UNIVERSITY OF CAMBRIDGE
Electricity Policy Research Group

Integrating short-term demand response into long-term investment planning

EPRG Working Paper 1113

Cambridge Working Paper in Economics 1132

Cedric De Jonghe, Benjamin F. Hobbs, Ronnie Belmans

Abstract Planning models have been used for many years to optimize generation investments in electric power systems. More recently, these models have been extended in order to treat demand-side management on an equal footing. This paper stresses the importance of integrating short-term demand response to time-varying prices into those investment models. Three different methodologies are suggested to integrate short-term responsiveness into a long-term model assuming that consumer response can be modelled using price-elastic demand and that generators behave competitively. First, numerical results show that considering operational constraints in an investment model results in less inflexible base load capacity and more mid-range capacity that has higher ramp rates. Then, own-price and cross-price elasticities are included in order to incorporate consumers' willingness to adjust the demand profile in response to price changes. Whereas own-price elasticities account for immediate response to price signals, cross-price elasticities account for shifting loads to other periods. As energy efficiency programs sponsored by governments or utilities also influence the load profile, the interaction of energy efficiency expenditures and demand response is also modelled. In particular, reduced responsiveness to prices can be a side effect when consumers have become more energy efficient. Comparison of model results for a single year optimization with and without demand response shows the peak reduction and valley filling effects of response to

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

EPRG WORKING PAPER

real-time prices for an illustrative example with a large amount of wind power injections. Additionally, increasing demand elasticity increases the optimal amount of installed wind power capacity. This suggests that demand-side management can result in environmental benefits not only through reducing energy use, but also by facilitating integration of renewable energy.

Keywords Wind power generation, power generation planning, load management, energy efficiency

JEL Classification L94

EPRG WORKING PAPER

Contact Cedric.DeJonghe@esat.kuleuven.be
Publication March 2011
Financial Support Flexnet - U.S. National Science Foundation

Integrating short-term demand response into long-term investment planning[1]

Cedric De Jonghe[2]
Visiting Researcher, Electricity Policy Research Group, University of Cambridge
Energy Institute (Electa), K.U. Leuven

Benjamin F. Hobbs
Electricity Policy Research Group, University of Cambridge
Whiting School of Engineering, The Johns Hopkins University

Ronnie Belmans
Energy Institute (Electa), K.U. Leuven

March 2011

Abstract
Planning models have been used for many years to optimize generation investments in electric power systems. More recently, these models have been extended in order to treat demand-side management on an equal footing. This paper stresses the importance of integrating short-term demand response to time-varying prices into those investment models. Three different methodologies are suggested to integrate short-term responsiveness into a long-term model assuming that consumer response can be modelled using price-elastic demand and that generators behave competitively. First, numerical results show that considering operational constraints in an investment model results in less inflexible base load capacity and more mid-range capacity that has higher ramp rates. Then, own-price and cross-price elasticities are included in order to incorporate consumers' willingness to adjust the demand profile in response to price changes. Whereas own-price elasticities account for immediate response to price signals, cross-price elasticities account for shifting loads to other periods. As energy efficiency programs sponsored by governments or utilities also influence the load profile, the interaction of energy efficiency expenditures and demand response is also modelled. In particular, reduced responsiveness to prices can be a side effect when consumers have become more energy efficient. Comparison of model results for a single year optimization with and without demand response shows the peak reduction and valley filling effects of response to real-time prices for an illustrative example with a large amount of wind power injections. Additionally, increasing demand elasticity increases the optimal amount of installed wind power capacity. This suggests that demand-side management can result in environmental benefits not only through reducing energy use, but also by facilitating integration of renewable energy.

Keywords
Demand response, generation investment, electricity resource planning models

---

[1] The authors are grateful for suggestions by Adriaan Hendrik van der Weijde (VU Amsterdam) and for helpful comments from the anonymous reviewers. The corresponding author would like to thank the Electricity Policy Research Group (EPRG) at the University of Cambridge for hospitality, advice and comments, as well as the Research Foundation – Flanders (FWO-Vlaanderen), Belgium. Partial support for B.F. Hobbs was provided by the U.K. EPSRC Supergen Flexnet program and the U.S. National Science Foundation through EFRI Grant 0835879.
[2] Corresponding Author – Cedric.DeJonghe@esat.kuleuven.be, Tel. +32 16 32 17 22, Fax +32 16 32 19 85 - University of Leuven (K.U.Leuven) Energy Institute, ELECTA branch (Electric Energy and Computer Architectures), Kasteelpark Arenberg 10 box 2445 / B-3001 Heverlee / Belgium.

1. INTRODUCTION

Large-scale deployment of renewable energy sources has been promoted by the European Commission in an effort to improve the sustainability of the electric power industry [1]. Most new renewable generation has been in the form of wind power. Large-scale wind power development impacts both short-term operation of the electricity system, as well as long-term investment planning. In operations, the integration of wind power significantly increases the variability of the generation output. Fluctuations in the amount of wind power fed into the grid require compensating changes in the output of other, flexible generators in the system. Flexibility can also be provided by international interconnections and energy storage. In case of insufficient flexibility, wind power curtailment, also referred to as wasted wind [2], can help to instantaneously balance generation and demand when over-generation is a problem. As flexibility of conventional generation technologies is restricted by technical constraints, such as ramp rates and minimum run levels, long-term investment planning should consider the increasing need for flexibility of generation units. Furthermore, occasional high demand situations with very little wind power injections require back-up capacity [3]. System reliability requirements and the joint distribution of load, wind output, and thermal unit forced outages determine the extent to which new wind development contributes to the ability of the system to serve peak loads. This concept is often referred to as the capacity credit of wind [4]. Generally, wind's capacity credit is significantly less than its average output, and depends on the extent of wind penetration into the market.

The above mentioned sources of flexibility are offered by the supply-side of the power system. However, integration of smart grid technologies in the electric power system [5], for example though smart meters, creates opportunities to more efficiently balance supply and demand. Smart meters resolve one of the demand-side market failures mentioned in [6] as without such meters, there is a lack of real-time billing. This prevents consumers from seeing and responding to real-time prices, resulting in perfectly inelastic demand in short-term. Consequently, traditional planning models suggest the optimal generation or transmission investment decisions for given projected load levels, neglecting the potential for short-term demand elasticity to trim peak loads and manage renewable energy fluctuations [7].

With the advent of smart meters that allow consumers to respond to real-time system conditions, investment planning models need to be enhanced in two ways in order to identifying the net benefit maximizing mix of generation, transmission, and demand-side investments. The first enhancement is representation of price-elastic demand. This representation should include cross-price elasticities, since the response to a higher price in one hour can both reduce demand by forgoing consumption without consumption recovery in the hour in question, as well as shift load to other times, also referred to as the substitution effect [8]. The second enhancement is the inclusion of dynamic operating constraints, especially ramp rate limitations, in order to appropriately value the flexibility contributed (or not) by alternative resources in the face of increased penetration of renewables.

This paper proposes such a generalization, and is organized as follows. First, in Section 2 we review how investment decisions are represented in long-term planning models available. Then in Section 3, a linear programming (LP) based long-term investment planning model is developed that represents system flexibility through the inclusion of chronologic dispatch constraints. Three methods are suggested to integrate short-term demand response into the planning model, assuming real-time varying rates that reflect the instantaneous marginal cost of generating electricity. The model accounts for both own- and cross-price elasticities. Hourly demand functions are defined with these characteristics. The model is also extended to account for investments in energy efficiency, whose effect can be viewed as a shifting of inverse demand curves to the left across a number of hours. Such investments can decrease the amount of demand response as well, if the slopes of those curves also steepen. This represents a negative interaction between the two types of demand-side measures. Results of an example application are presented in Section 4, followed by conclusions in Section 5.

2. Literature review

In this section, several long-term resource planning modelling approaches are reviewed. These models support decision making with a relevant time horizon of more than 20 years [9]. Optimization models, as described in [7] and [10], have been used for several decades. They offer solutions relevant in a regulated market or central planning context. Equilibrium models, as well as long-term market simulation models, typically used to represent market participants' (agents') strategic behaviour, show a similar market outcome under the assumption of perfect competition and perfectly inelastic demand [11]. Therefore, optimization models are often used as a benchmark of market prices or investment levels that could be expected if everybody behaved rationally and as price takers (perfect competition).

Long-term planning models support cost-minimizing investment decision making given long-term demand growth projections. These models are often referred to as Generation Expansion Planning (GEP) models. The LP formulation of this cost minimization problem was first presented in [12] and [13]. By minimizing the present worth of investment and operational costs, the optimal timing, location and type of newly commissioned plants is defined [14]. The basic model formulation has been extended in the past two decades by including variables and constraints that account for the following features: optimal plant scheduling, system security and reliability requirements (e.g., installed reserve margins [15]), and regulatory constraints such as emissions targets or caps [16]. Different resource attributes such as must-run capacity, operating reserve capabilities [7], and requirements for periodic maintenance [17] can also be added. An example of a commercial model of this type is the Integrated Planning Model.[3]

LP models have been successful because of their ability to model large and complex problems, but simplifying several assumptions was required. In order to improve the representation of several operational and investment related aspects of utility planning, alternative model structures have been proposed [18]. Alternative techniques for solving non-linear generation expansion models are suggested in [19]. Mixed integer programming is especially relevant when binary variables are associated with relevant investment projects or non-linear operational elements, such as minimum run levels and minimum up- and downtimes. In contrast, LP models assume that capacity and other variables can be varied continuously. Multi-criteria programming has been used in order to introduce additional objectives, such as

[3] Developed by ICF, Inc. and widely used in the U.S.; see www.icfi.com.

environmental impacts [20]. An overview of models including stochastic elements for demand or supply variables is given in [21]. Probabilistic production costing models, which account for the effect upon expected generation costs and customer outages or random plant forced outages, have been incorporated into the LP approach by decomposition methods [22], [23]. This approach is used by the commercial Electric Generation Expansion Analysis System (EGEAS).[4] By adding uncertainties (standard deviations) and correlations of different cost categories, optimal risk-cost portfolios can also be found [24]. This last technique has been widely applied within the financial sector, focussing on the benefits of diversification.

Dynamic programming (DP) is another useful programming methodology for electric utility planning, in spite of the scalability problems associated with these methods. It is a multi-stage optimization methodology that focuses on both the medium- and the long-term impact of decisions. Its advantage is that DP recognizes the binary nature of investment decisions and allows for many decision stages. Previously installed capacities and their possible decommissioning are integrated with capacity additions to derive a dynamic resource plan instead of a static, single year optimization used in many LP applications [25]. DP is used in commercial generation expansion packages, such as PROVIEW/PROSCREEN[5] and the Wien Automatic System Planning Package (WASP).[6]

In general, the above mentioned planning models present an investment plan based on a sophisticated supply-side analysis while demand-side options, such as energy efficiency programs or demand response, are remarkably simplified or even entirely neglected. The demand profile is typically described by a load duration curve, which is constructed by sorting the load in order of increasing hourly values, or by using three to six discreet load steps [16]. This representation loses information about the critical low load and high load situations, as well as chronologic hourly variability, which is crucial for assessing system flexibility in the face of varying demand and renewable production.

The chronological sequence of hourly load levels impacts interperiod operating constraints, which have, for many years, been assumed to be unimportant in the context of investment decision making. This assumption is, however, no longer tenable when there is large penetration of intermittent energy sources into the power system and the amount of required operating reserves and flexibility increases. In these cases, the rampability of existing and newly commissioned thermal generation types has to be explicitly taken into account in order to properly value their worth. Special attention is paid to ramp rates in [26] and [27].

In the 1970's, the energy crisis triggered public awareness of energy conservation, and utilities recognized that demand-side options could be seen as an alternative for satisfying customers' demand. The challenge in the 1980's for the electric utilities was to integrate the concept of influencing the electricity demand into traditional supply planning models [28]. As a result, the paradigm of Integrated Resource

[4] Developed by MIT under Electric Power Research Institute (EPRI) sponsorship [22], [23].
[5] Developed by the New Energy Associates (NEA). See discussions of the model in [79] and [80], and [81] for an application.
[6] Developed by the Tennessee Valley Authority (TVA) and Oak Ridge National Laboratory (ORNL) of the United States of America [82].

Planning (IRP) was developed.[7] Resource planning models to implement IRP were developed. These models intended to treat demand-side management (DSM) programs on equal footing with generation options [17]. DSM programs modify the timing and the amount of electricity demand through non-price mechanisms. Six different types of load shape objectives are identified and discussed in [28] and [29]: peak clipping, valley filling, load shifting, strategic conservation (also known as energy efficiency), strategic load growth and flexible load shaping. Each of those results in load shape changes, meaning that the electricity demand curve is shifted.

The most widely pursued type of DSM program, energy efficiency, focuses on energy [MWh] reductions. Energy efficiency refers to permanent changes to electricity usage through installation of, or replacement with, more efficient end-user devices (e.g., driven by subsidies for efficient air conditioning and lighting equipment [30]) or more effective operation of existing devices that reduce the quantity of energy needed to perform a desired function or service. Energy efficiency can be driven by consumers actively managing their energy costs, or result from DSM subsidies from utilities, or government regulations concerning equipment or building efficiency. In the former case, it can be considered as long-term demand response, because consumers respond to prices by adjusting their capital stock. In the case of utility subsidies or government rules, it is regulatory based. A wide variety of energy efficiency programs have been developed to subsidize and incent consumer investment in more efficient energy using equipment and buildings. The Demand Conservation Incentive [31] and the white certificates [32] are both examples of mechanisms supporting consumers to avoid electricity consumption.[8]

Most of these programs are economically justified by reductions of generation variable or investment costs, for instance as quantified by the California Standard Practice for benefit-cost analysis of DSM [33]. Strategic conservation and load management programs are sometimes included in IRP models as an alternative for minimizing costs [29]. However, cost-based paradigms underlying IRP models overlook how DSM can alter the value that consumers receive from consuming energy services [34]. Also, the interactions of energy efficiency investments with the ability of demand to respond to real-time prices were neglected by these models [8]. Furthermore, the cost-minimization objective is inappropriate for evaluating the net benefits of programs that influence demand by varying prices. Instead, the effects on consumer value and surplus (the difference between the value of energy services and expenditures) need to be considered in addition to resource costs [35].

---

[7] The fundamental differences between traditional planning and integrated resource planning are described in [55].

[8] The other types of DSM programs, which are not considered in this paper, are defined as follows. Strategic load growth attempts to increase energy [MWh] demands in order to provide value to consumers. In contrast, three basic forms of load management focus on instantaneous demand adjustments [MW] or load shapes. First, peak clipping (such as active controls of air conditions or water heaters) emphasize instantaneous demand reductions, allowing the system operators to deal with critical system situations. Peak clipping is considered by utilities as a means to reduce peaking capacity or capacity purchases. Second, valley filling involves building off-peak loads during hours with marginal cost below the average price of electricity. Third, load shifting encompasses moving loads from peak to off-peak periods. This last form of load management is often facilitated by thermal storage applications for space cooling and heating applications. The final type of DSM program, flexible load shaping focuses on reliability. Programs allowing interruptible or curtailable load are considered as sources of flexibility in the planning horizon of power supplier.

In contrast to traditional DSM and energy efficiency programs, demand response has the objective of integrating consumers into the spot electricity market, allowing them to interact with supply and express directly their willingness to pay for electricity over time and (in the case of locational pricing) space. At the same time that IRP was growing in influence, researchers at MIT developed the theory of dynamic or spot pricing [36]. Ironically, the developers originally anticipated that it would be most valuable for incenting consumers to modify the timing and amount of loads in response to system conditions; in actuality, it has instead provided the intellectual foundation for locational marginal pricing-based markets for coordinating power generation, where demand response has not yet played an important role. Caramanis [36] illustrates the different elements that compose a spot price. He distinguishes three tariff structures, depending on the frequency of metering and communication.

However, interest in demand response has grown considerably in the last decade. Several tariff options have been promoted, including time-of-use pricing (ToU), critical peak pricing (CPP), peak time rebate (PTR), as well as full real-time pricing. Prices or incentives can be based upon real-time wholesale prices, local congestion, or predetermined prices that are triggered by critical system situations. Examples of program options for implementing demand-response are interruptible load service, demand bidding, emergency demand response programs, capacity market programs, and Ancillary Services Market Programs [37], [38]. Changes in electric usage by end users in response to tariff changes is sometimes generally referred to as demand response, but for our purposes we define demand response more narrowly as response to real-time prices that are linked to spot bulk power prices.

Although demand response has heretofore not had a large impact on electricity markets, a number of researchers have analyzed its potential impact on market efficiency. A few early IRP models included long-run response to changes in average price levels [39], without examining in detail the impacts of hourly varying prices upon hourly loads. Another model considered the time lag or response gap until the next invoice period, resulting in consumers making medium-term adjustments in their consumption [40]. More recently, long run efficiency gains from implementing demand response along with real-time tariff structures are calculated in [41]. The impact of short-term demand response on the long-term optimal mix of generation technologies is also discussed in [42] and [43]. The latter uses the supply function equilibrium approach to model oligopolistic competition as well as a more traditional net benefit maximizing approach to modelling competitive markets. Both papers disregard short-term operational constraints and interperiod constraints.

This review of the literature reveals that there have been no long-run planning models that simultaneously integrate energy efficiency programs, demand response to hourly varying prices, and generation variable and investment costs, while considering the dynamic operating constraints whose importance is increasing in the face of increased renewable penetration. In the next section, we propose such a model.

3. Model description

3.1 Notation

Sets		***Indices***
I	Generation technologies	i
J	Periods (hours)	j

Parameters

A^+, B^+	Positive balancing linearization parameters
A^-, B^-	Negative balancing linearization parameters
BAL^+_j	Positive balancing requirement [MW]
BAL^-_j	Negative balancing requirement [MW]
CC	Cost of wind power curtailment [€/MWh]
DEM_j	Initial demand level in hour j [MWh]
EFF	Pump and turbine efficiency [%]
FC_i	Fixed investment cost for generation technology i [k€/MW/yr]
MR_i	Must-run parameter for generation technology i [%]
PM_i	Periodic maintenance parameter for generation technology i [%]
$PUMP\_CAP$	Maximum amount of energy pumped up or generated from the storage reservoir [MWh]
RR_C_i	Ramp rate on committed capacity for generation technology i [%]
RR_NC_i	Ramp rate on non-committed capacity for generation technology i [%]
STO_CAP	Maximum amount of energy stored [MWh]
T_CAP	Transmission capacity [MW]
VC_i	Variable generation cost for generation technology i, including both fuel and non-fuel components [€/MWh]
WP_j	Wind generated power output in hour j per MW installed [%][9]

Non-negative decision variables

cap_i	Installed capacity of generation technology i [MW]
$export_j$	Amount of energy exported during hour j [MWh]
$flex^+_{i,j}$	Upward output flexibility from generation technology i in hour j [MW]
$flex^-_{i,j}$	Downward output flexibility from generation technology i in hour j [MW]
$g_{i,j}$	Electric energy generation from generation technology i in hour j [MWh]
$spump_j$	Amount of energy pumped up in the storage reservoir in hour j [MWh]
$stored_j$	Amount of energy in the storage reservoir at the end of hour j [MWh]
sg_j	Amount of energy generated from the storage reservoir in hour j [MWh]
$wcap$	Level of installed wind power capacity [MW]
$wcurt_j$	Wind power curtailment in hour j [MWh]
$wind_j$	Amount of wind power injected in hour j [MWh]

[9] The wind generated power output profile is a fixed time series. Consequently, stochasticity or uncertainty about the output profile is not included. In order to account for variability, a representative time series is used in the model.

3.2 Basic LP generation planning model with operational flexibility constraints

This section presents a single node, LP resource planning model in which demand is fixed. In later sections, we elaborate it to include demand response. It is a static, single-year optimization based on [26]. Future technological, economic, and policy uncertainties with different possible future scenarios are not accounted for in this model[10]. In order to deal with those uncertainties related to the planning timescale, this model could be embedded into a stochastic (multiple scenario-based) planning model. A theoretical discussion of two stage and multistage programs is given in [44]. A multiple scenario-based electricity transmission planning model is used in [45] in order to accommodate renewables integration.

The static, single-year optimization structure is similar to the classic LP models in [10], except that we include chronological operating constraints that account for the need for operational flexibility. Annualized system costs are minimized, distinguishing between installed capacity (cap_i) and hourly electricity generation ($g_{i,j}$) with fixed (FC_i) and variable (VC_i) costs. A cost of wind power curtailment (CC) is added for each MW of the reduced wind power output during a full hour. The model defines the optimal installed capacity of different generation technologies as well as the hourly energy generation per type of technology.

As a single stage, static optimization model is used, the impact of an existing generation fleet with previously installed capacities are neglected. There is no particular obstacle to including the existing fleet. Pre-existing plants could be considered simply by including their fixed capacity in the models. The model can easily be extended with a positive lower bound for a particular technology type. Furthermore, decommissioning of installed capacities in the existing fleet can be added. Making this model dynamic, starting from an existing generation fleet and taking decommissioning of older generation plants into account would be a valuable extension to this model. It could help illustrating how a transition toward more renewables

Lumpiness of generation capacity investments is neglected. However the model is readily generalized to include linearized dc transmission constraints, lumpiness, operating reserves, and other complications present in other LP generation models. Those aspects are simplified in order to focus on the issues of modelling demand response together with operational flexibility.

The model can be viewed either as a simulation of a perfectly competitive market in which all market parties are price-takers, or as a planning model for a vertically integrated utility. Equivalence of both market formulations is argued in [46]. The full model is presented below.

[10] Presently, these structural uncertainties are not accounted for, but they will be dealt with in further research. It would be interesting to consider whether there are important interactions between those uncertainties, demand-response, and generation technology choice; considerations of construction lead times and option value [83] would then become important.

$$Min\ Cost = \sum_i FC_i * cap_i + \sum_{i,j} VC_i * g_{i,j} + \sum_j CC * curt_j \qquad (1)$$

subject to:

$$\sum_i g_{i,j} + wind_j = DEM_j + spump_j - EFF * sg_j \qquad \forall j \in J \qquad (2)$$
$$g_{i,j} \le PM_i * cap_i \qquad \forall i \in I, j \in J \qquad (3)$$
$$g_{i,j} \ge MR_i * cap_i \qquad \forall i \in I, j \in J \qquad (4)$$
$$wind_j = WP_j * wcap - wcurt_j \qquad \forall j \in J \qquad (5)$$
$$stored_j = stored_{j-1} + spump_j * EFF - sg_j \qquad \forall j \in J \qquad (6)$$
$$stored_j \le STO\_CAP \qquad \forall j \in J \qquad (7)$$
$$spump_j \le PUMP\_CAP \qquad \forall j \in J \qquad (8)$$
$$sg_j \le PUMP\_CAP \qquad \forall j \in J \qquad (9)$$
$$export_j \le T\_CAP \qquad \forall j \in J \qquad (10)$$
$$flex^{+}{}_{i,j} \le RR\_C_i * g_{i,j-1} + RR\_NC_i * \left(cap_i - g_{i,j-1}\right) \qquad \forall i \in I, j \in J \qquad (11)$$
$$flex^{+}{}_{i,j} \le cap_i - g_{i,j-1} \qquad \forall i \in I, j \in J \qquad (12)$$
$$flex^{-}{}_{i,j} \le RR\_C_i * g_{i,j-1} + RR\_NC_i * \left(cap_i - g_{i,j-1}\right) \qquad \forall i \in I, j \in J \qquad (13)$$
$$flex^{-}{}_{i,j} \le g_{i,j-1} \qquad \forall i \in I, j \in J \qquad (14)$$
$$g_{i,j} \le g_{i,j-1} + flex^{+}{}_{i,j} \qquad \forall i \in I, j \in J/\{1\} \qquad (15)$$
$$g_{i,j} \ge g_{i,j-1} - flex^{-}{}_{i,j} \qquad \forall i \in I, j \in J/\{1\} \qquad (16)$$
$$BAL^{+}{}_j = A^{+} * WIND_j * wcap + B^{+} * wcap \qquad \forall j \in J \qquad (17)$$
$$BAL^{-}{}_j = A^{-} * WIND_j * wcap + B^{-} * wcap \qquad \forall j \in J \qquad (18)$$
$$\sum_i g_{i,j} + BAL^{+}{}_j \le \sum_i g_{i,j-1} + \sum_i flex^{+}{}_{i,j} \qquad \forall j \in J \qquad (19)$$
$$\sum_i g_{i,j} - BAL^{-}{}_j \ge \sum_i g_{i,j-1} - \sum_i flex^{-}{}_{i,j} \qquad \forall j \in J \qquad (20)$$

The cost objective function (Eq. *(1)*) is minimized subject to the system energy balance constraint (Eq. *(2)*), the capacity constraint and operational constraints (Eq. *(3)*-*(20)*). Wind power is modelled as a generation unit with an hourly profile [%] time series, multiplied by the installed capacity, expressed in [MW]. This way of integrating wind power incorporates hourly variability but assumes that future wind output is perfectly predictable. The state of the art in assessment of short-term wind forecast error on system operation is discussed in [47]

In reality, the wind output in a given future hour *j'* is not known when making dispatch and commitment decisions in an earlier hour *j < j'*. This simplification likely understates the value of flexible generation. No optimization-based planning model in the literature includes such planning uncertainty. This paper is an improvement on existing planning models that exclude chronological constraints on output. However, this aspect of uncertainty is still a subject for future research.

The installed wind power capacity can be exogenously defined in response to a regulatory mandate. Alternatively, the amount of wind power capacity can be made endogenous by attributing a fixed investment and variable generation cost to wind power. Wind power capacity is then included as a decision variable. The latter way of integrating wind power is used in this model.

An alternative methodology to incorporate wind power is called the load modifier which reduces the net demand profile [48]; however, that would not allow wind capacity to be treated as a variable.

Wind power curtailment is proposed in order to reduce the hourly injections into the system when the system is in an over-generation situation. This amount of discarded energy (*$curt_j$*) will become substantial for increasing wind power generation [49]. Wind power curtailment is economically optimal when system marginal costs or prices are negative, and allows increasing power generation from a technology with low variable costs but limited flexibility. How wind curtailment is modelled depends on the particular market context. One context would involve a cost minimizing generator with some wind generation who operates in a perfectly competitive market. This market party faces an opportunity cost when his wind power output is reduced if wind subsidies are paid per MWh of wind production. In that case, this opportunity cost (*CC*) should be based on the feed-in tariff or the green certificate price in the region of curtailment.[11] Use of *CC* in the planning model simulates the outcome of the competitive market that is subject to such subsidies. On the other hand, assuming a central planning setting, total costs for the society are minimized, and such subsidies would be viewed as income transfers from consumers to generators and so *CC* would be set to zero. In this paper, we assume a perfect competition market setting. Adopting hourly time intervals and assuming that wind power curtailment applies for at least one hour, the cost of curtailing power *CC* is expressed in [€/MWh].

Turning from the objective function to the constraints, Eq. *(2)* shows that total energy generated from conventional and wind power facilities meets demand in *j* (with system losses regarded as demand). For each technology type *i*, the generated energy in hour *j* is restricted by the installed capacity (Eq. *(3)*). The available capacity is downscaled by technology specific periodic maintenance parameter (PM_i). An energy storage facility (e.g., pumped hydro) is included in order to improve system flexibility. The dispatch of this unit helps to balance generation and demand bearing in mind its efficiency and the need to satisfy the stored energy balance Eq. *(6)* where the amount of energy stored at the end of hour *j* equals the previous hour's storage (*j-1*) plus the net energy injected into the reservoir (pumpage minus generation) during hour (*j*). Energy storage results in energy losses during pumping (charging), as well as generation (discharging). Consequently, the net efficiency of the entire storage cycle from pumping to turbining equals EFF^2. The amount of energy stored is restricted by the total storage capacity, which here is fixed (Eq. *(7)*) but could in general be a decision variable. Also the amount of energy pumped up or generated from the energy storage reservoir is restricted by a maximum turbine capacity, respectively given by Eq. *(8)* and *(9)*. Again, the turbine capacity could be integrated as a decision variable.

In general, power systems are connected to neighbouring control areas. Those transmission interconnections create cross-border trade opportunities which improve system flexibility. Importing or exporting power during critical system situations impacts the operation and optimal mix of generation capacity in the exporting as well as in the importing region. Note that relaying on neighbouring systems for flexibility is limited because demand and wind power injections in neighbouring regions are highly correlated with those in the system being modelled. A positive variable (*$export_j$*) is added to Eq. *(3)*, allowing the export of excess power during high wind situations. The amount of power exported is restricted only by thermal limits of the

[11] Negative prices can already be seen in high wind regions. On the German spot market (European Energy Exchange: EEX), negative prices up to -119 €/MWh occur on the day-ahead market. Similar negative price occur on the Danish real time market.

conductor, indicated by the available transmission capacity in Eq. *(10)*, We assume that no price would be paid for nor received from those exports.

In this model, only export of excess power is permitted. Importing power during moments of high demand could easily be integrated by also adding a positive variable ($import_j$) to Eq. *(3)*, again restricted by the transmission capacity. This variable would also have to be included into the objective function linked with a price of imports. This practice is similar to the creation of a new generation technology at a limited (only variable) cost with a restricted total capacity. Comparable to total storage capacity, transmission capacity could also be included as a decision variable.

Operational constraints such as periodic maintenance [17] and must-run levels are included into the model. A technology specific periodic maintenance parameter (PM_i) is introduced to downscale the available capacity. The parameter PM_i can also be chosen in order to account for expected forced outages. It is well known that treating forced outages in this way decreases expected generation costs relative to true operating costs under random outages.

A full stochastic model with random plant outages is, in theory possible. Due to Jensen's inequality (the expected value of a convex function is more than the function evaluated at the expected input value, where the inputs are available capacity), expected production costs are still underestimated by this derating approach. However, it is more accurate than ignoring outages entirely [50]. It is possible that demand-response will become more valuable if outages were considered in a stochastic manner, as there would be a (small) probability of very extreme conditions with many plants unavailable.

Papers have been published on generation expansion planning under random outages, but they use non-chronologic convolution methods that cannot consider ramp rate limitations or pumped hydro storage units [22]. An alternative would be to consider many more days in the operating subproblems, with different configurations of outages (as in [51]), which would greatly increase the problem size.

Less flexible generation technologies will not be operated below certain output levels. Therefore a technology specific must-run (MR_i) level is integrated as in Eq. *(4)*. This constraint is typically not imposed for peaking generation technologies such as combustion turbines. Other unit-commitment constraints, such as start-up costs, minimum up and down times and minimum output levels are not considered in this model as these constraints require the use of integer variables. A linear programming approximation of start-up constraints and partial load levels is given in [52], and could be incorporated into this model.

The constraints above (Eq. *(3)* and *(4)*), determined by the parameters PM_i and MR_i, have no inter-period characteristic. Inclusion of ramp rate limits for conventional generation units can have a more complicated impact on the optimal mix when combined with a chronologic representation of hourly load and wind power production. Because the need for flexible thermal generation to make up for wind variability is of increasing economic importance, this aspect should be included in planning models.

In this model, ramp rates limits represent the output flexibility of generation. Operating a base load plant efficiently requires a higher yearly energy output per installed capacity. A deviation from the optimal point of operation harms efficiency, and frequent cycling can increase maintenance costs and reduce the lifetime of these units. Consequently lower ramp rates are assigned to base load technologies. Less stringent ramp limits are also assumed for mid and peak load generation units.

Ramp rates typically express flexibility as a percentage of the total installed capacity of a specific generation technology. In order to reduce the generation output flexibility when a plant is operated at lower output levels, a distinction is made between committed and non-committed capacity. A lower flexibility in terms of percentage is attributed to non-committed than committed capacity. Non-committed capacity is approximated by the amount of capacity that is not used for generation. This is respectively indicated by the parameters RR_NC_i and RR_C_i. Only for high peak load capacity is 100% flexibility assumed for upward as well as downward ramping.

The following restrictions are introduced into the model. Upward ($flex^{+}_{i,j}$) and downward ($flex^{-}_{i,j}$) flexibility of each generation technology are calculated in Eq. *(11)* and Eq. *(13)* for each hour. The former is restricted by the total amount of non-committed capacity Eq. *(12)*, meaning that the generation output can never ramp-up more in one hour than the non-committed capacity in the previous hour. This constraint is needed when the positive balancing requirement Eq. *(19)* is introduced. Correspondingly, the latter is restricted by the output level or committed capacity in the previous hour (Eq. *(14)*). This constraint is needed when the negative balancing requirement Eq. *(20)* is introduced.

For every hour, constraints *(15)*-*(16)* are introduced into the model to restrict the generation fluctuations of each technology type. These generation fluctuations constraints do not require the inclusion of Eq. *(12)* and *(14)*, given that $g_{i,j}$ is a non-negative decision variable, restricted by the available capacity in Eq. *(3)*.

The balancing requirement constraint is also important for a system with large-scale integration of wind power. This way of increasing ancillary service requirements in proportion to day-ahead wind schedules is inspired by [53]. Additional wind power injections require additional balancing power for both positive and negative regulation. Levels of balancing power are assumed to increase linearly with increasing hourly wind power injection [54]. These balancing power requirements have to be fulfilled by the aggregated upward or downward flexibility of conventional generation units (Eq. *(19)*-*(20)*).

3.3 Representing short-term demand response

In this subsection, we illustrate how a demand response function can be constructed; then in Section 3.4, we integrate it into the above planning model. Short-term demand elasticity must be contrasted with medium and long-term demand. Given long-term demand elasticity, consumers adjust their capital stock (equipment such as refrigerators, washing machine, lights, etc.) in response to shifts in electricity price levels. This last aspect is integrated as lagged consumption terms in [31]. Given medium-term demand elasticity, consumers respond to prices reflected on monthly bills [40], by changing their behaviour. Medium and long-term response is not accounted for in this paper, but will be dealt with in future work.

Given short-term demand elasticity, a distinction is made between residential, commercial and industrial customer classes in [31]. The characteristics of installed metering and communication infrastructure are one way that different classes are distinguished. At one extreme, 5-minute spot prices require advanced, real-time communication between the consumer and the utility. At the other extreme, use of day-ahead forecasts of 24-hour spot prices only requires a daily price update [55]. The involved metering and communication costs, as well as political and social aspects, mean that different consumers prefer different price or tariff structures. A single hourly demand function is used in this model, assumed to represent the demand response aggregated over different customer categories; sensitivity analyses could explore the impact of alternative assumptions concerning the extent of participation of different classes. The impact of varying the share of customers under real-time pricing is explored in [41].

In order to represent short-term demand response and include cross-elasticity price effects in one hour upon demand in others, elastic demand functions have to be calibrated for each hour. We express quantity demanded as a function of the bulk energy portion of the electricity price (that is, fixed customer charges, such as for billing, transmission and distribution, are excluded).[12] This is done by defining a reference price and quantity demanded for each hour, and then using elasticity assumptions to fit a demand curve through that price-quantity pair. The reference quantity demanded is based on a demand forecast. Then the reference price is obtained by applying the LP resource planning model to the reference demand, assuming fixed short-term demand levels. The model defines the optimal generation technology mix as well as the hourly generation output of each technology category subject to operational constraints. The reference price is assumed to be the same in all hours (no real-time pricing). This uniform price (P_0) is the quantity weighted average of the hourly (marginal) energy prices P_{oj} over the entire time horizon without demand response (Eq. *(21)*). It is assumed in this reference case that consumers do not yet face the hourly energy price as they are still assumed to be under a uniform pricing structure.

$$P_0 = \frac{\sum_j P_{0j} * DEM_j}{\sum_j DEM_j} \qquad (21)$$

A similar methodology could be applied in order to calculate weighted average prices for a double tariff structure, distinguishing between specific blocks such as peak and off-peak, as shown in [31].

The reference price-quantity pair composed of the weighted average (uniform) price and the fixed demand level $\{P_0, DEM_j\}$ is considered to be the anchor point of the linear demand function (Figure 1). The slope of the function is determined by the

[12] This is a common assumption in bulk power market models. The user therefore needs to be careful when using price elasticity values from the literature to calibrate such models, because those express percent changes in quantity demanded as a function of percent changes in retail prices. If instead expressed as a function of percent changes in bulk prices, the elasticities would be smaller (since bulk power prices are smaller than retail prices, so a given percentage change in retail prices would be a larger percentage of bulk prices).

price elasticity assumptions with own-price elasticities ($\varepsilon_{j,j}$) and cross-price elasticities ($\varepsilon_{j,k}$) being exogenously provided, based upon values from the literature. The addition of price elasticities results in a short-term demand response function $\boldsymbol{D_j}$ (Eq. *(22)*) which expresses quantity demanded (d_j) as function of relative deviations of hourly prices from the reference level; the simplified form in the right side of *(22)* is used. Inverting $\boldsymbol{D_j}$ results in the inverse demand function $\boldsymbol{P_j}$ (Eq. *(23)*), with parameters E_k and $F_{j,k}$. The inverse demand function is used in the generation planning model. Figure 1 shows the portion of that function that relates price to demand in its own period.

$$\boldsymbol{D_j}: d_j = DEM_j + \sum_k \varepsilon_{j,k} * \frac{DEM_j}{P_0} * (p_k - P_0) \equiv A_j + \sum_k b_{j,k} * p_k \tag{22}$$

$$\boldsymbol{P_j}: p_j = E_j + \sum_k F_{j,k} * d_k \tag{23}$$

With parameters:

$$B_{j,k} = \varepsilon_{jk} * \frac{DEM_j}{P_0} \tag{24}$$

$$A_j = DEM_j - \sum_k B_{j,k} * P_0 \tag{25}$$

The elastic linear demand function is constructed separately for each hour. At least some of the consumers are assumed to be participants in tariff systems in which they face real-time or spot prices; the assumed elasticities reflect the extent of participation of consumers in that tariff. (Smaller elasticities would correspond to less participation.) A real-time price higher than the weighted average electricity price ($P_{0j} > P_0$) results in decreased consumption of electricity. Correspondingly, real-time prices lower than the reference weighted average electricity price result in increased levels of electricity consumption.

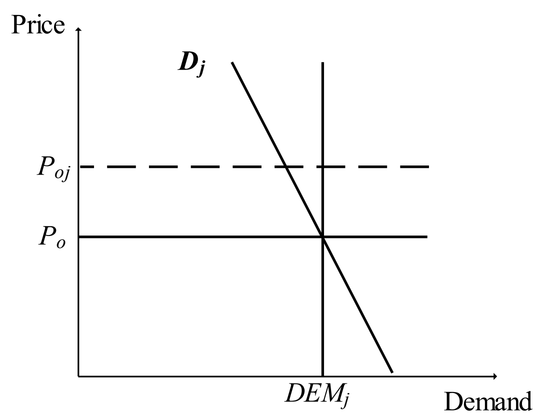

Figure 1: Construction of a short-term elastic demand function

It is suggested in [36] that this hourly energy price can be seen as a function of two components when abstracting from network constraints. The first one is the marginal operating cost determined primarily by the incremental fuel cost of the most expensive unit currently loaded in the system. The second component is called the energy balance 'quality of supply' premium. This premium is zero at times of surplus generation capacity. When all generators are in use at full capacity, which is termed a 'scarcity condition', the premium can be positive. The hourly energy price corresponds to the dual variable, or shadow price (λ_j) associated with market clearing requirement. With constraining capacity, the hourly energy price can rise above the marginal cost of the last unit operation in a given hour. For those hours *j*, a positive balance quality of supply premium ($\mu_{i,j}$) can be found, corresponding to the dual variable of the capacity constraint of the respective technology *i*. This premium is equal to the difference between the marginal cost of the last unit in operation and the energy price (λ_j) in that given hour *j*.

Dual variables are alternatively termed multipliers representing the marginal price on changes to the corresponding constraint. This is the case for both the elastic demand function ($\boldsymbol{D_j}$) and the inelastic demand, corresponding to initial demand levels (DEM_j) in Figure 2, illustrating the different spot price components. This graph shows that in case of deficient generation capacity, the spot price increases above the marginal fuel costs. Consumers under real-time pricing, with elastic demand function ($\boldsymbol{D_j}$), face increasing electricity prices and adjust their level of consumption. On the other hand, if demand is perfectly inelastic (vertical demand function), insufficient capacity would mean that instantaneous levels of consumption would have to be reduced by means of rationing. This practice refers to enforcing rotating black-outs in order to balance supply and demand. The amount of rationing (*R*) is also shown in Figure 2.[13] In both cases, it is said that the market is cleared at the demand-side of the system. However in the first case consumers define autonomously their welfare maximizing level of consumption. In the second case, system operator intervention is required, resulting in imposed outage costs on all users. This might not be the optimal outcome when outage costs differ from marginal consumer willingness to pay electricity.

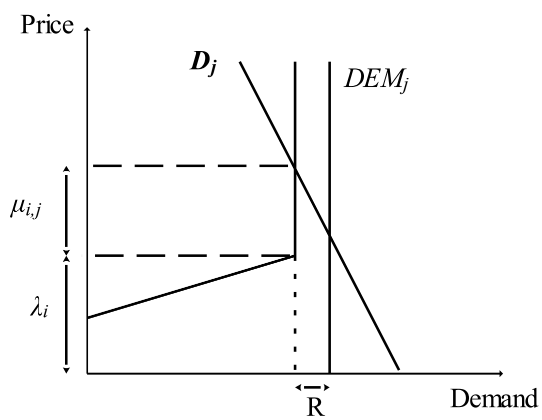

Figure 2: Rationing and spot pricing of electricity

3.4 Methodologies to include short-term demand response into a resource planning model

Resource planning models with fixed demand profiles pursue the reduction of system costs. When short-term demand response is integrated into the model, minimization of generation costs does not yield sensible results, because that would disregard the benefits consumers receive from electricity consumption. By definition, an equilibrium solution must be found between generation supply and demand. In

[13] Rationing is not considered in our model of Section 3.2, although it could be. Generally, in a long run planning model, rationing would be cost-minimizing for the number of hours per year determined approximately by VOLL/FC_{peaker}, where VOLL is the value of lost load (in €/MWh) [6] and FC_{peaker} is the capital cost of a peaking power plant (in €/MW/yr). A specific component for rationing linked to rotating blackouts was included to total system costs in [36].

abstract, we might characterise this equilibrium using supply ($\boldsymbol{S_j}$) and demand ($\boldsymbol{P_j}$) price functions (in €/MWh) that depend on Q (in MWh) as follows:

$$\boldsymbol{P_j}(Q) = \boldsymbol{S_j}(Q) \tag{26}$$

With:

$$\boldsymbol{P_j}(\boldsymbol{Q}) = k+KQ \tag{27}$$
$$\boldsymbol{S_j}(\boldsymbol{Q}) = m\text{-}MQ \tag{28}$$

The demand price function is Eq. *(22)*; the supply price function is instead an implicit function that is calculated by the LP. Three different methodologies to integrate short-term demand inelasticity into an LP resource planning model are presented below. These three methodologies yield the same solution under restrictive conditions, which is demonstrated below. For clarity of presentation, we show how demand response is included in a greatly simplified version of the LP model of Section 3.2 (Eq. *(29)-(31)*), omitting operational constraints, the storage unit, export of energy using transmission interconnection and wind power curtailment. Furthermore, initial demand levels in hour j (DEM_j) are replaced by demand levels (d_j) as a decision variable in Eq. *(30)*.

$$Min\ Cost = \sum_i FC_i * cap_i + \sum_{i,j} VC_i * g_{i,j} \tag{29}$$

subject to:

$$\sum_i g_{i,j} = d_j \qquad \forall j \in J \tag{30}$$
$$g_{i,j} \leq cap_i \qquad \forall i \in I, j \in J \tag{31}$$

The same three approaches are straightforwardly applied in the same manner to the full model of Section 3.2 in our case study of Section 4.

3.4.1 Complementarity programming model

A first way to model short-term demand response in a planning model is by using a complementarity program model structure, representing the competitive equilibrium solution. The competitive equilibrium represents a situation in which energy suppliers and consumers are each maximizing their individual profits and consumer surplus, respectively, subject to market prices, and the market clears (supply equals demand). Under certain conditions concerning price elasticities, this can be shown to be equivalent to maximization of market surplus (or 'social welfare'), equal to the sum of producer profits and consumer surplus [55]. The complementarity model solves a system of conditions including each market player's first-order optimality conditions or Karush–Kuhn–Tucker (KKT) conditions, plus the market clearing condition, one per period j [56]. As this model minimizes the cost of meeting a particular quantity demanded (during representative period) and accounts for demand response to prices, it should be viewed as a planning model.

Assuming integrable supply and demand functions, total consumer surplus (value received from consumption $\boldsymbol{U}$ minus expenditures) and profit (revenue minus costs $\boldsymbol{C}$) are given below. Correspondingly, the equilibrium solution is defined in the following general way for consumers maximizing their welfare ***U(Q)*** and generators simultaneously maximizing profits.[14]

$$Consumer\ Surplus\ =\ \boldsymbol{U}(Q) - PQ = \Sigma_j \int_0^{Q_j} P_j(q)dq - \Sigma_j \lambda_j Q_j \qquad (32)$$
$$Profit =\ PQ - \boldsymbol{C}(Q) = PQ - \Sigma_j \int_0^{Q_j} S_j(q)dq = \sum_j \lambda_j Q_j - \Sigma_j \int_0^{Q_j} S_j(q)dq \qquad (33)$$

Note that the price equals λ_j, in this simplified model representation. In this formulation, the price in period *j* is a function only of Q_j, neglecting cross-price impacts. In a more general formulation, the price in period *j* considers both own-price and cross-price elasticities. In that case, the price in period *j* is a function of Q_k with *k* equal to *j*, accounting for own-price elasticities and *k* not equal to *j,* accounting for cross-price elasticities.

The total costs ***C(Q)*** for the simplified model are given in Eq. *(29)*. The KKT conditions of the profit maximizing generator (Eq. *(33)*) for decision variable $g_{i,j}$ and cap_i are respectively given by Eq. *(34)* and *(35)*. The KKT conditions of the surplus maximizing consumer (Eq. *(32)*) is given by Eq. *(36)*.

The capacity constraint of this simplified resource planning model is included as Eq. *(37)*, with the dual variable of this constraint given by $\mu_{i,j}$. The market clearing condition is given by Eq. *(38)*. The dual variable associated with the market clearing condition (λ_j), is equal to the hourly (marginal) energy price in each period. Whenever the generation capacity of technology *i* during hour *j* is binding, the marginal cost of generation capacity $\mu_{i,j}$ can be positive. The KKT condition for consumer demand, given by decision variable d_j, corresponds to inverse demand function $\boldsymbol{P_j}$, representing consumer response. Eq. *(34)* indicates that the hourly (marginal) energy price λ_j is equal to the sum of variable generation cost and the marginal capacity cost if a plant is generating power. This can be interfaced with the inverse demand function by supposing $p_j = \lambda_j$ in (32), if demand is strictly positive.

$$0 \le VC_i - \lambda_j + \mu_{ij} \perp g_{i,j} \ge 0 \qquad \forall j \in J, \forall i \in I \qquad (34)$$
$$0 \le FC_i - \sum_j \mu_{ij} \perp cap_i \ge 0 \qquad \forall i \in I \qquad (35)$$
$$0 \le -E_j - \sum_k f_{jk} * d_k + p_j \perp d_j \ge 0 \qquad \forall j \in J \qquad (36)$$
$$0 \le cap_i - g_{i,j} \perp \mu_{i,j} \ge 0 \qquad \forall j \in J, \forall i \in I \qquad (37)$$

with market clearing condition:

$$0 = \sum_i g_{i,j} - d_j,\ \lambda_j\, free \qquad \forall j \in J \qquad (38)$$

with:
λ_j = dual variable or price associated with market clearing
$\mu_{i,j}$ = dual variable of the capacity constraint

This Linear Complementarity Program (LCP) uses the perpendicular operator "⊥" in order to indicate that at least one of the adjacent inequalities must be satisfied as an equality. This operator is defined as follows. A complementarity condition between a non-negative variable *x* and a non-negative function *f(x)* can be expressed as:

[14] Given perfect competition or a central planning approach, social welfare is maximized. The generation side is assumed to maximize profit given prices.

$$x \geq 0;\ f(x) \geq 0;\ x * f(x) = 0. \tag{39}$$

Using the "⊥" operator, Eq. (39) can be more compactly written as:

$$0 \leq x \perp f(x) \geq 0 \tag{40}$$

This model is a general representation of a linear complementarity problem, more specifically a Mixed LCP (MCP). The term "mixed" refers to the existence of both non-negative ($g_{i,j}$, cap_i, d_j and $\mu_{i,j}$) and free (λ_j) variables, associated with inequality and equality conditions, respectively [57]. As the number of all conditions gathered with the market clearing condition (Eq. *(34)-(38)*) equals the number of variables, it is called a "square problem" [56].

The sum of the capacity prices over different hours represents the cumulative value of additional generation capacity. This value provides a long-term signal for the optimal capacity, being a decision variable in the model. In equilibrium, the value of additional generation capacity is equal to the cost of installing additional capacity, Eq. *(35)*. No installed reserve margin constraint has been included into this model. A reserve margin (RM) is a constraint, requiring that the sum of installed capacity should be greater than or equal to (1+R) times the peak demand level (*PEAK)*. A reserve margin requirement could be integrated as Eq. (41), as suggested in [16]. This constraint deals with the uncertainty about peak demand levels and therefore improves the security of supply.

$$\sum_i cap_i \geq (1 + M) * Peak\_load \tag{41}$$

Such a condition results in an additional shadow price that represents the value of capacity for meeting that condition, and can make the gross margin (revenue minus variable cost) positive even if a generator never produces at capacity. The same accounts for the periodic maintenance requirement (PM), in which the generation output cannot exceed (1-PM) times the total installed capacity. This constraint is also suggested in [7] in the context of forced outages and results in an additional shadow price. The inclusion of the dual of the capacity constraint in Eq. *(34)* shows that price spikes appear at times of scarcity when no excess generation capacity is available. This practice of allowing prices to rise above marginal cost, also called "scarcity pricing", ensures that system energy balance is met by sending a signal to consumers that power is expensive and demand should be reduced.

The LCP methodology is used in [42] without inclusion of short-term operational constraints for thermal generation units. Adding the operational constraints (Eq. *(5)-(20)*) results in extra dual variables, as well as the new primal variables related to output flexibility ($flex^+_{i,j}$, $flex^-_{i,j}$), the energy storage unit ($spump_j$, $stored_j$, sg_j), transmission interconnection usage ($export_j$) and wind power injections ($wcap$, $wcurt_j$, $wind_j$). Including those constraints can impact the marginal cost (price) of electricity if any of the operational constraints are binding.

Including short-term demand response into a resource planning model using a complementarity formulation has the disadvantage that no 0-1 binary variables can be introduced. Such variables are required when including unit commitment related constraints, such as minimum run levels or minimum up and down times. This also means that complementarity models cannot represent discrete investments in new generation plants.

3.4.2 Quadratic program

A second formulation to integrate short-term demand response in a planning model is Quadratic Programming (QP). QP was applied in [58] to the problem of spatial price equilibrium calculation with linear supply and demand functions. In this context, the problem must be seen as a market equilibrium problem among producers and consumers, each maximizing their total surplus.

It is argued in [59] that continuous QP models (without binary variables) are a subset of LCP models because the KKT conditions for a quadratic program create a LCP or mixed LCP problem. However, the reverse is not true; not all LCPs can be formulated as QPs. If a LCP problem can be formulated as a QP, then standard QP or nonlinear programming software can be used to solve the problem, whereas LCP problems need specialised complementarity solvers.

The ability to reformulate a LCP, such as *(34)*-*(38)*, as a QP is valid under restrictive conditions, as is proven below. By definition, a fundamental LCP must find vectors w and z satisfying the following conditions:

$$w = q + M * z \tag{42}$$
$$z' * w = 0;\ z \geq 0;\ \omega \geq 0 \tag{43}$$

The variable w is a slack variable, typically added to a constraint in order to write an inequality as an equation. The slack variable is positive when the constraint $0 \leq M * z + q$ is non-binding, meaning that this constraint does not restrict the solution. Eq. *(43)* can be written more compact, by using the perpendicular operator, corresponding to the formulation in Eq. (39) and *(40)*.

$$0 \leq z \perp \omega \geq 0 \tag{44}$$

In order to prove the relationship between a QP and a fundamental LCP, a general QP is defined (Eq. *(45)*-*(47)*):

$$Max\ \boldsymbol{L}(x) = c^T * x + \frac{1}{2} * x^T * Q * x \tag{45}$$

subject to:

$$A * x \leq b \tag{46}$$
$$x \geq 0 \tag{47}$$

for which KKT conditions are derived.

$$u = -c + \frac{1}{2} * (Q + Q^T) * x + A^T * \lambda \tag{48}$$
$$v = b - A * x \tag{49}$$
$$0 \leq u \perp x \geq 0 \tag{50}$$
$$0 \leq v \perp \lambda \geq 0 \tag{51}$$

Let:

$$w = \begin{bmatrix} u \\ v \end{bmatrix}, \, q = \begin{bmatrix} -c \\ b \end{bmatrix}, \, M = \begin{bmatrix} \frac{1}{2} * (Q + Q^T) & A^T \\ -A & 0 \end{bmatrix}$$

By defining matrices *w*, *q* and *M*, it is clear that the KKT conditions of the QP (Eq. *(48)-(51)*) are equivalent to the fundamental LCP formulation (Eq. *(42)-(44)*). The definition of matrix *M* indicates that the equivalence is only valid when Q is symmetric and positive semi-definite:

$$\frac{1}{2} * (Q + Q^T) = Q \; if \; Q = Q^T \tag{52}$$

Correspondingly, the LCP with integrated short-term demand response can be reformulated as a QP if demand and/or supply functions are linear and the coefficient matrix is symmetric. An increase in the price of electricity must result in a reduction of electricity consumption, and analogously for supply considerations. This assumption is referred to as the integrability condition [60]. If the demand function is not symmetric, the integrability condition is not satisfied and the social welfare function cannot be constructed [61]. The QP is a representation of the equilibrium problem in Eq. *(26)*, in the form of a welfare maximization problem.

$$Max \; Welfare(Q) = \boldsymbol{U}(Q) - \boldsymbol{V}(Q) \tag{53}$$

This results in the simplified resource planning model with demand response:

$$Max \; Welfare = \left\{ \sum_j [d_j * e_j + \frac{1}{2} * d_j * \sum_k f_{j,k} * d_k] \right\} - \left\{ \sum_i cap_i * FC_i + \sum_{i,j} g_{i,j} * VC_i \right\} \tag{54}$$

subject to:

$$\sum_i g_{i,j} - d_j = 0 \qquad \forall j \in J \tag{55}$$

$$g_{i,j} \leq cap_i \qquad \forall i \in I, j \in J \tag{56}$$

According to the integrability condition, the coefficient matrix *F*, introduced in Eq. *(23)* with parameter $f_{j,k}$, must be symmetric to use the QP approach.

On the one hand, the QP has the advantage that adding more constraints does not require introducing more dual variables into the model. Additionally, formulating this problem as a QP is motivated by the wide availability of nonlinear optimization software. On the other hand, when the demand system does not satisfy the symmetry condition, an equivalent QP cannot be formulated, and the model should be solved as a complementarity problem. Alternatively, a symmetric matrix could be constructed as a close approximation of the actual matrix and used in the QP optimization.

3.4.3 Piecewise integration

Linear demand functions can be used together with a LP supply model to calculate a market equilibrium by reformulating the problem as an LCP or QP, as described in the previous subsections. However, it might be difficult to add consumer benefits to models with more computational complexities such as non-linear constraints or objective functions or binary variables. Therefore an alternative computational procedure is suggested in this subsection in order to find an equilibrium solution for a given supply and price elastic demand function.

Non-zero, cross-price elasticities can be added, assuming dominance of own-price elasticities. ‘Dominance’ means that own-price elasticities are larger than the sum of cross-price elasticities. The methodology is based on the procedure in [62] and convergence is mathematically proven in [61].[15] It is known as the ‘PIES algorithm’ because it was used to solve the Project Independence Evaluation System [63], the first comprehensive energy-economic model used by the US government. Project Independence was initiated by U.S. President Nixon in 1973, in response to the OPEC oil embargo.

Starting from the cost minimization objective function, a piecewise approximation of the welfare function is created that accounts for the marginal effects of changes in quantity upon market welfare. The optimal hourly demand levels d_j are chosen such that market surplus (the integral of the demand function, minus the generation and investment costs) is maximized. An iterative procedure solves the LP until the algorithm converges to the equilibrium solution. This equilibrium solution is the same as the one obtained by the LCP and QP approaches, with own-price elasticities higher than the aggregated cross-price elasticities, using a symmetric matrix as a close approximation of the actual matrix.

In order to find an appropriate solution to this problem, perturbations $y^+_{j,n}$ and $y^-_{j,n}$ are introduced, defined as the difference between initial demand level DEM_j for each hour and a new demand level d_j. These continuous, positive variables $y^+_{j,n}$ and $y^-_{j,n}$ allow building a partition of the interval around the anchor point with the initial demand level DEM_j. Given set N ($n= 1,\dots,m-1,m$) $y^+_{j,n}$ constructs m steps in the demand function approximation on the right-hand side of initial demand level DEM_j and $y^-_{j,n}$ constructs m steps at the left-hand side of initial demand level DEM_j (Figure 3). Variables $y^+_{j,n}$ and $y^-_{j,n}$ are constrained by $U^+_{j,n}$ and $U^-_{j,n}$ respectively (Eq. *(57)* and *(58)*), being the maximum step size on the right- and left-hand side. The step size can arbitrarily be chosen. It does not have to be the same on the left and on the right of DEM_j, neither does it have to be the same for each step m.

$$0 \leq y^+{}_{j,n} \leq U^+{}_{j,n} \tag{57}$$

$$0 \leq y^-{}_{j,n} \leq U^-{}_{j,n} \tag{58}$$

For each step around the initial demand levels, the inverse demand function $\boldsymbol{P_j}$ gives the resulting approximation to the price level $P^+_{j,n}$ and $P^-_{j,n}$.

[15] The computational procedure can also be found in a world oil market model [84] and in [85]. The formulation is summarized in this paper. For a detailed description of the procedure, the reader is referred to the appendix of [62]. More information on the convergence of the PIES algorithm can be found in [61].

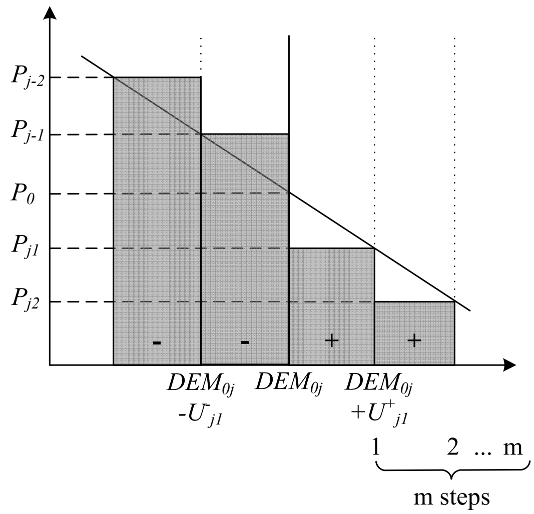

Figure 3: Partitioning for piecewise integration

$$P^{+}{}_{j,n} = \boldsymbol{P_j}\left(DEM_j + \sum_{n=1}^{m} U_{j,n}^{+}\right) \tag{59}$$
$$P^{-}{}_{j,n} = \boldsymbol{P_j}\left(DEM_j - \sum_{n=1}^{m} U_{j,n}^{-}\right) \tag{60}$$

Then, the integral calculating consumer value of consumption is approximated by a piecewise summation of the one-dimensional integrals.

$$\int_0^{DEM_j+y_j} \boldsymbol{P_j}(q)\, dq \approx \int_0^{DEM_j} \boldsymbol{P_j}(q)\, dq + \sum_{n=1}^{m}\left[\left(P^{+}{}_{j,n} * y^{+}{}_{j,n}\right) - \left(P^{-}{}_{j,n} * y^{-}{}_{j,n}\right)\right] \tag{61}$$

With:

$$y_j \approx \sum_{n=1}^{m}\left(y^{+}{}_{j,n} - y^{-}{}_{j,n}\right) \tag{62}$$

Eq. *(61)* indicates that increasing the demand level, when $y^{+}{}_{j,n}$ is greater than zero, increases the welfare. Correspondingly, decreasing demand levels, when $y^{-}{}_{j,n}$ is greater than zero, results in a decreasing consumer welfare. The resulting approximations to the changes in consumer value (integral of the demand curve) are illustrated by the gray rectangles in Figure 3. This equation is added to the objective function in order to maximize total welfare. Since the integral on the right-hand side of *(61)* is a constant, it can be omitted. The system energy balance requirement is then changed as well, shown in Eq. *(63)* and *(65)*. Given that the left-hand side of Eq. *(63)* is concave, the pieces of the piecewise linear approximation will come into the solution in the correct order and (approximately) find the welfare maximizing equilibrium. The performed perturbation does not influence the formulation of operational constraints. The resulting model is:

$$Max\ \sum_j \sum_{n=1}^{m}\left[\left(P^{+}{}_{j,n} * y^{+}{}_{j,n}\right) - \left(P^{-}{}_{j,n} * y^{-}{}_{j,n}\right)\right] - \sum_i cap_i * FC_i + \sum_{i,j} g_{i,j} * VC_i \tag{63}$$

subject to

$$g_{i,j} \le cap_i \qquad \forall i \in I, \forall j \in J \tag{64}$$
$$\sum_i g_{i,j} - \sum_k \left(y^{+}{}_{j,n} - y^{-}{}_{j,n}\right) = DEM_j \qquad \forall j \in J \tag{65}$$

A solution to this model will yield optimal values for decision variables $y^{+}_{j,n}$ and $y^{-}_{j,n}$, the integrated demand level, and this will be an approximate solution to the welfare maximization problem. The adjusted demand levels are recalculated by Eq. *(66)*

$$d_j = DEM_j + \sum_{n=1}^{m} \left(y^{+}{}_{j,n} - y^{-}{}_{j,n} \right) \qquad \forall j \in J \qquad (66)$$

Concavity of the demand curve integral ensures that it cannot be welfare maximizing to simultaneously increase and decrease the initial demand levels. Consequently, the positive and negative perturbation variables $y^{+}_{j,n}$ and $y^{-}_{j,n}$, respectively, will not be simultaneously be different from zero in an optimal solution. The following constraint is automatically satisfied by the optimal solution and does not need to be explicitly put into the model:

$$y^{+}{}_{j,n} * y^{-}{}_{j,n} = 0 \qquad \forall j \in J, \forall n \in N \qquad (67)$$

Furthermore, if $y^{+}_{j,n}$ is positive, it can be equal to or less than the maximum step size $U^{+}_{j,n}$. If the perturbation variable equals the step size, the welfare maximizing equilibrium solution might not yet be found. If the perturbation variable is less than the step size, the optimal number of steps n^* has been found, given convexity of Eq. *(63)*. The dual variable of the system energy balance constraint λ_j is an estimate of the (marginal) energy price. The quality of the estimate depends on the degree of dominance of own-price elasticities in the inverse demand function $\boldsymbol{P_j}$ [61], as well as the width of the steps in the approximation.

In case of zero cross-price elasticities, λ_j equals the energy price, subject to an approximation error. In case of non-zero cross-price elasticities, an iterative procedure is suggested by [62], which is guaranteed to converge to the solution with equilibrium supply and demand quantities, as well as the market price, if the dominance condition of own-price elasticities is satisfied.

If K and M in Eq. *(27)* and *(28)* are $j*j$ real matrices, with invertible K, and D the diagonal of the $j*j$ matrix M. The global convergence condition is given as follows [61]:

$$\left\| I - D^{-1/2} B D^{-1/2} \right\| < 1 \qquad (68)$$

The flow of this procedure is schematically represented in Figure 4, and the steps are summarized below.

1. *Assume initial demand levels DEM_j:*

These demand levels are typically given for the model without inclusion of demand elasticity.

2. *Calculate specific price levels $P^{+}_{j,n}$ and $P^{-}_{j,n}$ for each step:*

Given the most recent estimated demand levels, price levels are calculated for different steps around that demand using the inverse demand function as in Eq. *(59)* and *(60)* for use in the supply LP model. If desired, the step size could be changed, e.g., reduced with each iteration. The inverse demand function approximation does not take into account cross-price elasticities. Only own-price elasticities are included.

3. Solve supply LP for optimal values $y^{+}_{j,n}$, $y^{-}_{j,-n}$ *and marginal price estimate* λ_j*:*
Based on the integration approximation described above, optimal values for the perturbation decision values are defined. If the dual variable of the system energy balance constraint λ_j equals $P_{j,n}$ for the optimal number of steps n, [62] suggests that the equilibrium solution is found. If not, use λ_j to calculate new demand levels using the actual inverse demand function with inclusion of cross-price elasticities. Then replace the initially chosen demand levels with the new demand levels and return to step 2.

4. Define the equilibrium solution:
By using Eq. *(66)*, the optimal demand levels can be calculated given $y^{+}_{j,n}$, $y^{-}_{j,n}$.

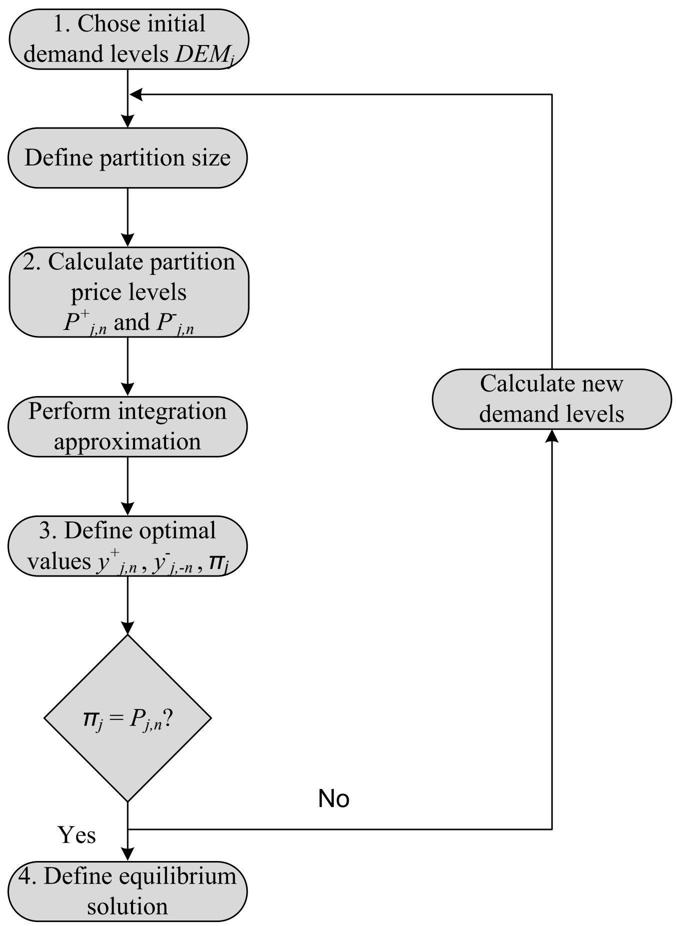

Figure 4: Flowchart iterative procedure

This algorithm can approximate a nonintegrable problem by a sequence of integrable problems. When applying the piecewise linearization approximation method, the QP or mixed complementarity problem can be solved by reformulating it as a linear problem, for which very efficient optimization software is available. This method has been applied in large-scale applications, exhibiting excellent computational characteristics [64]. Additionally, 0-1 binary variables for unit commitment or new plants can be included, in contrast to the complementarity method.

3.5 Impact of energy efficiency programs

In Section 3.4, short-term demand response has been integrated into the generation planning model in the form of elastic demand functions, allowing customers to change their consumption behaviour in response to real-time price signals. Extensive investment in energy efficiency can impact the energy demand function as well.

In this paper, long-term demand response is neglected, assuming that energy efficiency is fully driven by regulatory programs, implying utility and governmental spending. Pursuing energy efficiency reduces hourly electricity consumption. Additionally, extensive regulatory energy efficiency programs can also impact the responsiveness of demand. On the one hand, positive overlaps can be seen, e.g., when consumers become more conscious of their energy consumption or buy appliances offering more demand response flexibility [65]. On the other hand, energy efficiency and demand response might also have counteracting effects. Responsiveness of a consumer's load profile can be reduced, as switching off more energy efficient appliances in response to higher spot prices will result in smaller load reductions. Another potential conflict arises when customers participating in demand response programs are paid on the basis of the amount of load reduced when called upon. If reductions are measured from their average consumption level, they face an incentive to increase their baseline levels and a disincentive to become more energy efficient [66], [67].

Therefore, it is desirable to extend the elastic linear demand functions used in the models in order to account for interactions with energy efficiency programs. In this section, the elastic demand function is simplified to account only for own-price elasticities,[16] neglecting cross-price elasticities. Our starting point is to view the linear, hourly demand function $\boldsymbol{D_j}$ (Eq. *(22)*) as a Taylor series approximation with the second and higher order terms being dropped. The general form of the Taylor series approximation is given:

$$\boldsymbol{D_j}(p_j) = \sum_{n=0}^{\infty} \frac{D_j^n(P_{0,j})}{n!}\left(p_j - P_{0,j}\right)^n \tag{69}$$

The Taylor series representation is now extended by having a nonzero second-order terms that account for utility or government expenditures energy efficiency (*EE*), as a percentage related to current expenditures. The current energy efficiency expenditure equals 100% and is referred to as EE_0. The new demand function can be approximated by Eq. *(69)*.

$$\begin{gathered}\boldsymbol{D_j}(p_j, EE) = D_j(P_{0,j}, EE_0) + \\ \frac{\partial D_j(P_{0,j}, EE_0)}{\partial p_j}\left(p_j - P_{0,j}\right) + \frac{\partial D_j(P_{0,j}, EE_0)}{\partial EE}(EE - EE_0) + \\ \frac{\partial^2 D_j(P_{0,j}, EE_0)}{\partial {p_j}^2}\frac{\left(p_j - P_{0,j}\right)^2}{2} + \frac{\partial^2 D_j(P_{0,j}, EE_0)}{\partial EE^2}\frac{(EE - EE_0)^2}{2} + \\ \frac{\partial^2 D_j(P_{0,j}, EE_0)}{\partial p_j \partial EE}\frac{\left(p_j - P_{0,j}\right) * (EE - EE_0)}{2}\end{gathered} \tag{70}$$

[16] The inclusion of cross-price elasticities in demand functions that account for interactions with energy efficiency will be the subject of future work.

Disregarding cross elasticities with respect to prices at other times, the first derivative with respect to price and energy efficiency expenditure is negative. The second own derivatives of the demand function can introduce scale effects, but are assumed to be equal to zero. The last term, the cross second partial, is used to account for interactions. Based on this logic, we replace the following derivatives with the price elasticity of demand ε_j, efficiency elasticity of demand γ_j, and the cross-price-efficiency elasticity of demand δ_j:

$$\frac{\partial D_j(P_{0,j},EE_0)}{\partial p_j} = \varepsilon_j \frac{DEM_j}{P_{0,j}} \tag{71}$$

$$\frac{\partial D_j(P_{0,j},EE_0)}{\partial EE} = \gamma_j \frac{DEM_j}{EE_0} \tag{72}$$

$$\frac{\partial^2 D_j(P_{0,j},EE_0)}{\partial p_j \partial EE} = \delta_j \frac{DEM_j}{P_0,j} \frac{DEM_j}{EE_0} \tag{73}$$

The efficiency elasticity of demand γ_j shows the impact of energy efficiency expenditures. Said differently, this parameter indicates to what extent those expenditures affect electricity demand. The cross-price-efficiency elasticity of demand δ_j indicates to what extent increased energy efficiency expenditures affect the short-term responsiveness (elasticity) of demand.

We assume zero second own derivatives of the demand function and efficiency expenditures (*EE*) included as a parameter instead of a decision variable. This results in the following final form of the demand function, including interactions with efficiency:

$$\begin{gathered} \boldsymbol{D_j}(p_j) = D_j(P_{0,j}, EE_0) + \\ \varepsilon_j \frac{DEM_j}{P_0}(p_j - P_{0,j}) + \gamma_j \frac{DEM_j}{EE_0}(EE - EE_0) + \\ \delta_j \frac{DEM_j}{P_0} \frac{DEM_j}{EE_0} \frac{(p_j - P_{0,j}) * (EE - EE_0)}{2} \end{gathered} \tag{74}$$

Finally, Eq. *(74)* can be simplified as a short-term demand response function or an inverse demand function, corresponding to Eq. *(22)* and *(23)* respectively. Consequently, Eq. *(74)* can be implemented in any of the three solution methods outlined in section 3.4.

4. Case Study and Results

4.1 Data and assumptions

In this paper, four generation technologies are taken into account, i.e., base, mid, peak, and high peak load, each having different costs. Ordering the technologies in terms of decreasing capital cost and increasing operating cost, the first two technologies are nuclear and coal units, respectively, whereas peak and high peak load technologies correspond to Combined Cycle Gas Turbines (CCGT) and oil- or gas-fired open cycle gas turbines.

Table 1: Generation technology type costs

		Base	Mid	Peak	High Peak
Cost category					
Investment	[k€/MW/yr]	155	65	25	15
Fixed O&M	[k€/MW/yr]	65	35	15	10
Fuel	[€/MWh]	10	20	35	65
Variable O&M	[€/MWh]	5	10	10	10

Their assumed costs are inspired by data from the International Energy Agency [68]. Although these values are low by today's standards, the relative cost levels for different technologies and cost categories are representative and serve to illustrate the methodology.

The described model uses historical wind power and demand data on an hourly demand step.[17] The operational cost minimization is done for a 4 week period, corresponding to 672 consecutive hourly load and wind power levels. Annualized fixed conventional generation costs (Table 1) are scaled considering these 672 hours. For this period wind generated power output (WP_j) is on average 30%, between 0.5% and 94% as minimum and maximum output levels, respectively. Average demand levels are 5200 MW, fluctuating between a minimum of 3050 MW and a maximum of 7600 MW. The amount of wind power capacity installed (*wcap*) is a decision variable, depending on the investment costs. Annualized investment costs range from 40 k€/MW/yr up to 100 k€/MW/yr. In correspondence with conventional generation costs, wind power investment costs are scaled considering the 672 hour period.

For this illustrative example, a cost of 100 €/MWh for wind power curtailment is included, inspired by negative prices observed in the German and Danish energy market. A 250 MW pump/turbine capacity is assumed as well as a 250 MW transmission interconnection. In an interconnected market, however, the export price would be determined by the simultaneous interplay of supply and demand in all markets. Future work will extend this to a transmission constrained model of multiple markets. For the purpose of illustration, however, this interconnection is only used to export energy at the price of 0 €/MWh. This means that no wind power curtailment occurs unless the interconnector is fully used. Different ramp rates, levels of transmission interconnection and pump/turbine capacities are used to illustrate the impact of these parameters on the model results. The sensitivity analysis also compares model results for different interconnector capacities.

For each optimization it is assumed that total energy storage capacity corresponds to 5 hours pumping up water at nominal capacity. 90% efficiency is considered whenever the storage unit is used to pump or generate.[18] Periodic maintenance (*PM*) is set to 90% and a must-run requirement of 10% of the total installed capacity is included for base and mid load generation technologies. Although this is a low requirement on a per unit basis, it is nevertheless realistic when the total installed capacity represents several units. In that case, with some units being decommitted, the 10% must-run

[17] Hourly data for electricity demand and a wind power profile are published on the website of the Danish system operator, Energinet: www.energinet.dk.

[18] The efficiency of the total cycle of energy storage is 81%.

requirement for total capacity would correspond to a higher per unit requirement for the committed capacity.

Ramp rates are introduced as a percentage of the total installed capacity of a technology (Table 2). Ramp rates are included in [69], allowing coal units to completely ramp up or down in four hours. Reducing the ramping time is allowed by considering a ramping penalty as the cost of shortening the service life of the turbine rotor. Full ramp up or down times of 3 hours are used in [27] for coal and nuclear facilities and 2 hours for gas and petroleum fired combined cycle generators. Additionally, a distinction can be made between plant equipment vendors, who are typically more optimistic about ramp rates, and what plant operators actually do. Some plant operators report ramp rates being 2.5 up to 5 times lower than those suggested by the plant vendors [70].

Furthermore, actual ramp rates depend on plant loading and reportedly have as little as 60% of the full ramping capacity when a plant is operated at lower output levels [70]. Note that the model works on an aggregated technology level, rather than on a power plant basis. Hence, the introduced ramp rates should not be directly compared to ramp rates of individual power plants, which might be higher, but should be viewed as feasible variations in the output of a collection of plants of the same technology. This reasoning is supported by considering the impact of start-ups. Start-up costs, minimum run and minimum on/off times might make plant operators averse to starting up all individual units for generating during only a few or even one single hour. Consequently, the aggregated generation output flexibility could be considered to be lower than implied by individual ramp rates. The ramp rates used (Table 2), are inspired by a literature review ([27],[69],[70]) and discussions with experts.[19] Balancing power requirement data, accounting for the uncertainty in the wind and demand profile, are summarized in Table 3, and are based on [71] and [54].

Average and maximum balancing power requirements are suggested in [71] for positive as well as negative regulation and illustrated in Figure 5. For positive regulation, a balancing power requirement (BAL^+_j) equal to 9% of the wind power capacity should be kept available on average, corresponding to 20% wind power injections of the installed wind power capacity. During high wind situations, the positive balancing power requirement increases up to 19% of wind power capacity (*wcap*) (Eq. *(17)*). For negative regulation, a balancing power requirement (BAL^-_j) of 8% of the wind power capacity should be kept available on average. During high wind situations, the negative balancing power requirement increases up to 15% of wind power capacity (Eq. *(18)*).

Table 2: Technology specific ramp rates

Technology	Ramp rate committed capacity [%/hour] ($RAMP\_C_i$)	Ramp rate non-committed capacity [%/hour] ($RAMP\_NC_i$)
High peak load	100	100
Peak load	80	80*60%
Mid load	50	50*60%
Base load	16.7	16.7*60%

[19] The authors are grateful for suggestions by and discussions with Daniel Kirschen (The University of Manchester) and Yann Rebours (EDF).

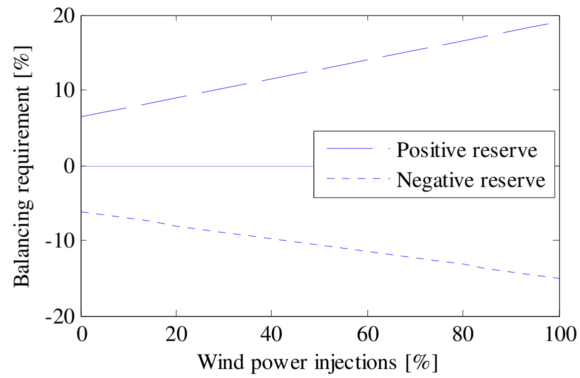

Figure 5: Balancing power requirement as a function of hourly wind power injections

Table 3: Balancing power requirements

	Average [%]	Maximum [%]
Positive regulation	9	8
Negative regulation	19	15
Linearization	*A* [%]	*B* [%]
Positive regulation (_*POS*)	12.5	6.5
Negative regulation (_*NEG*)	8.75	6.25

In this analysis, own-price elasticities of demand of -5% and -10% are tested (the low and high elasticity cases in the figures). These are complemented by positive cross-price elasticities. These numbers are comparable to data in [72] and the overview given in [73], after rescaling for transmission and distribution charges. We also consider a range of cross elasticities, with magnitudes of 0%, 0.5% or 1% in each of the previous and subsequent 4 hours to ensure symmetry. For each scenario, own-price elasticity is assumed to dominate, meaning that own-price elasticities are larger than the sum of cross-price elasticities of the previous and subsequent 4 hours. This assumption is in correspondence with empirical data, suggesting that cross-price elasticities are typically larger than own price-elasticities [74]. Above mentioned hourly cross-price elasticities result in an aggregate cross-price elasticity of 0%, 4% and 8% respectively, inspired by [75]. Cross-price elasticity allows consumers to shift a part of their consumption behaviour in time. As symmetry conditions must be satisfied for the QP model, the corresponding coefficient matrix is made symmetric. Thus, there are cross elasticities for the four previous hours of the same magnitude; this represents a situation in which consumers have foreknowledge of hourly prices and reschedule their loads earlier as well as later to avoid high prices.

As mentioned above, utility or governmental expenditures for energy efficiency (*EE*) are indicated as a percentage related to current expenditures. The current level of expenditures (EE_0) equals 100%. In this paper, a constant as well as a 50% increased level of expenditures is assumed, linked to a 2.5% and 5% efficiency elasticity parameter γ_j. Evidence concerning the impact of energy efficiency expenditures on consumption can be found in [76] and [77]. In order to account for interactions between demand response and energy efficiency, a 0% and 0.5% cross-price-efficiency elasticity of demand δ_j has been assumed. This parameter is also referred to as the mixed derivative parameter in Eq. *(70)*.

4.2 Reference scenario

First, the optimal generation technology mix is calculated for a reference scenario (Table 4). This scenario does not account for real-time demand response, assuming that demand is fixed. As the market is entirely cleared by supply-side measures with no DSM or demand response, more generation investments have to be made than in the demand response scenarios.

As an example of the demand and wind assumptions, initial load and wind power generation levels are shown for a representative week in the upper graph of Figure 6. The wind power generation profile is multiplied by the optimal installed wind power capacity, expressed in MW. The optimal capacity levels are shown in Table 4 for different levels of wind investment costs. The lowest cost assumption (40 k€/MW/yr) incents the most optimal installed wind capacity, indicated as "low wind cost" (dashed line) for each of the graphs in Figure 6. A much higher investment cost (100 k€/MW/yr), shown as "high wind cost" (full line) for each of the graphs in Figure 6, reduces optimal wind capacity by almost half. Those are two illustrations of the four scenarios in Table 4. The results for the other two scenarios are in between the full and the dashed line, but are left out of the figure for clarity reasons. The bottom graph in Figure 6 shows the corresponding real-time electricity price, without allowing consumers to respond to that price.

By subtracting wind power generation from initial load levels, a net demand profile is found (the middle graph in Figure 6). Wind power curtailment is allowed in order to eliminate excess injections during high wind periods, e.g., around hour 75 in the high wind scenario when price plunges to the curtailment cost of -100 €/MWh. The optimal generation technology mix is given in Table 4, both with and without ramp rates of the different technologies taken into account. In the scenario that disregards ramp rate limits, additional wind capacity serves mainly to displace base load capacity. This corresponds to the findings in [42]. Including ramp rates and thus decreasing generation flexibility results in further reductions of base load generation capacity. Thus, as we hypothesized, accounting for generator flexibility affects long run investment decisions. Base load technologies are not even part of the optimal mix after including ramp limits under the lowest cost wind power scenario. Compensating for the loss of base load capacity, the amount of mid load technologies increases as it offers more flexibility than base load generation technologies.[20] This greater flexibility is required to accommodate the high variability of net demand after subtracting wind power. Finally, the total installed generation capacity increases for lower wind power investment costs, as wind has a relatively low average capacity factor. The total installed conventional generation capacity as well as the optimal wind power capacity is comparable with or without taking into account ramp rates.

[20] If carbon emissions were capped, however, the story would become more complex, because it would not be possible to simply replace low emission nuclear capacity with high emission coal capacity; it is likely that the natural gas technologies would increase their capacity as much or more than the coal technologies in those cases.

Table 4: Optimal generation technology capacity in the reference scenarios [MW]

Wind Cost Scenario	Base	Mid	Peak	High Peak	Wind	Total
No ramp rate limits						
40 k€/MW/yr	1,414	2,359	2,067	1,225	5,889	12,954
60 k€/MW/yr	1,726	2,179	1,967	1,230	5,424	12,527
80 k€/MW/yr	2,046	2,034	1,831	1,236	4,874	12,021
100 k€/MW/yr	3,046	1,577	1,565	1,092	3,233	10,511
With ramp limits						
40 k€/MW/yr	0	3,797	2,042	1,225	5,897	12,962
60 k€/MW/yr	901	3,058	1,916	1,231	5,391	12,496
80 k€/MW/yr	1,728	2,420	1,786	1,221	4,776	11,931
100 k€/MW/yr	2,969	1,678	1,581	1,067	3,125	10,421

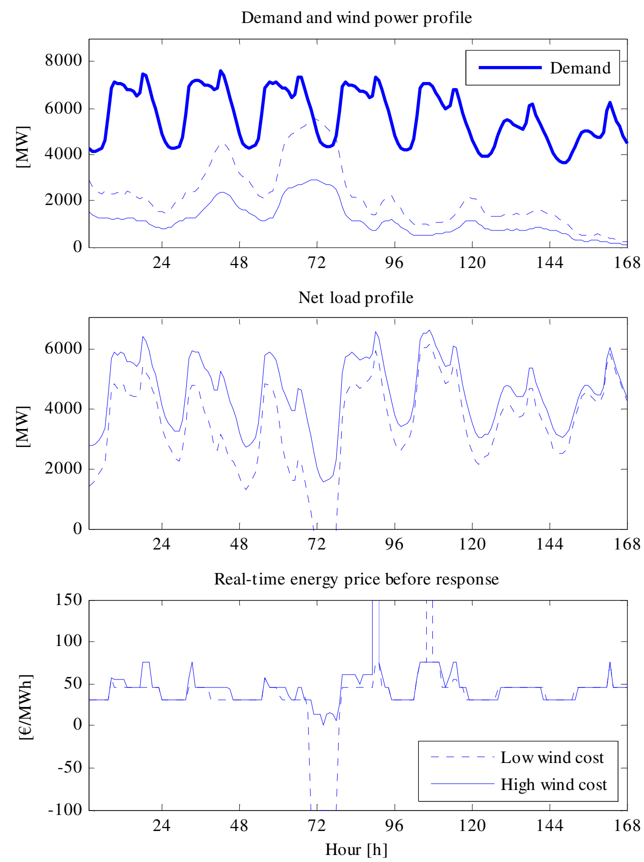

Figure 6: Load, wind generation, and prices assuming no demand response under alternative wind cost assumptions

4.3 Impact of demand elasticity

The present analysis focuses on the long run implications of demand-response under conditions in which all capacity is variable (static, single period optimization). This may impact the estimated benefits of demand-response relative to what they would be if a large fraction of capacity is predetermined as an existing generation fleet. The long run optimal mix is nevertheless useful as a benchmark.

By integrating demand elasticity, consumers are able to adjust their consumption in response to real-time price signals. The three approaches for including short-term demand response into a resource planning model have been tested and gave the same results, tested for a 168 hour period given -10% own-price elasticity with a fixed wind power capacity. The QP and LCP model only require one optimization instead of several optimization iterations when the piecewise integration is used. As additional constraints can more easily be added to the QP then to the LCP model, the results, presented in this paper, have been calculated by using the QP model. However, the limitation of the QP is that it requires symmetric elasticity assumptions, which we have made.

With demand responsive consumers, lower and higher net demand levels result in lower and higher electricity prices, respectively, in a particular hour. The MWh weighted average of the corresponding price levels for the above reference case is calculated and shown as the flat tariff in Figure 7. Demand responsive consumers benefit from increasing their electricity consumption during low price hours, e.g., between hour 20 and 32, and from decreasing their consumption during high price hours, e.g., between hour 42 and 45. Demand levels are increased during low demand hours, a phenomenon referred to as valley filling. Similarly, demands are reduced during peaks in response to high prices, which is called peak shaving.

The effect of including own- and cross-price elasticities is illustrated in Figure 8 for an example two day period. Given the initial flat tariff in Figure 7, the net demand/no response profile is shown in the bottom graph. During peak demand moments, e.g., hour 43, price spikes can be seen. Price-responsive consumers react to real-time prices rising above the flat tariff, by reducing initial demand levels. On the other hand, during moments with high wind power injections, e.g., hour 25 until 30, negative prices are observed. Correspondingly, demand levels are significantly increased as a matter of integrating excess wind power generation.

As a result of own-price elasticities, demand increases during nighttimes with low price periods, as indicated by the full line. Additionally, the complex effects of cross-price elasticity (dashed line) become apparent. Sometimes, consumer demand response is weakened with the dashed line lying between the no response and own elasticity only scenarios. At other times, cross elasticities increase the aggregate response, with the net load being pushed further from the no response case than with own elasticities only; consumer demand response is strengthened. The former situation would occur when the price in hour j as well as in the subsequent hours is above the weighted average price. Own-price elasticity effects counteract cross-price elasticity effects. The latter situation would occur when the price in hour j is lower than the weighted average price and the price in the subsequent hours $j+1$ until $j+4$ is higher than the weighted average price; consumer demand response is strengthened in hour j.

Peak demand reductions shown in Figure 8 are consistent with values found in literature. Based on [72], a potential for peak reduction from demand response up to 16% can be expected. Higher peak demand reductions would occur in regions with hot summers, high saturation of air conditioning systems and deficient capacity, resulting in high price signals. These numbers refer to the reduction potential for aggregate peak demand. Actual peak demand reductions between 8.5% and 18.5% are documented in [78], given different customer characteristics.

The three variables that affect energy reduction the most are air conditioning ownership, college education, and annual income. People with more education or income show a higher percent impact. During off-peak periods, minor demand increases between 0 and 4% were observed.

A sensitivity analysis of the effect of demand elasticity upon installed capacity is summarized in Table 5. As ramp limits have been included in these runs, a comparison with Table 4 shows the impact of including demand response. Firstly, most remarkable is the 50% reduction of the installed high peak generation capacity. Demand response often clears the market during peak periods, significantly reducing peak demands and the need for such capacity. With an own-price elasticity of -10% without the cross-price effect, the installed high peak capacity falls to as low as 312 MW. In contrast, without demand response, more than 1,000 MW was required for the lowest installed wind power capacity.

Secondly, for the least installed wind capacity case, the optimal base load capacity is increased by about 3 to 8%. This corresponds to an absolute increase of about 80 to 240 MW. The variability of the net demand profile is reduced, as well as the need for system flexibility. Thirdly, higher price elasticities yield a much higher optimal installed wind power capacity. The optimal installed wind power capacity can increase by more than 3% for higher investment costs (100 k€/MW/yr) and by up to more than 17% for the low investment cost scenario (40 k€/MW/yr). This illustrates the contribution of demand response to the integration of intermittent renewable energy generation.

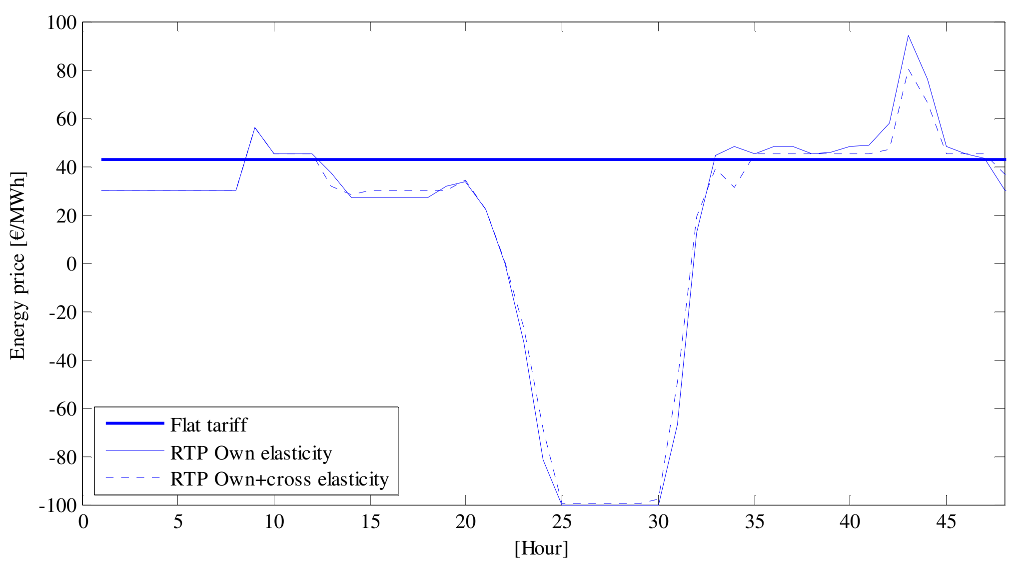

Figure 7: Price comparison: -10% own/+2% cross-price elasticity (wind investment cost: 40 k€/MW/yr)

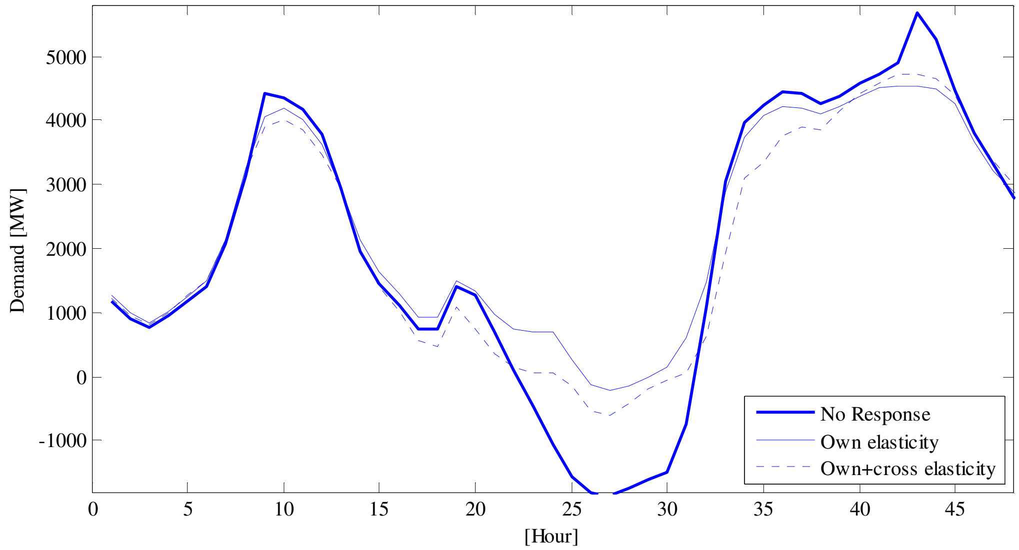

Figure 8: Net demand response comparison: -10% own/+2% cross-price elasticity (wind investment cost: 40 k€/MW/yr)

Table 5: Price elasticity sensitivity analysis

Scenario	Base	Mid	Peak	High Peak	Wind	Total
- Own-price elasticity -5% / cross-price elasticity 0%						
40 k€/MW/yr	0	3,668	1,668	560	6,381	12,276
60 k€/MW/yr	469	3,357	1,604	529	5,853	11,812
80 k€/MW/yr	1,557	2,498	1,490	483	5,182	11,210
100 k€/MW/yr	3,097	1,565	1,267	335	3,176	9,441
- Own-price elasticity -5% / cross-price elasticity 1%						
40 k€/MW/yr	0	3,760	1,741	567	6,112	12,180
60 k€/MW/yr	605	3,303	1,664	533	5,637	11,743
80 k€/MW/yr	1,533	2,548	1,571	485	5,085	11,223
100 k€/MW/yr	3,054	1,594	1,350	355	3,171	9,524
- Own-price elasticity -10% / cross-price elasticity 0%						
40 k€/MW/yr	0	3,535	1,357	565	6,951	12,408
60 k€/MW/yr	0	3,719	1,314	529	6,336	11,898
80 k€/MW/yr	1,433	2,547	1,233	463	5,490	11,166
100 k€/MW/yr	3,207	1,475	972	312	3,246	9,212
- Own-price elasticity -10% / cross-price elasticity 1%						
40 k€/MW/yr	0	3,625	1,502	549	6,581	12,258
60 k€/MW/yr	233	3,558	1,453	509	6,047	11,800
80 k€/MW/yr	1,476	2,544	1,357	465	5,320	11,163
100 k€/MW/yr	3,151	1,523	1,108	316	3,204	9,302
- Own-price elasticity -10% / cross-price elasticity 2%						
40 k€/MW/yr	0	3,717	1,506	531	6,356	12,110
60 k€/MW/yr	141	3,705	1,458	500	5,950	11,754
80 k€/MW/yr	1,425	2,636	1,349	458	5,280	11,148
100 k€/MW/yr	3,142	1,531	1,113	325	3,179	9,290

Finally, for a given level of own-price elasticity, an increased cross-price elasticity reduces the above mentioned effects because now a price increase in a given hour results not only in a load decrease in that hour, but some compensating load increases in earlier and later hours, given symmetry of cross-price effects. When several consecutive hours have similar prices, this means that the net effect of higher prices in those hours is less than if only own-elasticities are under consideration. A reduced net effect of price fluctuations results in reduced demand flexibility. Correspondingly, increasing the cross-price elasticity reduces the optimal installed wind power capacity, compared to the scenario without cross-price elasticities.

The impact of demand elasticity on the weighted average bulk electricity price is shown in Table 6. These results exclude transmission and distribution charges. The reference price level corresponds to the no response scenario. Consumers facing real-time electricity prices are encouraged to consume more during low price hours and less during high price hours. Consequently the weighted average price decreases 1.5 to 3.5 €/MWh.

Table 6: Weighted average electricity price impact [€/MWh]

Scenario	40 k€ /MW/yr	60 k€ /MW/yr	80 k€ /MW/yr	100 k€ /MW/yr
No response	36.87	39.18	41.27	43.01
Own -5	34.80	37.67	40.17	42.14
Own -5%/cross +1%+1%+1%+1%	35.32	37.95	40.32	42.23
Own -10	33.38	36.73	39.62	41.74
Own -10%/cross +1%+1%+1%+1%	34.14	37.20	39.88	41.91
Own -10%/cross +2%+2%+2%+2%	34.41	37.31	39.93	41.91

4.4 Impact of energy efficiency

Increased demand-side elasticity influences the optimal generation technology mix, as well as the weighted average electricity price. Table 7 shows the optimal generation technology mix for different levels of efficiency elasticity of demand, gamma (γ_j). In this analysis, a 10% own-price elasticity of demand is assumed with zero cross-price elasticities, and the budget spent on energy efficiency programs is assumed to be increased by 50%. The cost of the energy efficiency DSM program is not analyzed, nor how this cost can be recovered. The welfare implications of such programs are analyzed in [34]. The emphasis here is upon the analysis of their interactions with demand response.

Considering first just the first-order effect of efficiency expenditures on loads, we assume a 5% elasticity for the effect of efficiency expenditures upon demand. Then if the budget for energy efficiency is increased by 50%, this elasticity causes a reduction in demand of 2.5% on average for each wind case, corresponding to a reduction of 100 up to 150 MW. As a result, fewer conventional and renewable energy generation capacity additions are needed (a reduction of precisely 2.5%). The total installed capacity is reduced from 12,408 MW to 12,111 MW in the low wind investment cost scenario.

We now turn to the interaction of energy efficiency expenditures and demand response. The mixed derivative parameter δ_j is a measure of their conflicting interaction. That parameter reduces the responsiveness of demand when more is spent on energy efficiency. Because this impact of efficiency expenditures upon price elasticities has only been discussed qualitatively in the literature [65], we arbitrarily assume a value of δ_j of 0.5% to illustrate the potential impact of this interaction. In that case, when energy efficiency expenditures are increased 50%, the optimal amount of wind capacity is reduced even further. Comparing Table 5 with Table 7 shows that the optimal wind power capacity is reduced from 6,951 MW without energy efficiency expenditures to 6,812 MW with expenditures as a result of reduced demand (given 40 k€/MW investment cost and -10% own-price elasticity). When including a positive value δ_j of 0.5%, reduced short-term demand responsiveness results in an even lower optimal installed wind power capacity of 6,154 MW. This result shows that considering the interaction between energy efficiency and demand elasticity can significantly affect optimal generation mixes.

Figure 9 shows the load impact of demand response combined with energy efficiency expenditures, compared with the original, no response load profile. The original load profile is indicated by the bold full line. With an own-price elasticity of -10%, assumed in Figure 9, peak demand is reduced around hours 9 and 43, and some valley filing occurs circa hour 27. Additionally, if energy efficiency expenditures are increased by 50% (efficiency elasticity 5%), demand levels are slightly reduced (by about 300 MW or 2.5% on average). This is indicated by the dashed line just below the thin full line. If an overlap is assumed between the effects of demand response and energy efficiency (efficiency-price elasticity 0.5%), the responsiveness of demand is reduced. Consequently, peak load reduction and valley filling are noticeably less pronounced than without this counteracting effect.

Table 7: Effect of an energy efficiency impact: budget increase of 50% / -10% own-price elasticity

Scenario	Base	Mid	Peak	High Peak	Wind	Total
- Efficiency elasticity 2.5% / efficiency-price elasticity 0%						
40 k€/MW	0	3,486	1,333	559	6,882	12,260
60 k€/MW	0	3,668	1,291	523	6,272	11,755
80 k€/MW	1,409	2,518	1,212	458	5,433	11,029
100 k€/MW	3,171	1,453	952	309	3,208	9,091
- Efficiency elasticity 2,5% / efficiency-price elasticity 0,5%						
40 k€/MW	0	3,560	1,583	552	6,546	12,242
60 k€/MW	268	3,460	1,520	523	5,987	11,759
80 k€/MW	1,471	2,492	1,412	481	5,267	11,124
100 k€/MW	3,114	1,488	1,188	330	3,186	9,306
- Efficiency elasticity 5% / efficiency-price elasticity 0%						
40 k€/MW	0	3,437	1,309	553	6,812	12,111
60 k€/MW	0	3,618	1,273	511	6,209	11,611
80 k€/MW	1,383	2,490	1,190	453	5,375	10,891
100 k€/MW	3,134	1,430	932	305	3,170	8,971
- Efficiency elasticity 5% / efficiency-price elasticity 0,5%						
40 k€/MW	0	3,596	1,891	654	6,154	12,295
60 k€/MW	469	3,266	1,829	628	5,673	11,865
80 k€/MW	1,469	2,475	1,701	614	5,067	11,325
100 k€/MW	3,008	1,524	1,467	507	3,128	9,635

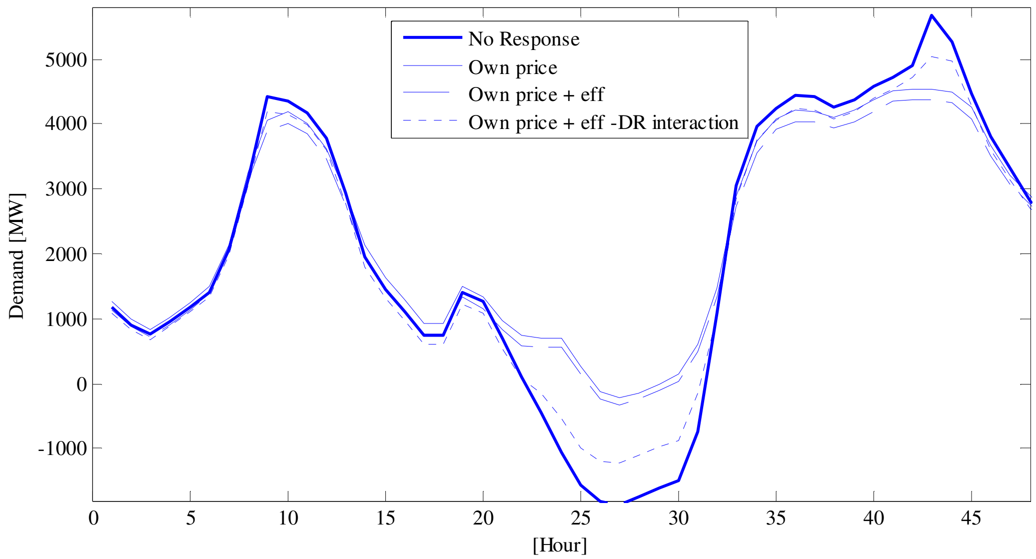

Figure 9: Comparison of demand response impacts under alternative elasticity assumptions

5. Conclusions

For many years, generation investment decision making has been supported by LP-based planning models. In this paper, these models have been extended to incorporate two considerations that are increasingly important as markets are restructured and increased amounts of intermittent renewable energy are provided. These considerations are operational constraints that limit the flexibility of thermal generation facilities to respond to demand and renewable energy fluctuations, and the ‘smart grid’ technology of short-term demand response to spot electricity prices.

The integration of demand response creates opportunities to more efficiently balance supply and demand. This paper has illustrated methods for integrating real-time price responsiveness into electric energy models. Elastic demand functions are constructed based on historic hourly demand levels and assumed levels of elasticities. These include own-price elasticity as well as cross-price elasticities with respect to prices in other hours in order to capture load shifting effects. Investment models, commonly LP-based cost minimizations, are expanded to account for consumer demand response. Three numerical approaches to accomplish this supply-demand integration are presented. In addition, the interactions of energy efficiency investments and demand responsiveness are also modelled by including those investments as first- and second-order terms in the demand function.

The integration of demand response decreases system peaks, decreasing the required investment in peaking generation capacity. Additionally, demand response creates valley filling effects, lessening over-generation problems during the night or high wind generation periods. Demand response also increases system flexibility, facilitating the integration of intermittent wind power generation. Simulations show that for higher demand elasticity, it is optimal to install a higher amount of wind power capacity.

Furthermore, price responsive consumers increase consumption during low price hours and decrease consumption during high price hours. As a consequence, the weighted average electricity price is reduced. However, the inclusion of cross-price elasticities reduces these effects as consumption during high price periods is shifted to other hours instead of being indefinitely postponed.

Then, the impact of energy efficiency is analyzed. Increased emphasis on energy efficiency reduces demand levels and therefore the total required installed generation capacity. This effect is reduced when the negative interaction of energy efficiency investments and responsiveness of demand is included. If this interaction is significant, the optimal amount of installed wind power capacity is reduced.

Demand-side aspects and the respective sensitivities within a long-term investment planning context are dealt with in this paper. Uncertainties on the planning timescale, such as future technological, economic, and policy uncertainties will need to be the subject of future research. Interactions between those uncertainties, demand-response, and generation technology choice would also be interesting to consider. Additionally, uncertainty with respect to generation plant availability and random outages is not emphasized in this paper, although it could be addressed in future research. Finally, making this model dynamic, starting from an existing generation fleet and taking decommissioning of older generation plants into account would be a valuable extension to this model. It could help illustrating how a transition toward more renewables and simultaneously a more responsive demand-side would occur.

6. References

[1] C. De Jonghe, E. Delarue, R. Belmans, and W. D'haeseleer, "Interactions between measures for the support of electricity from renewable energy sources and CO2 mitigation," *Energy Policy*, vol. 37, Nov. 2009, pp. 4743-4752.

[2] B. Ummels, "Wind integration: Power system operation with large-scale wind power in liberalised environments," *PhD Thesis, Technische Universiteit Delft*, 2009.

[3] P. Luickx, E. Delarue, and W. D'haeseleer, "Considerations on the backup of wind power: Operational backup," *Applied Energy*, vol. 85, Sep. 2008, pp. 787-799.

[4] P. Luickx, "The Backup of wind power: Analysis of the parameters influencing the wind power integration in electricity generation systems," *PhD Thesis, Katholieke Universiteit Leuven*, 2009, p. 276.

[5] V. Hamidi, "Smart grid technology review within the transmission and distribution sector," *Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010 IEEE PES*, Gothenburg: , pp. 1-8.

[6] S. Stoft, *Power system economics: Designing markets for electricity*, John Wiley & Sons, Inc., 2002.

[7] B.F. Hobbs, "Optimization methods for electric utility resource planning," *European Journal of Operational Research*, vol. 83, May. 1995, pp. 1-20.

[8] R. Earle, E.P. Kahn, and E. Macan, "Measuring the capacity impacts of demand response," *The Electricity Journal*, vol. 22, 2009.

[9] S.N. Talukdar and F.F. Wu, "Computer-aided dispatch for electric power systems," *Proceedings of the IEEE*, vol. 69, 1981, pp. 1212-1231.

[10] R. Turvey and D. Anderson, *Electricity economics: Essays and case studies*, The Johns Hopkins University Press, Baltimore, 1977.

[11] M. Ventosa, A. Baillo, A. Ramos, and M. Rivier, "Electricity market modeling trends," *Energy Policy*, vol. 33, May. 2005, pp. 897-913.

[12] F. Bessière and P. Massé, "Long term programming of electrical investments," *Marginal cost pricing in practice*, 1964, p. 266.

[13] P. Massé and R. Gibrat, "Application of linear programming to investments in the electric power industry," *Management Science*, vol. 3, Jan. 1957, pp. 149-166.

[14] J.L. Meza, "Multicriteria analysis of power generation expansion planning," *PhD Thesis, Wichita State University*, 2006, p. 158.

[15] E. Hirst, “A good integrated resource plan: electric utilities and regulators,” *Oak Ridge National Laboratory : ORNL/CON--354*, 1992, p. 77.

[16] B.F. Hobbs and P. Centolella, “Environmental policies and their effects on utility planning and operations,” *Energy*, vol. 20, Apr. 1995, pp. 255-271.

[17] D.T. Hoog and B.F. Hobbs, “An integrated resource planning model considering customer value, emissions, and regional economic impacts,” *Energy*, vol. 18, Nov. 1993, pp. 1153-1160.

[18] J. Pan, R.P. Broadwater, H.F. Vanlandingham, and H.D. Sherali, “MADM framework for strategic resource planning of electric utilities,” *PhD Thesis, Virginia State University*, 1999.

[19] J. Zhu and M.-yuen Chow, “A review of emerging techniques on generation expansion planning,” *IEEE Transactions on Power Systems*, vol. 12, 1997, pp. 1722-1728.

[20] B.F. Hobbs and P.M. Meier, “Multicriteria methods for resource planning: an experimental comparison,” *IEEE Transactions on Power Systems*, vol. 9, 1994, pp. 1811-1817.

[21] D. Anderson, “Models for determining supply in electricity investments,” *Bell Journal of Economics and Management Science*, vol. 3, 1972, pp. 267-299.

[22] J.A. Bloom, “Solving an electricity generating capacity expansion planning problem by generalized Benders’ decomposition,” *Operations Research*, vol. 31, Jan. 1983, pp. 84-100.

[23] J.A. Bloom, “Long-range generation planning using decomposition and probabilistic simulation,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-101, 1982, pp. 797-802.

[24] E. Delarue, C. De Jonghe, R. Belmans, and W. D’haeseleer, “Applying portfolio theory to the electricity sector: Energy versus power,” *Energy Economics*, vol. 33, May. 2011, pp. 12-23.

[25] D. Logan, C. Neil, and A. Taylor, “Modeling renewable energy resources in integrated resource planning,” *National Renewable Energy Laboratory: NREL/TP-462-6436*, 1994, p. 114.

[26] C. De Jonghe, E. Delarue, R. Belmans, and W. D’haeseleer, “Determining optimal electricity technology mix with high level of wind power penetration,” *Applied Energy*, vol. 88, 2011, pp. 2231-2238.

[27] J. Maddaloni, A. Rowe, and G. Van Kooten, “Wind integration into various generation mixtures,” *Renewable Energy*, vol. 34, Mar. 2009, pp. 807-814.

[28] C.W. Gellings, “The concept of demand-side management for electric utilities,” *Proceedings of the IEEE*, vol. 73, 1985, pp. 1468-1470.

[29] C.W. Gellings and W.M. Smith, "Integrating demand-side management into utility planning," *Proceedings of the IEEE*, vol. 77, 1989, pp. 908-918.

[30] EPRI, "Survey of utility lightning programs," *EM-5093, Palo Alto, CA*, 1987, p. 86.

[31] A. Paul, D. Burtraw, and K. Palmer, "Haiku documentation: RFF ' s electricity market model," *Resources for the Future*, 2009, p. 52.

[32] V. Oikonomou, M. Rietbergen, and M. Patel, "An ex-ante evaluation of a white certificates scheme in The Netherlands: A case study for the household sector," *Energy Policy*, vol. 35, Feb. 2007, pp. 1147-1163.

[33] California Energy Commission, "California standard practice manual: Economic analysis of demand-side programs and projects," 2001, p. 34.

[34] B.F. Hobbs, "The 'most value' test: Economic evaluation of electricity demand-side management considering customer value," *Energy Journal*, vol. 12, 1991, pp. 67-91.

[35] B.F. Hobbs and S.K. Nelson, "Assessing conservation payments: Least-cost, least-rates, or most-value?," *The Electricity Journal*, vol. 2, Jul. 1989, pp. 28-39.

[36] M.C. Caramanis, R.E. Bohn, and F.C. Schweppe, "Optimal spot pricing: Practice and theory," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-101, 1982, pp. 3234-3245.

[37] J.D. Kueck, B.J. Kirby, J.H. Eto, R.H. Staunton, C. Marnay, C.A. Martinez, and C. Goldman, *Load as a reliability resource in restructured electricity markets*, 2001.

[38] B.J. Kirby, "Spinning Reserve From Responsive Loads," *Oak Ridge National Laboratory*, vol. ORNL/TM-20, p. 41.

[39] W.L. Rutz, M. Becker, and F.E. Wicks, "Treatment of elastic demand in generation planning," *IEEE Trans. Power App. Sys.*, vol. PAS-104, 1985, pp. 3092-3097.

[40] E. Celebi and J.D. Fuller, "A model for efficient consumer pricing schemes in electricity markets," *IEEE Trans. Power Sys.*, vol. 22, Feb. 2007, pp. 60-67.

[41] S. Borenstein, "The long-run efficiency of real-time electricity pricing," *Energy Journal*, vol. 26, 2005, pp. 93-116.

[42] J. Bushnell, "Building blocks : Investment in renewable and non-renewable technologies," *Energy Institute at Haas Working Paper 202R, University of California, Berkeley*, 2010.

[43] R. Green and N. Vasilakos, “The long-term impact of wind power on electricity prices and generating capacity,” *Discussion paper, Department of Economics, University of Birmingham*, 2010, pp. 1-23.

[44] J.R. Birge and L. François, “Introduction to Stochastic Programming,” *Springer*, 1997, p. 448.

[45] A.H. Van Der Weijde and B.F. Hobbs, “Planning electricity transmission to accommodate renewables: Using two-stage programming to evaluate flexibility and the cost of disregarding uncertainty,” *EPRG Working Paper 1102 Cambridge Working Paper in Economics 1113*, 2011.

[46] P.A. Samuelson, “Spatial Price Equilibrium and Linear Programming,” *The American Economic Review*, vol. 42, 1952, pp. 283-303.

[47] E. Denny and M. O□Malley, “Wind Generation, Power System Operation, and Emissions Reduction,” *IEEE Transactions on Power Systems*, vol. 21, Feb. 2006, pp. 341-347.

[48] D. Logan, C. Neil, A. Taylor, and P. Lilienthal, “Integrated resource planning with renewable resources,” *The Electricity Journal*, vol. 8, Mar. 1995, pp. 56-66.

[49] G. Giebel, “On the benefits of distributed generation of wind energy in europe,” *PhD Thesis, Carl von Ossietzky Universität Oldenburg*, 2001, p. 97.

[50] B.F. Hobbs and Y. Ji, “A bounding approach to multiarea probabilistic production costing,” *IEEE Transactions on Power Systems*, vol. 10, 1995, pp. 853-859.

[51] B.F. Hobbs and Y. Ji, “Stochastic Programming-Based Bounding of Expected Production Costs for Multiarea Electric Power Systems,” *Operations Research*, vol. 47, 1999, pp. 836-848.

[52] L. Kuntz and F. Müsgens, “Modelling start-up costs of multiple technologies in electricity markets,” *Mathematical Methods of Operations Research*, vol. 66, Feb. 2007, pp. 21-32.

[53] R. Sioshansi and W. Short, “Evaluating the Impacts of Real-Time Pricing on the Usage of Wind Generation,” *IEEE Trans. Power Sys*, vol. 24, 2009, pp. 516-524.

[54] H. Holttinen, “The impact of large scale wind power production on the nordic electricity system,” *PhD Thesis Helsinki University of Technology*, 2004, p. 193.

[55] E. Hirst and C. Goldman, “Creating the future: integrated resource planning,” *Annual Review of Energy and the Environment*, vol. 16, 1991, pp. 91-121.

[56] B.F. Hobbs and U. Helman, “Complementarity-based equilibrium modeling for electric power markets,” *Modeling Prices Competitive Electricity Markets*, D.W. Bunn, ed., J. Wiley, Ch. 3, 2004.

[57] R.W. Cottle, J.S. Pang, and R.E. Stone, “The linear complementarity problem,” *Academic Press, New York*, 1992.

[58] T. Takayama and G.G. Judge, “Equilibrium among spatially separated markets : A reformulation,” *Econometrica*, vol. 32, 1964, pp. 510-524.

[59] R.W. Cottle and G.B. Dantzig, *Complementary pivot theory of mathematical programming*, Stanford University Press, 1967.

[60] T. Takayama, H. Hashimoto, and N.D. Uri, “Spatial and temporal price and allocation modeling: Some extensions,” *Socio-Economic Planning Sciences*, vol. 18, 1984, pp. 227-234.

[61] B.-H. Ahn and W.W. Hogan, “On convergence of the PIES algorithm for computing equilibria,” *Operations Research*, vol. 30, Mar. 1982, pp. 281-300.

[62] W.W. Hogan, “Energy policy models for Project Independence,” *Computers & Operations Research*, vol. 2, Dec. 1975, pp. 251-271.

[63] Federal Energy Administration, “PIES Documentation - The integrating model of the project independence evaluation system,” vol. 1, 1976.

[64] W.W. Hogan, “Project independence evaluation system: structure and algorithms,” *Proceedings of the symposia in applied mathematics*, vol. 21, 1977.

[65] C. Goldman, M. Reid, and R. Levy, “Coordination of Energy Efficiency and Demand Response,” *Lawrence Berkeley National Laboratory Report: LBNL-3044E, Berkeley, California*, 2010.

[66] D. York and M. Kushler, “Exploring the relationship between demand response and energy efficiency,” *American Council for an Energy Efficient Economy: Research report, Washington DC*, 2005.

[67] F. Wolak, “Residential Customer Response to Real-Time Pricing : The Anaheim Critical-Peak Pricing Experiment,” *Working paper CSEM WP 151*, 2006.

[68] IAE/NEA, *Projected costs of generating electricity*, 2010.

[69] M. Shahidehpour and C. Wang, “Ramp-rate limits in unit commitment and economic dispatch incorporating rotor fatigue effect,” *IEEE Transactions on Power Systems*, vol. 9, 1994, pp. 1539-1545.

[70] J. Ihle and B. Owens, “Integrated Coal and Wind Power Development in the U.S. Upper Great Plains,” 2004, p. 28.

[71] European Wind Energy Association (EWEA), *Large scale integration of wind energy in the European power supply: analysis, issues and recommendations*, 2005.

[72] A. Faruqui, R. Hledik, and S. Sergici, “Rethinking prices: The changing architecture of demand response in America,” *Public Utilities Fortnightly*, vol. January, 2010, pp. 30-39.

[73] M. Lijesen, “The real-time price elasticity of electricity,” *Energy Economics*, vol. 29, Mar. 2007, pp. 249-258.

[74] T.N. Taylor, P.M. Schwarz, and J.E. Cochell, “24 / 7 Hourly response to electricity real-time pricing with up to eight summers of experience,” *Journal of Regulatory Economics*, vol. 27, 2005, pp. 235-262.

[75] R. Boisvert, P. Cappers, B. Neenan, and B. Scott, “Industrial and commercial customer response to real time electricity prices,” *Neenan Associates*, 2004.

[76] K. Gillingham, R.G. Newell, and K. Palmer, “Energy Efficiency Economics and Policy,” *Annual Review of Resource Economics*, vol. 1, 2009, pp. 597-620.

[77] K. Palmer and D. Burtraw, “Cost-Effectiveness of Renewable Electricity Policies,” *Discussion paper*, vol. RFF DP 05-, 2005, p. 32.

[78] A. Faruqui and S. George, “Quantifying customer response to dynamic pricing,” *The Electricity Journal*, vol. 18, 2005.

[79] J.H. Eto, “An overview of analysis tools for integrated resource planning,” *Energy*, vol. 15, 1990, pp. 969-977.

[80] B.F. Hobbs, “Exploring trade-offs with OR/MS: a fertile source of challenges for the profession,” *OR/MS Today: Institute for Operations Research and the Management Sciences*, vol. 23, 1996.

[81] B.F. Hobbs, H.B. Rouse, and D.T. Hoog, “Measuring the economic value of demand-side and supply resources in integrated resource planning models,” *IEEE Transactions on Power Systems*, vol. 8, 1993.

[82] R.T. Jenskin and D.S. Joy, “WIEN Automatic System Planning Package (WASP),” *International Atomic Energy Agency, Vienna*, 2001, p. 284.

[83] A.K. Dixit and R.S. Pindyck, *Investment under uncertainty*, Princeton University Press, 1994.

[84] M. Kennedy, “An economic model of the world oil market,” *Bell Journal of Economics and Management Science*, vol. 5, 1976.

[85] T. Takayama and G.G. Judge, “Spatial and temporal price and allocation models,” *North-Holland, Amsterdam*, 1971.