Energy Storage Requirements for Achieving 50% Solar Photovoltaic Energy Penetration in California

Paul Denholm and Robert Margolis
National Renewable Energy Laboratory

Prepared under Task No. TM13.1514
2 Analysis Methods and Framework

To model the impact of high PV scenarios in California, we use the National Renewable Energy Laboratory’s Renewable Energy Flexibility (REFlex) model, a reduced-form dispatch model that calculates the supply/demand balance of an electricity system. REFlex performs a chronological dispatch of aggregated thermal and hydro units assuming generator flexibility limits, including ramp rates and minimum generation levels (Denholm and Hand 2011). It also performs chronological dispatch of demand response, energy storage, and vehicle charging to evaluate methods to improve utilization of variable generation resources (Denholm, Kuss and Margolis 2013).

As shown in Table 1 solar penetration in California in 2014 was about 6% (CEC 2014 and GTM and SEIA 2015).

Table 1. California Generation Mix in 2014

<table>
<thead>
<tr>
<th>Annual Generation in 2014</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Gigawatt-hours) GWh</td>
</tr>
<tr>
<td>Biomass</td>
<td>7,507</td>
</tr>
<tr>
<td>Concentrating solar power (CSP)</td>
<td>1,619</td>
</tr>
<tr>
<td>Fossil</td>
<td>151,037</td>
</tr>
<tr>
<td>Geothermal</td>
<td>13,030</td>
</tr>
<tr>
<td>Large Hydro</td>
<td>16,350</td>
</tr>
<tr>
<td>Nuclear</td>
<td>25,220</td>
</tr>
<tr>
<td>PV (Rooftop)</td>
<td>5,115</td>
</tr>
<tr>
<td>PV (Utility Scale)</td>
<td>10,932</td>
</tr>
<tr>
<td>Small hydro</td>
<td>2,787</td>
</tr>
<tr>
<td>Wind</td>
<td>23,913</td>
</tr>
<tr>
<td>Other (unspecified imports)</td>
<td>44,433</td>
</tr>
<tr>
<td>Total</td>
<td>301,943</td>
</tr>
</tbody>
</table>

Sources: GTM and SEIA 2015 for rooftop PV, and CEC 2014 for all other technologies. Imports are included in the respective generator category as described in CEC (2014).

3 The National Renewable Energy Laboratory developed REFlex to allow for analysis of large numbers of scenarios with fast solution times but a reasonable level of operational detail. The model allows rapid exploration of the surface space of operations with large amounts of renewables, but it considers only a limited number of aspects of system operation, including supply/demand coincidence, generator flexibility constraints, and the potential interface limits between California and surrounding states. This allows for the 450 yearly simulations described in the results section to be completed in a few hours. Alternatively, a detailed yearly simulation of the California/ Western Electricity Coordinating Council power system with the PLEXOS production cost model requires about seven days, which correspond to about 8.6 years of simulation time on a single machine.

4 This estimate is based on 12.6 terawatt-hours (TWh) of utility-scale (PV + CSP) generation (CEC 2014) and 5.1 TWh from rooftop PV (GTM 2015).
To this base mix of generators, we add PV to identify the increasing challenges and the role of storage as PV penetration in California approaches 50%. Hourly load, wind, and solar data are derived from the Low Carbon Grid Study (LCGS), which provides time-synchronized data for consistent modeling (Brinkman et al. 2016). As with LCGS, we use a base year of 2030, where we assume load grows to about 320 terawatt-hours (TWh), with a peak demand of 64.7 GW, which is about 6% higher than 2014 (CEC 2015). Load profiles are scaled from 2006 loads (i.e., weather and demand patterns are assumed to be the same in the future as they were in 2006). We assume wind generation grows to 35 TWh (or 11% of total demand before the addition of electric vehicle load) and concentrating solar power (CSP) generation grows to about 3.5 TWh (1% of total demand) based on existing capacity in 2016. We count the contribution of existing CSP toward our PV target; this is because the vast majority of CSP in California does not have thermal storage, so its output profiles are similar to PV’s. We fix the amount of annual generation from all other renewable resources at 2014 levels. We then model a series of cases with different grid flexibility assumptions and PV penetrations. Our approach is thus different from most of the studies cited above in that we do not model large mixed renewable portfolios to meet California’s renewable portfolio standard (RPS); in contrast, our objective is to isolate the operational impact of solar and its interaction with some mitigating policies and technologies, most notably energy storage.

PV profiles were based on a mix of 246 locations throughout California and surrounding states, with 92% of the capacity deployed in California. We also assume about 60% of the PV is utility-scale (with a 60%/40% split between tracking and fixed-tilt systems) and 40% is rooftop systems. Figure 1 shows the distribution of the solar capacity. Solar (and wind) profiles are based on 2006 meteorology to match the 2006 load profiles, and they are derived from LCGS (Brinkman et al. 2016).

---

5 We would expect load shapes to change over time as new appliances and devices are introduced, lighting becomes more efficient, and other changes occur. However, we are unaware of a data set that projects the overall change in load shape in California over time.

6 This value is based on scaling the wind values from the 2014 Long-Term Procurement Plan 40% RPS scenario by 25% (Liu 2014). For comparison, in 2014, California met about 8% of the total demand from about 24 TWh of wind generation (both in state and imported) (CEC 2014). Our assumption of 35 TWh results in a wind penetration of about 11% (before the addition of EV load) compared to our assumed demand of 320 TWh. This requires the addition of about 3.3 GW with an average capacity factor of 35%. For our simulations we use hourly generation data from the LCGS study (Brinkman et al. 2016) which assumes a mix of in-state and out-of-state wind generation.
The challenge of integrating large quantities of PV in the California system is illustrated in Figure 2, which shows the simulated net loads (normal load minus contribution from solar), assuming no PV curtailment is required even at very high levels of PV penetration for two 2-day periods (in spring and summer). In these examples, the increasing penetration values represent potential annual PV generation without curtailment.
Figure 2. Load and theoretical net load profiles for California during two days in the spring (a) and summer (b) where PV provides up to 50% of annual electricity, assuming no PV curtailment is required.

All data in the analysis have been shifted to Pacific Standard Time.
Figure 2 shows extreme changes in net load that are well beyond what can be accommodated in the current power system and fall below zero many days of the year at the highest levels of PV penetration. The overall ability of the grid to accommodate highly variable demand is based on its flexibility, which is driven by the ability of individual generators to change output to serve variations in electricity demand (Ela et al. 2014).

To model the impact of varying grid flexibility, we start with a base 2030 scenario and a conservative set of grid operational assumptions, including a system-wide minimum generation level of 15 GW. This limit is based on estimates of non-dispatchable capacity in California from the California Independent System Operator (CAISO) in 2014 (Bouillon 2014), but it assumes the Diablo Canyon nuclear power plant is retired before 2030 (PGE 2016). Because these estimates only apply to generators operating within the CAISO balancing area, we add an additional 2,000 MW of non-dispatchable capacity in the non-CAISO areas. This minimum generation limit is illustrated in Figure 3. Non-dispatchable capacity reflects the limited ability of generators to vary output, and it is based on a host of technical, economic, and institutional factors (Lew et al. 2015). Most thermal plants have a maximum ramp rate (the speed at which output can be changed) and ramp range (the difference between minimum and maximum output). The ramp range is limited by the plant’s minimum stable operating point, below which the plant cannot run. Because certain plant types cannot be started and stopped quickly, they are forced to remain operating at or above this minimum generation level. In addition, many generators have long-term contracts that create institutional limits to plant cycling.

In addition to the minimum generation constraint, the conservative base case also assumes zero exports of solar generation to surrounding states and no demand shifting. We assume that in the base case, the system has 4,427 MW of storage (3,100 MW of existing pumped storage plus 1,325 MW of storage being deployed under the California storage mandate) (CPUC 2013; Eichman et al. 2015), and the average roundtrip storage efficiency to be 80%. While the CPUC storage mandate requires a mix of transmission-, distribution-, and customer-sited storage (CPUC 2013), we also assume that all storage, including behind-the-meter and distributed storage, perform energy shifting, including storing otherwise curtailed PV generation. Overall, this case largely represents a business-as-usual scenario that assumes essentially no changes in the way the California grid is operated between now and 2030, with the exception of retiring Diablo Canyon and adding the California storage mandate. This example therefore represents a

---

7 In the 50% scenario, the net load is below zero for about 2,200 hours (25%) of the year.
8 This estimate was derived by scaling the thermal and hydro minimum generation levels in CAISO to be proportional to the state load. CAISO serves about 80% of the state’s annual electricity demand (CAISO 2015).
9 While many of these long-term contracts will be renegotiated by 2030, in order to develop a relatively conservative base case, we assume they persist in our base case. We relax this constraint in later scenarios.
10 The zero export assumption in the base case is based on the fact that California has historically been a net importer, and there are limited market mechanisms to allow California to sell excess renewable energy out of state. Further discussion is provided in Liu (2014a).
11 We assume existing storage has eight hours of storage capacity, which is actually less than much of the existing pumped hydro but more than the storage expected as part of the storage mandate. For a discussion of the storage mandate, including regional allocation of targets, size, and applications, see Eichman et al. (2015).
12 The CPUC storage mandate applies only to investor-owned utilities, so this value does not include any storage that may be developed by publicly owned utilities.
13 This latter assumption is used by the CPUC and CAISO when modeling future grid flexibility, which reflects the expectation that retail rates will be adjusted to reflect the changes in energy value (Eichman et al. 2015)
limiting case used in part to demonstrate the importance and impact of improved flexibility measures that can be implemented.

Figure 3 shows simulation results for two days in the spring (April 9 and 10) under a 20% annual solar scenario (requiring about 31 GW of PV capacity) and base assumptions described above. Figure 3a shows the overall system dispatch including additional detail about the source of the 15 GW minimum generation level, which uses estimates from Bouillon (2014). The black line shows the net load after the contribution from usable solar and wind. In the middle of each day, the amount of wind and solar exceeds what can be accommodated assuming the 15 GW minimum generation level, resulting in “overgeneration” where electricity supply exceeds demand requiring curtailment (CPUC 2015a).

Figure 3b shows the allocation of solar energy during this period of relatively low demand. Without existing storage, about 17% of potential PV generation would be curtailed. Most of the potentially curtailed energy is avoided because of the 4.4 GW of storage, which shifts this generation to the evening peak and reduces curtailment to about 5% on these two days (or about 7% when including storage losses). About 3% of the potential solar energy generation is curtailed during the entire year, which is about 5% when losses due to storage inefficiency are included. For comparison, simulations by the CAISO in a scenario with 18% PV and with similar assumptions of storage requirements and no exports (but including Diablo Canyon, which adds about 2,000 MW of non-dispatchable generation) estimated a 5.8% curtailment rate (CAISO 2014). Note that we assign all incremental curtailment to PV, as significant curtailment does not occur before the addition of PV, which means it is the added PV that results in curtailment. It is important to note that actual allocation of curtailment among different renewable generation sources (including existing or added generators) will be driven by various factors, including local grid conditions, underlying contractual agreements with suppliers, production tax credits, and other regulatory issues (Bird, Cochran, and Wang 2014).

---

14 Because of curtailment and storage losses, to derive 20% of the state’s energy from solar, a solar capacity potential equivalent to 21.4% of total energy would be needed, requiring about 2 GW of additional capacity.
Figure 3. System dispatch on April 9–10 in a scenario with 20% potential annual solar
Curtailment’s impact on PV economics can be measured as reduced PV value. Value declines due to curtailment because each unit of potential PV production no longer displaces one unit of fossil generation. As curtailment increases, the benefits of additional PV may drop to the point where additional installations are not worth the cost, creating an economic limit to deployment (Cochran et al. 2015). Previous studies have demonstrated the decline in PV value and how it can be partially mitigated with improved grid flexibility (including storage) (Mills and Wiser 2012a, 2015).

Curtailment’s impact on PV economics can also be measured as increased cost, here translated into “net LCOE,” which is defined as the cost of energy that can be used by the grid after considering curtailment. The net LCOE is calculated as follows:

\[
\text{net LCOE} = \frac{\text{base LCOE}}{1 - \text{curtailment rate}}.
\]

The base LCOE is defined using the “standard” method of calculating LCOE, which assumes all energy produced can be delivered to the grid.\(^\text{15}\) As curtailment increases, the net LCOE increases because the investment cost of the PV is divided over fewer units of useful energy. In the base 2030 scenario, with 20% solar, 4.6% of the solar energy is curtailed or lost in storage. Thus if the base PV LCOE in California is 6.0 cents per kilowatt-hour (kWh), this scenario would result in an average net LCOE of 6.3 cents/kWh. Potentially more important than the average curtailment rate and average net LCOE are the marginal curtailment rate and marginal net LCOE. The marginal net LCOE represents the cost of the final small amount of solar added, in this example the last unit of PV needed to achieve 20% solar. In this scenario, the marginal curtailment at 20% solar penetration is about 26% including storage losses. Thus, the net LCOE of the last unit of PV needed to achieve 20% solar is about 8.1 cents/kWh, assuming a base LCOE of 6 cents/kWh. Marginal curtailment rates can indicate the threshold at which PV becomes uncompetitive with alternative resources. If the addition of an alternative resource can provide the same amount of low-carbon energy for less than the marginal net LCOE of PV, then that alternative resource would likely be preferred.

Figure 4a shows the marginal and average curtailment rates of PV as a function of annual energy contribution from PV, using our 2030 base assumptions. Figure 4b translates this into the marginal and average net LCOE of PV assuming a base LCOE of 6 cents/kWh. As shown in the figure, under these assumptions the marginal curtailment rate increases rapidly once PV penetration rises above 20%. Reducing the base LCOE of PV would help, but the shape of the marginal curve in Figure 4b means even very low-cost PV would require additional grid-flexibility measures to achieve penetrations beyond 25%.

---

\(^{15}\) For a discussion of the calculation of LCOE, see Gilman et al. (2008).
When calculating curtailment rates, we do not distinguish between utility-scale and distributed PV. In reality, how curtailment is implemented will depend on a combination of technical and regulatory factors. Technically, curtailment requires having physical control over the operation of the plant. This type of control is readily available for utility-scale systems, but it will require advances in communications and controls to be applied effectively to distributed-generation PV. Substantial consideration by policymakers will also likely be required to establish appropriate rules, rate structures, and compensation mechanisms related to implementing curtailment more broadly in high variable renewable energy scenarios.
3 The Impact of Adding Non-Storage Grid-Flexibility Options

In the previous section, we demonstrated that achieving 50% PV in California is not economically feasible without substantial changes to the existing grid. Energy storage provides a potential solution, but several non-storage flexibility options are also available to system planners and operators (Cochran et al. 2014). Given the relatively high cost of energy storage (especially for long-duration batteries), we first explore the potential contribution of other lower-cost flexibility options. For example, improved forecasting and power system scheduling are likely less expensive than energy storage, and they have already been implemented or are being implemented in the United States and in countries with significant wind and PV penetration (FERC 2012; IEA 2014). CPUC (2015a) discusses many of these options in a California-specific context and describes many of the analytic needs to assess more fully the role of different flexibility options in integrating renewables at substantially increased penetration levels.

Here we specifically discuss four non-storage flexibility options that would impact the amount of storage required to achieve a 50% PV scenario: more flexible generation, exports to neighboring regions, demand response, and schedulable loads from electric vehicles. It should be noted that there are other sources of flexibility not considered here, such as adding concentrating solar power with integrated thermal energy storage. Sections 3.1–3.4 detail each of these flexibility options and their impact on PV curtailment and marginal costs. Section 4 integrates these options into a set of scenarios that evaluate energy storage requirements.

3.1 Flexible Generation and Reduced Minimum Generation Levels

Reducing the minimum generation from conventional generators is essential to integrating solar at levels beyond 20%. As discussed previously, minimum generation constraints are derived from a variety of technical, economic, and institutional factors. For California, we use estimates from Bouillon (2014), who identifies several17 categories of non-dispatchable plants, and we estimate the minimum generation level for each type (for a total of about 15,000 MW as illustrated in Figure 3):

1. Customer-owned cogeneration: ~6,000 MW
2. Geothermal: ~1,500 MW
3. Non-dispatchable imports: ~2,000 MW
4. Small hydro: ~1,500 MW
5. Thermal and hydro: ~2,000 MW
6. Non-Dispatchable plants outside CAISO: ~2,000 MW.

16 For example, analysis of CSP in the Western Interconnection found substantial flexibility benefits from both shifting loads via storage and reducing minimum generation constraints by providing very flexible capacity (Denholm and Mehos 2011)

17 Bouillon (2014) includes the Diablo Canyon nuclear power plant, which we assume is retired by 2030.
Determining the feasibility of gaining additional flexibility from these generators would require a plant-by-plant analysis of hundreds of plants in California and surrounding states. Because many constraints are likely contractual or based on non-technical factors, it is impossible to precisely identify the fleet’s minimum generation level. However, it is possible to identify types of generators that can or cannot provide additional flexibility.

First, although geothermal plants can vary output, reducing their output to accommodate PV does not provide significant economic or environmental benefits. Doing so simply replaces one carbon-free source for another, and thus, we do not consider them for reducing system minimum generation levels. Likewise, reducing hydro output to avoid PV curtailment only makes sense if the generation can be shifted to other periods. Many hydro plants have minimum streamflow requirements and so have technical limits on reducing output. However, this leaves over 8,000 MW of potentially dispatchable resources, including co-generation and thermal plants. Customer-owned cogeneration plants are typically used at industrial facilities to provide heat and electricity. They are normally considered non-dispatchable for various reasons including long-term contracts but also because reducing output or turning these plants off would reduce the production of process heat the facility needs. However, low-cost midday PV electricity might incentivize industrial facilities to reduce output, depending of the amount of process heat needed.

Several factors drive thermal plant minimum production levels, including start-stop constraints and the need to keep some plants online to provide reserves or grid-stability services (Hummon et al. 2013a). Reserves are needed to maintain frequency stability (the ability to maintain a constant frequency) and are often provided by partially loaded thermal generators (Kundur et al. 2004). While increased reserves may be needed with increasing PV because of the uncertainty of supply (CAISO 2013), improved variable generation integration practices can reduce these requirements. With more accurate solar forecasts, system operators can ensure adequate generation capacity is available without keeping excessive spare capacity online (spinning), which otherwise contributes to the minimum-generation problem. Furthermore, modern wind and solar plants can provide reserves and help regulate grid frequency, replacing the use of partially loaded thermal generation (GE 2014; Milligan et al. 2015). For PV, this requires operating below full output to provide the upward reserve (Gevorgian and O'Neill 2016). At low penetration, PV’s energy value greatly exceeds the value of reserves, so providing reserves via curtailment is not economic. However, at increased penetration, selective PV curtailment for reserves might be economic. These strategies—along with the replacement of old conventional generators with new, more flexible technologies—should enable lower minimum generation levels. New gas-fired generation typically can provide significant ramp rates and multiple operating reserves (Venkataraman et al. 2013). Certain new gas-fired combustion turbines and reciprocating engines can start very quickly, and these units could even provide some traditional synchronized (spinning) reserves without being online (Wärtsilä and Energy Exemplar 2014).

18 As solar forecasts are integrated into market operations (Tuohy et al. 2015), visibility and forecasts of all PV systems—including distributed rooftop PV—will be required (Wu et al. 2015).
19 Wind can provide inertial response by drawing stored energy from the spinning rotor. Because PV has no such inherent storage, any increase in output must be obtained by employing curtailment (GE 2014).
20 Provision of reserves from PV will require new mechanisms—market incentives, interconnection requirements, or other means—to ensure variable generation can reliably provide the services previously provided by thermal plants (Ela et al. 2014).
Given the uncertainties involved, we examine a range of minimum generation levels. Figure 5 shows curtailment as a function of PV penetration at three different minimum generation levels. The base case is 15 GW, as evaluated in the previous section. The 10 GW case could be achieved from a 50% reduction in minimum generation level from cogeneration plants and the elimination of “must take” contracts with out-of-state generation. The 7.5 GW case represents a further 50% reduction in minimum generation levels from thermal plants, and it represents a case where nearly all fossil generation in the state can be turned off during periods of high solar output. Both of these lower minimum generation scenarios substantially reduce the marginal curtailment rates.

Figure 5. Impact of system minimum generation level on curtailment with increasing PV penetration

Figure 6 shows the resulting marginal net LCOE values. Figure 6a assumes a base (pre-curtailment) LCOE of 6 cents/kWh, while Figure 6b assumes 3 cents/kWh. A conservative comparison for a future PV net LCOE would be 7 cents/kWh, which is equivalent to the variable cost of operating a combined-cycle gas generator with a fuel price of $6.2 per million British thermal unit (BTU), a heat rate of 7,500 BTU/kWh, and a social cost of carbon of $52/ton. The most aggressive flexibility case achieves this net LCOE at greater than 25% PV penetration. This result follows previous analysis demonstrating that such deployment is possible without huge additions of energy storage or other enabling technologies (Denholm et al. 2016). Increased flexibility and very low-cost PV could achieve penetrations greater than 30%. However, even in the least-expensive-PV and highest-flexibility scenario, the marginal net LCOE increases rapidly beyond 35% PV penetration, so additional measures likely are needed to enable such deployment.

---

21 We compare PV LCOE only to the variable cost of a combined-cycle generator because at high penetration of PV, the marginal capacity credit of PV (ability to offset conventional generation) is approximately zero (Denholm et al. 2016, Mills and Wiser 2012b). The price of natural gas is based on the EIA reference case projections for electric-sector natural gas in 2030 (EIA 2016). The social cost of carbon value is derived from EPA (2015).
3.2 Exports

Among the significant limitations of the base case is the inability to export energy from California to the rest of the western United States. California is highly connected with its neighbors, with transmission capacity to neighboring states that exceeds 10 GW (CAISO 2015). Expanding the geographical scope of the study beyond California could help achieve the 50% PV scenario by exploiting the spatial diversity of loads and resources. Greater regional interchange could also help address local minimum generation issues by exporting surplus PV generation to surrounding regions and thus reducing curtailment (Nelson and Wisland 2015; Eichman et al. 2015). Exports from California to surrounding regions are limited by physical transmission capacity, institutional and market barriers, and the potentially limited value of exported energy to surrounding regions under high solar penetration scenarios. Because California has historically never been a net exporter, and market mechanisms to allow exports have historically been limited, previous analysis has often limited exports. For example the CAISO LTPP model (40% RPS study) enforced a zero net export constraint (CAISO 2014). The E3 50% RPS study assumed up to 1,500 MW of exports in its base scenario, but evaluated a sensitivity case with up to 6,500 MW of export capacity (E3 2014). Alternatively, the LCGS study assumed new market mechanisms will allow even greater utilization of existing transmission, along with construction of several new lines throughout the Western Interconnection. As a result, the LCGS simulations allowed for exports from California exceeding 10 GW (Brinkman et al. 2016).

The existing western energy imbalance market is currently restricted to real-time energy transactions and is therefore unable to fully capture the opportunities for regional dispatch improvements. However, it is reasonable to evaluate scenarios with greatly enhanced coordination across the West. A key challenge will be evaluating impacts of greater regional adoption of PV throughout the entire Western Interconnection (in particular, in Arizona and Nevada), which represents a large fraction of the potential export capacity from California. Because we do not simulate the generation and demands of states outside California, we examine the sensitivity of PV curtailment and net LCOE to various export assumptions.
In allowing exports, a key question is whether or not exported PV from California can be counted toward meeting the percent generation from PV in California. California currently allows up to 10% of its RPS (after 2016) to be met with unbundled renewable energy credits (CPUC 2011). Given the uncertainty about the regulatory process and establishing which renewable generators might have already been allocated to this limit, we make the conservative assumption that exports from California to neighboring states do not count toward in-state generation. As a result, exports help integrate PV in California in a different manner than reduction of in-state minimum generation levels. For example, in a case with a 10 GW minimum generation level, installed solar capacity of about 34 GW produces an annual energy penetration of about 23.3% with a marginal curtailment of 17.6%. Reducing the minimum generation level by 2.5 GW (to 7.5 GW) with the same PV capacity increases PV penetration to 23.6% and reduces marginal curtailment to 9.6%. Alternatively, keeping the minimum generation level at 10 GW and adding 2.5 of export capacity does not increase the PV penetration, because none of the exported PV generation counts toward California generation. Doing so does however reduce curtailment of PV (to 17.6%, or about the same amount as the minimum generation case). As a result, adding exports shifts the curtailment curves, but each GW of export capacity is less effective than each GW of minimum generation reduction.

This effect is illustrated in Figure 7, which shows curtailment as a function of PV penetration for two cases of increased export capacity. The first case (red lines) adds 5 GW of export capacity to the 10 GW minimum generation scenario, while the second case (green lines) adds 10 GW of export capacity to the 7.5 GW minimum generation scenario. The 5 GW export line is somewhat similar to the 7.5 GW minimum generation case (reducing minimum generation by 2.5 GW), demonstrating the relative effectiveness of the two flexibility options.

![Figure 7. Increase in PV penetration resulting from increased exports](image-url)
3.3 Demand Response and Shiftable load

Low-cost PV will likely increase the use of demand response (DR) through rate structures or contracts that incentivize demand shifting to periods of high solar output. There are two general categories of DR mechanisms. First, price-based DR programs vary the electricity price to encourage changes in electricity use. These include time-of-use pricing (which assigns prices for different blocks of time) and real-time pricing (which varies rates in response to wholesale market prices) (FERC 2015). Second, incentive- or event-based DR programs compensate customers who allow program administrators to control certain electricity-consuming equipment directly and/or reduce their electricity demand upon request, such as during grid emergencies (Ma et al. 2013).

Demand response can help integrate PV and reduce curtailment in multiple ways, including by reducing dependence on partially loaded synchronous generators for frequency stability and operating reserves. Several U.S. regions already derive significant resource adequacy capacity and operating reserves from DR (FERC 2015). More important for avoiding PV curtailment is the use of DR to shift load to periods of high solar output. Additional discussion of the changes needed to retail rate structures and other regulatory reforms needed to increase consumer participation in DR programs is provided by CPUC (2015a) and Cappers, MacDonald, and Goldman (2013).

To provide context for the potential role of DR under high-PV scenarios, Figure 8 shows hourly curtailment in a scenario with enough PV to satisfy 50% of annual demand (requiring about 80 GW of PV capacity) with a 7.5 GW system minimum generation level and 5 GW of export capacity. Curtailments peak in the spring and are relatively limited from mid-June to mid-September, largely because this is the period of highest load.

![Figure 8. Hourly curtailment in a system with enough PV to achieve 50% PV with a 7.5 GW system minimum generation level and a 5 GW export capacity](image-url)
To model the opportunity for load shifting, we must quantify the ability to reduce demand in periods of low solar output and increase demand in periods of high solar output. To estimate load shifting potential in California, we use two assessments of DR potential in the western United States. The first, performed by Lawrence Berkeley National Laboratory (LBNL), estimates shifting of 13 classes of loads primarily in residential and commercial buildings, but it also includes certain industrial “non-manufacturing” loads such as municipal and agricultural water pumping loads (Olsen et al. 2013). The second, performed by Oak Ridge National Laboratory (ORNL), estimates shifting potential for certain industrial manufacturing applications (Starke, Alkadi, and Ma 2013). The two studies estimate the “projected” availability of DR using filters that estimate acceptability and controllability of various loads. In the base case assessments, using estimated “current” levels of consumer acceptability for demand response, the LBNL study identifies an average of 248 MW of potential load reduction in California with a peak of 964 MW, while the ORNL study identifies the potential for an average of 95 MW of load reduction with a peak of 145 MW.22

Both studies also provide a higher “technical potential” estimate for load reduction. The LBNL estimate averages about 2,100 MW of potential load reduction in each hour with a peak of about 8,500 MW. The ORNL study estimates an average of about 450 MW and a peak load reduction of about 650 MW. The LBNL and ORNL data were also used to estimate the hourly “headroom” or ability to increase demand.23 Of the hourly load reduction potential, less than half of the reduction in load is assumed to be able to be shifted over multiple hours. As a result, even using the full technical potential from these studies, the amount of shiftable loads during periods of high solar curtailment (in the spring) is relatively low. Most load shifting potential is assumed to be driven by heating and cooling demands, which are lowest in the spring.

Figure 9 shows the estimated hourly load reduction potential from the LBNL DR assessment (Olsen et al. 2013) for two days in the spring and summer. Each figure shows the total load as well as the load evaluated by the LBNL study. The large amount of load that was not evaluated includes residential lighting, manufacturing, and all “plug” loads including appliances, computers, etc. Of the load that was evaluated, only a relatively small amount is assumed to be shiftable, as indicated by the green area. Figure 10 adds the ORNL estimates and translates the results into the fraction of daily load that can be shifted. The top curve is the estimated hourly shiftable demand divided by the hourly load. The bottom curve applies the additional filter that considers the limits to multiple hours of shifting. Only a relatively small fraction of load can be shifted using the technical potential estimates, particularly during non-summer periods. Based on these assumptions, during days of high curtailment in the spring (March 1–June 1), on average only about 2% of demand is assumed to be shiftable.

---

22 The LBNL projected availability is similar to the existing demand resource availability in California. The three investor-owned utilities in the CAISO territory can shift up to about 1.3% of peak demand (as much as about 900 MW during periods of peak demand) via “price response” DR programs (CPUC 2015b).
23 Most DR research to date has focused on evaluating potential for load shedding, as opposed to increased loads, which will require new smart-grid technologies (e.g., programmable communicating thermostats) that will allow devices to increase load during periods of PV curtailment. Estimates for headroom are not provided by the LBNL or ORNL data sets; DR headroom values were derived by subtracting the actual load in each hour from the maximum annual load for each load type. The accuracy of this approach is unknown.
To estimate the curtailment reduction possible with these DR assumptions, we incorporate the energy-shifting and headroom estimates into the REFlex model following methods described in Hummon et al. (2013b) and assuming load is shifted with 100% efficiency.24

Figure 11 shows the marginal curtailment as a function of PV for the two minimum generation level and export levels evaluated in the previous section with and without DR. Adding DR shifts the curtailment curves by as much as about two percentage points. Thus, additional strategies are needed to shift supply/demand coincidence further and integrate 50% PV. A number of load-shifting opportunities are not analyzed here, including the large fraction of demand not evaluated

24 Compensation for load shifting will likely require new market mechanisms (Hogan and Paulos 2014).
in the LBNL and ORNL studies as well as new demand sources that could be flexible; these include electrification of thermal demand such as water heating and industrial processes.

Figure 11. Impact of added DR on marginal curtailment in scenarios with 20%–40% annual PV penetration with two minimum generation level and export capacity scenarios

3.4 Additional Load from Electric Vehicles

Electric vehicles (EVs), including plug-in hybrid EVs and battery EVs are a potentially significant new source of flexible load. Previous analysis has demonstrated that EVs can provide a significant source of load flexibility for integrating PV (Denholm, Kuss, and Margolis 2013). The ability of EVs to integrate PV and the amount of flexible load added by EVs will depend on several factors, including the:

1. Added electricity demand per EV, which includes each EV’s daily distance traveled and electrical efficiency (kWh/mile)
2. Load profile and controllability of EV charging
3. Total number of EVs.

Because of the uncertainty surrounding potential EV deployment, we simplify the analysis by assuming a fixed demand per vehicle of 12.1 kWh/day. This assumes the average EV will travel 57 km (35.4 miles) per day, and the fleet average efficiency is 0.21 kWh/km (0.34 kWh/mile) based on the mix of California vehicles and driving patterns in 2013 (DOT 2014) as well as estimates of EV efficiency (EPA 2016). We do not assume any “vehicle-to-grid” capability where vehicles can discharge the batteries to provide grid services (Tomić and Kempton 2007).

The load profile and charging controllability of EV are based on multiple factors, including vehicle driving patterns, charging availability, rate structures, and the size of the vehicle battery.
In our analysis, we explore the potential impact of three different charging strategies on PV curtailment:

1. “At home charging,” which assumes that vehicle charging only occurs at the end of the day when the vehicle is at home
2. “Opportunity charging,” which assumes that vehicle charging begins whenever the vehicle arrives at its destination, and assumes widespread availability of charging stations
3. “Optimized charging,” which assumes that vehicle charging uses as much PV output as possible that would otherwise be curtailed.

Figure 12 shows the profiles (derived from Brinkman et al. 2016) for the “opportunity” and “at home” scenarios (fraction of EV charging in each hour under the two scenarios). The figure also shows the average daily PV output (fraction of energy generated in each hour). The “at home” uncontrolled charging pattern would substantially decrease PV’s ability to meet load, because the additional load would occur almost entirely at night. The “opportunity” charging scenario is better for integrating PV, with about half of the demand occurring during periods of significant PV output. However, under this scenario, peak charging demand occurs in the early evening when PV output is declining rapidly. If charging could be optimized, a large fraction of the added EV demand could be shifted to periods of high PV output and use otherwise curtailed energy.

Figure 12. EV load profiles associated with uncontrolled charging and average PV output profile

To examine the impact of the size of the EV fleet and the amount of optimized charging, we added EV demand to the REFlex model. During each day, the model must provide 12.1 kWh of additional demand per vehicle, using a predetermined combination of the fixed profiles or

---

25 This pattern could, however, be useful for wind integration.
optimizing charging by timing demand to periods of lowest net load. We limit the instantaneous power consumption of the average EV to 6.5 kW, which assumes a mix of charging circuits.\textsuperscript{26}

Figure 13 shows the impact of different charging profiles on the load in a 10% EV scenario, in which 2.6 million light-duty vehicles are electrified; this figure assumes 100% of the charging would occur with one of the three profiles. The non-optimized charging profiles simply add the profiles from Figure 12 to the normal load. The optimized charging profile is the output from the model, optimized against a scenario with enough PV to meet 50% of total (normal plus EV) demand (without curtailment or storage losses).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure13.png}
\caption{Impact of charging on load in a scenario with 10\% EVs and three different charging profiles on April 1}
\end{figure}

Because of their different charging patterns, EVs can have varying impacts on PV integration and curtailment. Figure 14, which assumes a 7.5 GW minimum generation level and full DR availability, shows the impact of different EV charging patterns on PV curtailment. The “base” curve assumes no EVs. To illustrate bounding cases, we then add a 25% EV scenario (6.4 million electric vehicles)\textsuperscript{27} with the assumption that all vehicles charge with one of the three charging assumptions. The case with only at-home charging shifts the curtailment curve to the left, increasing PV curtailment and making it more difficult to integrate PV. As the fraction of optimized charging increases, curtailment rates drop, enabling higher PV penetration.

\textsuperscript{26} This mix is 40\% 120V/15A, 40\% 240 V/20A, and 20\% 480V/40A. This limit was not a binding constraint. The maximum demand in the 100\% optimized charging scenario was typically less than 4 kW/vehicle.

\textsuperscript{27} This results in up to 6.2 GW of added demand in the all-opportunity charging case, 8.9 GW in the all-optimized case, and 19.6 GW in the all-optimized charging case. In the optimized charging case, the additional load always occurs during periods of lowest net demand, so the added load, while significant, never increases the peak demand of the system. For a discussion of the relationship between PV generation, EV charging and peak demand patterns, see Denholm et al. (2013).
As Figure 14 illustrates, it is not possible to state definitely that EVs will help integrate PV and reduce storage requirements. However, the availability of low-cost midday PV electricity likely would incentivize optimally timed charging and widespread charging station availability. With a large fraction of optimized charging, the 25% EV scenario provides a substantial benefit. In the following section, we develop a set of scenarios combining EV charging, along with reduced minimum generation levels and increased DR to determine the total storage required to achieve 50% PV.
4 Energy Storage Scenarios

The quantity of energy storage needed to meet renewable energy or low-carbon policy goals will be a function of many parameters, including the amount of the four grid-flexibility options discussed above (as well as other options not considered here), desired PV penetration, and the acceptable curtailment level. To explore the range of options, we establish three scenarios combining the range of flexibility options shown in Table 2. The low-flexibility scenario requires only limited changes in fundamental grid flexibility, rate structures, or demand patterns. The biggest change in this scenario is a reduced minimum generation level, which would be accomplished largely by changing institutional barriers to ramping generators. The mid-flexibility scenario adds additional generator flexibility, exports, DR, and significant EV deployment. The high-flexibility scenario uses the full DR potential described in Section 3.2 and most of the physical transmission capacity that exists between California and surrounding states. It also represents significant electrification of the vehicle fleet (25% of all vehicles), with ubiquitous charging stations available to maximize midday charging. This additional load (from 6.4 million vehicles) is about 28.2 TWh, which would increase California electricity demand by about 9%.

<table>
<thead>
<tr>
<th>Table 2. Characteristics of Flexibility Scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum generation level (GW)</td>
</tr>
<tr>
<td>Export capacity (GW)</td>
</tr>
<tr>
<td>DR availability (GW peak/avg. daily GWh)</td>
</tr>
<tr>
<td>EV penetration (% of California light-duty vehicles)</td>
</tr>
<tr>
<td>EV charging profile (optimized-opportunity—at home)</td>
</tr>
</tbody>
</table>

*a These values represent the peak and average shiftable load during months of highest PV curtailment (March–May), with the high-flexibility scenario using the full LBNL technical potential, which assumes about 2% of the average daily demand is shiftable.

For each scenario, we calculate marginal curtailment and net LCOE for different levels of PV penetration with additional storage. We assume all additional storage has eight hours of usable capacity, which would allow the added storage to replace new peaking capacity.28 We also assume roundtrip efficiency is 80% (so 20% of energy placed into storage is lost) and is optimized by the system operator to provide maximum benefits to the system as a whole. This critical assumption would require optimization either (1) directly by a system operator in the case of utility-scale storage or (2) indirectly through real-time pricing or other mechanisms that would optimize behind-the-meter storage. This in turn would require new rules to optimize utility-scale and behind-the-meter storage (Koritarov et al. 2014; Sioshansi, Denholm, and Jenkin 2012).

---

28 By usable capacity, we mean the battery can discharge at rated capacity for eight hours, so a 1 MW battery would be able to store eight MWh of usable energy. For many battery types, this would require more than eight hours of physical capacity, as they are typically operated at less than 100% of full capacity to preserve battery life (Neubauer and Simpson 2015).
Market changes proposed or underway will address some, but not all, of the limitations in how storage can help integrate variable generation.

Figure 15 summarizes marginal curtailment as a function of PV penetration for each flexibility scenario. Each curve within the separate scenarios represents a different amount of storage, ranging from the base storage capacity of 4.4 GW to 30 GW. Figure 15a demonstrates that 30 GW of storage and low flexibility result in marginal curtailment exceeding 60% at 50% PV. Curtailment decreases substantially as flexibility and storage increase. Each curve shows a lower limit to marginal curtailment independent of the amount of storage added, because we count storage losses as curtailment. At lower levels of flexibility, more PV energy must be cycled through storage. In the low-flexibility scenario, for example, when PV reaches 30% penetration, about 60% of the incremental PV is placed into storage and about 40% is used directly. This means the marginal curtailment is 20% of 60%, or about 12%. As flexibility increases, less PV energy must be cycled through storage, which lowers this minimum curtailment level.

In the mid-flexibility scenario (Figure 15b), with 50% PV penetration and 30 GW of storage, the marginal curtailment rate drops below 40%. The high-flexibility scenario (Figure 15c) with 30 GW of storage achieves 50% PV at marginal curtailment approaching 20%.
The marginal net LCOE for PV can be calculated from these figures by multiplying the base LCOE by 1/(1-marginal curtailment rate), as discussed above. However, evaluating the amount of energy storage required to achieve 50% PV penetration requires comparing the marginal net LCOE with a target LCOE. Figure 16 translates the results from Figure 15 into marginal net LCOEs as a function of storage capacity for each flexibility scenario as well as two “base” PV costs: 6 cents/kWh and 3 cents/kWh. Figure 16 also shows a line at 7 cents/kWh, which we previously used as a target net LCOE (approximating the variable cost of a future combined-cycle gas turbines, including carbon costs). Achieving 50% PV appears to require a combination of medium and high flexibility scenarios.
of lower-cost PV, aggressive flexibility measures, and substantial storage. Achieving the 50% goal would require roughly 19 GW of storage under this case. For very low cost PV with a less flexible system, reaching 50% PV penetration could require 25–30 GW of storage.

![Graph](image-url)

**Figure 16. Marginal net LCOE as a function of energy storage capacity at 50% PV penetration for each flexibility scenario and two “base” PV costs: 6 cents/kWh and 3 cents/kWh**

Figure 17 examines the 50% PV case in detail, focusing on the relationship between base PV LCOE and storage requirements. In this figure, the amount of storage needed is examined as a function of base PV LCOE, using the target of 7 cents/kWh for the net LCOE. This figure illustrates that both grid flexibility and low-cost PV appear critical to reducing storage requirements.

![Graph](image-url)

**Figure 17. Energy storage required to achieve a marginal net PV LCOE of 7 cents/kWh as a function of base PV LCOE at 50% PV penetration and three levels of grid flexibility**
The energy storage values shown previously include existing storage, as well as storage expected by 2020 as part of the California storage mandate. Figure 18 shows the amount of additional storage that would be needed to achieve the 7 cents/kWh net-LCOE target. It includes the capacity needed for both 40% PV and 50% PV. The top bar is the high flexibility case with the low (3 cents/kWh) base cost for PV. With this highly flexible system, the storage that will be installed by 2020 is sufficient to support 40% PV. To achieve 50% PV would require about 15 GW of additional storage capacity to be built by 2030. We also consider cases where we change two of the more aggressive flexibility assumptions. The second bar reduces the EV penetration from 25% to 5% (or reaching a total EV fleet of 1.3 million vehicles in California by 2030). The third modifies the base case by increasing the base PV LCOE to 5 cents/kWh, which assumes only modest reductions in PV costs beyond those expected by 2020. Finally, we combine both of these reduced flexibility cases. As a result, if California can substantially increase grid operational flexibility, but not achieve either wide-scale deployment of EVs, or a substantially decreased PV cost, about 10 GW of new storage capacity would be required to achieve 40% PV, or about 28 GW of new storage to achieve 50% PV.

![Figure 18. Additional energy storage needed to achieve a marginal PV net LCOE of 7 cents/kWh for the high flexibility case and three reduced flexibility cases](image-url)

Based on these assumptions, a 50% PV scenario would achieve 66% to 68% from qualifying renewables depending on the EV penetration scenario as discussed in Section 4. Adding existing large hydro would result in greater than 70% generation from low-carbon sources in California. The rapidly increasing amount of storage needed at penetrations beyond 40% PV suggests the need to examine both the feasibility of large-scale energy storage deployment and the optimal mix of low-carbon generation resources. For example, the last unit of PV needed to achieve 50% in the high-flexibility case with 19 GW of storage has a curtailment rate of about 58%. Alternatively, replacing this last unit of PV with wind produces a marginal wind curtailment rate of about 20%, implying a more balanced mix of resources may produce a lower overall system cost.
References


This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.


This report is available at no cost from the National Renewable Energy Laboratory (NREL) at [www.nrel.gov/publications](http://www.nrel.gov/publications).


