario, we fit the same functions given state-wise data to produce GDP forecast at the same resolution. GDP per capita at state-level is computed using the projected GDP data and state level population projections<sup>12</sup>.

**GDP dependence and limitation.** Relating growth in electricity demand to GDP is a strong generalization, however it is not a novel one in the case of India. Strong correlation between economic growth and energy consumption has been established in the Indian context in this study and other studies<sup>13</sup> given data from the past two decades<sup>6</sup>. We recognize that GDP as a metric of economic growth has several limitations particularly related to projecting how economic growth is distributed among society within a state or nation. This may be the strongest limitation of the data we are presenting in the manuscript. However, lack of historical record and long-term projections of alternative open-access economic data at the desired spatial and temporal resolution limit the development of a framework to project energy consumption with other metrics. While GDP and energy consumption growths may differ in the long-run, there is an evident correlation between the two that can be used to estimate long-run energy consumption growth. Deviating away from linear regression may yield better results, however, data scarcity is again a limitation to the development of more complex models. Furthermore, this manuscript motivates the need for more bottom-up projections and not just regression models because historical consumption cannot infer consumption trends from new demand sources such as cooling and EVs.

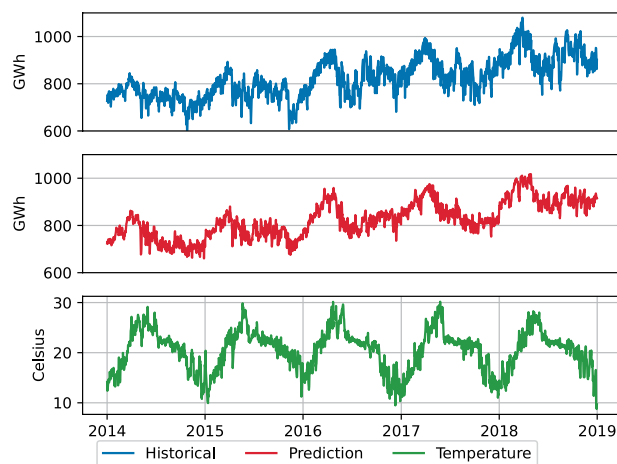
Additionally, since the Future of Cooling study by the International Energy Agency relies on GDP forecasts developed by the International Monetary Fund<sup>3</sup>, we elected to use a similar metric. We intentionally develop a large bandwidth of projection scenarios to mitigate the limitation of an individual snapshot representing a singular assumption. The motivation behind presenting the described results is ability to compare different scenarios and post-analyze the demand growth and the trade-offs. To produce a large bandwidth of growth scenarios we needed to use a straightforward metric that has enough historical data to produce various fitted curves for projections.

**Business-as-usual model.** The business as usual projections are modeled with a linear regression considering weather and economic growth features. The ground truth historical daily peak and total consumption for each electric grid were obtained from the Power System Operation Corporation (POSOCO) for 2014–2019<sup>14</sup>. The GDP used in the model was obtained, as explained in the previous section. Weather data was secured from the NASA Merra-2 data set<sup>15</sup>. The choice of features for the regression model is limited to GDP and weather variation due to the limitation in availability of data, both historical and future projections, at the desired spatial and temporal resolution. GDP is identified as a long-term parameter driving growth in year over year demand projections as highlighted in Fig. 5. Weather data is identified as a short-term parameter driving seasonal variation within a year's demand projections as highlighted in Fig. 6. Previous parametric analysis on these features and their coefficient for short and long term demand forecasting in both time and frequency domain<sup>16</sup> reinforce their use as features for the business-as-usual regression model. We present detailed outcomes for the Southern region, with further details available in<sup>16</sup>.

**NASA Merra 2 data acquisition.** For each of the five electric grid demand regions highlighted in right panel of Fig. 1, the largest cities in each region were identified using population data made available by the United



**Fig. 5** Southern region back test annual demand growth given GDP projection.



**Fig. 6** Southern region back test seasonal demand variation given weather data.

Nations<sup>17</sup>. Then, the city's latitude and longitude were used to pull down the corresponding environmental data from the Nasa Merra-2 data set. The cities used for each of the five regions are listed here:

- Northern: Delhi, Jaipur, Lucknow, Kanpur, Ghaziabad, Ludhiana, Agra
- Western: Mumbai, Ahmadabad, Surat, Pune, Nagpur, Thane, Bhopal, Indore, Pimpri-Chinchwad
- Eastern: Kolkata, Patna, Ranchi (Howrah was ignored because the environmental factors are the same as Kolkata)
- Southern: Hyderabad, Bangalore, Chennai, Visakhapatnam, Coimbatore, Vijayawada, Madurai
- Northeast: Guwahati, Agartala, Imphal

From the NASA set, 11 variables were included for each city: specific humidity, temperature, eastward wind, and northward wind (all 2 m above the surface and 10 m above the surface - eight total variables), precipitable ice water, precipitable liquid water, and precipitable water vapor. In particular, the instantaneous two-dimensional collection “inst1\_2d\_asm\_Nx (M2I1NXASM)” from NASA was used. Detailed descriptions of these variables are available in the Merra-2 file specification provided by NASA<sup>15</sup>. The environmental variables available from the NASA MERRA-2 dataset were given on an hourly basis. The daily minimum, daily, maximum, and daily average was calculated for each of the 11 variables for each day.

*Forecasts.* The business-as-usual demand forecasting problem was divided into ten separate problems, corresponding to one problem each peak and total consumption for each of the five regional grids shown in Fig. 1. To

ensure the model would not overfit the data, the model was trained with Elastic Net<sup>18</sup> to regularize results, and validated on held out 2019 data. An L1 ratio (Lasso) of 0.9 was chosen to minimize error in 2019 as the validation set. Then all of the models were trained with 0.9 L1 ratio on the full dataset.

**Addition of natural variation.** This step aimed to match the statistical characteristics of an actual load year with the projected year. 2019 was used to derive the differences. Natural variation was estimated by a distribution characterized by the mean and standard deviation of the differences (in absolute value). Then, a natural variation adjustment was added to that day (with a random true/false bit for positive or negative variation). The noise was calculated for each region and peak demand and daily consumption separately. The natural variation (noise) vectors used are on the Github repository for this paper<sup>19</sup>. This part of the process is non-deterministic and replication of the results requires using the same natural variation vector used in our projections.

**Hourly profiles.** The statistical inference model presented above forecasts daily consumption driven by state-level economic parameters and weather data. The produced projections are at a daily resolution. We downscaled the data to hourly load profiles based on the 2015 hourly load profile data<sup>7</sup>. The result of the regression model is at regional level, breaking it down state-wise is pro-rated based on state-wise to region-wise GDP per capita projections ratios for the respective year. To do so, we tag each day of the year by the month it corresponds to and whether it is a weekday or weekend. We cluster demand for each hour by month and day. Each hour of the day then has its own cluster of demand data from 2015 based on the assumption that the same hour of the day for a given month and the same day type will exhibit similar demand behavior. This biases the construction of the profiles to demand patterns from 2015 only. To minimize the impact of this bias, we use the historical weather data<sup>15</sup> of the testing data years (2014–2019) for each day to simulate daily temperatures variations that are reflected in higher or lower demand. We sample weather data for each day and compare it to 2015, and subsequently use normalized the difference to scale the demand on a daily basis. Finally, we sample demand for each hour of the year from the corresponding cluster (defined by month and weekend or weekday) and scale it accordingly. Constructing the hourly load profile and fitting them to match the projected daily consumption and the projected daily peak demand then becomes a trivial exercise of sampling and fitting from the corresponding clusters and weather data space. The 2015 hourly demand data used in this study is documented in detail elsewhere and has been used in projecting demand for supply-side modeling efforts<sup>8</sup>. Limited availability of complete hourly data at state and regional level in India biases the hourly profiles to the 2015 datasets. However, the business-as-usual projections are for existing demands composed mainly of lighting and appliance at the residential level and large daytime loads at the commercial level<sup>20</sup>. Our approach implicitly assumes that energy consumption trends for these loads will follow historical patterns and therefore sampling from a given year with post-processed noise variation can yield reasonable results.

**Impact of Climate change on business-as-usual demand.** As per the International Energy Agency (IEA) World Energy Outlook (WEO) 2019<sup>21</sup> only 5% of households in India currently own air conditioning units and 2.6% of commercial building energy use is from space cooling. Historically, electricity consumption in India has been driven by lighting and appliances in the residential sector<sup>20</sup> with commercial and industrial sector contributing via larger daytime loads. Since cooling demand is not historically available in the data that the business-as-usual regression model is learning from, there is no parametric value to projecting increase in temperatures since there is no evident correlation between temperature increase and lighting or appliance use. Moreover, since space cooling is a small percentage of current electricity demand in India, no major trends can be identified given the limited daily training data that is being used for the business-as-usual regression. It is then safe to assume that weather remains constant for the business-as-usual demand.

**Technology model.** Since a regression model can only produce forecasts of data it can learn from, additional bottom-up processing must be carried out to get a full picture of India's demand in the future. We identify trends and data points at the state level of the country to build a regional profile as well as the national one.

**Cooling.** Cooling is divided into two main categories: residential and commercial. The ratio of commercial to residential consumption is computed from state-level data<sup>22</sup> and is used as the ratio of commercial to residential cooling demand. Using the IEA's baseline and efficient cooling projections from the Future of Cooling study<sup>3</sup>, we use the annual sales and unit types to calculate the energy consumption and growth rate at a national level and pro-rate it down to state level given GDP per capita. Surveyed hourly demand profiles<sup>20</sup> are indicators of behavioral cooling energy consumption patterns as exemplified in Supplementary Figs. 2 and 3. The survey produce various profiles given climate seasons, household income and size. We apply a time-domain convolution of these profiles to generate a representative profile for each state for the various climates and seasons.

We can generate the air conditioning demand profiles for two weather seasons (winter and summer) by convolution of the sample profiles to generate a smooth aggregated demand profile. Moreover, coincidence factors must be applied to properly estimate the simultaneity of the demand and its peak. Two coincidence factors are identified: weekday and weekend, values are extracted from a Reference Network Model Toolkit<sup>23</sup>. We break down the national cooling demand to residential and commercial at state level by identifying state-level sector size and growth trends. Scaling the profiles to match the projected cooling energy demand produces hourly energy consumption profiles from residential and commercial cooling. Aggregating the appropriate states together will produce the same results at the regional level.

More importantly, the IEA's future of cooling study<sup>3</sup> stresses the usage of Cooling Degree Days (CDD) to project cooling demand dependency on temperature. The unit consumption pattern and projections of capacity for

India's share of global cooling demand is based on growth in electrification, urbanization as well as Purchasing Power Parity. The IEA future of cooling study estimates that a 1-degree Celsius increase in decadal average temperature in 2050 will lead to 25% more CDD and a 2-degree Celsius increase will lead to 50% more CDD. Climate change impacts are considered in the unit sales and energy consumption data used from the IEA's future of cooling study. In our analysis, we use IEA's 50% increase in CDD to model cooling demand in 2050. For prior periods, we interpolate CDD between 2018 and 2050 to model cooling demand. The increase in CDD and the addition of noise variation are introduced for the purpose of modeling the projected increase in peak demand due to climate change. Specifically, this analysis does not consider frequency nor forecast of extreme weather events.

*Electric vehicles.* The second component of the technology model projects EV demand in India. The data presented here considered electric two, three, and four-wheel vehicles. Two-wheelers, being the dominating vehicle in terms of annual sales in India<sup>24</sup>, are expected to be electrified first, followed by the three-wheelers and regular cars<sup>25</sup>. The Indian government has set a goal of converting 100% of two-wheeler sales and 30% of all vehicle sales to electric by 2030<sup>26</sup>, so the starting point is vehicle sales at the state level<sup>24</sup>. Using the regression equations of the corresponding GDP growth scenarios, we can project car sales with the EV targets by 2030 met in the rapid growth scenario. From vehicle sales and conversion rates, we get an estimate of the number of EV that will require charging. From a market survey on the average commute distance of vehicles in urban areas and rural areas<sup>25</sup>, long and short-range battery capacity and EV energy can be estimated. We introduce a mix of EV sales starting with short-range as the dominant market product and shifting to long-range, a market-dominant market in 2050. This trend reflects the current economic competitiveness of short-range EVs vs. existing internal combustion engine vehicles as well as the long-term competitiveness of long-range EVs with declining battery costs.

Similar to the construction of the cooling profiles, a coincidence factor must be implemented, so as to not over-predict peak EV charging demand. Since this is a new consumption behavior and given the relatively small batteries of two-wheelers and three-wheelers, it is assumed that every vehicle needs to charge every other day on average for urban drivers and every day for rural ones. This yields an average daily consumption from EV charging. As shown in Supplementary Fig. 4, three different charging profiles — home, work, public — are identified in an EV pilot project study in Mexico City<sup>27</sup>. While Mexico and India differ greatly in many socio-economic aspects. The different hourly EV charging profiles collected were for a pilot project to deploy electric two-wheelers and small sedans in the metropolitan area of Mexico City. This presents two synergies enabling the usage of the charging profiles in India. Under the assumptions that EV deployment will be more prevalent in urban areas in India with initial conversion of smaller vehicles (two-wheelers and three-wheelers), the charging data collected<sup>27</sup> is a suitable fit for potential EV charging schemes in India. Energy consumption is computed from vehicle sales, projections, and electrification conversion. That calculated number is then fitted under the chosen charging profile. Time domain convolution of the profiles is applied to smoothen the peakiness of the total constructed hourly time series.

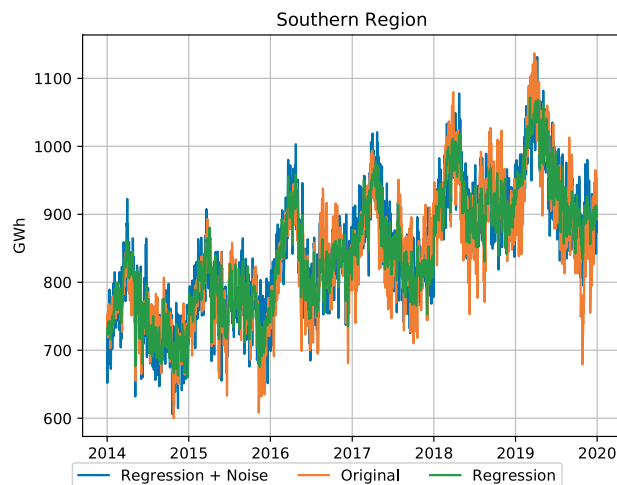
*Data dependence.* The technology model relies heavily on surveyed data to produce the representative hourly profiles for cooling and electric vehicle demands at state levels. This is indeed a limitation, and our projections assume that future technology adopters will behave just like initial adopters. In the absence of a better alternative at a similar spatial and temporal resolution, the bottom-up modeling effort provides a reasonable estimate of temporal patterns expected from these new demand sources. For the hourly sample cooling profiles, the main assumption is that cooling demand consumption is only dependent on weather patterns and econometric patterns. Specifically, we apply a weighted sum convolution of the income level cooling profiles based on the states' GDP per capita ranking. For the total cooling demand at national level, we depend on the air cooling unit sales projection as well as break down of unit energy consumption under baseline and efficient scenarios of the IEA's Future of Cooling report<sup>3</sup>. We pro-rate residential cooling at state level using the GDP per capita projections. For commercial cooling we use the state-wise sector growth trends<sup>28</sup>. A sanity check for this break down is to sum both residential and commercial state-wise cooling demand and compare to the IEA's all India cooling demand annual electricity consumption projections to 2050, the difference is highlighted in Supplementary Figs. 5 and 6. Regarding the EV profiles, while there are alternative choices of charging schemes, we identified the synergies with the Berkeley study<sup>27</sup> to be best reflective of the bookend EV charging scenarios across India.

## Data Records

The data is uploaded on Zenodo<sup>29</sup> and is available to download at <https://doi.org/10.5281/zenodo.4564581>. The path leading to a CSV file indicate the scenario corresponding to the results of that file. Breakdown of the folder hierarchy listed as:

1. GDP Growth: slow, stable, rapid
2. EV charging: home, work, public
3. Cooling: baseline, efficient
4. Type: detailed, summary

The *detailed* results are tables of the itemized hourly demand profile of each considered scenario; all files will produce 8760 rows (number of hours in a year). The *summary* are tables of the itemized annual energy consumption for the considered years; all files will produce seven rows (number of considered future years). Both file types are itemized the same way as per Table 3. The path of each file is the reference to the specific scenario the data in the tables represents. For example *SR.csv* file under *slow/home/efficient/summary* is the summary file of the case of slow economic growth, home EV and energy efficient air conditioning consumption.



**Fig. 7** Back test result for Southern Region regression model.

Column Header	Description
DateTime	Hourly or yearly time resolution
Base	Business-as-usual model resulting demand
Com AC	Commercial Air Conditioning demand
Res AC	Residential Air Conditioning demand
E2W	Electric Two-Wheelers demand
E3W	Electric Three-Wheelers demand
E4W	Electric Four-Wheelers demand

**Table 3.** Output data headers descriptor.

Region	Regression	Regression + Noise
Eastern	0.709	0.798
Northeastern	0.608	0.722
Northern	0.691	0.784
Southern	0.744	0.825
Western	0.680	0.778

**Table 4.** Business-as-usual Regression R-squared consumption results.

### Technical Validation

The Business-as-usual statistical model is validated using standard statistical metrics when backtesting is applied. Further details on the backtesting are available elsewhere<sup>16</sup>. For the technology model, we compare our estimates to the IEA's WEO<sup>1,21,30,31</sup> and Brookings India<sup>5</sup>. Furthermore, our projections compare favorably against the EV projections to the IEA's Global Electric Vehicle Outlook 2020<sup>2</sup>.

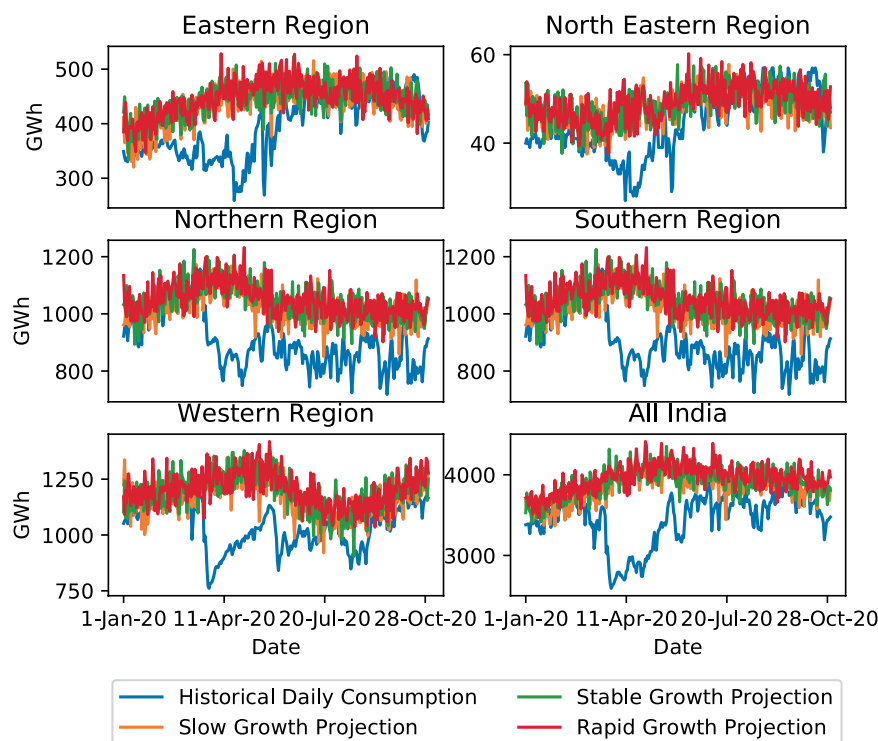
**Back testing.** Daily consumption and peak are projected for all five regions, we show the daily consumption back tests of the Southern Region in Fig. 7. More results can be found on the GitHub repository. It is important to note that the regression model captures the organic growth of the historical demand as well as the seasonal variation in demand but is not accurate at predicting daily variation. This shortcoming can be attributed to the small training dataset that is available. To compensate for this short-coming, we add additional noise variation as discussed earlier in the Methods section. We compare the R-squared value of the regression only versus the regression and noise time series as shown in Table 4. Additionally, selected parameter performance metrics of the model for the Southern Region are presented in Table 5. The model's independent variables are the 2 meters and 10 meters elevation historic temperature and humidity data for the selected cities and GDP data for the state. Various weather parameters will have a higher coefficient than GDP since the latter is not as granular as a metric but will still be factored in for longer term growth as interpreted by its Fourier component<sup>16</sup>.

**Cross-comparison.** Supplementary Figs. 7 and 8 compare the forecasting results to the WEO 2020 projections of India's Energy Demand to 2040. Our band of projections is notably wider due to the large number of scenarios that are combined to forecast energy demand. We further compare our results to Brookings India's study



Parameter	Coefficient	Standard Error	t value	P >  t
GDP	13.3630	34.663	0.386	0.700
Bangalore t2m max	931.0014	353.052	2.637	0.008
Chennai h2m min	1.736e + 04	1.65e + 04	1.050	0.294
Hyderabad t10m max	-502.7828	356.558	-1.410	0.159
Vijayawada h10m min	-2.229e + 04	7671.803	-2.905	0.004

**Table 5.** Business-as-usual Southern Region consumption Regression performance of select parameters.



**Fig. 8** 2020 year-to-date demand comparison with projections.

in Supplementary Fig. 9. We also compare our electric vehicle projections to those of the Global EV Outlook in Supplementary Fig. 10. Finally, we compare our air conditioning demand contribution to the peak demand to the Future of Cooling study in Supplementary Fig. 5.

**COVID-19 pandemic impact on year 2020.** The COVID-19 pandemic has drastically affected the global population in various ways. Energy consumption dropped severely as people were advised to stay at home. While it is not possible to project such “Black Swan” events from historical data, their long-term effects can be modeled as delayed growth under various recovery schemes. Figure 8 shows that our projections for the month of January 2020 align with the realized demand, which is prior to the global outbreak of COVID-19. Evidently, there is a strong mismatch in the following months as the outbreak developed into a global pandemic. However, in the later part of the year, signs of recovery are noticed where the historical daily consumption once again reaches projected levels.

The impact of extreme events on energy consumption are difficult to predict at a granular level. Our projections are at a five year increment so that such yearly variations are smoothed out and the regression towards the mean phenomenon is observed. Moreover, the recovery from extreme events and their long-term impact can depend on many factors: economic, social, scientific and more. Without modeling those events in detail, projected growth can model the long-term average growth rate. In case of a negative extreme event, a smaller growth rate can model the long-term impact caused by the slow down. Similarly, a positive extreme event can be modeled as larger growth rate to include the long-term impact by the rapid growth. With signals of a fast recovery in total daily consumption for most regions, we elected to disregard projections that model long-term COVID-19 pandemic impact to avoid confirmation bias. Moreover, there is little data to support projections modeling a long-term impact on Indian energy consumption. We believe that the model and data presented in this paper are valid beyond the COVID-19 pandemic.

## Usage Notes

The format of the results is comma-separated values (CSV). All the results are available on the Zenodo Open-Access repository<sup>29</sup>.

## Code availability

The code used in the generation of the data sets is open-sourced on Github repository<sup>19</sup>.

Received: 16 December 2020; Accepted: 20 May 2021;

Published: 15 July 2021

## References

1. International Energy Agency. *World Energy Outlook 2020* (IEA, 2020).
2. International Energy Agency. *Global EV Outlook 2020* (IEA, 2020).
3. International Energy Agency. *The Future of Cooling* (IEA, 2018).
4. Spencer, T. & Awasthy, A. *Analysing and projecting Indian electricity demand to 2030* (The Energy and Resources Institute, 2019).
5. Ali, S. *The future of Indian electricity demand: how much, by whom and under what conditions?* (The Brookings Institute, 2018).
6. Ministry of Statistics and Programme Implementation. *National accounts statistics 2019*. <http://mospi.nic.in/> (2019).
7. Shakti Sustainable Energy Foundation, GE India Exports Pvt. Ltd. *Modelling with power sector planning for India* (2018).
8. Garc a, I. R. *Decarbonizing the Indian power sector by 2037: evaluating different pathways that meet long-term emissions targets*. Master's thesis, Massachusetts Institute of Technology (2019).
9. Sankhe, S. *et al. India's turning point: an economic agenda to spur growth and jobs* (2020).
10. PricewaterhouseCoopers. *The long view how will the global economic order change by 2050?* (2017).
11. Winsor, C. P. The Gompertz curve as a growth curve. *Proceedings of the National Academy of Sciences of the United States of America* **18**, 1–8 (1932).
12. Datta, P. Population projection of EAG states of India: vision until 2051. SSRN <https://doi.org/10.2139/ssrn.1861769> (2011).
13. United States Energy Information Administration. *India's economic growth is driving its energy consumption* (2013).
14. Ministry of Power. *Power system operation corporation: load dispatch center monthly reports*. <https://posoco.in/reports/> (2020).
15. Gelaro, R. *et al.* The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate* **30**, 5419–5454 (2017).
16. Alsup, M. *Forecasting electricity demand in the data-poor Indian context*. Master's thesis, Massachusetts Institute of Technology (2020).
17. United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects 2019: Highlights*. Report No. ST/ESA/SER.A/423 (United Nations, 2019).
18. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).
19. Barbar, M. Source code for: Demand forecasting india. *Github*. <https://github.com/barbarmarc/demand-forecasting-india/> (2020).
20. Garg, A., Maheshwari, J. & Upadhyay, J. *Load Research for Residential and Commercial Establishments in Gujarat*, vol. ECO-III-1024 of *Energy Conservation and Commercialization* (Indian Institute of Management Ahmedabad, 2010).
21. International Energy Agency. *World Energy Outlook 2019* (IEA, 2019).
22. Datanet India Development Team. *Electrical energy consumption by industry*. <https://www.indiastat.com/> (2020).
23. Mateo Domingo, C., Gomez San Roman, T., Sanchez-Mirallas, A., Peco Gonzalez, J. P. & Candela Martinez, A. A reference network model for large-scale distribution planning with automatic street map generation. *IEEE Transactions on Power Systems* **26**, 190–197 (2011).
24. Society of India Automobile Manufacturers. *Automobile domestic sales trends 2015–2020*. <https://www.siam.in/> (2020).
25. Batra, R. *et al.* *Standing up India's EV ecosystem - who will drive the charge?* (Ernest and Young, 2018).
26. Government of India National Institution for Transforming India. *Zero emission vehicles: towards a policy framework*. [https://niti.gov.in/writereaddata/files/document\\_publication/EV\\_report.pdf](https://niti.gov.in/writereaddata/files/document_publication/EV_report.pdf) (2018).
27. von Meier, A. *Ev infrastructure planning and grid impact assessment: a case for Mexico*. Technical Report No. UCB/EECS-2018-70 <http://www2.eecs.berkeley.edu/Pubs/TechRpts/2018/EECS-2018-70.html> (University of California, Berkeley, 2018).
28. Ministry of Statistics and Programme Implementation. *Energy statistics 2019*. <http://mospi.nic.in/> (2019).
29. Barbar, M., Mallapragada, D., Alsup, M. & Stoner, R. Electricity demand forecasting of india. *Zenodo* <https://doi.org/10.5281/zenodo.4564581> (2020).
30. International Energy Agency. *World Energy Outlook 2018* (IEA, 2018).
31. International Energy Agency. *World Energy Outlook 2017* (IEA, 2017).

## Acknowledgements

This research was supported by MIT Energy Initiative's Low Carbon Energy Centers and the Future of Storage Study.

## Author contributions

R.S. collected the data, M.B. built the Technology Model, M.A. built the Business-as-usual Model, and D.M. validated the results. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41597-021-00951-6>.

**Correspondence** and requests for materials should be addressed to M.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

The Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files associated with this article.

© The Author(s) 2021, corrected publication 2021