Title

Long term load forecasting accuracy in electric utility integrated resource planning

Authors
Carvallo, Juan Pablo
Larsen, Peter H
Sanstad, Alan H
et al.

Publication Date
2018-08-01

DOI
10.1016/j.enpol.2018.04.060

Energy Policy 119 (2018) 410–422

Contents lists available at ScienceDirect

Energy Policy

journal homepage: www.elsevier.com/locate/enpol

Long term load forecasting accuracy in electric utility integrated resource planning

Juan Pablo Carvallo*, Peter H. Larsen, Alan H. Sanstad, Charles A. Goldman

Ernest Orlando Lawrence Berkeley National Laboratory, 1 Cyclotron Road, MS, 90R4000, Berkeley, CA 94720-8136, United States

ARTICLE INFO

Keywords:
Resource planning
Forecast accuracy
Load
Retrospective analysis
Resource expansion
Electric utility

ABSTRACT

Forecasts of electricity consumption and peak demand over time horizons of one or two decades are a key element in electric utilities' meeting their core objective and obligation to ensure reliable and affordable electricity supplies for their customers while complying with a range of energy and environmental regulations and policies. These forecasts are an important input to integrated resource planning (IRP) processes involving utilities, regulators, and other stake-holders. Despite their importance, however, there has been little analysis of long term utility load forecasting accuracy. We conduct a retrospective analysis of long term load forecasts on twelve Western U. S. electric utilities in the mid-2000s to find that most overestimated both energy consumption and peak demand growth. A key reason for this was the use of assumptions that led to an overestimation of economic growth. We find that the complexity of forecast methods and the accuracy of these forecasts are mildly correlated. In addition, sensitivity and risk analysis of load growth and its implications for capacity expansion were not well integrated with subsequent implementation. We review changes in the utilities load forecasting methods over the subsequent decade, and discuss the policy implications of long term load forecast inaccuracy and its underlying causes.

1. Introduction

From the origins of the U. S. electricity industry in the 19th century with Thomas Edison's first power-generation plant in New York City, electric utility planning and operations have become highly complex, multi-faceted processes. Vertically integrated[1] U. S. utilities or load-serving entities (LSEs)[2] operating in states with a regulated electricity sector must determine how to provide electricity services to customers while complying with a range of energy and environmental regulations and policies, and respecting the economic objectives of both the utility and customers. These functions entail the use of a range of quantitative analytical methods, including computational modeling and statistical analysis. LSEs' core obligation is to ensure reliable, clean, and affordable electricity supplies for their customers. It follows that forecasts of electricity consumption (GWh) and peak demand (MW) over the time horizons of one or two decades are a cornerstone of LSE's planning process.

Long term load forecasts are a key input to integrated resource planning (IRP), which has become the core process whereby many U.S. LSEs, in consultation with regulators and other stakeholders, determine portfolios of electricity resources to meet demand over the long term. Such forecasts form the basis of utilities' capacity expansion planning, which consists of building or acquiring power generation plants, purchasing power from other sources, and other means of securing electricity supplies and services for their customers. Because energy and environmental policy goals are also a major element of IRP in many states, these forecasts also influence efforts to achieve larger social objectives. An important example is the consideration of energy efficiency and other demand side measures into utility planning, which has become a high policy and regulatory priority in much of the U.S.

Load forecast horizons employed in the electric industry often range from hours to decades. Hour and up to yearlong forecasts are categorized as short and medium term and are commonly used for operational efficiency. Decades long forecasts are categorized as long term and are the type used in utility planning. Short and medium term electric load forecasting has been and continues to be the focus of considerable research, and is the subject of a sizable literature. Hong and Shahidehpour (2015) provide a comprehensive overview. In contrast, there has been relatively little study of long term load forecasting. Willis and Northcote-Green (1984) compared methods and accuracy of

* Corresponding author.
E-mail address: jpcarvallo@lbl.gov (J.P. Carvallo).

[1] Vertical integration refers to the combination of different stages of production or segments in a value chain under a single company.

[2] "Load-serving entity" is a more precise term than "utility" to refer to firms that sell electric power to end-use customers. However, in this paper these terms will be used interchangeably.

https://doi.org/10.1016/j.enpol.2018.04.060
Received 14 September 2017; Received in revised form 29 March 2018; Accepted 27 April 2018

14 distribution system load forecasts. Nelson and Peck (1985) analyzed load forecasts from the 1970s prepared by the National Electricity Reliability Corporation (NERC), which combine individual utility service territory and regional level forecasts into a national level forecast. They found systematic over projection of demand. Mitchell et al. (1986) retrospectively evaluated the accuracy of long term energy and peak demand forecasts by utilities, government agencies, and academic researchers.

This paper aims to help fill the knowledge gap on long term forecasting by focusing on forecast performance or accuracy. It reports the results of a retrospective analysis of load forecasts produced in the mid-2000s by twelve utilities in the western United States. It also reviews the utilities' forecast methodologies and sensitivity analyses. This analysis is the companion paper to Carvallo et al. (2017), which studies the relationship between utilities' planning processes – including load forecasting – and their actual resource procurement decisions.

This paper is organized as follows. We report the sources of data used in the analysis in Section 2, followed by a discussion of the LSEs' forecasting methodologies in Section 3. In Section 4 we describe our quantitative analysis of forecast error. We then turn in Section 5 to a discussion of the effects of economic growth assumptions on forecast accuracy. We present an analysis of the LSEs' approaches to load sensitivity analysis in Section 6, followed by a discussion of changes to LSEs' load forecasting methods and inputs over time in Section 7. We conclude with Section 8, which holds a summary, discussion of policy implications, and suggestions for further research.

2. Data sources and methods

2.1. Information on forecasts

We collect forecasts from IRP produced from 2003 to 2007 by twelve LSEs across the Western Electricity Coordinating Council (WECC). We focus on WECC because this territory includes the largest U.S. LSEs that were required to file resource plans during this period (Wilkerson et al., 2014). Three large California investor-owned utilities (IOUs) were excluded because they did not use IRP during the analyzed timeframe. Aside from the California IOUs, the LSEs selected for this study are the twelve largest in WECC representing 34% and 32% of customers and retail sales in 2014, respectively.

The vintage years for the IRPs, which correspond to the base years for the forecasts, were selected for several reasons. These plans were created sufficiently long ago that their forecasts could be compared to actual[3] values over periods long enough to allow substantive analysis – to the year 2014, the most recent year for which these values were available at the time this study was conducted[4] (see Table 1). Depending on the LSE, between seven and eleven years of observed energy and peak demand are available to be compared to the original forecast. In addition, in reviewing plans older than those selected we found several shortcomings, including limited data and documentation of the type needed for this analysis. As discussed in Section 7, we also review one recent plan (produced between 2011 and 2015) for each LSE to understand whether and how the methodologies and techniques used to produce forecasts have changed over time.

The analysis period includes the 2008/2009 economic recession, which would be expected to have a substantial or even disproportionate effect on the accuracy of load forecasts made prior to its onset. It is a truism that all forecasts, including those of electricity use, are subject to error due to unforeseen circumstances. As we discuss later in the paper, the documentation indicates that the LSEs view economic and demographic variables as the primary drivers of demand, and the inevitable but always uncertain timing of events such as recessions means that such events are essentially guaranteed to affect long term load forecasts in not fully predictable ways, regardless of the forecast interval. Thus, an analysis period including the downturn that began in 2008, which was unusually severe, can if anything allow greater insight into the nature of load forecast accuracy and how forecast errors are addressed in the IRP process than might be available from studying a period without such an event. Put differently, the 2008/2009 recession provides an interesting "stress test" of LSE load forecasting procedures in the context of IRP.

We collect three basic types of numerical forecast information from each IRP: electricity use, peak demand, and the demand side resources of energy efficiency (EE) and demand response (DR).[5] For the forecast to actual comparison we used the base or reference case load forecast in each resource plan (all 12 LSEs produced these cases for energy and 11 for peak). We use high and low load forecasts where these were available for sensitivity analysis. LSEs account differently their energy efficiency and demand response measures, with some subtracting projected savings from these resources into their load forecasts, and some reporting them separately. For the forecasts that had not already done so, we subtract these savings from the raw energy and peak demand forecasts in order to calculate net load.[6] The use of net forecasts is appropriate for comparison with actual energy and peak demand, since the latter have embedded within them the effects of demand side programs and other acquired energy efficiency over the periods considered in the analysis.

2.2. Information on actual energy use and peak demand

Data on energy consumption and peak demand is obtained primarily from the Velocity Suite system supplied by ABB-Ventyx—an online database system that compiles publicly-available data and also contains proprietary values for variables that are not always publicly-available, including retail fuel prices and marginal costs (ABB-Ventyx, 2016).

The Velocity Suite system contains load data as measured by retail sales, which is typically reported through the Energy Information Administration (EIA) Form 861 (EIA, 2016). In order to compare forecasts to actual values, it was necessary to identify the types of sales that utilities themselves considered as part of the *position*[7] for the resource planning process. All 12 IRPs in our sample accounted for retail sales to ultimate consumers when creating their forecasts, and most (10 LSEs) included transmission and distribution losses to reflect demand at the generation level. For the remaining two cases, we added transmission and distribution losses.

In addition, we review the IRP documentation to determine which LSEs accounted explicitly for selected wholesale sales for which they had firm contracts at the time of the forecasts, and use data from EIA Form 412 and FERC Form 1 to identify and include appropriate wholesale sales as necessary. Finally, we use historical load information when available in the most recent LSEs plans to check our estimates for actual values.

[3] We refer to these also as "realized" or "observed" through the paper.

[4] In the case of PNM and PGE we selected the oldest plans we were able to find that included the required data. PNM filed its first resource plan in 2005 but it did not include most of the quantitative data required for the analysis.

[5] By the time the IRP documents we analyze were issued, adoption of demand side resources such as distributed generation or storage was very limited and usually not considered. Therefore, we limit our analysis to EE and DR.

[6] By doing this, we implicitly include in our assessment the performance of energy efficiency and demand response forecasts. We recognize that the actual demand side resources may differ from these forecasts, but we lack the data to test this.

[7] The position is a term used to describe the annual expected difference between load and resources to meet it. When load is expected to be higher than the available resources, it is referred to as a negative position.

Table 1
Load serving entities (LSEs) and integrated resource plans analyzed in this study.

LSE short name	LSE name	First plan year	Recent plan year	Intermediate plan year(s)	Reference
Avista	Avista Corporation	2005	2013	2006–2011	(Avista, 2013, 2005)
COPSC[a]	Public Service Company of Colorado (Xcel Energy)	2003	2011	None	(COPSC, 2011, 2004)
Idaho	Idaho Power Company	2006	2013	2008, 2009, 2011	(Idaho, 2013, 2006)
LADWP	Los Angeles Department of Water and Power	2006	2012	2011	(LADWP, 2012, 2006)
NVPower	Nevada Power Company	2006	2012	2007, 2010	(NVPower, 2012, 2006)
NW	NorthWestern Corp. dba NorthWestern Energy	2004	2013	2007, 2011	(NW, 2013, 2004)
PacifiCorp	PacifiCorp	2004	2015	2007, 2009, 2011	(PacifiCorp, 2015, 2005)
PGE	Portland General Electric Company	2007	2013	2009, 2011	(PGE, 2014, 2007)
PNM	Public Service Company of New Mexico	2007	2011	None	(PNM, 2011, 2007)
PugetSound[a]	Puget Sound Energy, Inc.	2005	2013	2007, 2009, 2011	(PugetSound, 2013, 2005)
Seattle[a]	Seattle City Light	2006	2012	2008, 2010	(Seattle, 2012, 2006)
SierraPacific[a]	Sierra Pacific Power Company	2004	2013	2007, 2010	(SierraPacific, 2013, 2004)

[a] These LSEs are also known as PSCo (COPSC), PSE (PugetSound), SCL (Seattle), and SPP (SierraPacific). We use our own short names through this paper.

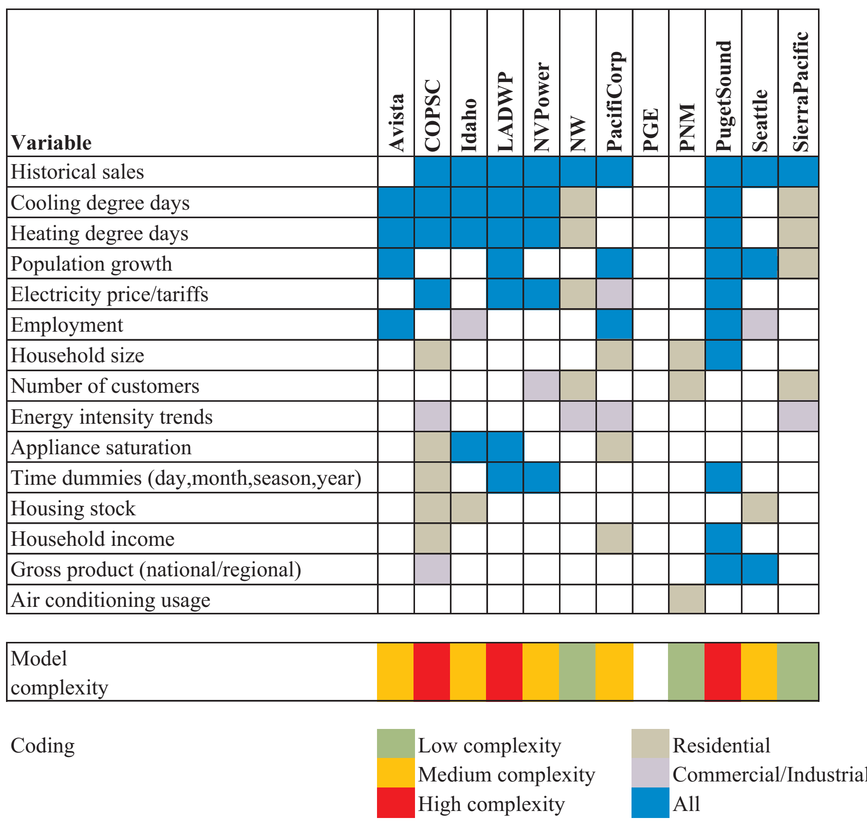

Fig. 1. Variables used for residential and commercial/industrial load forecasts, and model complexity. There is no information available for PGE in their 2007 plan. Blank spaces in the table indicate that the variable was not documented or formally employed in the forecast.

3. Description of forecasting methodologies

In this section we provide an overview of the LSEs' load forecasting methods.

Economic-demographic projections, historical sales data, and weather variables are employed by all forecasts (see Fig. 1). Most methods rely on historical data for each variable, but some LSEs develop variable projections (see Table A-1 and A-2 in the Appendix). Roughly three quarters of the LSEs in our sample relied on externally-developed demographic and economic forecasts, from a mix of public and private sources, including universities, state/federal agencies, and consulting firms. Most of these sources were proprietary.

Load forecast complexity varies widely across the LSEs examined in this study. We qualitatively assess the complexity of the load forecast by comparing the number of variables used to forecast residential and commercial/industrial demand and the analytical methods employed[8] (see Table 2). For example, utilities including PNM, NW and SierraPacific use simpler models compared to the models employed by COPSC, LADWP, and PugetSound.

Four types of modeling approaches, of varying degrees of complexity, were used by the various LSEs to create energy, peak demand, and hourly load forecasts: time-series regression, cross-sectional regression, engineering or bottom-up, and statistically adjusted end-use (SAE). Time series and cross-sectional regressions consistently use historical sales and weather variables as determinants of electricity demand. SAE models have a hybrid structure combining engineering end-

[8] We create an index that is the sum of squares of the number of variables used for forecasting in each segment (residential, commercial, and industrial). We then rank LSEs using this index and classify the bottom third as low complexity, the middle third as medium, and the top third as high complexity. Therefore, the complexity classification is relative, not absolute.

Table 2
Modeling approaches for residential and commercial load forecasting.[a]

	Time series regression (AR[b], MA[c])	Cross section regression	Engineering model	Statistically Adjusted End-use (SAE)
Avista		RC		
COPSC				RC
Idaho				RC
LADWP		RC		
NVPower	RC	RC		
NW	C	R		
PacifiCorp				
PGE				
PNM			RC	
PugetSound		RC		
Seattle		RC		
SierraPacific				

R: Residential; C: Commercial.
[a] There is no information available for PGE in their 2007 plan.
[b] Auto-regressive.
[c] Moving average.

use technology models with econometric equations. This type of data intensive model represents demand in terms of a saturation component (for appliance ownership), an engineering component (for appliance energy intensity), and a behavioral component (Hirst et al., 1977; Hirst and Carney, 1978; Sanstad et al., 2014). SAE models were employed by three of the twelve LSEs in our sample, while the two pure regression models are used by the majority of the LSEs.

LSEs developed customized forecast for three general customer classes: residential, commercial, and industrial. Residential and commercial forecasts were typically split into numbers of customers and use per customer, using different methods to forecast each separately. For example, PacifiCorp used a moving average method for short term forecasting and an SAE for long term projections while SierraPacific used an ARIMA (autoregression, integrative, moving average) method for the number of customers and a regression method for the use per customer. These methods were applied to both residential and commercial customer classes. In contrast, industrial consumption forecasts were mostly based on direct feedback from the largest customers, complemented by regional or sectoral market research reports. Finally, only a few of the LSEs evaluated in this paper reported load forecast results by customer class.

A key aspect of load forecasting frameworks is the extent to which they represent customer reductions in electricity consumption, fuel switching, or both, in response to changes in electricity prices. About half of the LSEs in our sample reported specific information about price elasticities, but only Avista reported cross-price elasticities with natural gas, although Idaho also pointed out the importance of relative fuel price changes for electricity demand. In contrast, NW and SierraPacific reported that they found no empirical evidence for statistically significant price elasticities for electricity or for natural gas.

Finally, we note that the breadth and depth of technical documentation on load forecasting varied widely among the LSEs in the older IRPs. In some cases, detailed information – including input types and values, mathematical formulae, and parameter estimates - were provided; in others, there was only narrative description. In no case, however, was there sufficient information to actually replicate the individual LSE forecasts or to test their sensitivity to the error in the input parameters.

4. Quantitative analysis of forecast error

To estimate forecast errors, we compare forecasts to actual results using two metrics: sum of errors and annual average growth rate, as defined in Hyndman (2006):

Sum of errors: Annual forecast errors for each LSE were calculated as the differences between that LSE's forecasted value and the actual value for each year of the forecast. We divide the sum of these errors by the corresponding sum of total load that was actually realized by the LSE during the forecast period. This serves to normalize the metric in order to compare forecast performance across LSEs of varying sizes. This technique averages out positive and negative deviations, which is useful for identifying systematic error that is expected given the variability of loads.

- *Annual average growth rate (AAGR):* We compare the first and last year forecast and actual values to estimate an average annual growth rate for each. The AAGR represents the rate at which the first year forecast or actual value would need to grow to match the final year assuming a compound growth rate. This relationship is captured in equation (1) and (2) below:

$$Y_{t+n} = Y_t \times (1+AAGR)^n \tag{1}$$

where Y_t is a forecast variable of interest (in our case energy and peak demand). Rearranging terms we have:

$$AAGR = \left(\frac{Y_{t+n}}{Y_t}\right)^{\frac{1}{n}} - 1 \tag{2}$$

The sum of errors is a relative metric, so a larger % difference implies larger forecast error over the time period of analysis. The AAGR captures, on average, the implied growth rate for a given variable. The difference between a forecast and an observed growth rate is, in this case, a measure of forecast error.

4.1. Energy consumption

Fig. 2 shows forecast and observed energy consumption normalized using the corresponding first year value (forecast or observed). This normalization essentially yields growth rates for both forecasted and observed values. The analysis time frame corresponds to the range of years between the first year in the forecast and 2014 (the most recent year for which we have observed values). For some LSEs, sales seem relatively inelastic to the 2008/2009 economic crisis and actual energy sales are close to the base forecast. However, sales for most LSEs slowed during and after the recession.

Table 3 compares the sum of errors in each LSE's analysis period with the sum of actual energy consumed in that period. The table shows that proportional forecast error varies considerably across the LSEs in our sample, ranging from a small negative value[9] to almost 20%. We note that with this metric, forecasts from earlier plans have greater chance of having larger proportional error. However, we do not find a correlation between age of plans and error in our analysis, probably because the plans are at most 3 years apart.

In Fig. 2, the depicted growth rate for each LSE reflects the ratio between a given year's forecasted value and the first year forecast (and similarly for observed values). Growth rates are relevant because we can compare growth expectations of different LSEs regardless of their current consumption levels. Some LSEs' growth rates were less sensitive to the recession than others', and several have not shown signs of recovery in energy sales growth rates by 2014.

We calculate the AAGR to facilitate comparison of growth estimates (Table 4). The AAGR condenses the medium-term accuracy of the forecast when compared to observed values as it is calculated by taking 2014 as the end year for all samples.

Utilities in our sample expected between 0.6% and 2.6% average growth rates for energy net of demand side resources. Observed growth

[9] Northwestern (NW) had − 2% error because it was the only utility that under-forecasted energy consumption.

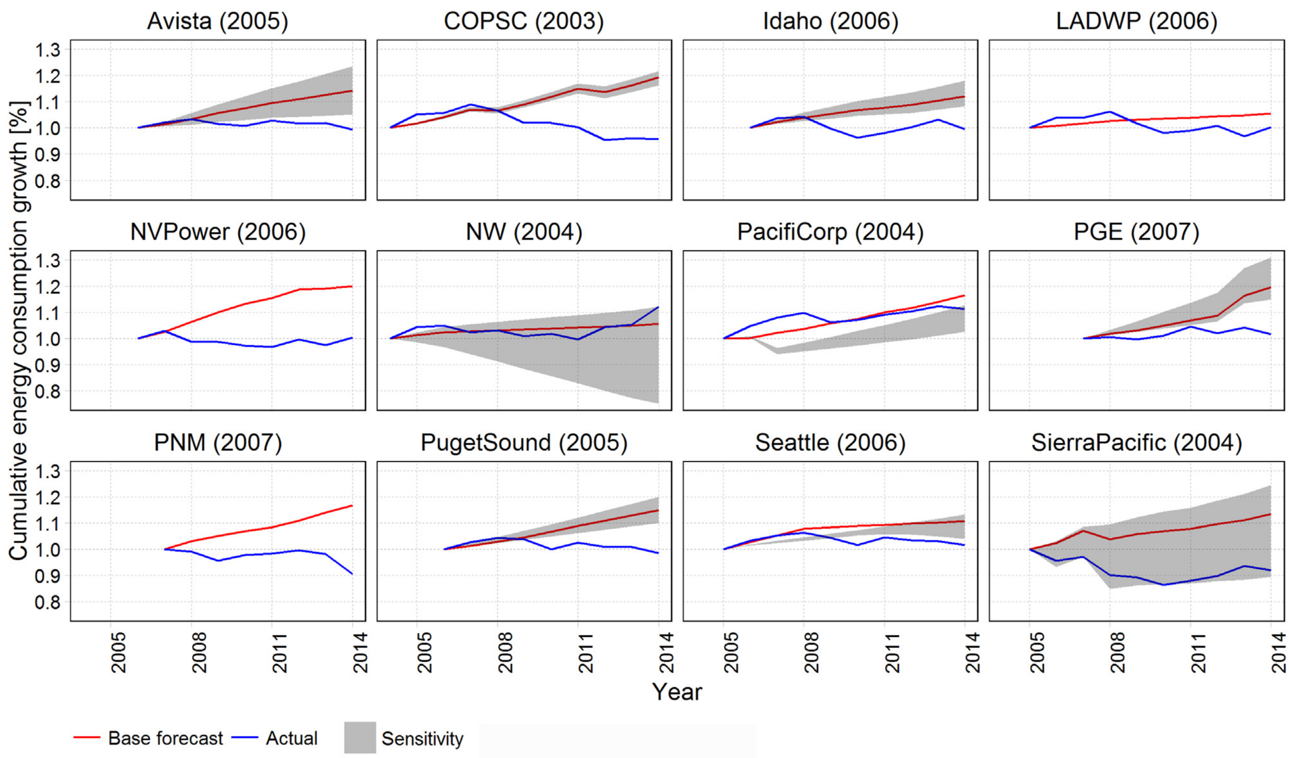

Fig. 2. Forecasted and actual energy consumption growth with alternative load growth forecast (sensitivity).

Table 3
Sum of errors as a proportion of total load for forecast horizon of 2014.

LSE	Sum of errors (1) [TWh]	Sum of actual load (2) [TWh]	Proportional Error (1)/(2)
Avista	14.73	85.36	17%
NVPower	26	199.01	13%
SierraPacific	10.57	89.37	12%
PGE	16.1	164.37	10%
Idaho	13.47	138.43	10%
PNM	5.64	85.17	7%
COPSC	21.41	365.05	6%
LADWP	13.04	236.45	6%
PacifiCorp	33.43	580.63	6%
Seattle	5.15	100.48	5%
PugetSound	2.09	206.15	1%
NW	− 1.29	68.5	− 2%

Table 4
Average annual growth rate for actual and forecast energy consumption.

	Energy AAGR (%)		
LSE	Base forecast	Actual	Difference
PNM	2.2%	− 1.4%	3.6%
PGE	2.6%	0.2%	2.4%
SierraPacific	1.4%	− 0.9%	2.3%
COPSC	1.8%	− 0.4%	2.2%
NVPower	2.3%	0.1%	2.2%
PugetSound	1.7%	− 0.2%	1.9%
Avista	1.7%	− 0.1%	1.8%
Idaho	1.4%	− 0.1%	1.5%
Seattle	1.1%	0.2%	0.9%
LADWP	0.6%	0.0%	0.6%
PacifiCorp	1.9%	1.3%	0.6%
NW	0.6%	1.2%	− 0.6%

rates for energy were much smaller, averaging close to zero across our sample of LSEs. About half of our sample shows negative observed AAGR and the ones that show positive AAGR are just above zero. The two exceptions are PacifiCorp, which grew at roughly two thirds of its expected rate, and NW, whose observed growth doubled its forecast growth for energy consumption. Comparing the results in Tables 3 and 4 shows that, in general, LSEs with smaller proportional errors also had more accurate forecast AAGR.

Accuracy of energy and peak demand forecasts differs across LSEs. The emphasis placed on energy forecast accuracy compared to peak demand forecast accuracy suggests the greater importance of the former for utilities that own or rely on hydropower to cover large portions of their retail obligations. Systems with high presence of reservoir hydropower are sensitive to hydrological conditions, as rain and snowfall determines how much energy will be available to generate on a given basin. These systems are energy-constrained but not capacity-constrained, since they can almost always meet peak demand by dispatching reservoir hydropower at peak load hours. This short term flexibility lowers the pressure for accurate peak demand forecast, but makes the system sensitive to long term energy consumption forecast accuracy.

4.2. Peak demand

Peak demand forecasts are qualitatively different from energy consumption forecasts, particularly due to their greater sensitivity to weather variation. The accuracy of energy consumption forecasting for a given utility does not necessarily correlate with the accuracy of its peak demand forecasts. In addition, several utilities reported progressive reductions of load factors (the ratio of average load to peak load) in their residential customer base. This means that historical hourly profiles and load factor assumptions may be less informative for peak demand forecast and make the latter more difficult to assess.

Forecasting results for several LSEs (COPSC, PGE, and NVPower) are mixed — for some years underestimating and for others over-estimating. Other LSEs (Avista, Idaho, SierraPacific, NVPower, and Seattle) consistently over-estimated in the period after the financial crisis,

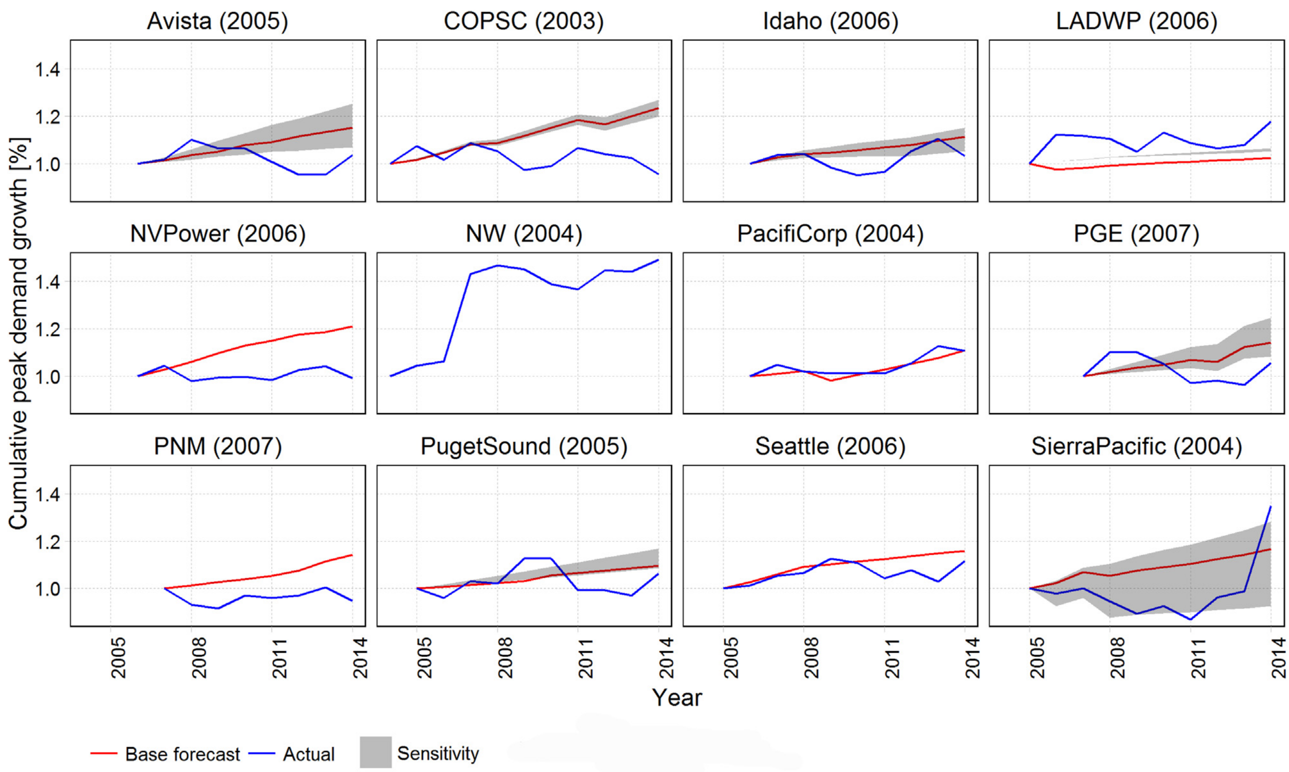

Fig. 3. Forecasted and actual peak demand growth, with alternative load growth forecasts.

Table 5
Average annual growth rate for actual and forecast peak demand.

	Demand AAGR (%)		
LSE	Base forecast	Actual	Difference
PNM	1.9%	− 0.8%	2.7%
COPSC	2.1%	− 0.5%	2.6%
NVPower	2.4%	− 0.1%	2.5%
Avista	1.8%	0.4%	1.4%
PGE	1.9%	0.8%	1.1%
Idaho	1.4%	0.4%	1.0%
Seattle	1.7%	1.2%	0.5%
PugetSound	1.1%	0.8%	0.3%
PacifiCorp	1.3%	1.3%	0.0%
LADWP	0.3%	1.8%	− 1.5%
SierraPacific	1.7%	3.4%	− 1.7%
NW	NA	4.1%	NA

which is symptomatic of a slower than expected recovery. Finally, some LSEs (PacifiCorp, LADWP, and PNM) had small systematic under or over-estimation of energy and peak load, but reasonably accurate average growth rate forecasts. This occurred with forecasts that underestimated and overestimated actual values in different periods. The average over longer periods of time yields reasonably accurate growth rates, but still shows errors in energy and/or peak load forecast.

As with energy consumption, we calculate and compare the implicit growth rates in both forecasted and observed peak demand values (see Fig. 3). We also calculate the AAGR to facilitate comparison (Table 5). Peak demand growth rates generally show a slowdown after the economic crisis, but not for all LSEs. Seattle, Avista, PGE, and PugetSound – all in the Pacific Northwest – show a lagged halt in growth compared to other utilities (e.g. COPSC, NVPower, and PNM) whose growth rates reflect an immediate impact. PacifiCorp, LADWP and NW were relatively less affected by the crisis. Peak demand growth rates are more resilient when compared to energy consumption growth rates, which is consistent with the LSEs reporting reduced load factors after the economic crisis.

Utilities in our sample expected growth rates between 0.3% and 2.4% for peak demand net of demand side resources. Growth rates for peak demand are much higher than for energy. In addition, several utilities reported higher peak demand growth than forecasted, even in the presence of the 2008/2009 crisis. This is consistent with statements in recent IRPs that report a reduction of load factors among residential and commercial customers. In addition, comparison of energy and peak demand observed values indicates that peak demand forecast error shows much larger variance across utilities. This supports the notion that it is more difficult to forecast long term peak demand than energy consumption.

A comprehensive and exhaustive assessment of load forecast error would require weather normalization of actual values. Unfortunately, we neither have access to the data nor the resources to perform this level of analysis. It is important to note, however, that weather normalization typically has a larger effect on short term forecast performance. Normalizing weather has less of an impact on long term forecasts like those analyzed in this paper.[10] To show this, we examine the historical record for cooling and heating degree-days (CDD and HDD) for the Pacific U.S. region to characterize the weather in the period analyzed. We find that 2014 and 2015 were warmer than average, but that all other years in our period of analysis were considered normal. For this reason, we believe that our findings would be largely unchanged if we included weather normalization for each observed value and for every LSE. Interestingly, LSEs do not report weather as a primary source of long term forecast error, which also supports the aforementioned point.

Notwithstanding the general pattern of forecast inaccuracy, forecasts for some LSEs performed significantly better than others, even in

[10] Climate change will impact CDD and HDD in the very long term. However, changes in CDD and HDD will not be large enough in the 10–20 year periods used for resource planning to be a relevant source of forecast error.

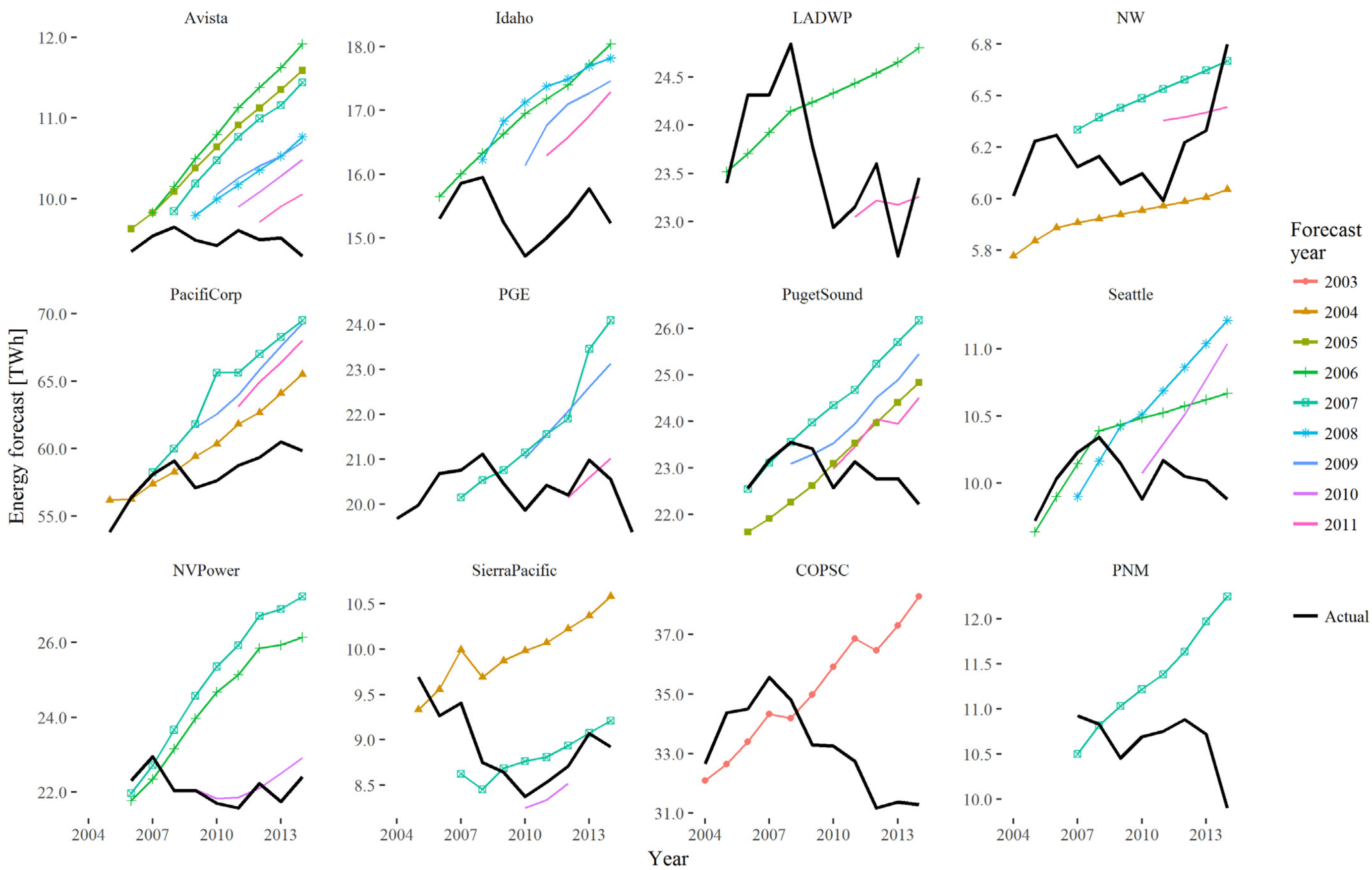

Fig. 4. Intermediate energy consumption forecasts.

the presence of the economic recession. Our results suggest that this was not correlated with the size of the utility or with its geographical location, although the relative ranking in Table 3 indicates that larger utilities have intermediate forecast error (see Appendix B). This is perhaps a consequence of having a diverse and large pool of customers that smooths economic impacts on forecast. This very preliminary assessment suggests that load composition may have an important effect on the planning strategy and load sensitivity analyses. For example, we find that LSEs with lower forecast error tend to have lower sales to industrial customers in proportion to their total sales. This makes intuitive sense: industrial customers are probably the most elastic customer class in relation to economic growth.[11] This makes their load is hard to forecast and its lumpy nature has a significant impact on forecast results. As some LSEs report in their plans, industrial customers commonly communicate their intention to move in to their service area or to increase load, but they rarely report an impending termination of operations or downsizing.

5. Economic forecasts and revisions to load growth forecasts

5.1. Economic forecasts

Long term energy modeling, including load forecasting, is unavoidably subject to considerable uncertainty. As noted above, a key issue for the present analysis is the U. S. national recession that began in 2008. Although the macroeconomic business cycle is an established phenomenon, predicting the timing and magnitudes of economic downturns remains an inexact process, and moreover the magnitude and duration of the recession that began in 2008 are widely (though perhaps not universally) recognized to have been unusually severe.

However, despite the *ex-ante* unpredictability of the exact macroeconomic details, in the case of load forecasting it might be expected that the frequent revision and updating within the LSE's ongoing IRP processes would have served to progressively reduce forecast errors by accounting for dramatically reduced economic activity and its effects on electricity use (along with other influences on load growth subsequent to the year the original forecasts were created).

To investigate this, we examine load forecasts in IRPs for certain years following those in which the above-discussed forecasts were made. We call these IRPs "intermediate" in the context of this analysis since they were produced between the "older" and "recent" IRPs employed throughout this study. Dates for intermediate IRPs can be found in Table 1. Intermediate IRPs were available for all but two of the utilities in our sample. These plans reveal that the LSEs themselves devote varying levels of attention to retrospective examination, evaluation, and correction of their own load forecasts and forecast errors. In some cases, there is both considerable analysis of this type and also improvements in forecasting methods in order to obtain greater accuracy. In others, while forecasts are updated, there is little or no retrospective discussion in the documents we examined.

In those cases in which forecast errors are discussed *ex post*, the LSEs highlight reduced economic activity as the key factor for previous overestimation of load growth along with resulting reductions in growth rates of population and the numbers of customers. In some cases, the underestimation of the effects of demand side management programs to promote energy efficiency are also cited as reducing growth more than had been anticipated. As we note previously, the available documentation is not sufficient to replicate the load forecasts and fully determine the quantitative importance of different inputs.

[11] The intuition behind this is that industrial activity is highly correlated with economic performance, as reported in several of the IRP documents revised. In addition, Paul et al. (2009) and Ros (2015) report that industrial customers have the highest short and long run elasticities among electric customer classes.

However, in the plans that cite these demand side effects, they are reported as significantly secondary to those of reduced economic growth.

5.2. Revisions of load growth rates in subsequent forecasts

One particular interest is the extent to which forecast errors are reduced during the planning periods corresponding to our forecasts to actuals comparisons. Among other reasons, this in turn facilitates comparison of load forecasts with capacity expansion decisions over these periods. As discussed above, we are focusing on the years up to and including 2014 since that is the most recent year for which estimates of actual load are available from the EIA. Thus, consider a load forecast made in 2005 that extends to 2014 or beyond. Although the forecast may, in retrospect, embody non negligible errors over the 2005–2014 horizon, updated forecasts made after 2005 might have reduced these errors and thus mitigated their potential impact on capacity expansion. Economic factors contributed significantly to the lower load growth than was forecast in the older IRPs. Moreover, significant economic recovery, and therefore a possible return to higher load growth rates, was not forthcoming for a number of years. However, review of the intermediate forecasts shows that most of the LSEs continued to expect some degree of economic recovery and predicted increased load growth rates (see Fig. 4).

For most LSEs their errors remain non negligible. Indeed, in most cases there is a sustained overestimation of load growth to 2014 even as the year in which the forecast was conducted approaches 2014. Specifically, actual load growth to 2014 was in most cases small or even negative as the forecast years approached 2014, but the forecasts themselves continue to project positive growth at rates that have turned out to be higher than actual rates and in some cases of the opposite sign (negative rather than positive).

Given the LSE's reliance on macroeconomic forecasting services such as Global Insight, these findings may reflect the fact that forecasters and economists did not anticipate the very slow and partial recovery from the recession that began in 2008. In any case, given the considerable apparent impact of economic factors on load growth over the period we are studying, these facts highlight the importance of economic and demographic forecasting in the IRP process, including the *relative* impact of these factors compared to others

This pattern of forecast errors highlights the importance of sensitivity analysis. In the following section, we explore load growth sensitivities reported in older plans to understand the methods and strategies they developed and planned for to deal with this inevitable uncertainty.

6. Load forecast sensitivities in resource planning

The pattern of actual load growth rates underestimation for most of the LSEs in our sample raises the question of whether there was a risk of building excessive capacity if expansion plans were not revised after the initial IRP was filed. This risk of acquiring more resources than needed – either by overbuilding capacity or through power purchase agreements – may translate to higher costs to consumers if these resources were actually procured and included in the rate base. This is a reason to analyze the low and high load sensitivities from older IRPs to understand whether and how utilities were required to analyze *ex ante* the implications of different forecast input assumptions. A summary of load sensitivity methods can be found in Table 6.

6.1. Review of load forecast sensitivity analysis

An analysis of older IRP plans reveals that at the time they were produced about half of the LSEs in our sample were not required to either perform sensitivity analysis, or examine changes in their projected preferred portfolios in light of alternative load growth assumptions. Across the sample, we find three general approaches to load forecast sensitivity analysis in older plans. First, some LSEs did not

Table 6
Summary of load forecast sensitivity methods by LSE.

LSE	Source of alternative forecast	Assessment method	Horizon	Results	Strategy	Change from older to recent IRP
Avista	Economic model; Statistical (Distribution)	Scenarios; Stochastic	Long term for energy, short term for peak demand	Capacity adjustment; timing and resource mix not changed.	React to new information	Quantitative instead of qualitative scenario analysis; improved load model.
COPSC	Statistical (Percentile)	No information	Long term for energy, short term for peak demand	No information	No information	None
Idaho	Statistical (Percentile)	Scenarios	Short term for peak demand	Capacity and timing adjustment	Procure small, flexible resources	Stochastic instead of scenario analysis
LADWP	Statistical (Percentile)	No information	Short term for peak demand	No information	No information	None
NV Power	No information	No information	No information	No information	No information	No information
NW	Market prices elasticity	Scenarios	Short term for peak demand	Operational cost reassessment	No information	Stochastic instead of qualitative scenario analysis
PacifiCorp	Statistical (Distribution)	Stochastic	Long term for energy, short term for peak demand	Operational cost reassessment	No information	Add scenario analysis.
PGE	Statistical (Percentile)	Scenarios; Stochastic	Long term for energy, short term for peak demand	Capacity and timing adjustment	Use market purchases/sales as buffer	None
PNM	Statistical (Percentile)	Scenarios	Short term for peak demand	No information	No information	Improved load model
Puget Sound	Economic model	Scenarios	Long term for energy.	Capacity adjustment; timing and resource mix not changed.	No information	Only additional scenarios
Seattle	Economic model	Scenarios; Stochastic	Long term for energy	No information	No information	Improved load model
Sierra Pacific	Economic model	Scenarios	Long term for peak demand	Capacity and timing adjustment	No information	None

Table 7
Average annual growth rate for actual and forecast energy consumption, with sensitivities.

LSE	Energy AAGR (%)			
	Low forecast	Base forecast	High forecast	Observed
Avista	0.3%	1.7%	2.9%	− 0.1%
COPSC	1.6%	1.8%	2.0%	− 0.4%
Idaho	1.5%	1.7%	2.3%	− 0.1%
LADWP	–	0.6%	–	0.0%
NV Power	–	2.3%	–	0.1%
NW	− 1.7%	0.6%	1.9%	1.2%
PGE	1.2%	2.6%	3.1%	0.2%
PNM	–	2.2%	–	− 1.4%
PacifiCorp	1.1%	1.9%	2.1%	1.3%
Puget Sound	1.2%	1.7%	2.3%	− 0.2%
Seattle	0.3%	1.1%	1.9%	0.2%
Sierra Pacific	− 0.2%	1.4%	2.5%	− 0.9%

Table 8
Average annual growth rate for actual and forecast peak demand, with sensitivities.

LSE	Peak demand AAGR (%)			
	Low forecast	Base forecast	High forecast	Observed
Avista	0.3%	1.8%	2.9%	0.4%
COPSC	1.9%	2.1%	2.5%	− 0.5%
Idaho	1.5%	1.7%	2.3%	0.4%
LADWP	–	0.3%	1.1%	1.8%
NVPower	–	2.4%	–	− 0.1%
NW	–	NA	–	4.1%
PGE	1.3%	1.9%	2.9%	0.8%
PNM	–	1.9%	2.4%	− 0.8%
PacifiCorp	–	1.3%	–	1.3%
Puget Sound	0.9%	1.1%	1.8%	0.8%
Seattle	–	1.7%	–	1.2%
Sierra Pacific	− 0.8%	1.7%	2.8%	3.4%

perform any sensitivity analyses, even when estimating alternative load forecasts. Second, some LSEs performed the analysis, but did not produce an alternative portfolio. Third, some LSEs analyzed the effects of alternative load forecasts on their preferred resource portfolios. The difference between the second and third approaches is that the second holds investments as fixed to test the impact of load deviation on operational costs/savings in their portfolios to verify that their preferred portfolio remained as the least-cost solution. The third approach, in contrast, produces an adapted portfolio that can be the basis of an adjustment strategy to alternative load conditions. In all cases, most LSEs used percentiles or deviations from the base forecast as their sensitivity metric.

In most of the cases in which preferred resource portfolios were reassessed in response to load forecast sensitivity analysis, the result was a substantial change in the projected timing and magnitude of required resource additions. As part of IRP, LSEs are required to develop and compare several resource portfolios to find the least cost and lowest risk (i.e., preferred) portfolio. We call this an “inter-portfolio” comparison, as it is performed keeping all other assumptions fixed. In addition, LSEs assess the revenue requirement effects from varying assumptions in key variables including load growth, natural gas prices, capital costs, etc. We call this an “inter-scenario” comparison. We find that in both older and recent IRP inter-scenario utility revenue requirement differences are much larger than inter-portfolio revenue requirement differences. In some cases, the inter-portfolio valuation difference was small enough that it could be statistically insignificant. In contrast, several LSEs reported adjustments up to ± 20–40% of capacity under low or high load conditions.

The development of sensitivity scenarios was not always accompanied by a strategy to deal with the effects of these uncertain outcomes. In the older IRPs, most LSEs did not report any type of analysis on the effects that alternative load growth scenarios would have on their planning outcomes. For those plans that did report these analyses, we identify two approaches to deal with this uncertainty: resource flexibility and market transactions. Resource flexibility refers to the procurement of smaller and quick-deployment supply or demand side technologies to adjust rapidly to new conditions (e.g. Idaho and Avista). LSEs report that they would expedite or defer deployment of these smaller and modular (flexible) resources in response to unexpected load conditions. Market transactions pertain to purchases or sales using non-firm transactions as a buffer for long term and structural adjustment due to unexpected customer load (e.g. PGE).[12] LSEs report using market transactions to sell their output to the market if load conditions were lower than anticipated and purchase if load was higher.

Both of these strategies have limitations. The focus on flexible resources restricts the types of technologies that would be deployed and reduces opportunities for larger capital intensive projects. The use of market transactions, as suggested by some LSEs, assumes that market purchases are always on the margin, which is not necessarily accurate in all cases. Also, national or global economic performance will jointly affect electricity market conditions as well as load growth. Economic downturn may create surplus on electricity markets due to load contraction and therefore make market purchases more attractive. The use of market purchases or sales as buffers may not recognize this strategy. Finally, relying on market purchases as a strategy for long term

[12] Other LSEs did mention in their IRPs market purchases as a hedging tool for short term supply-demand mismatches, but these market purchases are not discussed within the context of a load sensitivity analysis.

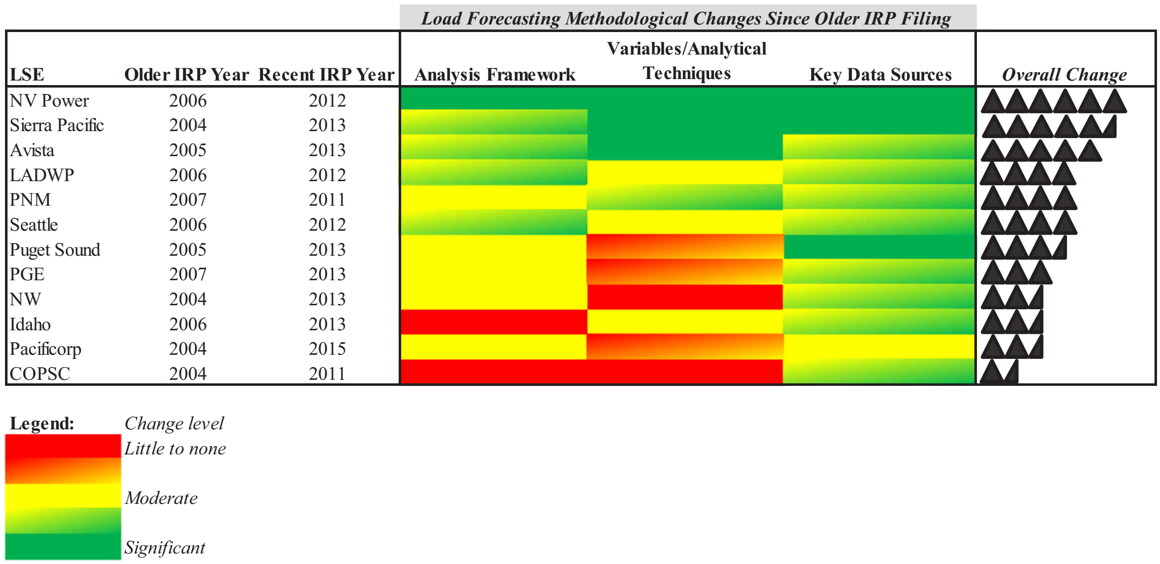

Fig. 5. Load forecasting methodological changes since earlier IRP filing.

adjustment implies coupling electricity price uncertainty with load growth uncertainty. This makes the entire strategy formulation much more complex.

6.2. Quantitative analysis of load sensitivities

In this section we discuss the base forecast, the range covered by the high and low load growth forecast estimates, and the actual load.[13]

Two LSEs, Northwestern and Sierra Pacific, developed very large envelopes, or spreads, around their base forecast that encompassed their actual retail energy sales and obligations (Figs. 2 and 3). All other LSEs, including those with a smaller forecast error, did not produce alternative forecasts that encompassed actual outcomes for energy sales. Most of the LSEs developed symmetrical and narrow forecast envelopes with a low AAGR forecast boundary that was significantly higher than the observed average annual growth rate for energy (see Tables 7 and 8). The preceding is an example of the challenges of producing alternative forecasts that can span a wider range of possible future outcomes. It also reflects the tradeoff between the span of alternative forecasts and the complexity of the strategies to address them: a larger span requires a more sophisticated sensitivity analysis and strategy development.

The results for the peak demand forecasts are different than the results for the energy forecasts. Observed energy consumption growth is generally less than anticipated, but peak demand growth exhibits mixed results with both over and underestimation of actual peak demand. In addition, in most cases the spread of the forecast envelope is wider for peak demand than for energy (e.g., COPSC, PGE, Puget Sound, and Seattle). This wider spread may reflect the simultaneous consideration of short term (e.g. weather) and long term (e.g. growth) uncertainty in the sensitivity analysis (energy sensitivity only considers long term). Sierra Pacific was the only utility whose forecast envelope consistently encompassed the observed load over time, but it was also the sensitivity with the largest spread (see Fig. 2).

[13] In the case of Pacificorp, which does not provide point estimates for its alternative load growth forecast but a distribution of values, we use the 10th and 90th percentiles as the low and high values, respectively. No alternative energy forecast information was reported for LADWP, NVPower, and PNM, and no alternative peak demand forecast were available for NVPower, NW, PacifiCorp, and Seattle.

7. Comparison between older and recent plan load forecast methodologies

Over time, electric LSEs often make adjustments to their load forecasting analysis frameworks such as the mix of customer classes evaluated and makeup of forecast scenarios; choices of variables; analytical techniques such as time-series regression and SAE models; and sources of key economic and demographic assumptions such as IHS Global Insight Inc., EPRI, and Moody's Analytics Inc. These changes are made in an effort to ultimately improve forecast accuracy in light of evolving market and regulatory conditions; development of novel analytical techniques; and access to more accurate forecast assumptions.

Older and newer forecasting methods, as documented in the utilities' IRPs, can be compared to determine the degree of change between filing dates as a possible response to forecast errors. Indeed, LSEs that had relatively large errors may have had the most incentive to make changes to their forecasting inputs and methods. Fig. 5 summarizes the extent of changes made for each of the LSEs considered in this study and the three categories related to forecast methodology described in Section 3.

Overall, nearly all of the LSEs considered in this study found new data sources for key modeling assumptions such as population or regional economic activity. Half of the LSEs made changes to all three components of their load forecasting methodology (analysis framework, technique, and source of data). Some LSEs made significant changes to load forecasting related variables and analytical techniques, but a larger share of LSEs made very small or no changes within this specific category. Most LSEs did not make significant changes to the analysis framework between filings. NV Power, Sierra Pacific, and Avista made the most significant methodological changes between plan filings. Colorado Public Service Corporation, PacifiCorp, and Idaho Power made the least number of changes.

7.1. Changes to analysis framework

Changes to load forecasting frameworks involved incorporating additional sets of load forecasts based on a wider range of growth scenarios (NV Power, PGE, LADWP, PacifiCorp, Seattle City Light) or changing the mix (or number) of customer classes considered in the analysis (Seattle City Light, LADWP, Avista, Puget Sound, NV Power). In some cases, LSEs assumed that future load growth was lower than the low growth rate reported in the earlier plan. Sierra Pacific and Puget

Sound Energy are two examples of LSEs that made changes to customer classes to reflect the importance of new types of electric energy services such as electric vehicles.

7.2. Changes to variables and analytical techniques

A number of changes to load forecast variables and analytical techniques involved migrating from one modeling technique to another. Both time series and cross sectional regressions have become the typical analytical framework to produce base case forecasts for energy and peak demand. For example, Sierra Pacific switched from an econometric and time-series based modeling approach to a SAE modeling approach. Conversely, Avista indicated that their load forecasting methodology is "undergoing significant restructuring [and] involves using an Auto Regressive Integrated Moving Average (ARIMA) technique" (i.e., time-series based econometric modeling). Other LSEs simply incorporated new variables including those used to capture adoption of electric vehicles (Idaho Power, LADWP, PNM, Seattle City Light, Avista, NV Power) or saturation of energy efficiency initiatives (PGE, NV Power, Idaho). Notwithstanding these various changes, however, LSEs continue to report that economic and population growth rates are the main drivers of their load forecasts.

7.3. Changes in sources of key exogenous assumptions

There has been a significant consolidation in the source of external data used in the production of LSE load forecasts. A number of LSEs used IHS Global Insight, Inc. in their earlier plans for demographic and regional economic growth estimates, and the majority of the LSEs now do so (COPSC, Sierra Pacific, PGE, NV Power, Avista, PacifiCorp, and Seattle). A smaller number of LSEs relied on Moody's Analytics, Inc., local/state/federal government agencies, or post-secondary educational institutions for regional demographic and economic assumptions. The Electric Power Research Institute (EPRI) and Itron, Inc. were consistent sources of assumptions about customer responses to prices and end use saturation and efficiency projections.

7.4. Changes in sensitivity and stochastic analysis

The recent IRP documentation indicates that, in contrast to their earlier approaches, most of the LSEs now develop comprehensive future scenarios that reflect the interactions of several different fundamental variables such as economic and population growth and alternative technology adoption, among others. The LSEs' methodologies have evolved to consider the risks due to uncertainty of certain key variables, including future customer load, and to analyze joint variation in such inputs rather than sensitivity analysis of single variables. Most of the LSEs that perform sensitivity or stochastic risk assessments also develop new portfolios that are different than their original and preferred base case. A number of the utilities use analytical techniques to measure how robust resource portfolios are to exogenous changes to these key variables. These analysis techniques are classified as scenario-based or sensitivity and probabilistic or stochastic risk assessments (see e.g. Wilkerson et al., 2014).

While the design of future scenarios remains challenging, adopting these approaches should provide a better basis for robust planning processes. However, there is still a general absence of methods to produce and follow up with clear strategies that respond to higher or lower realized load. In one of the few examples of regulatory implementation of adjustment strategies, the Utah Commission requires PacifiCorp to produce resource acquisition paths (UT PUC, 1992). These paths transparently lay out responses to specific potential outcomes of relevant variables in the planning process and act as an extension of the typical action plan included in most IRPs.

8. Summary and policy implications

We analyze load forecasting methods, performance and sensitivity analyses using a set of electric IRPs created by utilities across the Western U.S. A comparison of forecasts to actual energy use and peak demand reveals that all but one of the LSEs overestimated energy consumption growth over planning periods beginning in the mid-2000s and ending in 2014, and that eight of the eleven LSEs that forecast peak demand also overestimated this quantity, although to a lesser degree. In addition, we find that most of the LSEs that had the highest expected growth rates also experienced the lowest actual – in some cases negative – demand growth.

Furthermore, examination of forecasts from subsequent IRPs reveals that while the utilities did adjust their forecasts of load growth downward in response to much lower than expected demand growth, in most cases, there continued to be over-estimation in subsequent planning periods. For most LSEs, IRP documentation suggests that there was an expectation among macroeconomic forecasters that the national and regional economies would follow a historical pattern of relatively quick recovery from the recession. It was expected that load growth would also recover relatively quickly. The actual, slower-than-expected economic recovery thus contributed significantly to persistent overestimates of future load.

There is some correlation between the complexity of forecast methods and the accuracy of forecasts. LSEs with relatively more complex models generally had less forecast error than those that employed simpler models. However, among the more complex techniques, SAE models did not perform significantly better than other load forecasting methods and models. In addition, the LSEs that had the most accurate peak demand forecasts were also among the most conservative in terms of their expected peak demand growth. These results suggest that there may be relatively small marginal benefits from employing more complex models, but that there are other confounding variables that influence forecast error beyond model complexity.

There are structural reasons that may also explain the relative accuracy of load forecasts. For example, utilities with a larger share of industrial load in their mix generally had larger forecast error. We believe that this may be caused by the highly elastic and lumpy nature of industrial customer load as well as the difficulty in predicting entry and exit of industrial customers from an LSE service area. This suggests that industrial loads should be modeled and risk-assessed separately from other customers for a comprehensive evaluation of the potential impacts of losing these large customers.

For most of the LSEs in our sample, load forecast methodologies have evolved over the past fifteen years with respect to analysis frameworks, techniques, and data sources. We find that some LSEs improved their forecasting methodologies and achieved smaller forecast errors in more recent periods. This suggests an active effort to at least react to forecast error, although it remains to be seen whether these changes will lead to long term improvements in accuracy.

Load sensitivity analysis is an important component of risk assessment and management in IRP. In the context of our study, it is especially important because strategies derived from load sensitivity analysis may adjust and impact resource plans as new information comes in. Over time, LSEs have improved the breadth and sophistication of their sensitivity analysis of load forecasts. However, in most of the cases we study actual load growth exceeded or was less than the sensitivity bounds of what was considered possible in the original IRP analyses. This is concerning because both older and more recent IRPs generally lack an adaptive component that details how utilities would respond in practice if actual conditions more closely followed the sensitivity analyses rather than the conditions assumed in the base case analysis. More importantly, we find that the difference in revenue requirement across resource portfolios for a given IRP is much smaller than the cost impact of variation in load. This means that rates are much more likely to be impacted by changes in load rather than by resource selection. IRP devotes considerable effort in selecting and testing a preferred portfolio. It follows that the IRP and procurement processes should increase their focus on their strategic response to load forecast accuracy at a level similar to the effort involved in determining the preferred portfolio.

Long term load forecasts are a key input to utility IRP. Given the increasing importance of IRP as a locus, not just for utility operations, but also for implementation of energy and environmental policies and regulations, the forecasts have become a driver of the evolution of the electric power system. To give just one example, by determining the rate and magnitude of resource capacity additions, the forecasts help to determine the extent to which energy efficiency can be used to meet increasing demand, and therefore the need for new capacity. This in turn may affect the deployment rate of renewable generation and thus affect progress toward clean energy goals. In short, load forecasts—and their accuracy—have important implications for energy and environmental policy—and, in some cases, economic development; the role of long term load forecast accuracy in the development of a large hydroelectric project in British Columbia is one salient example.[14]

There is a disconnection between the strategic response to load forecasting sensitivity and the actual procurement of resources. First, processes involving the procurement of new resources may not be as open to public scrutiny as IRP. This means that adjustment strategies undertaken during the procurement phase may not fully reflect the interests of the public as expressed during the long term planning process. Second, the lack of integration between load forecasting sensitivity and subsequent adjustment strategies may reflect a general absence of information flows between the planning and procurement phase (see Carvallo et al., 2017). Considering the extent, depth, and complexity of risk analysis, regulators and utilities should revisit the benefits of including comprehensive and costly risk-based analysis within the long term planning process.

As discussed by Carvallo et al. (2017), the increasing number and complexity of the issues evaluated by utilities is leading to increased detail and complexity in modeling and other quantitative methods in the long term planning process. Against this background, it is noteworthy that the LSEs report that economic growth and demographic change assumptions were the dominant influences on long-term load forecasts and that load growth is generally the most important assumption in sensitivity analyses conducted by the utilities. More sophistication and complexity in load forecasting modeling and other analytical methods cannot improve the accuracy of macroeconomic or demographic forecast inputs, which are determined independently from the IRP process. Moreover, the latter are likely to remain a source of significant uncertainty in electricity planning, along with the introduction of unforeseen technologies, and energy and environmental policies.

An important implication of these findings is that both regulators and utilities should no longer assume that more complexity and detail in load forecasting necessarily constitutes improvement in load forecasting or other forms of modeling. To address these issues, regulators and utilities should undertake applied research to: identify and rigorously analyze and quantify the gains from greater model complexity both in load forecasting and in other IRP processes; determine specific planning functions that would benefit from the transparency and accountability of simpler models and tools; and develop, test, and implement these models and tools in the IRP process. Such a program of research could yield substantial benefits in expanding and improving the informational output, and increase the usefulness of integrated resource planning.

[14] The controversial “Site C” project on the Peace River has implications not just for the province's energy and environmental policies but also bears on employment and fiscal as well as on social policy (regarding the First Nations indigenous population). It has been argued that systematic overestimation of long term load was an important factor in justifying the project (Hendricks et al., 2017).

For jurisdictions around the world that may be considering IRP or are in the early stages of its development and implementation, a stronger conclusion can be drawn: careful consideration should be given as to whether to follow the example of complex and elaborate modeling. It may be the case that simpler and more transparent and accessible tools will not just suffice, but in fact be preferable for fulfilling integrated resource planning functions as energy systems and regulatory processes evolve.

Acknowledgments

The authors would like to thank Caitlin Callaghan and Matthew Rosenbaum (DOE OE) for their support of this project. The authors would also like to thank Galen Barbose, Joe Eto, Andrew Mills, and Lisa Schwartz (LBNL); Steve Johnson (Washington UTC); Phillip Popoff and Villamor Gramponia (Puget Sound Energy); James Gall and Grant Forsyth (Avista Corp.); Alison Lucas (Portland General Electric); Rakesh Batra and Caitlin Callaghan (DOE OE); and Aliza Wasserman (National Governors Association) for their thoughtful comments on earlier drafts. All remaining errors and omissions are the responsibility of the authors.

Funding sources

The work described in this article was funded by the National Electricity Delivery Division of the U.S. Department of Energy's Office of Electricity (OE) Delivery and Energy Reliability under Lawrence Berkeley National Laboratory Contract No. DE-AC02-05CH11231.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.enpol.2018.04.060.

References

ABB-Ventyx, 2016. Ventyx Velocity Suite Energy Mapping Software. ABB-Ventyx, Boulder, CO.
Avista, 2013. 2013 Electric Integrated Resource Plan. Avista Utilities Inc, Spokane, WA.
Avista, 2005. 2005 Electric Integrated Resource Plan. Avista Utilities Inc, Spokane, WA.
Carvallo, J.P., Sanstad, A.H., Larsen, P.H., 2017. Exploring the Relationship between Planning and Procurement in Western U.S. Electric Utilities. Lawrence Berkeley National Laboratory, Berkeley, CA.
COPSC, 2011. 2011 Electric Resource Plan, CPUC Docket No. 11A. Xcel Energy.
COPSC, 2004. 2003 Least-Cost Resource Plan. Public Service Company of Colorado, Denver, CO.
EIA, 2016. Electric power sales, revenue, and energy efficiency Form EIA-861 detailed data files [WWW Document]. US Energy Inf. Adm. URL ⟨https://www.eia.gov/electricity/data/eia861/⟩ (Accessed 18 March 2016).
Hendricks, R., Raphals, P., Bakker, K., 2017. Reassessing the Need for Site C, Program on Water Governance. University of British Columbia, Vancouver.
Hirst, E., Carney, J., 1978. Ornl Engineering-Economic Model of Residential Energy Use (No. ORNL/CON-24). Oak Ridge National Lab., TN (USA).
Hirst, E., Lin, W., Cope, J., 1977. Residential energy use model sensitive to demographic, economic, and technological factors. Q. Rev. Econ. Bus.
Hong, T., Shahidehpour, M., 2015. Load Forecasting: White Paper and Case Study. EISPC, NARUC, and US Department of Energy, Washington, DC.
Hyndman, R., 2006. Another look at forecast accuracy metrics for intermittent demand. Foresight Int. J. Appl. Forecast. 43–46.
Idaho, 2013. 2013 Integrated Resource Plan. Idaho Power, Boise, ID.
Idaho, 2006. 2006 Integrated Resource Plan. Idaho Power, Boise, ID.
LADWP, 2012. 2012 Power Integrated Resource Plan. Los Angeles Department of Water and Power, Los Angeles, CA.
LADWP, 2006. 2006 Integrated Resource Plan. City of Los Angeles - Department of Water and Power, Los Angeles, CA.
Mitchell, B.M., Park, R.E., Labrune, F., 1986. Projecting the Demand for Electricity (Product Page).
Nelson, C.R., Peck, S.C., 1985. The NERC Fan: A Retrospective Analysis of the NERC Summary Forecasts. J. Bus. Econ. Stat. 3, 179–187. http://dx.doi.org/10.2307/1391589.
NVPower, 2012. 2013 Integrated Resource Plan. Nevada Power Company, Las Vegas, NV.
NVPower, 2006. 2006 Integrated Resource Plan (2007–2026). Nevada Power Company, Las Vegas, NV.
NW, 2013. 2013 Electricity Supply Resource Procurement Plan. NorthWestern Energy, Bozeman, MT.
NW, 2004. Electric Default Supply Resource Procurement Plan. NorthWestern Energy,

Bozeman, MT.
Pacificorp, 2015. 2015 Integrated Resource Plan.
Pacificorp, 2005. 2004 Integrated Resource Plan. Pacificorp Inc, Portland, OR.
Paul, A.C., Myers, E.C., Palmer, K.L., 2009. A partial adjustment model of US electricity demand by region, season, and sector.
PGE, 2014. 2013 Integrated Resource Plan. Portland General Electric Company, Portland, OR.
PGE, 2007. 2007 Integrated Resource Plan. Portland General Electric Company, Portland, OR.
PNM, 2011. 2011–2030 Electric Integrated Resource Plan. Public Service Company of New Mexico, Albuquerque, NM.
PNM, 2007. 2007 Electric Resource Plan. Public Service Company of New Mexico, Albuquerque, NM.
PugetSound, 2013. 2013 Integrated Resource Plan. Puget Sound Energy, Bellevue, WA.
PugetSound, 2005. 2005 Least Cost Plan. Puget Sound Energy, Bellevue, WA.
Ros, A., 2015. An Econometric Assessment of Electricity Demand in the United States using Panel Data and the Impact of Retail Competition on Prices. NERA Economic Consulting, Boston, MA.
Sanstad, A.H., McMenamin, S., Sukenik, A., Barbose, G.L., Goldman, C.A., 2014. Modeling an aggressive energy-efficiency scenario in long-range load forecasting for electric power transmission planning. Appl. Energy 128, 265–276. http://dx.doi.org/10.1016/j.apenergy.2014.04.096.
Seattle, 2012. Integrated Resource Plan. Seattle City Light, Seattle, WA.
Seattle, 2006. Integrated Resource Plan. Seattle City Light, Seattle, WA.
SierraPacific, 2013. Triennial Integrated Resource Plan. Sierra Pacific Power Company, Las Vegas, NV.
SierraPacific, 2004. 2004 Integrated Resource Plan. Sierra Pacific Power Company, Las Vegas, NV.
UT PUC, 1992. Docket 90-2035-01 Report and Order on Standards and Guidelines.
Wilkerson, J., Larsen, P., Barbose, G., 2014. Survey of Western U.S. electric utility resource plans. Energy Policy 66, 90–103. http://dx.doi.org/10.1016/j.enpol.2013.11.029.
Willis, H.L., Northcote-Green, J.E.D., 1984. Comparison tests of fourteen distribution load forecasting methods. IEEE Trans. Power Appar. Syst. U.S. 6 (PAS-103).