Hudson Valley Transmission Line Plan: Assessing Need & Alternatives

Gidon Eshel
www.environmentalcalculations.com
and
Research Professor of Environmental Physics, Bard College
geshel@gmail.com

Disclaimer: *The views and analyses presented in this document are solely those of the author. The author's long term and/or ad hoc affiliations and appointments do not imply endorsement of this document or the scientific findings it conveys by the Hudson Valley Smart Energy Coalition or Bard College. The author is the sole contributor to this document.*

1 Executive Summary

Focus: **The New York Public Service Commission's plan to add 1 GW of additional north-to-south transmission capacity into the downstate region by constructing new high voltage AC power lines roughly paralleling the Hudson River throughout the Hudson Valley's north–south extent.**

Key Questions:

1. **Given reasonably expected future peak downstate power loads, is there a need for additional north-to-south transmission capacity?**
2. **If the answer is No, does it change to Yes in the case of retiring Indian Point?**

Overall Approach:

The author—a geophysicist and applied mathematician by training, a Senior Scientist at Northwest Research Associates, and a Bard College environmental physics

research professor—has carried out this research on behalf of the *Hudson Valley Smart Energy Coalition* on a *pro bono* basis. The author then employed a mathematical-statistical approach characteristic of published geophysics and environmental physics scientific publications, such as the author's.

Main Findings: **Consistent with the proposed project rationale, downstate peak loads are expected to rise by about 10% of today's peak loads through about 2035, and stabilize subsequently.**

However, assuming no additional energy efficiency inroads beyond the 0.9% per year observed in recent years, the closure of the Indian Point nuclear power plant by late 2015, and the completion of only half of the projects currently in the NYISO "in queue" category, assets currently in place plus the gradually added half of the in queue assets easily exceed expected future peak loads at any point in time from now through 2040.

Further, assuming no efficiency gains (i.e., holding steady at the 0.9% energy efficiency increases per year observed in recent years) is unrealistically pessimistic relative to much faster gains expected by other northeastern states or recorded in response to recent specific policy changes.

No discernible evidence thus exists that additional generation or transmission capacity is needed in New York's downstate region.

The Emergent Message: **To the best of our knowledge, the analysis presented in the full report is the first to address quantitatively and rigorously the need for the proposed project. Like all scientific analyses, especially those not yet published in a reputed scientific journal, the presented analysis is neither definitive nor final. Yet it represents a quantum leap in public discourse about the plan—whose necessity has been merely *assumed* rather than demonstrated up until now—because it is rigorously scientific, impartial, and entirely transparent. All data sets used, all analytic methods, and all assumptions are unambiguously presented, and can thus be readily challenged. Logically, a proposed plan—especially one with significant societal costs—must follow, not precede, an existing, technically vetted and fully**

transparent scientific demonstration of need. If such an analysis has been carried out by the PSC (The New York Public Service Commission), it has never surfaced publicly. Absent such an analysis, the current plan fails to meet this criterion, and is thus premature until made to meet basic due process.

2 Preface and Background

2.1 How to Read this Document

This document summarizes some elements of a fact based approach to the issue of proposed Hudson Valley power transmission upgrades (hereafter interchangeably "*the problem*" or "*the plan*"). Saddled by no pre-conceived notions, this document is motivated by the desire to help New York State become a national leader in 21st Century energy issues by promoting an innovative approach to the analysis of energy assets and needs.

Because the problem is very complex (see below), and various aspects of it are active foci of work by many able researchers from numerous disciplines, it will be preposterous to strive for an exhaustive solution. Instead, this document only focuses on the issue of future need, and demonstrates the enormity of the challenge and its sensitivity to various assumptions and imperfectly known facts.

As a scientist, not an engineer, the approach I have taken, and thus the tenor of this document, is scientific rather than engineering based. This is not a coincidence, stemming instead from viewing the issue at hand not as an engineering problem. Clearly electrical engineering is reasonably deferred to for solving specifically posed transmission problems. Yet while transmission issues figure in the problem of downstate power demand, the latter is far broader than the former.

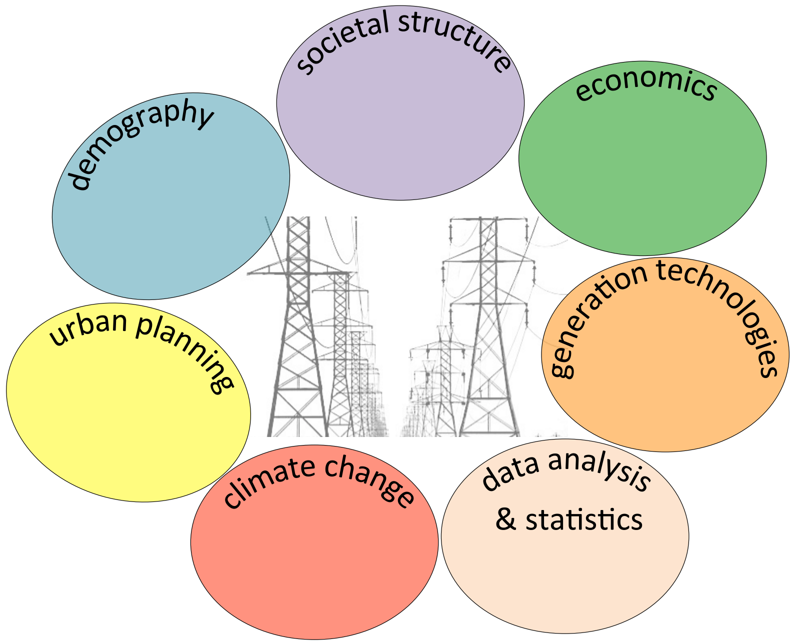

Figure 1: A schematic representation of some of the key disciplines and schools of thought that are intimately relevant to the Hudson Valley transmission problem.

2.2 The Complexity of the Multidisciplinary Problem

The downstate power delivery problem comprises a suite of mutually coupled problems falling into the domains of several disciplines, as Fig. 1 schematically depicts. Even assuming complete knowledge of the electricity market, to forecast downstate 2030 electricity needs, one must accurately forecast the population size of the City and neighboring counties, as well as such demographic attributes as median number of inhabitants per household or—because income has been associated with per capita energy demands[1]—the characteristic income of a typical household. Yet to predict those, one must rely on economic forecasts, themselves notoriously imperfect, to say nothing of assumptions about immigration, spreading rates of western family planning, among many other sociological phenomena.

Psychology is also crucially and nonlinearly important. Technophobia[2], the suspicion of new technologies, often initially limits deployment of new alternative energy sources[3]. Yet this effect is clearly cost dependent, with explosive transition to renewable energy sources once their costs become competitive and psychological barriers have run their course. For example, along with solar energy installation costs dropping by roughly 50% over 2010–2013[4] and by 15% in 2013's

Q4 alone[5], deployment rates have been rising steadily: installation rates increased by 41% from 2012 to 2013, and solar accounted for 29% of the total electricity generation capacity added in the U.S in 2013[6].

Anthropogenic climate change adds an additional layer of complexity to all of the above interactions. First, one must decide which global path of human action or inaction is likely to be followed. In addition, the path taken and all of the above are mutually intertwined through a complex web of feedbacks. For example, city economics and climate change are strongly coupled[7], and demographics are impacted by the climate refugee problem[8], which is expected to involve hundreds of millions if not billions[9] in coming decades.

This rich web of interactions renders the downstate power supply problem seemingly intractable. Indeed, comprehensively addressing the full problem will require decades of work by many individuals from several disciplines. This report most certainly does not come close to delivering this product. No party involved with this proposal—neither the current author, nor the PSC, or FERC (Federal Energy Regulatory Commission), or the Governor's office—has the intention or all the needed expertise to meet this challenge.

Instead, in this report I highlight quantitatively some key issues pertaining to predicting future peak loads that likely exert a disproportionate control over the overall estimate of future downstate power needs. I also emphasize uncertainty ranges, and explore in quantitative details the variability in the overall estimate of future downstate power demand to which these ranges give rise.

Finally, it is my hope that this report will form a basis for an open, rational debate over desirable options and their combination likely to yield the optimal overall solution, the solution that maximizes overall societal betterment under finite resources.

2.3 A Note on the Use of Mathematics

In writing this document, I deliberately decided to fully and entirely transparently describe the analyses carried out and the results obtained. Because the machinery used is inherently algebraic, the above goal has proven directly at odds with the desire to produce a widely readable document, a document accessible by all readers. I chose to sacrifice broad readability for technical completeness and transparency because I believe the power of this document is in its thorough, methodologically defensible treatment of the problem of estimating future downstate peak electricity loads. It is therefore my hope that in the future I will be able to synthesize the salient, essential elements of the analysis and translate them into a broadly readable document. For now, however, the technical details, the essence of the document, remain here and are derived in details in the following sections.

3 An Outline of the Demand Estimate Logic

There are many independent, complementary approaches to estimating future power demands of a given area. Most algebra based ones can be represented generically and symbolically as

$$e(t) = f\left[P_T(t), A(t), C(t)\right] + g\left[f_1(t), f_2(t) \cdots f_I(t)\right], \tag{1}$$

where $e(t)$ denotes downstate peak electricity demand in MW at time t.

While many choices of an addressed e are reasonable, here I define e as the combined peak summer (and thus annual) demand in NYISO[1] Zones[2] G–K. In Section 5 I elaborate on and unambiguously define the exact meaning of $e(t)$, its derivation and temporal behavior. Eq. 1 splits $e(t)$ into two parts. in the first right hand term, f is a function to be specified later of downstate total population P_T, some measure of the region's mean affluence A, and of the region's characteristic climate C. While f is not explicitly time dependent, it implicitly is, through the time dependence of all its three input arguments. In Eq. 1's second term, g is a function of the fraction of the total population P_T accounted for by I population sub-groups (indexed by i), chosen based on their distinct characteristic power demand patterns.

In the following sections, I choose explicit f and g, and systematically derive the right hand side terms. In sections 4 and 5 I derive the population estimates P_T and f_i. In section 6 I then identify a suitable I, and devise defensible estimates of A and C. Using those, in the latter parts of section 6 I finally derive the full model for $e(t)$, train it on past observations, and then use the coefficients thus derived to calculate future downstate electricity peak summer demand $e(t)$ for 2015–2040 in five year intervals.

4 Population Estimates

Fig. 2 shows expected population trends[1] in the downstate area. The Fig. shows two distinct regimes. The populations of Queens, the Bronx and Staten Island are expected to continue to rise into the 2040s and beyond. In sharp contrast, the populations of all other addressed counties are expected to either decline (western Long Island) or increase very slightly and insignificantly until roughly 2025-2030, and slowly decline subsequently. These expected population trends suggest that to fully address the challenge of meeting future downstate peak power demands, two distinct approaches must be devised, with the above New York City outer boroughs requiring one approach, and all other counties requiring a different one.

Because of continued population growth in Queens, the Bronx and Staten Isl., and because population totals tend to lag behind transient demographic trends, the overall population of the combined downstate area is expected to continue to rise (Fig. 2n), at a decreasing rate, through 2025 or so, and stabilize thereafter at 12.8 ± 0.1 million people. Importantly, the number of downstate women of reproductive ages (Fig. 2m) is expected to drop markedly, leading to decreasing net births.

While the above statements address medium range projections, recent near term trends and their spatial structure (Fig. 3) are similarly revealing. They show the clear distinction between dense urban counties, in which population is currently still modestly increasing, and exurban and rural counties, whose populations are already declining.

Because electricity consumption is age dependent[2], in addition to population size, the age distribution within the population crucially impacts future electricity demands. Fig. 2 affords a preliminary insight into this, delivering a mixed message. While Fig. 2m shows that reproductive age women population in the full downstate region has already declined, and is expected to continue to decline at a rapid clip, Fig. 2f, h show that the corresponding female populations in the Bronx and Staten Island are expected to continue to rise at least through 2040. Of particular interest among those is the Bronx (Fig. 2f), which currently accounts for 11% of the downstate combined population.

Fig. 4, presenting 2030 vs. 2010 population changes by age, affords a closer look at the issue.

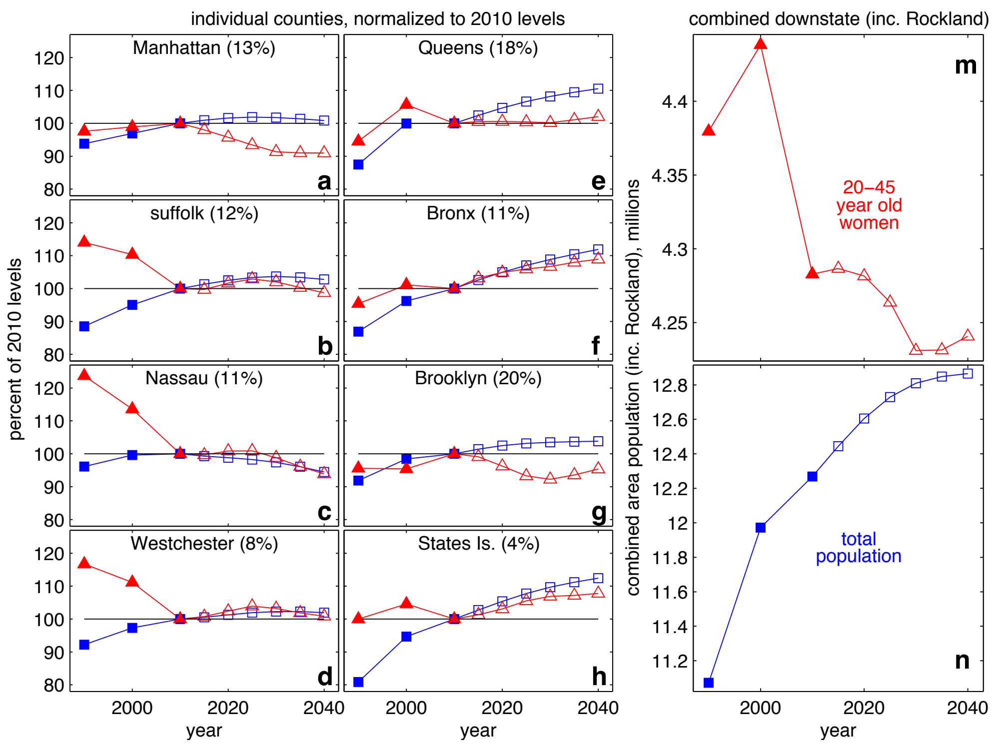

Figure 2: Observed (earliest three data points, solid symbols) and projected populations through 2040 of eight relevant downstate counties including New York's five boroughs (panels a, e–h), both Long Island counties (panels b, c), and Westchester (panel d). In each panel, both total population (blue squares; both genders, all age groups) and women of child bearing age (red triangles) are shown. In panels a–h, the data are represented as percentages of the respective 2010 values (so that stasis is represented by the solid horizontal black lines). The percentages of combined 2010 downstate total population (inc. Rockland Ct., which is not individually shown) the shown counties account for individually are given parenthetically in the panels' top center near the county's name. The combined down state area time series, the sums of the time series of the shown eight counties plus Rockland Ct., are given in panels m, n, where observations are again distinguished from projections by solid symbols. Data from the Cornell Program on Applied Demographics, http://pad.human.cornell.edu/index.cfm.

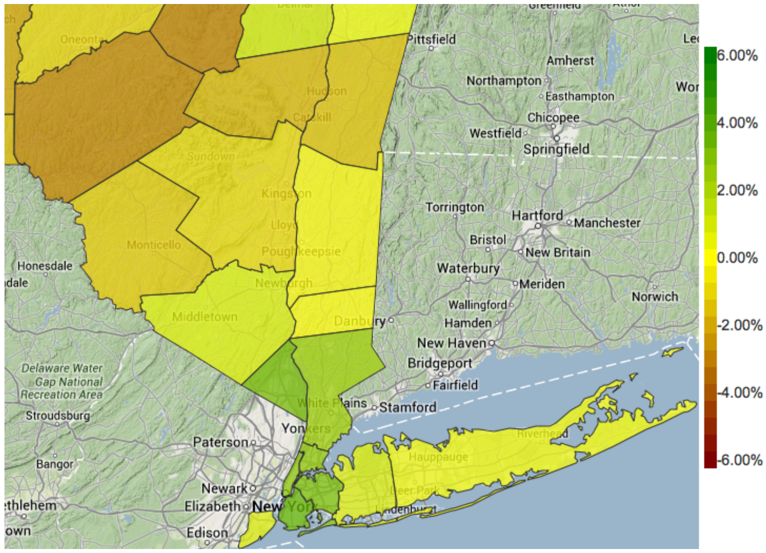

Figure 3: County level observed population changes between Apr. 2010 and 2013, from the Cornell Program on Applied Demographics.

In general and throughout the considered sub-regions, the changes are mostly roughly gender symmetrical throughout most ages save the >70 group, in which female gains outpace males' (in terms of absolute changes, Fig. 4h, i), increasing somewhat the female dominance of this age group. The population changes are summarized in Fig. 4d–o for three key age groups (following and modifying an earlier study[3] examining models of electricity consumption as a function of age and demographic parameters), delivering the following messages about expected downstate population changes:

1. Most notably, the >70 age group is expected to rise significantly for both genders.

2. Younger adult populations are expected to rise modestly in some urban counties while declining elsewhere, giving rise to slight female-dominated downstate declines (lowermost bars in Fig. 4h-i, n-o).

3. The middle age group is expected to increase modestly (decline somewhat) in urban (ex-urban) counties, yielding slight full downstate increases.

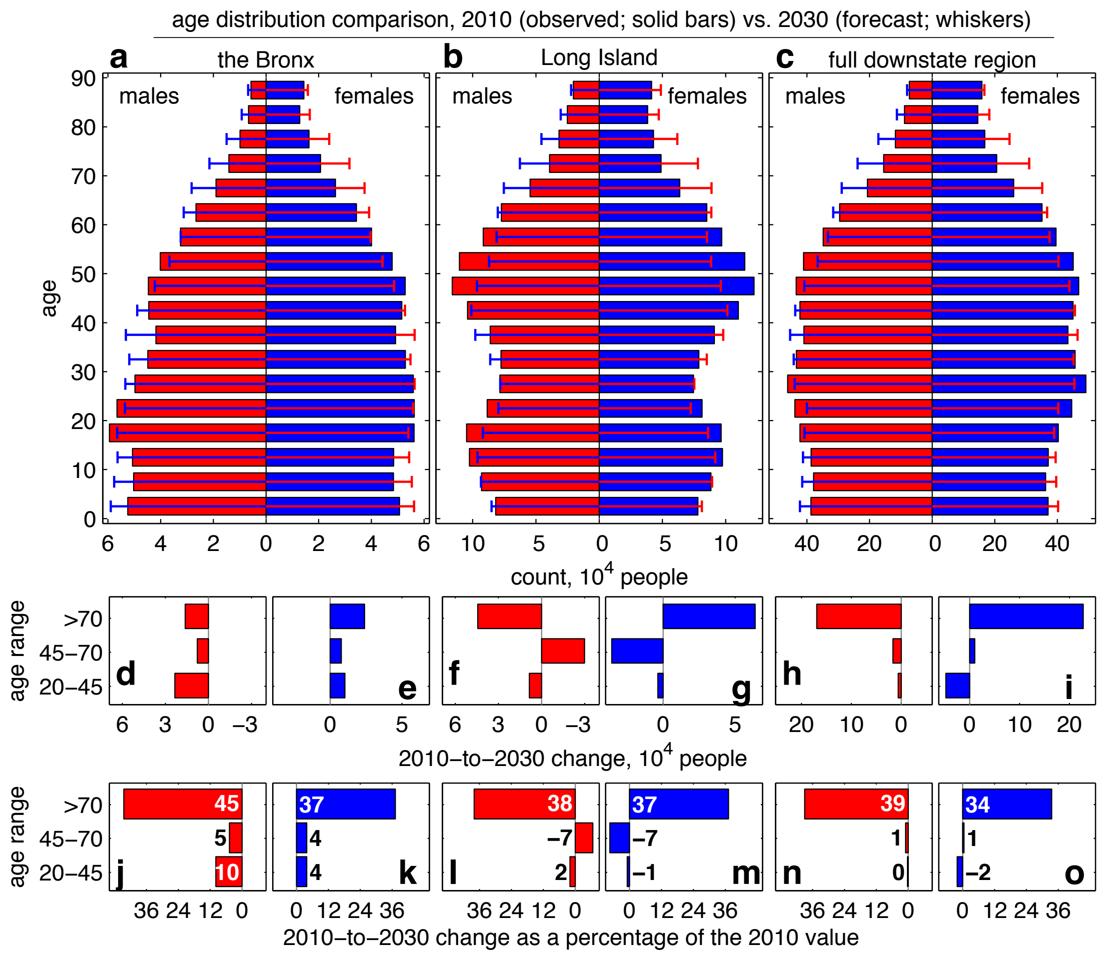

Figure 4: Changes in age distribution of male and female populations in the Bronx (panel a), Long Island (Nassau plus Suffolk counties, panel b), and the full downstate region (the sum of all 9 counties; panel c). Observed 2010 census age distributions are given in red (blue) solid bars for males (females). The predicted 2030 distributions are given by the overlaid blue (red) whiskers for males (females). Panels d–i present the changes in male (red) and female (blue) populations within the indicated age groups between 2010 and 2030, in 10^4 people. Note that while panels d–g share a horizontal range, the combined area panels (h, i) have a wider horizontal scale. The values in panels d–i are shown again as percentages of the 2010 observed values in panels j–o. Data from the Cornell Program on Applied Demographics.

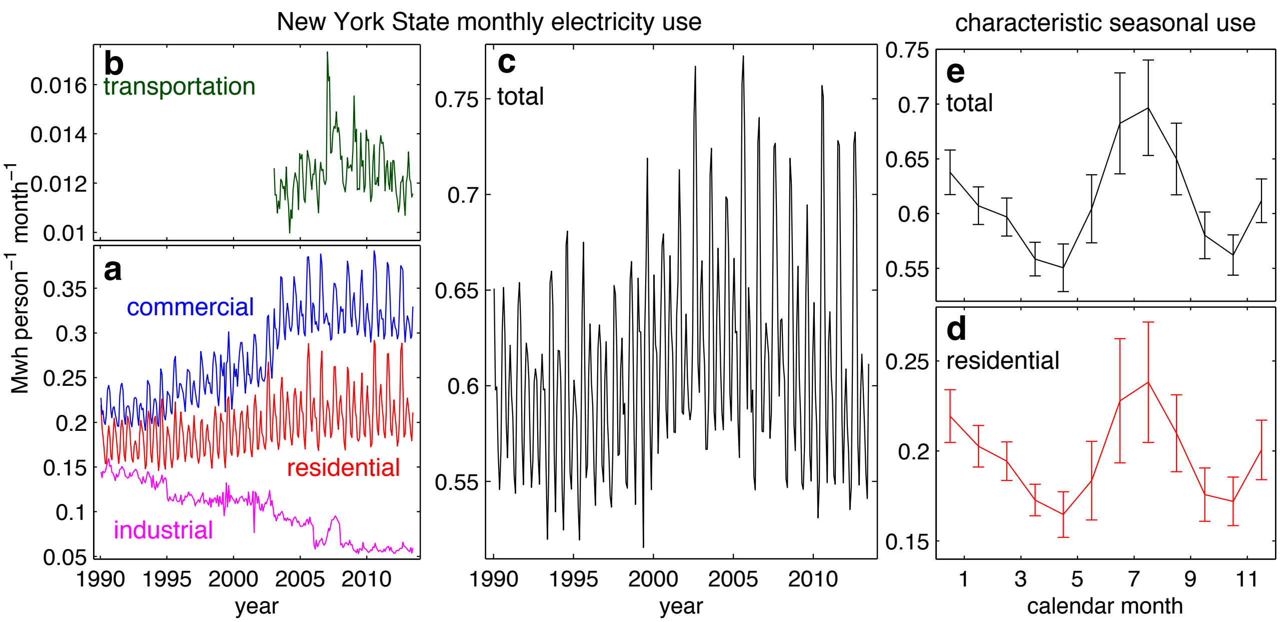

Figure 5: Recent monthly per capita New York State electricity use by sector. Panel a presents use by the industrial (magenta), residential (red), and commercial sectors, and panel b shows the very modest use by the budding electricity powered transportation sector. The combined all-sector statewide per capita electricity consumption is shown in panel c. Panels e, d present the seasonal cycle of monthly mean electricity use, with variability whiskers (showing the ± standard deviation range) calculated over the full annual cycles of all available years. They demonstrate that while summer is reproducibly the load "bottleneck", it is highly interannually variable, and that winter mean use is not far behind, statistically speaking.

5 Recent Trends in New York State Energy Use

As stated earlier, the key electricity use variable of interest is peak load (e); generation capacity in New York State is *on average* no more than roughly 60% tapped. Yet by way of requisite background, and because peak demand is not unrelated to overall load, mean (i.e., non-peak) use is also important.

Recent trends in mean New York State electricity use are given in Fig. 5a–c for both total use and use by key individual economic sectors. (Note that the information summarized in Fig. 5 is *not* ultimately used for e prediction, but rather is given as added background information.) Because of the characteristic northeastern hot, humid summer climate, and because much heating is oil based and air conditioning is ubiquitous in New York State, the State's power demand peaks are most often realized within the July–August time span, as hinted (but not explicitly shown) by Fig. 5d–e (see Fig. caption).

Returning to the main focus of annual peak power load e necessitates two methodological asides. First, the modeling and statistical work in the remainder of this report is based on treating peak loads, and by extension their sums, as random variables. While this is meant in the statistical sense, it is intuitive. Clearly a given Zone's peak load depends on weather variability, a quintessential random variable. But even for a given weather pattern, the actual power needs depend on the collective behavior of the Zone's residents, and thus on countless random choices. Consequently, forecasting future downstate peak loads is an inherently statistical—rather than deterministic—modeling problem.

The second issue is practical, having to do with unavailability of the ideal peak load data needed for the statistical modeling problem at hand. Two data sets exist, and both are highly pertinent, but neither is ideal. The first data set is from the NYISO[1]'s so-called Gold Books[2], comprising historical coincident and non-coincident peak loads for each individual load Zone during 1995–2013, and (in, e.g., Tables I-4a, b on pp. 21–22 of the 2011 Gold Book). The challenge with this source is that the reported peaks are not summable, as in general they are realized at different times. What is needed instead is the *actual* maximum combined downstate load. Understandably, such sums specifically for Zones G–K are not available in the Gold Books, nor kept by NYISO[3]

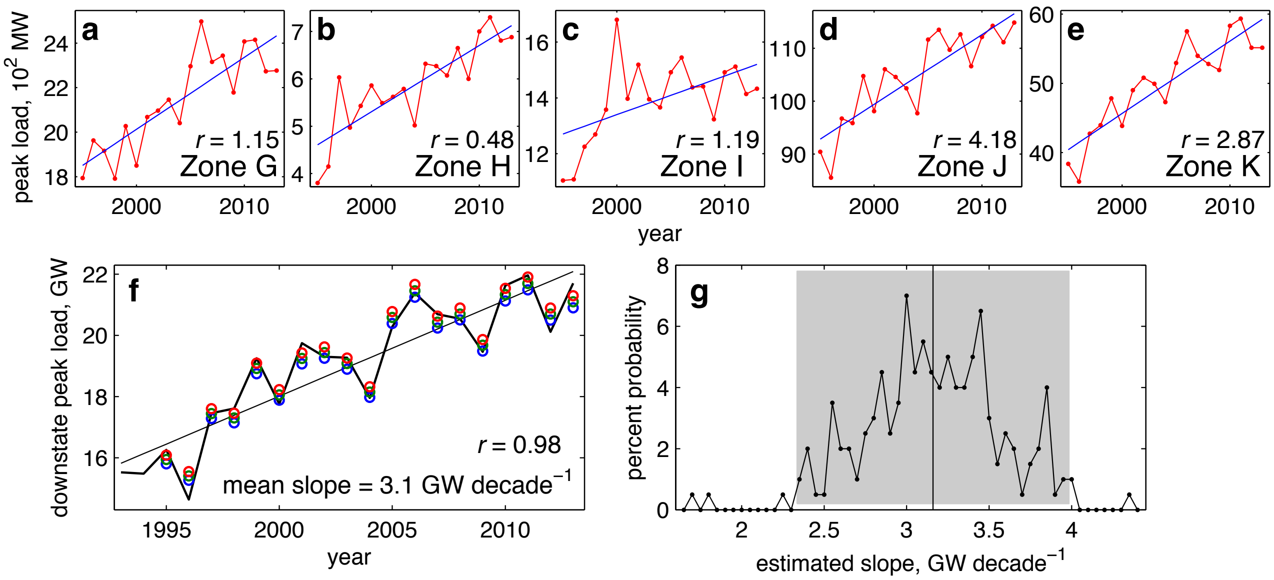

Figure 6: Recent summer (and thus annual) peak loads in the individual counties the downstate region comprises (panels a–e) and the combined downstate region as a whole (panel f). In a–e, the Zones are indicated, the individual Zone's peak load 1995–2013 time series is shown in red, and $r := \sqrt{(\hat{\mathbf{e}} - \mathbf{e})^T (\hat{\mathbf{e}} - \mathbf{e})/N}$ is the root mean squared misfit about the shown linear trends (blue), where $\mathbf{e}$ and $\hat{\mathbf{e}}$ are the observed and trend-predicted peak load 19-vectors, $\mathbf{q}^T$ denotes the transpose of an arbitrary real vector $\mathbf{q}$, and $N = 19$ is the number of annual peak data points included in panels a–e. Data from NYISO's 2014 Gold Book. In panel f, the above data are blended to yield a range of estimates using the methodology introduced in the text, shown by open circles. These are best thought of as imperfect estimates of the actual sought combined peak load in the addressed Zones. Also in f, the modeled blended estimates are augmented with additional data (solid black curve) addressing the sum of actual loads summed over Zones G–K during times of annual maximum *statewide* loads. Because the time of each annual peak load in the latter data set is the time of maximum load in the State as a whole, it too is imperfectly suited for the modeling task. This is further discussed in the text. The slope of the data shown in f (whose best estimate is 3.1 GW per decade, as indicated near the bottom of f) is addressed further in panel g, which shows the probability density function (or, more accurately, its finite approximation, the histogram) derived from 200 Monte Carlo realizations, in each of which the calculation is repeated anew with only 19 of the 21 available data used. The message of panel g is that peak loads have been rising, significantly and unequivocally, since the early '90s, justifying the concern that peak load shortfalls may soon follow.

The second data source—also by NYISO but not a part of the traditional Gold Books—was generously provided by NYISO upon personal request (Mr. Arthur Maniaci, by email communication over Aug. 15–21 2014). It comprises mutually additive Zone loads during times of statewide annual peak loads, with dates of occurrence of the peaks. Here the challenge is that while the area of interest here, downstate, is the largest power user in the state, it is still only a subset of the state. The times of statewide and downstate peak loads therefore need not temporally coincide.

To overcome the above data imperfections, I blend the two data sources and unify them despite the disparate time extents, as the follows. The starting point is the observation (implied by but not explicitly shown in Fig. 6) that the magnitudes of all Zones' peaks in a given summer are very similar across Zones if expressed as a fraction of the Zones' long term mean summer peaks. This is made all the clearer upon linear detrending (the removal the dominant linear trend), which was done but not presented pictorially here for brevity. Put differently, the downstate region is internally rather coherent. While very important, this is not surprising given the very modest spatial extent of the region—of order 10^2 km—relative to the spatial scale that governs weather variability, the so-called synoptic scale which is characteristically of order 10^3 km. Because the extent of the downstate region is roughly a tenth of the synoptic scale, when it is unusually hot and muggy in Brooklyn, say, it is highly improbable that Westchester is enjoying a crisp, cool day. Consistently, when load peaks in one of the downstate Zones, while other Zones' loads need not also peak at the same time, they surely are well above average.

Exploiting the similarity of individual normalized peak load fluctuations, I devise three estimates of combined downstate peak load in year i, e_i, as

$$\begin{pmatrix} e_i^{\text{low}} \\ e_i^{\text{med.}} \\ e_i^{\text{high}} \end{pmatrix} \approx e_{i,\text{J}} + \sum_{j=\text{G,H,I,K}} e_{i,j} \begin{pmatrix} 0.95 \\ 0.97 \\ 0.99 \end{pmatrix} \tag{2}$$

where here e is discrete in time, hence the replacement of $e(t)$ by e_i, in which $i = [1, 19]$ is the year index. In Eq. 2, I take note of the fact the New York City, Zone J, is the largest downstate power consumer. All three estimates thus assume peak load in New York City, rendering them conservative upper bounds. In Eq. 2 I also assume—also conservatively—that at the time of peak load in New York City, loads in the other four load Zones, are 95%, 97% or 99% of their own respective peak loads. These three estimates are shown as colored circles in Fig. 6f.

The three estimates based on blending individual Zone peak loads are then plotted in Fig. 6f along with an additional independent estimate of downstate combined peak load. That latter estimate (black solid curve closely tracking the circles in Fig. 6f) is based in simply summing the five Zones' observed loads during the time of statewide peak load.

As made clear by Fig. 6f, while based on distinct assumptions, and utilizing two distinct data sets, during their 19 year overlap, the two estimates clearly mutually agree very closely. Either one of them is thus a suitable choice of the peak load data to be modeled. In what follows I use the following as the predictand time series $e(t)$. For 1995–2013, I use the upper bound downstate peak load estimate, the sum of the asynchronous Zone peaks. For 1993–94, I use only the data obtained by the second of the above two methods. Because there appears no discernible discontinuity between 1994 and 1995 (as Fig. 6f's lower left corner makes clear), I simply suture together those two estimates as the full 1993–2013 training e timeseries.

It will prove useful to express the series as the vector

$$\mathbf{e} := \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_{21} \end{pmatrix} \in \mathbb{R}^{21} \tag{3}$$

in which the 1993–2013 temporal coverage accounts for the dimension.

6 A Simple Model of Peak Downstate Electricity Loads

As emphasized earlier, the problem addressed here—estimating a given area's electricity demands—can be handled many different ways. Ideally, all variables that have been historically cross-validated and proven predictive should be employed together to obtain the best possible estimate. A thorough implementation of this approach is very laborious and exceeds the scope of this—one man, uncompensated—effort. Instead, here I choose an approach that builds on earlier work on electricity demand and on the already introduced Eq. 1.

The recent work of Liddle[1] has built on and expanded a body of literature in which electricity demands on various spatial and temporal scales are modeled statistically. The methodologies Liddle and his predecessors have favored amount to stochastic extension[2] of the original IPAT[3] approach to resource consumption.

Below, while conceptually following the above approaches, I differ from them in one important way: While in the stochastic IPAT formalism explanatory variables (predictors) are multiplicatively related, thus giving rise to an overall log-linear optimization problem, here I favor the additive, linear, approach, for two reasons.

First, it seems unlikely that all predictors are mutually multiplicatively linked. For example, it is hard to see how affluence—a key element of the original IPAT formalism and its stochastic extensions—and climate will be multiplicatively related for an application that strives to address peak load rather than mean use. This is so because given how affordable air conditioning has become, it is hard to imagine households of even modest means not availing their denizens of the heat respite afforded by air conditioning during an extreme heat wave. In sharp contrast, it is rather intuitive that in the mean, averaging over many summer days far less extreme than the singularly worst summer day shaping the current problem, increased affluence will indeed imply more regular reliance on air conditioning.

The second reason I favor here the linear approach over the log-linear alternative is methodological robustness. Because the dynamic ranges of variability of both predictors and predictand in the current problem are rather small, giving rise to an almost ill posed inversion, taking the log only diminishes the range and further degrades the problem's posedness.

For both of these reasons I favor the linear approach. Neither unique nor necessarily superior to other methods, the linear representation yields an initial model for peak downstate electricity load at time t of the general form

$$e(t) = x_1 + x_2 P_T(t) + x_3 A(t) + x_4 C(t) + \sum_{i=1}^{I} (x_{4+i} f_i) + \xi(t), \tag{4}$$

to be further tested and perfected below. In Eq. 4, the $4+I$ x_is are the sought fitting (regression) coefficients (of which x_1, a fixed additive term, allows for e's nonzero mean), and in keeping with the original IPAT formalism and Eq. 1 above, Eq. 4's $P_T(t)$, $A(t)$ and $C(t)$ denote time dependent downstate total population, mean affluence and climate.

As a clearly imperfect measure of affluence, I choose per capita GDP for the states of NY, NJ, and CT, and that of York-Newark-Jersey City, NY-NJ-PA (Metropolitan Statistical Area, FIPS 35620), all taken From the U.S. Dept. of Commerce, Bureau of Economic Analysis[4] and averaged. There may be better choices of affluence characterizing variables, but I was able to find no sufficiently specific, spatiotemporally resolved such data sets. It is thus simply noted that the choice of GDP may represent a weakness in the model that will require follow-up work to overcome. While the data in each of the used sets are initially chained to different years' Dollars, I correct all to 2009 Dollars using U.S. Dept. of Labor, Bureau of Labor Statistics on-line inflation calculator[5].

As a simple measure of climate, I use annual maximum of daily maximum temperature $T_{\max}$. In the process of devising the model I also explored the humidity-mindful alternative of cooling degree days. Perhaps because those data were available over a shorter time span, their performance in the model was weaker than that of $T_{\max}$. While follow-up investigation may well improve on the current $T_{\max}$ choice, for the purpose of this analysis I represent climate variability C as annual maximum $T_{\max}$, as stated above.

Also as in Eq. 1, f_i is the fraction of the total downstate population P_T for which age group i accounts. Modifying Liddle's choices, here I consider the $I = 3$ age groups (consistent with Fig. 4) 20–45, 45–70, and >70 years old. Finally, allowing for misfit (in the regression sense), $\xi(t)$ denotes noise. Thus in Eq. 4 I replace Eq. 1's f and g by their simplest, least conjectural alternative, linear relations.

In deviation from Liddle's original 2011 formulation, in Eq. 4 I disregard electricity's share of total domestic energy consumption, originally included in Liddle's analysis as an explanatory variable. The exclusion is based on the fact that while in Liddle's original analysis this variable was included as a proxy measure of access to electricity (which was necessary because of the inclusion of some less developed OECD countries), in New York State such access is a non-issue. Because the formulation is linear, the exclusion poses no methodological difficulties.

I employ two additional steps to cautiously and conservatively develop the final form of the model from Eq. 4's skeletal form. In the first, I test individually each candidate predictor before I

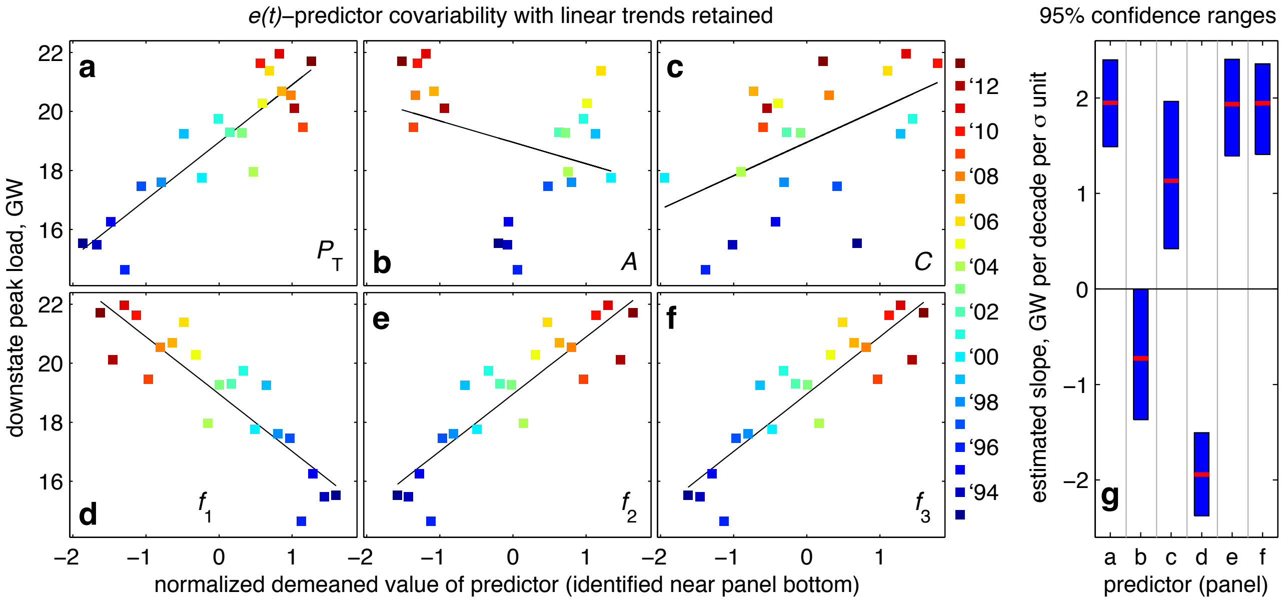

Figure 7: Dependence of downstate 1993–2013 peak loads (vertical axes of panels a–f) on individual predictors (indicated near each panel bottom). Predictor values (horizontal axes) are not addressed in their raw form. Rather, I first detrend in time each predictor time series (with the added effect of having zero mean), and then divide each of the 21 values in the detrended time series by the mean of the raw (pre-detrending) predictor value over the full period. Thus, e.g., a value along any of the horizontal axes of 0 means that the predictor value at that time exactly coincided with the linear trend value. A value of 1 along any of the horizontal axes, likewise, means that the deviation of the predictor value at that time from the linear time trend is exactly equal to the predictor's temporal long term mean. This, and the fact that all vertical axes span the same range, permits direct comparison of any of panels a–f to any other. Also in panels a–f, while time is not explicitly shown as an independent variable, its progression is shown by the dots' colors, with dark blue (brownish red) indicating 1993 (2013), and the full range of colors/years shown to the right of panels c and f. Linear trends of $e(t) = e[p_i(t)]$ (in which $p_i(t)$ is predictor i's time series) are presented as solid black lines. Panel g tests the robustness of those linear dependencies using a 200 member Monte Carlo boot strapping randomization, in each of which the trends are recalculated using 90% of available data. Sorting the trends thus obtained, the estimated trends in positions round(200×0.025) and round(200×0.975) (5 and 195) are the lower and upper bounds of the trend distribution, respectively. A trend is said to be $p < 0.05$ significant if both of those are of one sign, so that their spread excludes zero. Panel g thus shows that downstate peak electricity load is insignificantly related to affluence, but significantly related individually to the other five tested predictors.

retain in the final model. Then, I employ an R^2-based stepwise regression formalism to eliminate redundant predictors and retain only ones that offer additional explanatory power beyond the combined explanatory power of already retained predictors.

The first of these steps is addressed in Fig. 7. In Fig. 7, I test individually each of Eq. 4's six potential predictors. The tests are based on evaluating the presence or absence of linear dependence of e on each of the tested predictor individually (panels a–f), and then (panel g) Monte Carlo testing the significance of each of the trends. As Fig. 7g shows, downstate peak electricity load is unrelated to affluence, but significantly related individually to the other five tested predictors. The five predictors that prove significantly related to e, P_T, C, and $f_{1,2,3}$ are still potential contenders for the final model, while A is no longer considered; Eq. 4 is reduced to, at most,

$$e(t) = x_1 + x_2 P_T(t) + x_3 C(t) + \sum_{i=1}^{I} (x_{3+i} f_i) + \xi(t), \tag{5}$$

or, in vector form,

$$\mathbf{e} = \begin{pmatrix} \mathbf{o} & \mathbf{P}_T & \mathbf{C} & \mathbf{f}_1 & \mathbf{f}_2 & \mathbf{f}_3 \end{pmatrix} \mathbf{x} + \vec{\xi} \equiv \mathbf{A}\mathbf{x} + \vec{\xi}. \tag{6}$$

In Eq. 6, which implicitly defines $\mathbf{A} \in \mathbb{R}^{21\times 6}$, $\mathbf{A}$'s six columns, and $\vec{\xi}$, are 21-vectors, $\mathbf{o}$ is a vector of ones, $\mathbf{x}$ is the 6-vector of optimized coefficients, and $\vec{\xi}$ is the misfit vector.

Two additional conservative exclusions take place following the stepwise formalism and the algebraic analysis of $\mathbf{A}$. As given in Eq. 6, $\mathbf{A}$ is effectively if not formally rank deficient, with a trailing singular value, while not identically zero, 5 orders of magnitude smaller than the leading. This can be understood by reexamining simultaneously Fig. 7e,f. The panels clearly show nearly identical time trends, with unstructured residuals about the respective trends. If both variables are related primarily through the trend, and if the trends of both are essentially the same, the two are mutually redundant, and one must be eliminated. The reverse is true for Fig. 7d,e, whose trends are sign reversed but are otherwise similar, also with unstructured residuals about the respective trends. In recognition of these two redundancies, and the fact that f_3 accounts for a small portion of the population (under 10% and 13% in 2013 and 2040), f_3 is a natural first candidate for elimination in order to render the final $\mathbf{A}$'s singular spectrum less lopsided, In addition, I handle the partial mutual redundancy of f_1 and f_2 by using their ratio as a single predictor, $r_i := f_{i,1}/f_{i,2}$ and $\mathbf{r} = (\ r_1\ r_2\ \cdots\ r_{21}\ )^T \in \mathbb{R}^{21\times 1}$.

The final model is thus

$$\mathbf{e} = \begin{pmatrix} \mathbf{o} & \mathbf{P}_T & \mathbf{C} & \mathbf{r} \end{pmatrix} \mathbf{x} + \vec{\xi} \equiv \mathbf{A}\mathbf{x} + \vec{\xi}. \tag{7}$$

with $\mathbf{A} \in \mathbb{R}^{21\times 4}$ and $\mathbf{x} \in \mathbb{R}^{4\times 1}$. The singular spectrum of this final $\mathbf{A}$ satisfies $\sigma_4 \approx 0.15\sigma_1$ (where σ_i is $\mathbf{A}$'s ith singular value, with the spectrum arranged according to $\sigma_i \geq \sigma_{i+1}$). This final $\mathbf{A}$ thus possesses a stable, adequate singular spectrum.

7 The Final Predictions and Their Testing

I next test again the predictors of the final model given in Eq. 7. The test again consists of 200 Monte Carlo experiments, in each of which I withhold two randomly chosen data points (roughly 10% of the available 21), and recalculate the linear trends. Sorting these 200 trends, and choosing estimated trends in positions $200 \times 0.025 = 5$ and $200 \times 0.975 = 195$ as the lower and upper bounds of the trend distribution, respectively. A trend is said to be $p < 0.05$ significant if both of those are like-sign, so that their span excludes zero.

The results of these tests are summarized in Fig. 8a–d. All three predictors prove $p < 0.05$ significant (Fig. 8a), but the climate measure C is least significant (in repeated experiments, at times 1 or 2 of the estimated climate trends are negative, whereas the other two predictors are at least $p < 0.005$ significant as in those repeated experiments not a single one of the 200 MC realization has ever fallen across the zero line from the bulk of the distribution).

The final prediction comprises two steps. In the first, $\hat{\mathbf{x}}$ is obtained by solving Eq. 7 with $\mathbf{A}$ holding 1993–2013 observations. In the second step, which involves bootstrapping randomization, I use this $\hat{\mathbf{x}}$, denoted $\hat{\mathbf{x}}^{\text{past}}$, and an $\mathbf{A}$ (as defined in Eq. 7) holding future observations (2015–2040 in 5 year increments) to produce future e estimates,

$$\hat{\mathbf{e}}^{\text{future}} = \tilde{\mathbf{A}}^{\text{future}} \hat{\mathbf{x}}^{\text{past}}. \tag{8}$$

Here,

$$\tilde{\mathbf{A}}^{\text{future}} = \begin{bmatrix} \mathbf{o} & \mathbf{P}_T^{\text{future}} + \mathcal{N}(0, s_{P_T}^2) & \mathbf{C}^{\text{future}} + \mathcal{N}(0, s_C^2) & \mathbf{r}^{\text{future}} + \mathcal{N}(0, s_r^2) \end{bmatrix} \tag{9}$$

is a perturbed version of the 2015–2040 $\mathbf{A}$ in which $\mathbf{o}$ is a vector of ones, as before, and the latter three columns are P_T, C and r predicted for 2015–2040 plus perturbations randomly drawn in

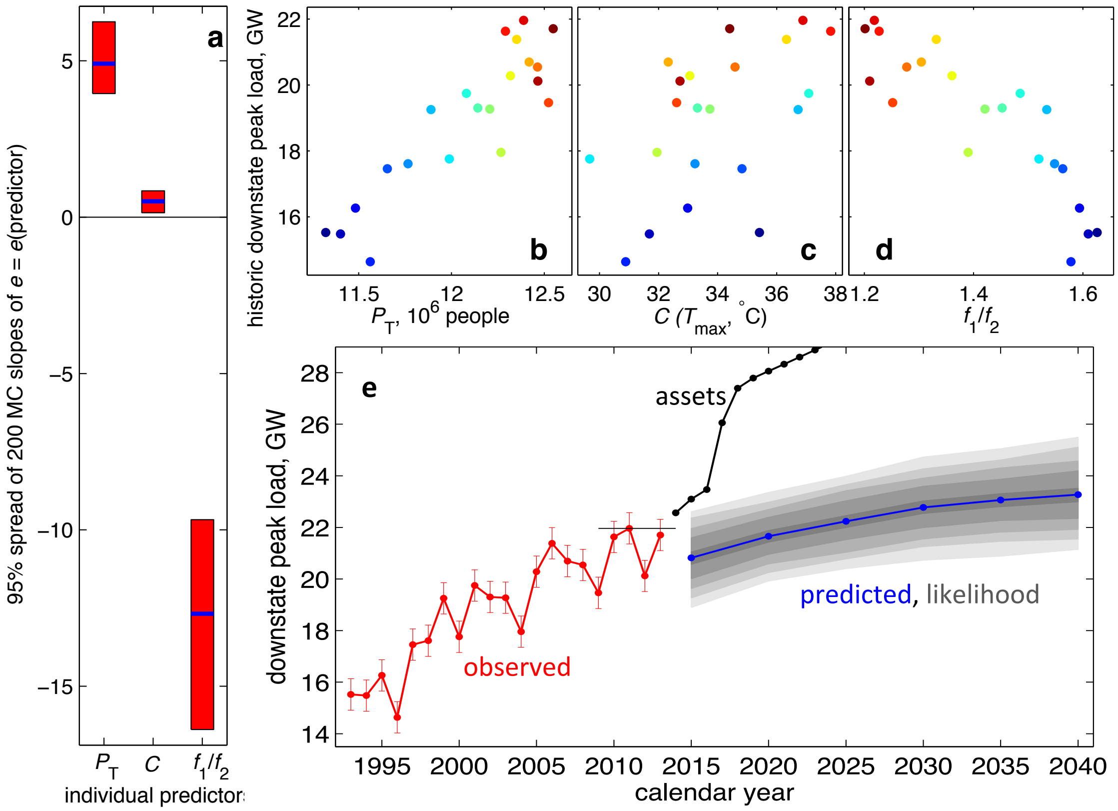

Figure 8: Some attributes of the final model and future downstate peak load prediction. In panel a, I test again the significance of the linear trends of e as a function of retained predictors, presenting the 95% spreads of 200 member Monte Carlo populations. As all ranges exclude zero, all are $p < 0.05$ significant. Panels b–d present e vs. a retained predictor (in the original values and units shown by the horizontal axis labels, unlike in Fig. 7a–f). As in Fig. 7a–f, time is implicitly shown by the colors; see scale there. Panel e presents the final prediction of future downstate peak loads. The red curve shows the 1993–2013 time series of annual downstate peak loads. The whiskers about the red curve show the cross validation skill obtained by training the model on 19 randomly chosen data points and using the model thus trained to "forecast" peak loads of the withheld 2 years. The most likely forecast of future downstate loads for 2015–2040 are shown in solid blue. The probability density function of the peak load forecasts is shown in gray shadings, with lightest to darkest corresponding to percentiles 2–98, 5–95, 10–90, 20–80, and 40–60. Downstate assets (comprising generations withing the region plus transmission capacity into it) are shown in solid black. The curve is based on (*i*) starting from the maximum peak load negotiated successfully, that of 2011 (thin horizontal black); (*ii*) adding half of the assets already in the NYISO "in-queue" designation and slated for completion in each year over 2015–2019; (*iii*) continuation of the recently observed efficiency gains of 0.9% yr^{-1} (several other states have attained 2 or even 3% per years savings in recent years). Details of these calculations are given in the text; and (*vi*) random interannual variability of all predictors according to their 1993–2013 observed variance.

each Monte Carlo realization from normal distributions with zero mean and the variances of P_T, $T_{\max}$, and f_i/f_2 obtained from the 1993–2013 observations. The addition of the random perturbations, along with the assumption of roughly time invariant variances, allows for simulating a range of future peak loads that reflects the full scope of possible future total population, climate, and population age distribution.

The use of Eqs. 8 and 9 for predictions hinges of future predictor data, obtained as follows. The two population related predictors P_T and $r = f_1/f_2$ are still from the Cornell Program on Applied Demographics[1], available for 2015–2040 in 5 yr increments (six points). The climate predictor $C = T_{\max}$ is taken from future climate model predictions of annual maximum of $T_{\max}$, daily maximum temperatures. The model outputs are taken from CMIP5 (the 5th Coupled Model Intercomparison Project[2]), a climate research project under the auspices of World Climate Research Programme in support of the quasi regular Assessment Reports (the most current of which is the 5th[3], compiled by the IPCC (Intergovernmental Panel of Climate Change[4]).

For the grid point containing New York City, the 17 available models predict a warming of 0.2–4.0 K by mid 21st century, with most falling in the 1.8–2.8 K range. I therefore assume here a conservative 3 K warming by 2050. Future downstate peak power loads e are shown by the blue curve in the right hand part of Fig. 8e.

Next, and crucially important, is testing the dependence of the future e predictions on the specifics of the available past data. That is, examining the red curve in Fig. 8e, one immediately recognizes that it comprises two parts. The first, the secular rise ("the trend") is clearly robust, as Fig. 6g has already established. Conversely, the up and down fluctuations about the trend of Fig. 8e's red curve are clearly random. Any robust estimate of future downstate e trajectory must therefore depend strongly on the trend, but be largely independent of the vagaries of the specific realization of past e trajectory that was actually observed. To ascertain that our e projections meet this criterion, I again carry out 200 Monte Carlo experiments. In each MC realization, I estimate $\hat{\mathbf{x}}^{\text{past}}$ using only randomly chosen 19 of $\mathbf{A}^{\text{past}}$'s 21 rows, and then use this deficient $\hat{\mathbf{x}}^{\text{past}}$ and a full, perturbed $\tilde{\mathbf{A}}^{\text{future}}$ to solve Eq. 8 for one $\hat{\mathbf{e}}^{\text{future}}$. Repeating this randomization 200 times, sorting the resultant predicted e trajectories, and retaining trajectories 4 and 196, I obtain the 96% confidence interval on future e projections. This interval is spanned by the full scope of gray shadings of Fig. 8e's right. Subintervals within the full span are derived by straightforwardly extending the above limits to other percentiles of the distribution.

The estimates cluster together very well, indicating that the model is indeed robust with respect to the random parts of Fig. 8e's red left curve. To remove ambiguity, a bad model—one that depends too strongly on the random, irreproducible parts of past observations (the red curve)—will produce confidence intervals that fail two important criteria. First, they will not coincide nicely with and bracket centrally the optimal prediction (the one based on using all 21 rows of $\mathbf{A}^{\text{past}}$). Instead, the gray shadings in Fig. 8e achieve both for the central blue curve. Second, a bad model's confidence interval will diverge widely with time, yielding a gray shaded region whose height increases wildly from left to right. Again, Fig. 8e shows quite the opposite for our model, a tight and minutely increasing shaded region. Together, these tests give every reason to interpret Fig. 8e as predicting a very gradual and slight increase in downstate future peak loads, rising to, at most, 26 GW by 2040 from the recent values, 21-22 GW.

For comparison and as a yardstick to which the current results can be compared, Fig. 8e presents existing downstate power assets and their expected growth in coming years. The estimate of peak load relevant assets starts with the recent 2001 peak load, at 21.96 GW the highest load ever recorded. From that point—which is our starting point because if such a load was encountered and negotiated successfully, assets of *at least* this magnitude are clearly available for the downstate region—forward, two types of added assets are considered. The first addresses load reductions due to energy efficiency gains, which I express as the equivalent addition of assets. (Because the variable of interest is the assets minus consumption, consumption reduction by added efficiency is entirely equivalent to added assets, as discussed further below.)

The second addition of assets available to meet downstate peak loads involves energy efficiency gains. While New York State recent energy savings are modest relative to other states, they are positive. In recognition of this, e.g., the NYISO Gold Book load forecasts take explicit note of projected energy efficiency savings attained through ratepayer-funded energy efficiency programs administered by NYSERDA, NYPA or LIPA. Additional savings are also expected as already enacted more rigorous building codes and appliance standards take wider hold (as the existing housing stock and currently used appliances are gradually replaced due to normal end-of-life by new standard compliant assets, an effect Aroonruengsawat et al.[5] have estimated to fall in the 3–5% of total residential electricity use per year). The NYISO near future forecasted annual savings, e.g., amount to 1.1% of total state electricity sales. For reasons not explicitly stated by NYISO, subsequent savings decline, amounting to roughly 0.7% (0.3%) of electricity sales in 2017 (2019-2021). A recent NYSERDA study[6] employed a different methodology to estimate achievable electricity savings potential of 18% by 2032, which amounts to an average savings rate of 0.9% per year. Another estimate, that of Woolf et al.[7] is 1.5% savings per year, which those authors deemed "clearly feasible" for New York State in the decade following their 2011 publication.

With the above estimates in mind, the 0.9% per year energy savings rate estimate used in Fig. 8e is suitably conservative. I represent this as an equivalent virtual added assets that would have maintained the same vertical distance between demand and expected assets. This is shown by the solid black curve in Fig. 8e's upper right. Expected peak loads, even their upper bounds (and even the not shown single highest realization of the 200 MC realizations), are well shy of expected assets at any time through 2040 (the solid black curve).

8 Summary

Downstate peak loads are indeed expected to rise modestly (by about 10% of today's peak loads) through about 2035, and stabilize thereafter. This may explain the assertion that additional transmission assets are necessary.

A closer examination of the facts, however, indicates that even assuming no further rise in the recently observed energy savings rate of 0.9% per year, existing assets handily exceed expected future peak loads with Indian Point closed.

At the same time, 0.9% electricity savings per year is probably a low estimate. This is so both relative to actually realized recent saving in New York State, as well as relative to other northeastern states. For example, Massachusetts and Vermont both plan—based on their recent records—to exceed 2% savings per year[1]. It is also far smaller, by as much as a factor of 2, than expected gains by measures already in place in the State (notably more stringent building efficiency code) whose effect is known to lag behind their enactment.

Thus I find no evidence that additional generation or transmission capacity is needed in New York's downstate region.

9 Notes

Section 2

1. which appears scale-independent and true in both developed and developing nations, e.g., respectively, Gough, I. et al. 2012: The distribution of total greenhouse gas emissions by households in the UK, and some implications for social policy, CASE/152, Centre for Analysis of Social Exclusion, London School of Economics; and Halicioglu, F. 2008: An econometric study of CO_2 emissions, energy consumption, income and foreign trade in Turkey, MPRA Paper No. 11457, Presented at the 31st IAEE Annual International Conference Istanbul-Turkey, June 18-20, 2008, http://mpra.ub.uni-muenchen.de/11457/.

2. Rosen, L. D. and M. M. Weil, 1995: Computer availability, computer experience and technophobia among public school teachers, *Computers in Human Behavior*, **11**(1), 931, DOI: 10.1016/0747-5632(94)00018-D.

3. Geelsa, F. W., M. P. Hekkert and S. Jacobsson, 2008: The dynamics of sustainable innovation journeys, *Tech. Analysis & Strategic Management*, **20**(5), 521-536, DOI:10.1080/09537320802292982.

4. *Progress Report: Advancing Solar Energy Across America*, http://energy.gov/articles/progress-report-advancing-solar-energy-across-america, retrieved 6/25/14; data from the National Renewable Energy Lab, U.S. Dept. of Energy.

5. Solar Energy Industries Association (SEIA; http://www.seia.org/), *Solar Energy Facts: 2013 Year in Review.*

6. *Solar Energy Facts: 2013 Year in Review* cited above.

7. Hallegatte, S., F. Henriet and J. Corfee-Morlot, 2011: The economics of climate change impacts and policy benefits at city scale: a conceptual framework, *Climatic Change* **104**(1), 51-87; Climatic Change Hunt, A. and P. Watkiss, 2011: Climate change impacts and adaptation in cities: a review of the literature , **104**(1), 13-49. .

8. Biermann, F. and I. Boas, 2010: Preparing for a warmer world: Towards a global governance system to protect climate refugees, *Glob. Environ. Politics*, **10**(1), 60-88

9. Biermann, F. and I. Boas, 2008: Protecting climate refugees: The case for a global protocol, *Environment: Science and Policy for Sustainable Development*, **50**(6), 8-17.

Section 3

1. http://www.nyiso.com

2. http://www.nyiso.com/public/markets_operations/market_data/maps/index.jsp.

Section 4

1. Data from the Cornell Program on Applied Demographics, http://pad.human.cornell.edu/index.cfm; see, e.g., Vink, J., 2014: Cornell Program on Applied Demographics, *Highlights of the US Census Bureau 2013 Estimates of County Population Characteristics*, June 2014. 5 pp.

2. e.g., Lariviere, I. and G. Lafrance, 1999: Modelling the electricity consumption of cities: Effect of urban density, *Energy Econ.*, **21**, 53-66; Liddle, B., 2011: Consumption-driven environmental impact and age structure change in OECD countries: A cointegration-STIRPAT analysis, *Demographic Res.*, **24**, 749-770.

3. Liddle, B., 2011: Consumption-driven environmental impact and age structure change in OECD countries: A cointegration-STIRPAT analysis, *Demog. Res.*, **24** (article 30), 749-770.

Section 5

1. http://www.nyiso.com.

2. e.g., http://www.nyiso.com/public/webdocs/markets_operations/services/planning/Documents_and_Reso or its most recent, 2013, counterpart.

3. as conformed by NYISO Customer Service Department, http://www.nyiso.com/public/contact/index.jsp, contacted on Aug. 12th, 13th and 14th, 2014.

Section 6

1. Liddle, B., 2011: Consumption-driven environmental impact and age structure change in OECD countries: A cointegration-STIRPAT analysis, *Demographic Res.*, **24**, 749-770.

2. Dietz, T. and E. Rosa, 1997: Effects of population and affluence on CO_2 emissions, *Proc. Nat. Acad. Sci. of the U.S.A.*, **94**(1), 175-179.

3. Ehrlich, P. and J. Holdren, 1971: The impact of population growth, *Science*, **171**(3977), 1212-1217.

4. http://www.bea.gov/regional/index.htm

5. http://www.bls.gov/data/inflation_calculator.htm

Section 7

1. http://pad.human.cornell.edu/index.cfm.

2. http://www-pcmdi.llnl.gov/projects/cmip/.

3. http://www.ipcc.ch/report/ar5/index.shtml.

4. http://www.ipcc.ch.

5. Aroonruengsawat, A.,, M. Auffhammer and A. H. Sanstad. 2012: The impact of state level building codes on residential electricity consumption, *Energy Journal-Cleveland*, **33**(1), 31.

6. New York State Energy Research and Development Authority, 2014: *Energy Efficiency and Renewable Energy Potential Study of New York State*, Vol. 2, NYSERDA Report 14-19, April 2014, Fig. 3, p. 13.

7. Woolf, T., M. Wittenstein and R. Fagan, 2011:*Indian Point Energy Center Nuclear Plant Retirement Analysis: Replacement Options, Reliability Issues and Economic Effects*, Synapse Energy Economics, Inc., Cambridge, MA, October 17 2011, 33 pp.

Section 8

1. Massachusetts Electric Efficiency Program Administrators, Three-Year Energy Efficiency Plan, October 2009; Vermont Efficiency Investment Corp, Vermont Energy Efficiency Utilities Investment Report, April 2011.