GE
Energy Consulting

Grid of the Future: Quantification of Benefits from Flexible Energy Resources in Scenarios With Extra-High Penetration of Renewable Energy

ARPA-e Project Report

Jovan Bebić
Gene Hinkle
Slobodan Matić
William Schmitt
GE Energy Consulting

Ver. 1.2
January 15, 2015

imagination at work

Legal Notices

This report was prepared by General Electric International, Inc. (GE) as an account of work sponsored by Advanced Research Projects Agency–Energy (ARPA-E). Neither ARPA-E nor GE, nor any person acting on behalf of either:

1. Makes any warranty or representation, expressed or implied, with respect to the use of any information contained in this report, or that the use of any information, apparatus, method, or process disclosed in the report may not infringe privately owned rights.
2. Assumes any liabilities with respect to the use of or for damage resulting from the use of any information, apparatus, method, or process disclosed in this report.

Contact Information

This report was prepared by General Electric International, Inc.; acting through its Energy Consulting group based in Schenectady, NY, and submitted to Advanced Research Projects Agency–Energy. Technical and commercial questions and any correspondence concerning this document should be referred to:

Technical
Jovan Bebić
Managing Director
Energy Consulting
General Electric International, Inc.
Building 53, Room 312
Schenectady, New York 12345
Phone: (518) 385-1951
bebic@ge.com

Commercial
Denise Benoit
Contracts Manager
Energy Consulting
General Electric International, Inc.
Building 53, Room 311A
Schenectady, New York 12345
Phone: (518) 385-0901
Fax: (518) 385-9529
denise.benoit@ge.com

Acknowledgements

This research was supported by the grant from Advanced Research Projects Agency–Energy, under assistance agreement number: DE-AR0000510.

Executive Summary

The main objective of this study is to quantify the entitlement for system benefits attainable by pervasive application of flexible energy resources in scenarios with extra-high penetration of renewable energy.

The *quantified benefits* include savings in thermal energy and reduction of CO_2 emissions. Both are primarily a result of displacement of conventional thermal generation by renewable energy production, but there are secondary improvements that arise from lowering operating reserves, removing transmission constraints, and by partially removing energy-delivery losses due to energy production by distributed solar.

The *flexible energy resources* in the context of this study include energy storage and adjustable loads. The flexibility of both was constrained to a time horizon of one day. In case of energy storage this means that the state of charge is restored to the starting value at the end of each day, while for load this means that the daily energy consumed is maintained constant.

The *extra-high penetration* of renewable energy in the context of this study means the level of penetration resulting in significant number of hours where instantaneous power output from renewable resources added to the power output from baseload nuclear fleet surpasses the instantaneous power consumption by the load.

At extra-high levels of penetration, three major technical barriers stand in the way of renewable energy deployment.

1. The renewable energy resource, particularly wind, is not correlated with the load, which leads to concentration of renewable generation sites at significant distances from load centers and limits their power output to the available transmission system capacity connecting them to the load.

2. Because their fuel is free, renewable sources of energy are typically held at maximum available output. Consequently, they are not counted-on to provide up-reserves to stabilize the power-system frequency in case of unplanned outages of generation sources. This limits their total penetration to a level where non-renewable generation is able to provide sufficient up-reserves for the entire power system.

3. At extra-high levels of penetration, instantaneous renewable power output may surpass the system load, which makes curtailment inevitable unless there is flexibility available on the load side. The flexibility can be provided by energy storage or by scheduling system load to correlate its consumption with the availability of renewable energy.

This study evaluates the benefits associated with removing, or partially removing, these barriers to help prioritize development of technologies for addressing the barriers based on their impact or, alternatively, to set the upper-bound for cost of technologies based on their value to the power system.

The findings are as follows. Increasing renewable penetration from 30 to 50% at the footprint of the United States, reduces energy consumed by thermal generation fleet by $\sim$4 quads (quadrillion BTU) and CO_2 emissions by $\sim$340 Mtons. At 50% penetration, renewable energy curtailment due to the three constraints are significant. Removing the transmission-imposed limits is the necessary first step. Today's transmission infrastructure is simply not sized to evacuate this scale of energy from regions where the renewable resources are optimal. With the transmission limits completely removed, the curtailments of renewable energy due to the other two constraints were still significant. Only 79% of available renewable energy was actually utilized, the remaining 21% was curtailed. Assuming that the role of providing up-regulation can be covered by flexible energy resources and removing the corresponding part of renewable curtailments resulted in additional thermal energy savings of $\sim$2.8 quads and reduction of CO_2 emissions by $\sim$250 Mtons. Finally, load-scheduling to achieve better correlation with available renewable energy brought the cumulative reduction of consumed thermal energy to between 3.2 and 3.5 quads and reduction of CO_2 emissions to between 290 and 315Mtons.

The study also briefly reviewed the gaps between the existing and novel technologies needed to achieve the estimated benefits. It was found that a new control architecture is needed to overcome frequency stability limit by managing large number of dynamically responding energy resources. Advantageously, however, since the dynamic load response can be provided relative to inherently available signal of system frequency, the deployment of dynamically responsive energy resources is not contingent on demanding communications. The benefits of load scheduling are most dominantly affected by accurate renewable energy forecasting and require system operators to adjust unit-commitment to yield savings in cost of dispatch. Although implementing these adjustments and the corresponding market mechanisms and settlements are not trivial, they too are not contingent on demanding communications.

Revision Notes

Authors	Date	Rev	Notes
JZB	Jan. 2015	1.2	Incorporated comments from NWM
JZB	Dec. 2014	1.1	Editorial changes from GAH
JZB	Nov. 2014	1.0	Defining barriers to renewables penetration more explicitly and structuring the narrative around those definitions. Minor edits throughout.
SMM	Oct. 2014	0.8	First draft

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The main objective of this study is to evaluate the benefits of the widespread and flexible participation of distributed energy resources (DERs) and loads to the overall power system. The key metrics of interest are the improvements in system operating efficiency and emissions. Both are dominantly affected by the operating performance of central-station thermal generation: currently 87% of 4058 Bn kWh of electric energy in the US comes from central-station thermal generation fleet [2]. This is likely to remain true over the next two decades; the 2040 projections by the EIA forecast the total renewable energy output to be 16% in the reference case and 25% in the GHG-25 case [3], leaving the substantial share of total energy to the central-station thermal fleet. Furthermore, the average conversion efficiency of the US fleet is 32.5% for coal and nuclear, and 42.4% for gas [4] – a blend of combined cycle and simple cycle plants. For added perspective, the total transmission and distribution losses are 6% [5]. Therefore, in this study we set out to quantify the improvement in emissions and operating efficiency for the central-station generation fleet that can be obtained by flexibly managing distributed energy resources, including loads, pervasively.

Over the course of each day, the central-station fleet is committed (brought on-line) and dispatched (brought to a desired output) to follow the system load. The system load is forecast at multiple time horizons with great care and accuracy to allow for precise unit commitment and to ensure that the committed generation hour-by-hour has adequate margin to serve the load according to NERC reliability standards [6]. The forecasting process does not, however, consider load scheduling – it predicts the load based on the weather forecast and, where appropriate, the forecast of energy output from the available renewable generation.

Based on the forecast, the unit commitment and dispatch are performed in merit order (see A.3) [1], subject to constraints of the transmission system. The transmission constraints are determined in off-line studies and are seldom due to thermal limits of transmission circuits or components (e.g., transformers.) These limits are most often a consequence of dynamic constraints, such as adequate system recovery after disturbances, which are stipulated by NERC planning standards (TPL-001-0.1, TPL-002-0b, TPL-003-0b, TPL-004-0a). Notably, the NERC planning standards consider the load as fixed and unable to contribute to dynamic system recovery.

These practices give rise to two major areas for improvement in operation of the central-station fleet: the proactive shaping of load over all relevant time horizons, and the relaxation of transmission limits due to the potential ability of load and DERs to positively contribute to dynamic system recovery. The available improvements can be quantified (independently or in aggregate) by performing production simulations, preferably at a scale representative of an interconnected power system to minimize errors due to region-specific characteristics of central-station generation.

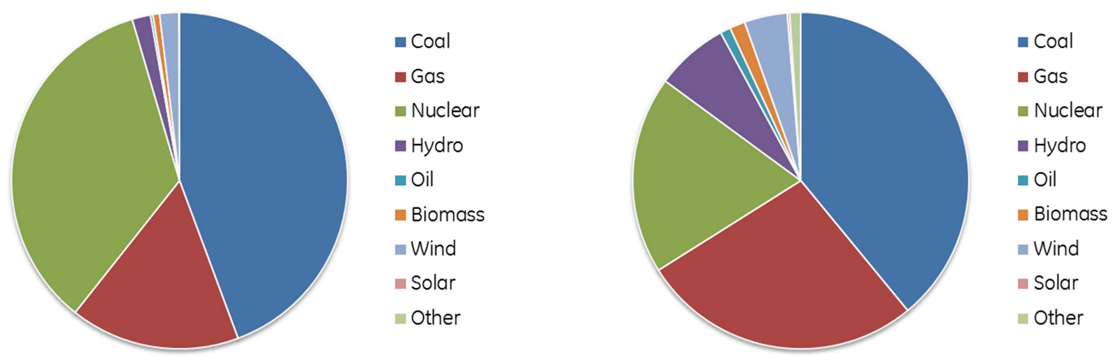
The first area of improvement can be quantified by shaping the temporal profiles of nodal loads and measuring the impact on operating performance of the central-station generation. In doing this, the energy of the load is maintained constant over a specified time horizon. This is an extension of the traditional demand response, which curtails the load to avoid using the most expensive peaking units. Making adjustments to load shapes over longer time horizons

[1] Short explanations of fundamental concepts from Power Engineering are provided in Appendix A.

affects commitment and dispatch beyond just peaking plants and leads to improved efficiency of the supply side fleet. In a more advanced scenario, consumers and central stations (both with advanced coordination control systems deployed) could adapt their operation to achieve system-wide energy efficiency targets.

The second area of improvement can be quantified by gradually relaxing known transmission limits and measuring the savings to total thermal energy consumed and total emissions of the central-station fleet. While it is widely recognized that active participation from the load and distributed energy resources can improve dynamic behavior of the system, such studies generally consider only the correlation between technology features and dynamic recovery of frequency or voltage. Recent work by GE Energy Consulting [7], goes further and connects the technology features of renewable energy with the operating performance of the central-station fleet. In this work we extended these insights beyond renewable energy and apply them to load and distributed energy resources. Clearly, a high-level study cannot analyze relaxing of the specific limits by specific technology features and quantifying the impact on central-station fleet. Instead, we measured the benefit to the system from relaxing the limits that could be effected by technology.

The benefits were measured relative to a baseline adopted from an earlier study by GE Energy Consulting [8], which explored technical and economic impacts of integration of 30% of renewable energy into PJM Interconnection. The earlier study assumed only traditional flexibility from the load (demand response) and was performed using production simulations in GE MAPS. All the analysis in this study was done on the model of PJM Interconnection because the baseline was well understood, the quality renewable-resource data was already developed, the scale of PJM is sufficiently large to minimize errors due to region-specific characteristics of central-station generation, and, finally, the generation mix of PJM is representative of the generation mix of the US, so the scaled differences between the base- and change-cases are indicative of the changes attainable at the scale of the entire US system.



(a) (b)

Figure 1.1: 2013 Energy production by fuel source: (a) PJM area [1], (b) entire US [2]

The last point is reinforced by Figure 1.1. As shown on the left of Figure 1.1 the 2013 distribution of energy produced by different types of generation in the PJM resembles that of the US.

Two use-cases for flexibility of energy resources were considered:

Dynamic response to help alleviate transfer limits imposed by dynamic stability limits of the transmission system, and reduce the up-reserves that would otherwise need to be provided by the thermal fleet.

Scheduling to increase utilization of renewable energy by aligning the flexible part of consumption with availability of renewable energy.

Taking a top-down approach and quantifying the resulting benefits enables the prioritization of control features necessary to effect changes with the highest impact, and the necessary degree of their adoption to achieve the desired scale of benefit. Thus, we start in Chapter 2 by defining the technical barriers to renewable integration at extra-high levels of penetrations. In Chapter 3 we describe the new tool developed by GE Energy Consulting that can model integration scenarios and the flexible load technologies that address these challenges. In

Chapter 4 we present the results of simulations that characterize the potential benefits of removing the penetration limits. We conclude with Chapter 5 by discussing the existing gap in technologies and proposing a few research paths forward.

Chapter 2

Technical Barriers to Renewables Integration at Extra-High Levels of Penetration

At extra-high levels of penetration, there are three major technical barriers that stand in the way of economical deployment of renewable energy by forcing curtailments.

1. The renewable energy resource, particularly wind, is not correlated with the load, which leads to concentration of renewable generation sites at significant distances from load centers and limits their power output to the available transmission system capacity connecting them to the load.

2. Because their fuel is free, renewable sources of energy are typically held at maximum available output. Consequently, they are not counted-on to provide up-reserves to stabilize the power-system frequency in case of unplanned outages of generation sources. This limits their total penetration to a level where non-renewable generation is able to provide sufficient up-reserves for the entire power system.

3. At extra-high levels of penetration, instantaneous renewable power output may surpass the system load, which makes curtailment inevitable unless there is flexibility available on the load side. The flexibility can be provided by energy storage or by scheduling system load to correlate its consumption with the availability of renewable energy.[1]

Our analysis shows that removing transmission constraints is the necessary first step in integration of renewables. Not removing transmission constraints forces development of renewable energy in locations with sub-optimal renewable resources, making their economics less favorable than developing at locations of optimal resource and covering the cost of transmission. As an illustration, the work in [8] shows that the amortized cost of transmission upgrades adds ∼$5/MWh relative to renewable energy value of ∼$80/MWh making it a worthwhile investment even at 30% penetration. Therefore, the analysis throughout this report assumes no transmission constraints.

The work presented in [7] studied frequency response with increasing levels of penetration of renewables. In particular, it studied the short-term response of the power-system frequency to large mismatches between generation and load, and showed how frequency nadir and settling frequency (see A.10) change in different scenarios. Based on the findings from that work, we will assume that this frequency stability barrier keeps instantaneous power output of renewable generation[2] limited to 50% of the load. In the rest of this report the scenario in which the output of renewables is curtailed at 50% of load the baseline case, as it represents the entitlement of today's technology.

The third barrier is associated with the instantaneous power of renewable generation plus the power from baseload generation is greater than the load. The US power system utilizes nuclear power plants to supply baseload and its desired operating condition is to hold the

[1]The flexibility is also helpful in dealing with demanding ramp-rates that are getting significant attention in California because of the high relative content of solar energy.

[2]sometimes referred to as system non-spinning penetration or SNSP

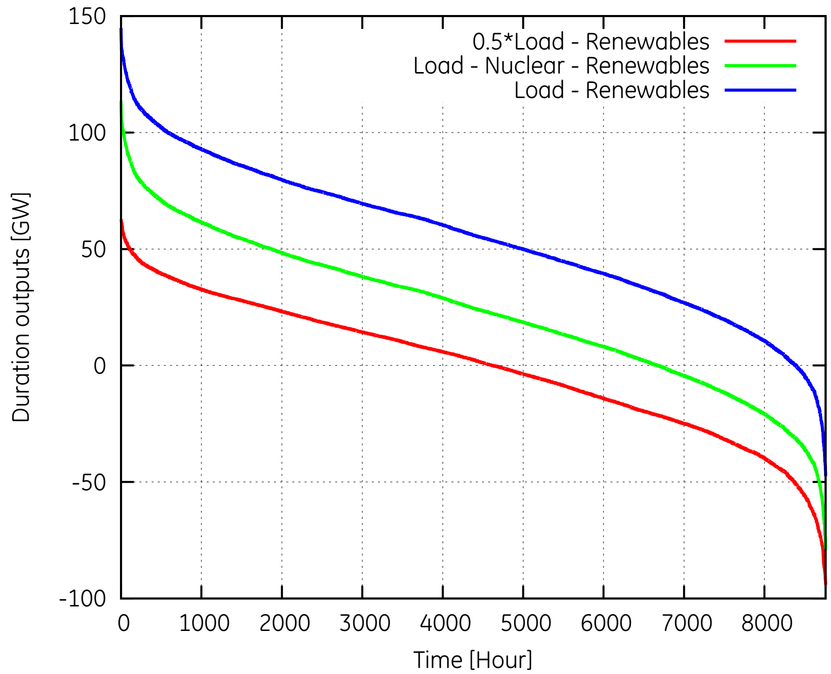


Figure 2.1: Load duration curves for PJM Interconnection

constant output. Consequently, the instantaneous output of renewables is limited by the difference between total system load and the total output of the nuclear fleet.

Installing sufficient capacity of renewable fleet to meet a large portion (e.g. 50%) of the system load energy results in many hours where the instantaneous output from renewables violates one or more limits mentioned above. This can be seen from the duration curves (see A.8) shown in Figure 2.1. The curves are shown using the data from the PJM Interconnection and assuming retirement of 10% of the 2014 nuclear fleet. The penetration level of renewables is %50 by energy, and the mix of renewables is such that 15% comes from distributed solar, 10% from central solar and 75% from wind plants.

The baseline case depicted by the red curve, shows the impact of "frequency stability limit" modeled here by limiting instantaneous renewable penetration to no more than 50% of the system load. Assuming installed renewable fleet capable of supplying 50% of load energy, the instantaneous output of renewable fleet surpasses the stability limit (50% of the load) in ∼4600 hours of the year. Thus, the area enclosed by the x-axis and the curve to the right of zero crossing represents energy of renewables that would have to be curtailed. Assuming that the limit of frequency stability defining the baseline case could be overcome by technology, the next binding constraint would come from combined output of renewables and baseload nuclear fleet becoming greater than the system load. This is depicted by the green curve showing ∼6600 operating hours without curtailment. Finally, if all of nuclear fleet was retired and all other parameters held the same, the curtailment would happen only in ∼350 hours, depicted by the blue curve.

By relying on the governor's response from the existing thermal generation, today's technology can reach the first limit. In other words, unlimited capacity of renewables can be installed, as long as their instantaneous output is curtailed to 50% of the load. One way to relax this limit is to force renewable generators to operate at lower-than-maximum output and thus enable their participation in providing up-reserves. This pre-emptive curtailment has a financial penalty in that it reduces the utilized renewable energy. An alternative is to provide frequency response by dynamically-responding energy storage and loads. The frequency events in the power system typically last less than 20 seconds, therefore partial participation of loads in response to system frequency could be fully automated and would not cause noticeable changes in most settings.

The other two constraints can be relaxed by load scheduling to correlate energy consumption with available renewable energy. While it is challenging to predict the percentage of system load that would be amenable to scheduling, the benefits of scheduling for any hypothesized percentage can be evaluated to glean first-order insights into this relationship. Scheduling is also helpful in dealing with demanding ramp-rates anticipated in systems with high relative content of solar energy. Implementing consideration of fleet ramping capability and constraining scheduling to accommodate it would, we felt, complicate the tool and interpretation of results so it was left out from this version of the tool.

In the next two chapters we review the benefits of relaxing the limits by two use-cases of load flexibility. This is followed by discussion of the most promising technology paths to get there in Chapter5.

Chapter 3

Modeling of Flexible Load and Storage

GE Energy Consulting has developed Flex, the new tool that estimates utilization of renewables and operating performance of conventional generation under target scenarios. The Flex tool cannot model the operation of the power system with the precision of advanced production simulation tools such as GE MAPS (Multiple Area Production Simulation Software.) However, due to its simplicity, Flex can efficiently perform sensitivity analyses that can help assess the benefits of different scenarios of integration of renewables. Moreover, Flex can model novel technologies that are not yet widely deployed: for instance, flexible loads with different scheduling algorithms or storage with different charge-discharge cycles.

3.1 The Flex Tool

Flex imports the data about conventional generation fleet, hourly profiles for available renewable energy and system load for the entire year under consideration, and parameters that describe the target level of renewables and technologies applied. For such a scenario the tool then computes hourly profiles of utilized renewable energy and dispatched conventional generation for that year. In addition, the tool provides a report file with total energy data summarized including the befits and costs of production (see A.1) and CO_2 emissions. The tool currently executes in a command window and produced files in the csv format. It also comes with the post-processing scripts that can generate plots as shown later in this section.

The set of parameters that characterizes the desired scenario consists of:

- Level of penetration of renewables, in terms of the utilized renewable energy with respect to system load (e.g. 50%).
- Mix of renewables. The renewable technologies assumed in the report are distributed solar, central solar and wind, and the mix parameter describes relative proportion of these three technologies (e.g. 40% distributed solar, 10% central and 50% wind).
- Flexible load technology and related settings. The following subsections describe a couple of technologies used to align the load with the renewables. This parameter specifies the technology under consideration or the algorithm applied, together with the necessary settings.
- Transmission losses. The distributed solar production, i.e. the production from photovoltaic installations collocated with the load, is not subject to transmission losses because the energy is used close to where it is being produced. The Flex tool uses this parameter to account for the avoided energy delivery losses. The typical value within the U.S. is 6% [5].
- Assumed retirements of nuclear fleet.

In the rest of the report we present the tool and the modeling approach on the data from the PJM Interconnection. As discussed earlier in the report, PJM is a representative example of the US fleet with respect to the fuel mix. In particular, Figure 3.1 shows a week of PJM input data: the system load (black curve) and the available renewable power (green curve).

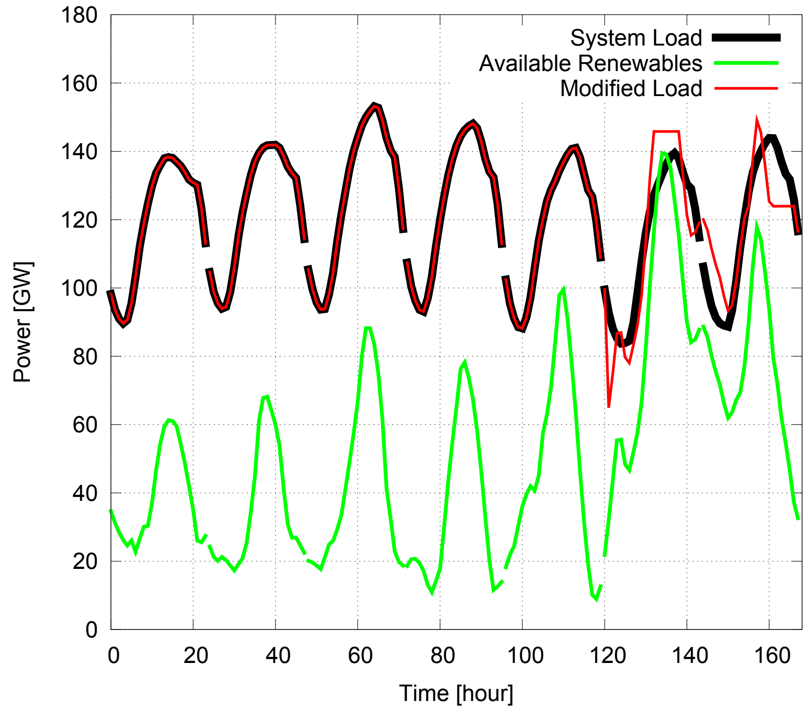


Figure 3.1: Flex tool input profiles

After discounting for the lower transmission losses due to the distributed solar, the first step that Flex does is scaling the profiles of different types of renewables with respect to the input parameters, namely level and mix of renewables. The Flex then computes the modified load (red curve) based on the specified technology and its settings. Note that towards the end of the considered week the renewable energy becomes larger than the load. As explained in the following sections, that is where the load gets to be modified and that is why there is a deviation of modified load compared to the original system load.

As a next step Flex computes the net load (see A.7), i.e. the difference between the system load and the amount of available renewable energy. More precisely, the net load is computed after the tool accounts for nuclear production since the assumption is that power produced by the nuclear fleet is used first. The net load is shown in the upper left plot of Figure 3.2. Note that towards the end of the week the net load is shown to be zero because at that time the available renewable generation was larger than the system load. The tool then simulates

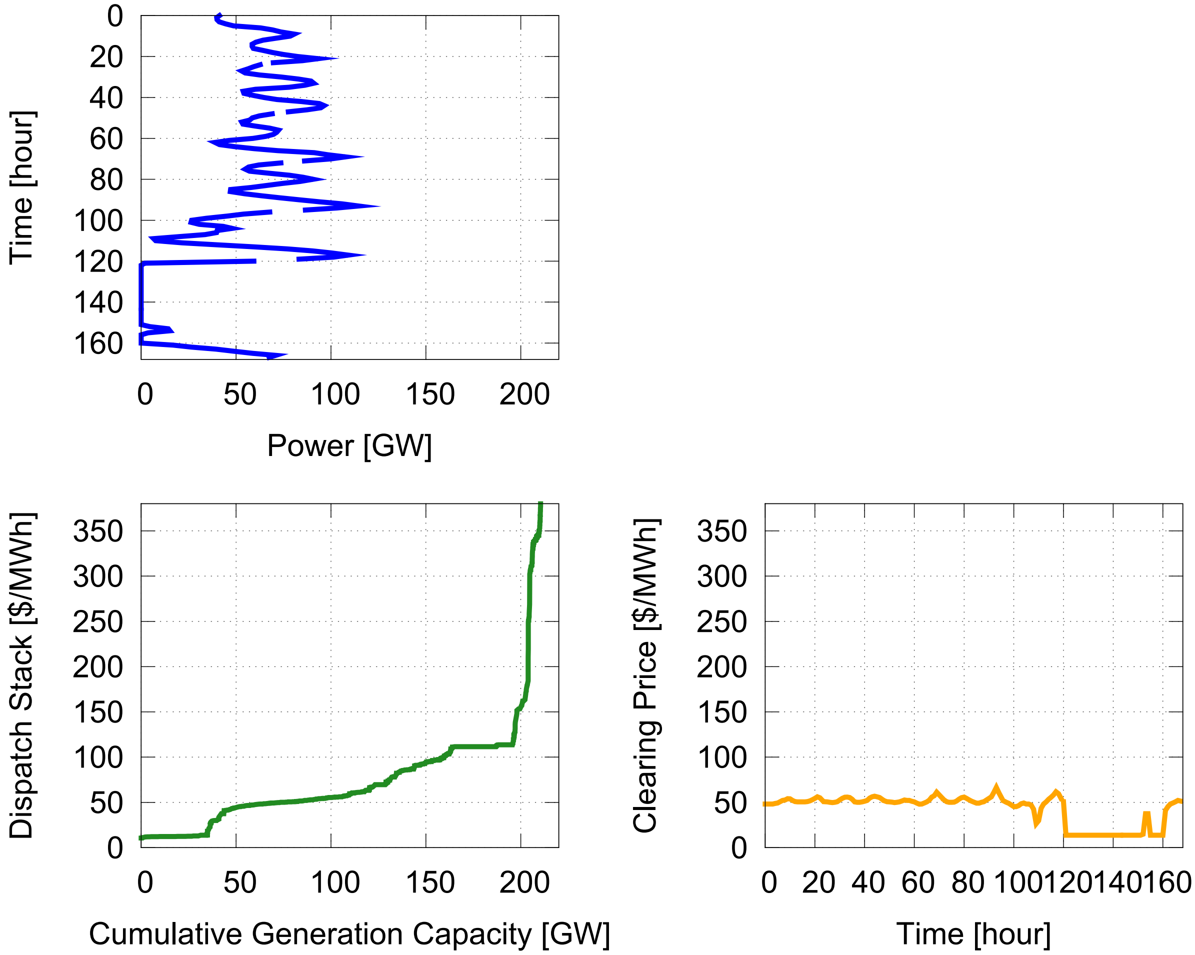


Figure 3.2: Flex tool processing flow

optimal dispatch, by considering for each hour the net load value and the dispatch stack (see A.3), green curve shown in lower left plot of Figure 3.2. It determines at what price of energy the market clears, i.e. it computes the profile of the clearing price (see A.2) as shown in the lower right plot of Figure 3.2.

The clearing price is determined by the mix of conventional generation since the units are dispatched in the order of increasing costs of their production of energy per produced MWh. In addition, the tool determines the mix of renewable energy first assuming that distributed solar is always utilized first, i.e. it cannot be curtailed. The mix between utilized central solar and utilized wind is then determined based on the relative proportions of the available central

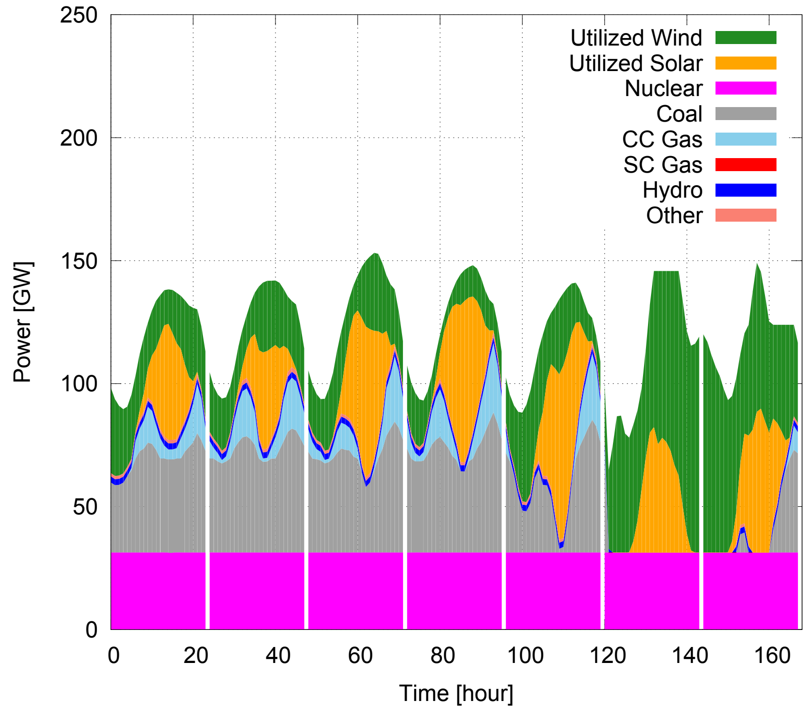


Figure 3.3: Generation mix output profile

solar and wind. Based on the computed activity of all units, conventional and renewable, Flex can output the plot of generation mix as in Figure 3.3. Note that for the week considered in this example the tool computes a significant amount of conventional generation used in most days of the week, except towards the end of the week where nuclear together with solar and wind is sufficient to produce the energy needed by the load.

Based on the determined mix of generation Flex can compute the fuel consumption and $C0_2$ emissions for each hour. Finally, the tool summarizes all the output data in the table format where the totals for the year are reported. Tab. 3.1 lists these numbers for all renewables as well as for each type of renewable generation individually. Tab. 3.2 does the same for conventional generation. For the definitions of reported quantities refer to the Appendix A, e.g. please see A.6 for capacity factor or A.4 for generation revenues.

Flex currently simulates two technologies, energy storage and flexible loads. The objective with both technologies is to use renewables as much as possible. In case of flexible loads

Table 3.1: Flex output report for renewables

Renewables Type	Available Energy [TWh]	Utilized Energy [TWh]	Percentage %	Capacity factor (available)	Capacity factor (utilized)
Total	485	463	95.592	0.290	0.277
Distributed solar	73	73	100.000	0.165	0.165
Wind	364	346	95.107	0.396	0.376
Central solar	48	45	92.616	0.155	0.143

Table 3.2: Flex output report for conventional generation

Conventional Type	Available Energy [TWh]	Delivered Energy [TWh]	Capacity factor (delivered)	Generation Revenues [$MM]	Production Costs [$MM]	Fuel Consumption [GBtu]	$C0_2$ emissions [Mton]
Total	1885	502	0.024	21817.255	12766.431	4903515.166	171.604
Nuclear	275	275	1.000	9979.509	3395.305	2914995.866	0.000
Coal	558	173	0.310	8967.386	7377.505	1674592.700	159.512
CCGT	334	32	0.095	1801.341	1656.224	226192.375	11.901
SCGT	566	0	0.000	2.863	2.734	319.980	0.017
Hydro	21	14	0.674	684.868	197.730	0.000	0.000
Other	89	8	0.088	381.289	136.932	87414.243	0.174

two store-release algorithms are implemented. How Flex simulates these technologies and algorithms is discussed in the following sections.

3.2 Algorithm A: Energy Storage

In this scenario the energy storage installed to support renewables is supposed to store energy when there is a surplus of renewable energy over load and release it when not. We assume that the total available storage is specified with the aggregated charge/discharge rate limit B (e.g. 30GW), which is given as the Flex input. The constraint in this scenario is that the state of charge of an equivalent energy storage device has to be maintained. In particular, it is assumed that charging and discharging cycles are such that the state of charge is kept the same at periodic instances of time. In the examples that follow we assume that such a period is one day.

Figure 3.4(a) shows how scheduling of charging and discharging is performed in this scenario. It shows the profiles of system load, available wind, available solar, available total renewable power (wind and all solar) and nuclear generation for one week of the year. Note that the power profiles are shown with the offset of nuclear production since the assumption is that nuclear power is always used first. According to the algorithm A, the equivalent energy storage is charged in hours when the total of nuclear and available renewable generation power is greater than the system load. The energy storage is charged with as much as the difference is, but not more than the specified charge/discharge rate limit. Similarly, the energy storage is discharged when nuclear plus available renewable generation power is smaller than the demanded system load.

Depending on the conditions on a particular day, i.e. between available renewable energy and system load, the total energy that energy storage can be charged with in a cycle period (e.g. one day) is not equal to the total energy that can be discharged. For instance, for the first day shown in Figure 3.4(a), i.e. for the time between hours 2280 and 2304, less energy can be charged than discharged. To maintain the constant state of charge, the scheduling algorithm has to select hours during which the energy storage should be discharged. For the second day, i.e. for the time between hours 2304 and 2328, the situation is opposite, i.e. more energy can be charged than discharged. In this case the scheduling algorithm has to select hours during which the energy storage should be charged.

The first objective of the algorithm is to utilize as much renewable energy as possible. To select which hours to charge or discharge the algorithm tries to satisfy the second objective: minimization of total costs of production. For instance, in case that discharging energy is larger than charging, the algorithm selects hours to discharge the energy storage by sorting net load power in each hour, and choosing the hours when net load is largest. This is done because it is likely that the clearing price of energy in such hours is high. Similarly, if charging energy is larger that discharging, the algorithm chooses to charge when net load is smallest

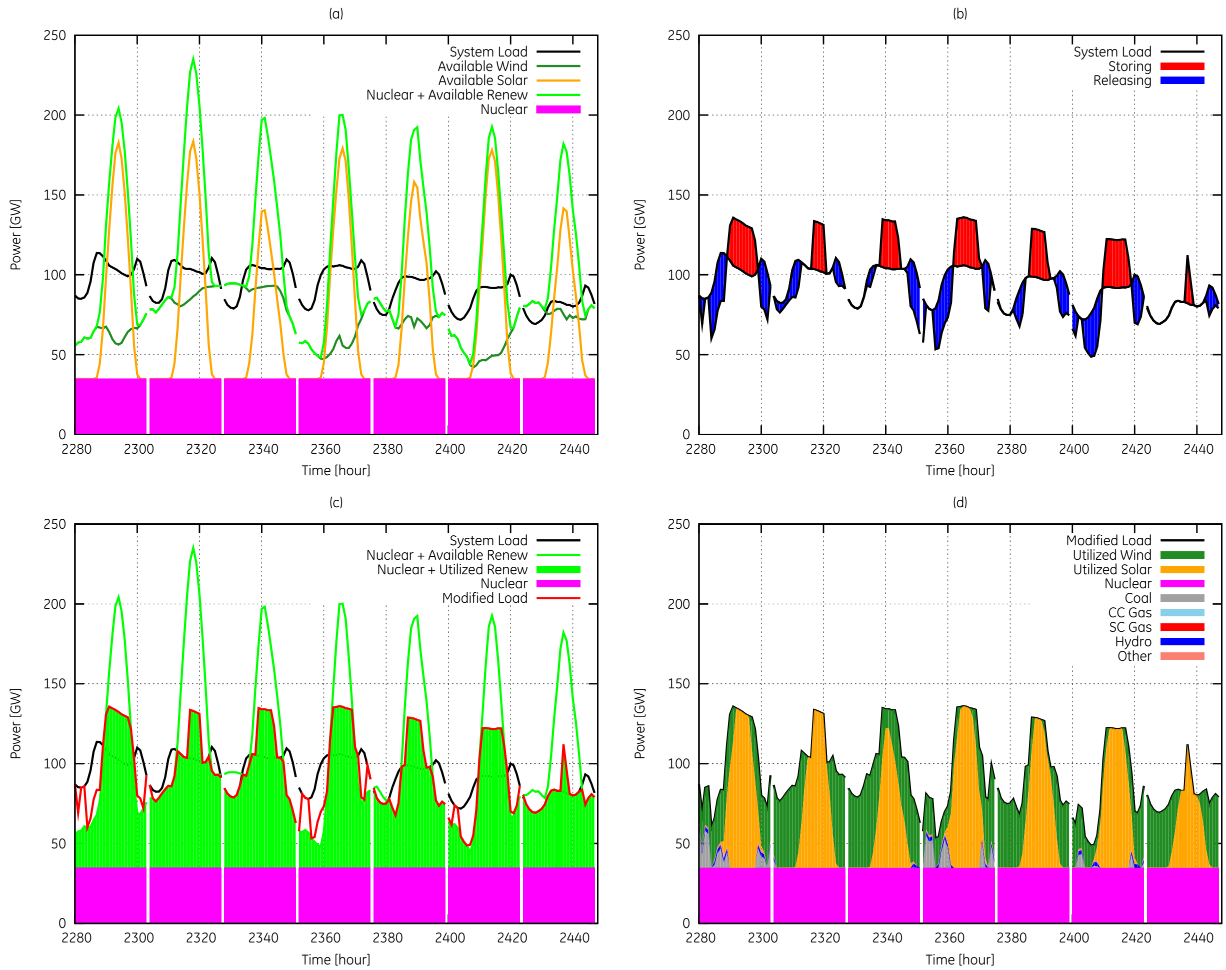


Figure 3.4: Algorithm A for r=50% ds/cs/w=40/10/50, week=14: (a) power profiles, (b) energy storage scheduling, (c) storing vs. releasing, (d) generation mix

because that is when energy price is likely to be low. The algorithm assumes existence of the perfect forecast for both the load and renewables. This is line with the premise of the study to evaluate the entitlements for savings.

This algorithm can symbolically be represented as follows:

1. if $L < R + N$ then $\delta L = min(B, R + N - L)$

   if $L > R + N$ then $\delta L = -min(B, L - R - N)$

2. if $E_C < E_D$ then discharge when $L - R - N$ maximal

if $E_C > E_D$ then charge when $L - R - N$ minimal

$M = L + sort\_select(\delta L)$

Figure 3.4(b) shows charging (red) and discharging (blue) hours for each day. One can see that during the first day discharging was not performed in all hours when it was possible just by considering the difference between the system load and available power of renewables. Analogous situation can be noticed for charging during the second day. Figure 3.4(c) shows utilized renewable energy as the area shaded light green. In one of its output tables Flex reports that in this case the total utilized renewable energy for the entire year is 453 TWh out of total available renewable energy of 485 TWh.

The plot in Figure 3.4(c) also shows the profile of the modified load (red), i.e. the original system load plus the power of charging/discharging. This is the load as seen by the ISO. When looking at the modified load profile for the first day, one can see that there exist some pockets between it (red curve) and the utilized renewable energy (green area). These amounts of energy have to be supplied by the conventional generation other than nuclear. This is evidenced in Figure 3.4(d) that shows the entire generation mix starting from nuclear at the bottom, other conventional generation in the middle and renewables at the top. Note that the selected week is week 14 of the year, i.e. the week during the late spring when the net load is usually smallest in a year. That is why the generation mix shows very little generation from conventional sources other than nuclear.

3.3 Algorithm B: Flexible Load

The algorithm B is used for one of the two scenarios with flexible loads, loads that can increase their energy consumption when there is a surplus of renewable energy over load and decrease it when not. It is assumed that the aggregate flexibility of all loads can be described by the flexibility limit input parameter F (e.g. 0.3): if the current system load is P, the control system can increase it up to $P*(1+F)$ or decrease it down to $P*(1-F)$. From the simulation purposes discussed in this section, one can think about flexible loads as about energy storage that have charging/discharging power limits variable with time and proportional with the load. The constraint in this scenario is that the total energy consumed in periodic intervals is not modified by the algorithm, i.e. the average system load remains the same. Again, it is assumed in the following examples that such a period is one day.

Similar to Figure 3.4(a), the flexible load scheduling for this scenario is shown in Figure 3.5(a). According to the algorithm B, the load is increased in hours when the total of nuclear and available renewable generation power is greater than the system load. The consumed power is increased as much as this difference is, but not more than what the load flexibility limit allows.

In order to satisfy the constraint of keeping the same average load, and after the load is increased for the amount allowed by the flexibility parameter, the algorithm B translates down the load profile for the amount needed in each periodic interval. The entire procedure can symbolically be summarized as:

1. if $L < R + N$ then $\delta L = min(F*L, R+N-L)$ else $\delta L = 0$

   $\bar{L} = L + \delta L$

2. decrease $\bar{L}$ uniformly to keep average

   $M = \bar{L} + avg(L)$ - $avg(\bar{L})$

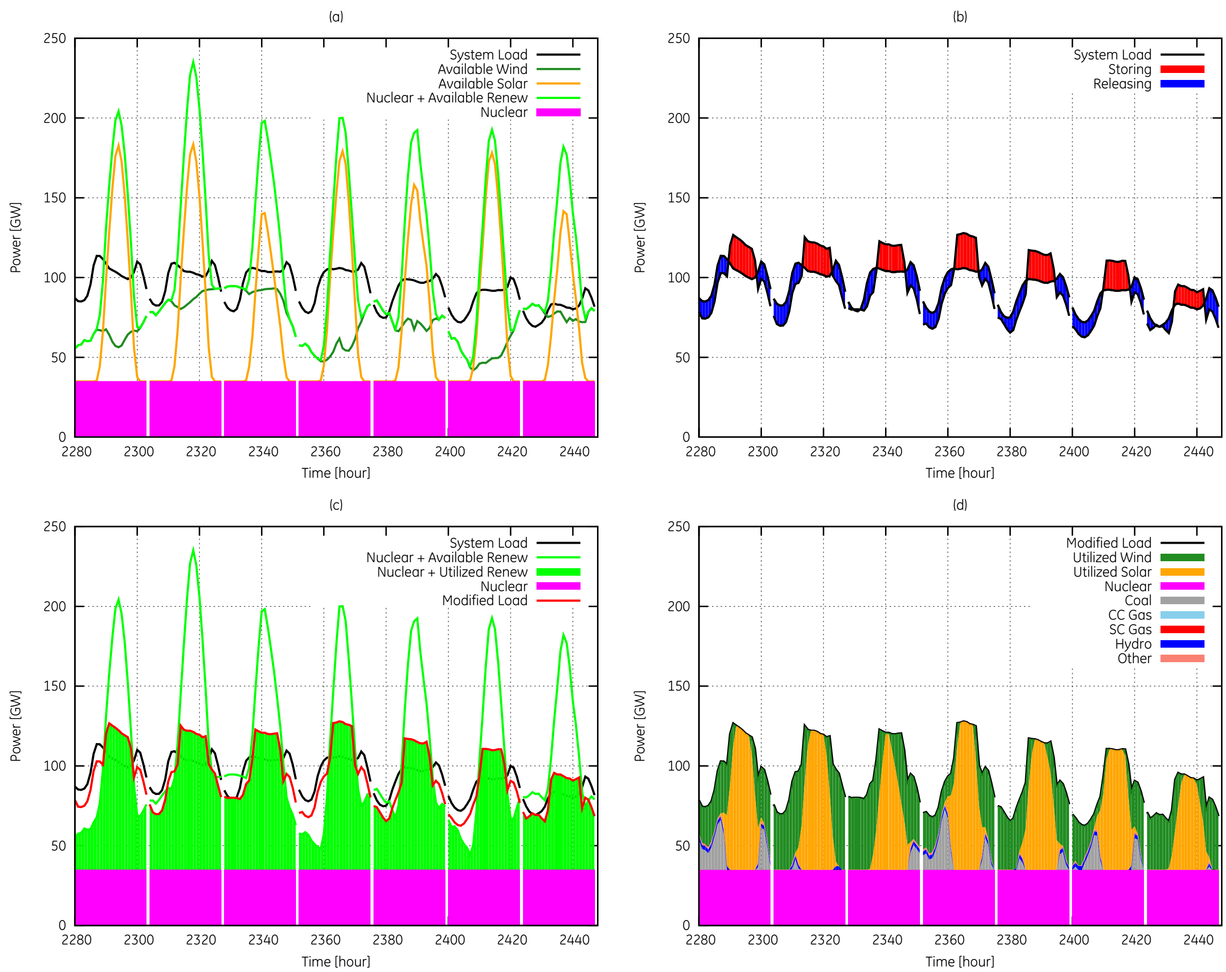


Figure 3.5: Algorithm B for r=50% ds/cs/w=40/10/50, week=14: (a) power profiles, (b) load scheduling, (c) storing vs. releasing, (d) generation mix

Figure 3.5(b), Figure 3.5(c) and Figure 3.5(d) show load scheduling, energy storing and releasing time intervals and the resulting generation mix respectively and as described in the previous subsection. One can note that the total utilized renewable energy is smaller than in the same conditions when the Algorithm A is implemented. This is the consequence of the step that uniformly decreases the profile for all hours of a day, i.e. decreases it for the same amount regardless of the net load amount. This way the total energy that can potentially be stored in the load is reduced. For this case Flex reports that the total utilized renewable energy for the entire year is 440 TWh (out of total available 485 TWh).

3.4 Algorithm C: Flexible Load with Scheduling

The Algorithm C is similar to the Algorithm B, but it avoids lower amounts of utilized renewable energy by applying different load scheduling. Similar to the previous algorithm, it is assumed that the aggregate flexibility is specified by the flexibility limit input parameter F, and that the constraint of keeping the average load the same is imposed.

The two algorithms are different in how they balance storing and releasing energy in the load. The total energy that can be stored within a cycle period (e.g. one day) is not equal to the total energy that can be released in the same period. In case when potential energy to store is larger, the Algorithm C computes what the modified load upper limit is, so that, when the limit is applied, the balance between stored and released energy exists. In case when potential energy to release is larger, the algorithm computes the modified load lower limit and then limits the original system load from below. Note that such an algorithm also tries to achieve the secondary objective of limiting the variation of the modified load.

The algorithm can symbolically be represented as follows:

1. if $L < R + N$ then $\delta L = min(F * L, R + N - L)$

   if $L > R + N$ then $\delta L = -min(F * L, L - R - N)$

2. if $E_S < E_R$ then compute M_{min}

   if $E_S > E_R$ then compute M_{max}

   $M = clip(L + \delta L, M_{min}, M_{max})$

Figure 3.6(a), Figure 3.6(b), Figure 3.6(c) and Figure 3.6(d) show power profiles, load scheduling, energy storing and releasing time intervals and the resulting generation mix respectively for the Algorithm C. Similar to what is discussed for the Algorithm A, one can note in Figure 3.6(b) and Figure 3.6(c) that during the first considered day, the amount of energy to store was smaller than the amount of energy to release, so the Algorithm C has limited the modi-

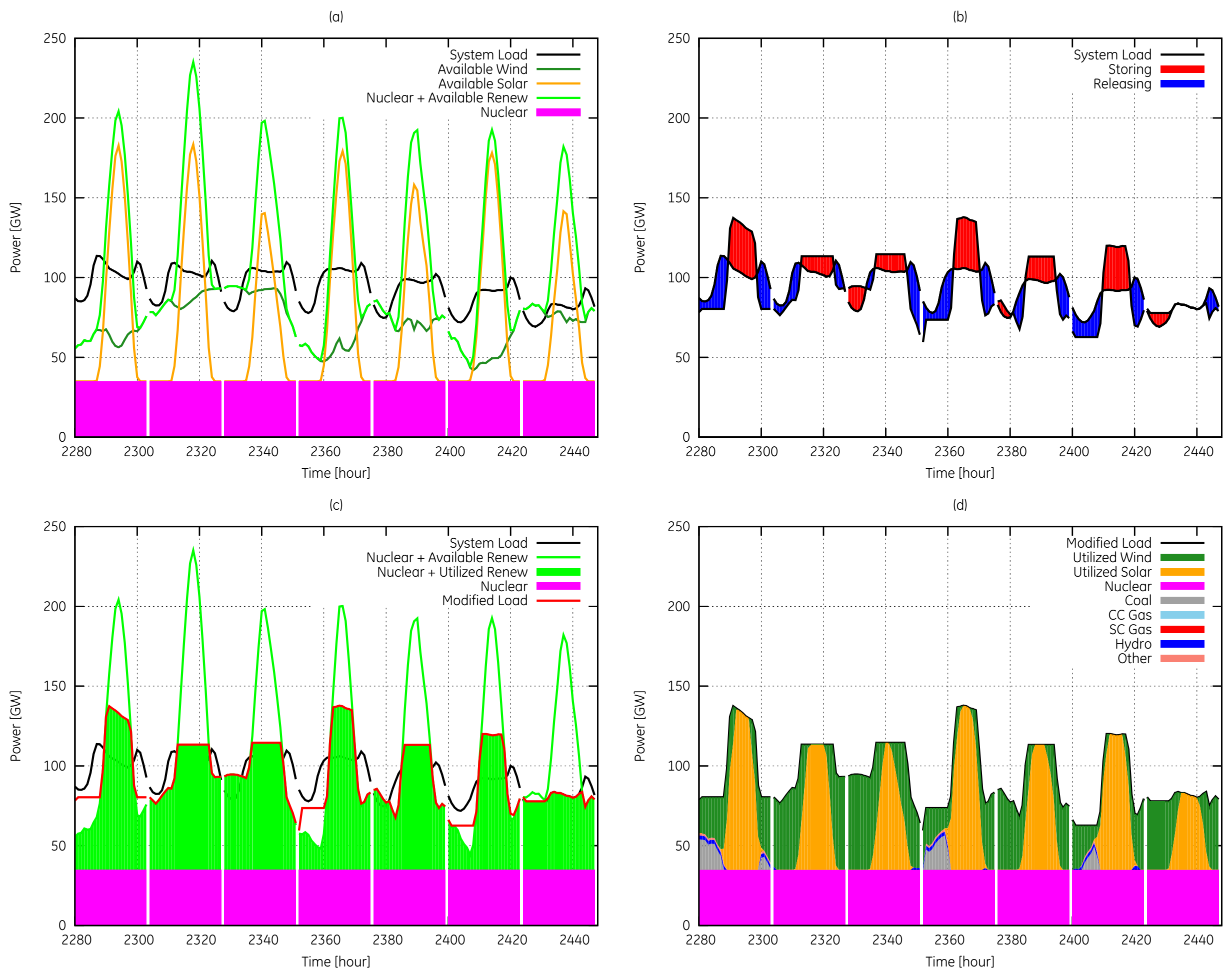


Figure 3.6: Algorithm C for r=50% ds/cs/w=40/10/50, week=14: (a) power profiles, (b) load scheduling, (c) storing vs. releasing, (d) generation mix

fied load from below. In contrast, during the second day, the amount of energy to store was larger than the amount of energy to release, so the algorithm has limited the modified load from above. This way, by considering the maximum amount of energy that can be stored in the load for each day, this algorithm achieves better results in terms of the utilized renewable energy than the Algorithm B. In fact, using the Algorithm C, Flex reports that the total utilized renewable energy for the entire year is 455 TWh (out of total available 485 TWh), which is larger than what is obtained with either the Algorithm A or the Algorithm B.

Chapter 4

Quantitative Analysis of System Benefits

4.1 Transmission Limits and Spinning Reserve

As discussed in Chapter 1, the unit commitment and dispatch are subject to constraints of the transmission system.

To examine how relaxing transmission limits could lead to these improvements a change-case was run using GE Multiple Area Production Simulation (MAPS) tool. The study was performed on the system model of the PJM Interconnection described earlier. GE MAPS simulation of PJM system was performed with all of the transmission congestion removed and the operational differences quantified relative to the case with transmission constraints.

In particular, for the scenario of 30% penetration of renewables, compared to the model with transmission constraints, the relaxed model exhibited:

- The total reductions in energy production costs was $10.1bn per year.
- The emissions of CO_2 dropped from 330 Mton to 260 Mton due to renewable energy substituting thermal energy.

- The capacity factor of coal plants dropped from 47% to 40%. The capacity factor of CCGT dropped even more, from 49% to 24%. The capacity factor of SCGT dropped from 1.2% to 0.6%.

The reduction of system spinning reserves, and consequently, the reduction of operating costs, could be achieved by providing the dynamic response to frequency from load. The benefits of such an approach were also quantified using GE MAPS by reducing the requirement for spinning reserve. The same scenario of renewables providing 30% of energy was considered. By eliminating 4.5GW of spinning reserves, the savings in production costs were $3.3bn per year.

4.2 Relaxing Renewable Integration Limits

In addition to GE MAPS, the Flex tool was used to evaluate the benefits of flexible loads for the integration of renewables. As discussed in 3.1, the tool simulates shaping of load over relevant time horizons in different scenarios. The goal was to perform sensitivity analyses with respect to different limits of integration discussed in Chapter 2. We focused on the following benefits: the increase of utilized renewable energy, the decrease of thermal energy of conventional generation fleet and decrease of CO_2 emissions.

In particular, in order to assess these benefits for different technologies, the following cases were considered in Flex simulations:

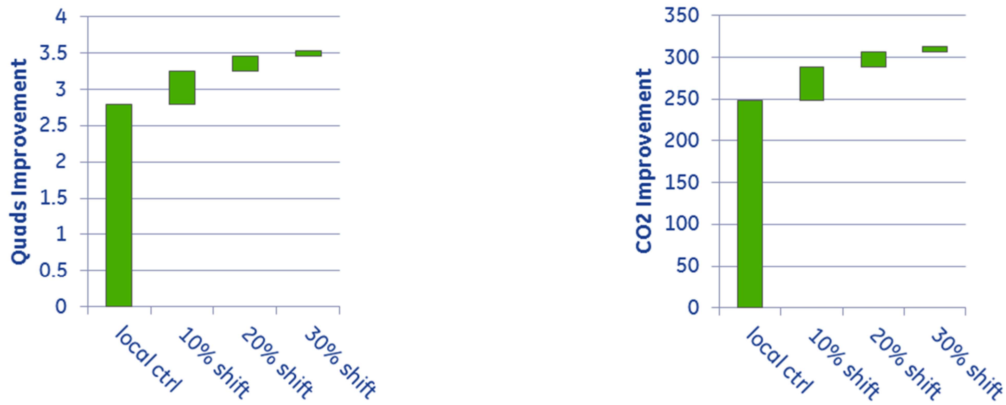
- **Reference.** This case is characterized with the level of 30% of renewable penetration and no flexibility in loads. It is used as a benchmark to GE MAPS cases discussed in the previous section.

- **Baseline.** No transmission limits, 50% renewable penetration, curtailment of renewable power when greater than 50% of load (the frequency stability limit discussed in Chapter 2.)

- **Local control.** The case characterized by distributed energy resources that apply local controls to provide dynamic response to frequency and eliminate the frequency stability limit in the baseline case.
- **Flexible loads.** The case in which loads exhibit flexibility and, thus, can help in aligning the system demand to the availability of renewable energy. The loads are scheduled according to the algorithm C described in Chapter 3.4 with three different values for the flexibility parameter: 10%, 20% and 30%.

These sensitivities were tested in scenarios with levels of renewables penetration equal to 50%. Two different mixes of renewable types were considered:

- **Scenario 1.** The scenario in which distributed solar is set to a maximum level that does not require utility control. In present-day systems distributed solar (e.g. roof-top solar) delivers available solar power to the load and the surplus of power to the utility system without any coordination with the utility. This lack of coordination sets the limit on total installed capacity of distributed solar to ensure that its instantaneous output is never greater than available system load. This scenario captures the maximum attainable level of distributed solar assuming that these practices continue. The renewables mix was set to: 15% distributed solar, 10% central solar, and 75% wind.
- **Scenario 2.** Distributed solar can be curtailed. In this scenario mix was chosen to be: 40% distributed solar, 10% central solar, and 50% wind.



(a) (b)

Figure 4.1: Scenario 1 (r=50% ds/cs/w=15/10/75): (a) quad improvements, (b) CO_2 improvements

Table 4.1: Scenario 1 (r=50% ds/cs/w=15/10/75)

Renewable Power Limit	Hours
50% Load	4141
100% Load - Nuclear	2091
100% Load	367

(a) Number of curtailment hours for limits

Penetration %	Mix ds/cs/w	Technology	URE %	Energy quads	CO_2 Mton
30	9/9/82	reference	99.6	+4.16	+338
50	15/10/75	baseline	79.0	—-	—-
"	"	local Ctrl	92.0	-2.79	-249
"	"	10% shift	94.2	-3.25	-288
"	"	20% shift	95.2	-3.46	-306
"	"	30% shift	95.6	-3.54	-313

(b) Estimates of utilized renewables, quads and CO_2

The Flex simulation results are summarized in tables Tab. 4.1 and Tab. 4.2 for scenario 1 and scenario 2 respectively. The results are organized into two subtables. The subtables (a) provide the general idea of the scale of installed renewable fleet by quantifying the estimated number of hours in a year in which the surplus of renewable energy would have to be curtailed as discussed in Chapter 2. The subtables (b) offer more details on impact of technology in analyzed sensitivity cases. First, we can observe that the reference case (30% of renewables) attainable with today's technology has high utilization of renewables but consumes approximately 4 quads more thermal energy than the baseline case, despite the lesser percentage of utilized renewable energy in the baseline case. The high curtailment in baseline case makes it a hypothetical measure of brute force approach where the renewables are added to reduce consumed thermal energy without regard for underlying economics.[1] The difference in consumed thermal energy also amounts to approximately 300 million tons of difference in CO_2 emissions. To make high-percentage of renewable deployment economically viable, the curtailment levels must be made comparable to curtailment levels in the reference case. We quantified the impact of load flexibility progressively by first removing the constraint of frequency stability labeled "local control" to signify that controls respond to a universally available signal of frequency. Next, we quantified the incremental benefits of load scheduling, or temporal "shifting". The results indicate that in both scenarios implementing local controls that eliminate the first limit of integration would reduce the energy of the thermal fleet by about 2.8 quads and CO_2 emissions by about 250 Mton. The subsequent technology improvements with increasing load flexibility result in proportionally smaller improvements in benefits. Note also that the incremental improvements are higher for scenario 2 in which distributed solar power is allowed to be curtailed. Figure 4.1 and Figure 4.2 give graphical representations of improvements when different technologies are applied. It is worth noting that the entitlement for improvement by load scheduling (shifting) would increase if the entitlement of local control was not fully realized.

[1]high curtailment rates would discourage investment into renewable energy making this a hypothetical case

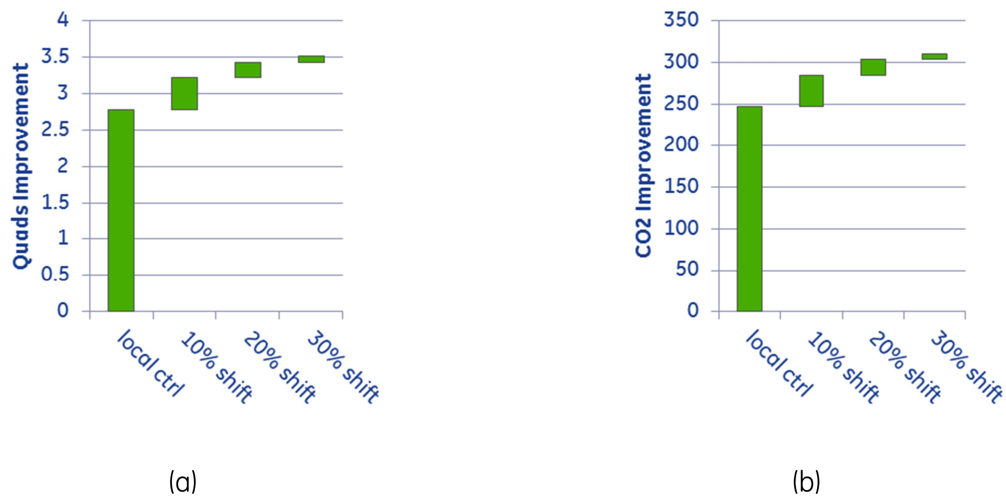


Figure 4.2: Scenario 2 (r=50% ds/cs/w=40/10/50): (a) quad improvements, (b) CO_2 improvements

Table 4.2: Scenario 2 (r=50% ds/cs/w=40/10/50)

Renewable Power Limit	Hours
50% Load	3569
100% Load - Nuclear	2036
100% Load	913

(a) Number of curtailment hours for limits

Penetration %	Mix ds/cs/w	Tech	URE %	Energy quads	CO_2 Mton
30	9/9/82	reference	99.6	+3.97	+325
50	40/10/50	baseline	72.6	—-	—-
"	"	local Ctrl	86.1	-2.89	-258
"	"	10% shift	89.9	-3.70	-326
"	"	20% shift	92.8	-4.30	-377
"	"	30% shift	94.7	-4.72	-414

(b) Estimates of utilized renewables, quads and CO_2

Chapter 5

Overcoming Integration Limits

5.1 Technology Gaps

In this section we discuss control and communications technology that needs to be developed for changes in operation of the power system discussed in previous sections. To do this we convert the desired load behavior into functional requirements for control technology and compare those requirements with the technology available today.

Here we assume that the overall control system is going to be hierarchical with the independent system operator (ISO) being on top and a large number of participating distributed energy resources (e.g. load and storage) at the bottom. The middle layers of control will be handled by entities that we call aggregators (e.g. utilities or curtailment service providers). In extreme implementations coordination of resources could be handled either at the ISO (centralized solution) or resource (peer-to-peer solution) level. Here we do not focus on the question of optimal level of distribution and further studies will be necessary to establish it. However, we do distinguish between two control architectures that address different functional requirements.

As discussed in Chapter 2, instantaneous output of renewable energy is limited by frequency

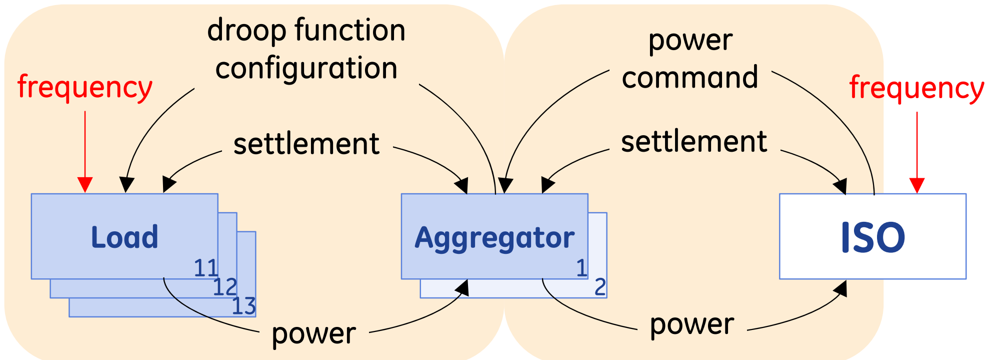


Figure 5.1: Local control functional requirements

stability to about 50% of the load. To overcome this with increasing penetration of renewables, loads will have to provide inertial and governor response. Figure 5.1 maps this feature into a high-level description of control and communications signals. The black arrows denote non real-time communication, whereas red denote real-time measurements or communication. Real-time in this context would typically mean periodic communication of one second order or below. The areas shaded in light-orange represent the gaps relative to today's operating practices.

Since frequency can be measured locally, at each participating load, in this case control could be implemented locally. An aggregator would still need to coordinate the response of the loads it services by occasionally setting the droop control parameters (see A.10) from the required reserve power command it obtains from the ISO. Compliance to the requested behavior, payments or warning notifications, etc., are achieved through settlement channels. Note that even though in current solutions for governor response control the communications shown in the figure do not occur, the related technical requirements (e.g. communication bandwidth or latency) would not have to be considerable.

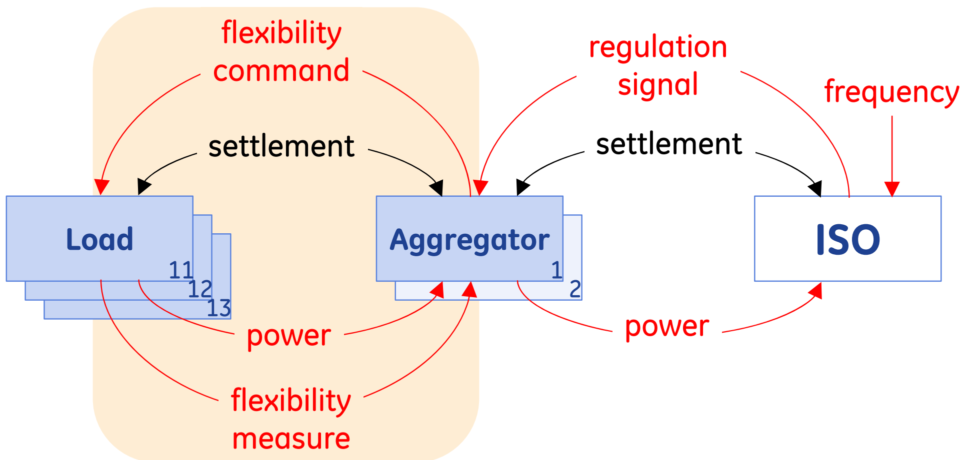


Figure 5.2: Global control functional requirements

To enable better temporal alignment of demand with available renewable energy, i.e. to relax the second or third limit from Chapter 2, we need flexible loads or storage technologies. These technologies can also be used if we want to improve other power system regulation mechanisms such as automatic generation control (AGC, see A.11) or optimal dispatch (see A.12).

Figure 5.2 addresses these use cases. The regulation signal such as AGC is communicated periodically from the ISO. Each aggregator computes flexibility (or power) command for each load it serves based on this signal and the operating state of the loads. For instance, for thermostatically-controlled loads (e.g. refrigerators) the distance from the desired cooling temperatures might be used as the state in this control schemes. Even though this architecture would require distributed controls, i.e. real-time communication, note that the interface between the ISO and an aggregator is the same as the current interface between the ISO and a thermal plant that takes part in regulation (e.g. the four second AGC communication interface).

5.2 Paths Forward

We conclude this report with a list of a few suggested areas of research that are most likely to address the technology gaps mentioned above:

- Study end-use consumption patterns and develop robust load behavior models. Design programs and incentives for flexible loads accordingly.
- Leverage system model built in this study (area operational characteristics and basic load models) by adding complex load behaviors to the model.
- Study different control interfaces of flexible loads: e.g. deferrable loads, loads with inertia, loads that can be shed, storage devices, etc.
- Develop methods for computation of aggregate properties of collections of loads (e.g. flexibility). For instance, develop procedures for estimation of equivalent energy storage parameters for a set of loads.
- Develop control methods for aggregators to compute power and/or flexibility commands.
- Develop control vocabulary and load behavior attribution (ramp up/down signals, load class/capability/flexibility/sensitivity, etc).
- Estimate telemetry and metering data bandwidth, data retention and security requirements in order to identify necessary communication technologies.
- Run simulations and hypothesis analysis to quantify performance and flexibility metrics:
  - Quantify responsiveness of load types and develop metrics expressing load class, ability to participate, inertia/governor response contribution.
  - Quantify heuristics of different control models across target use cases.
  - Quantify heuristics of the load behavior models.

Appendix A

Definitions of Power Engineering Terms

A.1 Operating Cost

A cost to deliver a unit of energy from a power plant; expressed in \$/MWh. Includes the cost of fuel and prorated annual maintenance to keep the facility in good working order. Operating cost depends on a number of factors such as: operating point, ambient temperature, number of starts per year, hours since the last start, etc. This is all incorporated into a bid delivered by the plant owner to the system operator.

A.2 Clearing Price

The price of the last (most expensive) MWh required to serve the system load. Equal to the operating cost of the most expensive power plant delivering energy to the system.

A.3 Dispatch Stack

A power engineering equivalent of cumulative generation capacity as a function of operating costs. Presented as a chart with cumulative generation capacity on the x-axis, and operating cost on the y-axis. An example dispatch stack is shown in Figure 3.2. Such visual offers a convenient insight into the relationship between the spot price and the level of system load. Because of their variable output, wind and solar plants are excluded from the dispatch stack. Wind and solar plans still affect the spot price because their output reduces the system load. As a result, the spot price is a function of net load. With reference to Figure 3.2, if the net load is ∼100GW the spot price will be ∼$55/MWh , while if the net load is 150GW the spot price is ∼$80/MWh .

A.4 Energy Revenues

Payments generators receive for delivered energy. Equal to the time integral of the clearing price multiplied by the generator output. Assuming that generators bid their true operating cost, they make a profit in the time intervals when the clearing price is greater than their operating costs.

A.5 Capacity Payments

Payments generators receive to be available for service; expressed in $/kW/year. Assuming the capacity value of $120/kW/year, a power plant that has the rated output of 500MW will receive $60M in capacity payments for the year.

A.6 Capacity Factor

A measure of average annual utilization of a power plant. It is calculated by dividing energy delivered by a facility with the facility's capacity, i.e. energy that could have been delivered if the facility ran all year long. The capacity factor of renewable power plants depends on availability of their fuel; wind is not steady over 8760 hours in the year, and solar resource is available in ∼10 hours of sunlight. Hydro power plants depend on availability of water, their capacity factors are lower in years of drought. Capacity factors are often averaged for the same technology fleet-wide. Average capacity factor of wind plants in the United States is ∼35% and of solar plants ∼16%. Capacity factors can get reduced below these averages by curtailment. In situation where curtailments are significant, two capacity factors are reported: one for available and another for delivered energy. Introducing renewable energy to the power system displaces energy supplied by the existing thermal fleet, which lowers the capacity factors of the thermal fleet and, consequently, lowers its energy revenues. As a result, for the thermal fleet to remain financially viable, capacity payments to the thermal fleet may need to increase with increased penetration of renewable energy.

A.7 Net Load

The system load minus the output from renewable generation fleet.

A.8 Duration Curve

A power engineering equivalent of a cumulative probability density function. It is created by plotting observations of a temporal variable in descending order as a function of time. Figure 2.1 shows an example of a load duration curve drawn by arranging hourly observations of system load [in GW] in descending order by value, for 8760 hours of a year. Such visual offers a convenient insight into how extreme the system peak load is and how long it lasts. For example, load of the system in Figure reffig:Durationoutput is greater than 100GW for $\sim$6000 hours, but greater than 150GW for only $\sim$500 hours.

A.9 Inertial Response

Overall, the response of the dynamic system described by the system of equations (1)-(4) to a loss of a generating plant, begins by the change of speed. To get to the change of speed, kinetic energy from the collective rotating mass of the system of generators must first be extracted and delivered to the electrical system. This is the fastest, inherent response, called inertial response. The generators begin to deliver their inertial energy into the system before any other control actions take place.

A.10 Governor Response

The control action of speed controllers (governors) in response to changing system frequency. The governors implement the so-called droop function that establishes the linear relationship (with negative slope) between the error in frequency and the steady state mechanical torque. Governor action is the main contributor to stabilizing the system frequency in transients initiated by large power imbalances. Their action takes place over a time interval from a few seconds to a few tens of seconds.

The transients of frequency have two components: the fall to the minimum value and the rise back to the steady state value. At the initiation of the transient the frequency error is insignificant so the governors are inactive and the dynamics of the transient are dependent on the collective inertial response from synchronous generators in the system. As the frequency error begins to increase, governor control systems begin to act and increase the output power (mechanical driving torque) of the fleet. The frequency nadir is a result of the total system inertia (reduced by renewable penetration) and of the weighted average speed of response of governors. Governor control dynamics dominantly depend on generation technology: hydro units are slower-acting than coal units, which are slower acting than gas units.

A.11 Automatic Generation Control (AGC)

AGC is the first control action orchestrated centrally. Whenever the system is operating at off-nominal frequency, up/down signals are sent to participating generators in 4 second intervals. This is used to cancel both the steady state error of frequency and the integral error of frequency.

A.12 Economic Dispatch

Economic dispatch is another centrally orchestrated change of operating points of all units but executed in one hour increments. This adjusts power outputs of in-zone fleet and power exchanges with the neighboring control zones to follow a schedule that results in the lowest load payments system-wide.

Bibliography

[1] Monitoring Analytics, "PJM State of the Market - 2013," http://www.monitoringanalytics.com/reports/pjm_state_of_the_market/2013/2013-som-pjm-volume2-sec3.pdfx, p. 66, accessed: 2014-10-23.

[2] U.S. Energy Information Administration, "What is U.S. electricity generation by energy source?" http://www.eia.gov/tools/faqs/faq.cfm?id=427&t=3, accessed: 2014-10-23.

[3] ——, "Annual Energy Outlook 2014 with projections to 2040, Figure IF7-2," http://www.eia.gov/forecasts/aeo/pdf/0383(2014).pdf, accessed: 2014-10-23.

[4] ——, "Average Operating Heat Rate for Selected Energy Sources," http://www.eia.gov/electricity/annual/html/epa_08_01.html, accessed: 2014-10-23.

[5] ——, "How much electricity is lost in transmission and distribution in the United States?" http://www.eia.gov/tools/faqs/faq.cfm?id=105&t=3, accessed: 2014-10-23.

[6] North American Electric Reliability Corporation, "United States Mandatory Standards Subject to Enforcement," http://www.nerc.com/pa/stand/Pages/ReliabilityStandardsUnitedStates.aspx?jurisdiction=United%20States, accessed: 2014-10-23.

[7] N. Miller, M. Shao, S. Pajic, and R. D'Aquila, "Eastern Frequency Response Study," GE Energy, Tech. Rep. NREL/SR-5500-58077, May 2013.

[8] GE Energy Management, “PJM Renewable Integration Study - Project Review (Task 3a),” http://www.pjm.com/~/media/committees-groups/committees/mic/20131028-impacts/20131028-pjm-renewable-integration-study.ashx, accessed: 2014-10-23.