The integration of variable generation and storage into electricity capacity markets

Citation for published version:
Zachary, S, Wilson, A & Dent, C 2022, 'The integration of variable generation and storage into electricity capacity markets', *The Energy Journal*, vol. 43, no. 4. https://doi.org/10.5547/01956574.43.4.szac

Digital Object Identifier (DOI):
10.5547/01956574.43.4.szac

Published In:
The Energy Journal

arXiv:1907.05973v1 [math.OC] 12 Jul 2019

The integration of variable generation and storage into electricity capacity markets

Stan Zachary*, Amy Wilson†, and Chris Dent†

July 16, 2019

Abstract

We show how to value both variable generation and energy storage in such a way as to enable them to be integrated fairly (from the point of view of capacity providers) and optimally (from the point of view of society) into electricity capacity markets, for example that which operates in Great Britain. We develop a theory based on balancing expected energy unserved against costs of capacity procurement, and in which the optimal procurement may be expressed as that necessary to meet an appropriate reliability standard. In the absence of variable generation and storage the entire theory reduces to that already in common use—both in the definition of a standard and in its economic justification. Further the valuation of both variable generation and storage in the proposed approach coincides with the traditional risk-based approach leading to the concept of an equivalent firm capacity. The determination of the equivalent firm capacity of storage requires particular care; this is due both to the flexibility with which storage added to an existing system may be scheduled, and also to the fact that, when *any* resource is added to an existing system, storage already within that system may be flexibly rescheduled.

1 Introduction

In order to ensure the adequacy of electricity supplies it is now necessary to provide a *capacity market*, or mechanism, in many countries of the world, including Great Britain (see, for example, the discussions in [9, 16, 12]). The purpose of such a market is typically to guarantee the availability of an appropriate level of capacity during particular years in the future, and to obtain this capacity at minimum cost. So as to operate such a market both fairly and optimally it is necessary to value correctly the contributions of the individual capacity providers, whether they provide conventional generation, variable generation, or storage. The present approaches to capacity market design have primarily been designed with conventional generation in mind, e.g. in GB [13] and in the various North American systems which operate

*Heriot-Watt University and University of Edinburgh. Research supported by EPSRC grants EP/I017054/1, EP/K002252/1, EP/K002252/1 and EP/R014604/1.

†University of Edinburgh. Research supported by EPSRC grants EP/K002252/1, EP/K002252/1 and EP/R014604/1.

such a market [3, 12]. Conventional generation is typically, and often naturally, treated analogously to *firm capacity*, i.e. as capacity which is able to supply energy as needed up to a given constant rate. (In order to do so the nominal capacity of any individual generator is usually multiplied by an appropriate "de-rating" factor to acknowledge the fact that it is occasionally unavailable—see [13, 3].)

When all capacity is firm, an economic theory of capacity markets is relatively straightforward, and is based on balancing investment or procurement cost against the cost of unserved energy—see, for example, the textbook treatment of [19] and, for more detail, [22]. The paper [2] considers the impact, within a linear programming formulation and in terms of reduction in societal welfare, of incorrectly de-rating variable generation in order to treat it as if it were firm capacity, but does not consider storage, and does not consider the mechanism of running a capacity market. Both variable generation—e.g. wind power and solar power—and storage now have important contributions to make to capacity adequacy. The present paper shows how, under appropriate circumstances which we identify, current approaches to capacity market design (as in GB and the US as referenced above) may be extended to give an integrated theory for the inclusion within a capacity market of all types of capacity provision. Again as at present, the theory is necessarily based on a probabilistic description of the electricity supply-demand balance process. However, storage in particular has a natural energy constraint and thus can supply energy only for a limited period of time before its energy requires to be replenished; subject to this constraint it may be scheduled flexibly. Hence, in order to understand both how to schedule storage and to determine its contribution to capacity adequacy, it is necessary to pay more attention than would otherwise be required to the sequential statistical structure (i.e. the time evolution) of the supply-demand balance process to which that storage is contributing—see, for example, [15] and also the more detailed discussion of Sections 2 and 3. The present paper extends and generalises theory which was developed by the authors in conjunction with National Grid ESO for the integration of storage in the GB capacity market—again see [15]. However, the theory is applicable wherever the capacity contributions of variable generation and storage need to be correctly assessed.

The determination of a volume of *capacity-to-be-procured* in a capacity market may be achieved either via the satisfaction of an appropriate *security-of-supply* standard defined in terms of some given system *risk metric* or via the minimisation of an appropriate economic cost. (In the latter case the capacity-to-procure may be variable and specified as a function of the clearing price in the capacity auction, as is currently the case in GB—see Section 5.) These two approaches are closely related; indeed the latter approach yields the former as a sub-problem—see Section 5. In either case, a key step in the development of an integrated theory of capacity markets is that of the provision of an appropriate definition of the *equivalent firm capacity* (EFC) of any capacity-providing resource. In essence, the EFC of any such resource is that firm capacity which makes an (appropriately defined) equivalent contribution to the overall supply-demand balance. Hence such an EFC is necessarily defined with respect to the pre-existing supply-demand balance process to which this resource is being added—see [21, 5]. When the set of capacity-providing resources contains significant storage, particular care is required in the determination of EFCs. One reason for this is the need to account for the flexibility of scheduling of additional storage added to an existing supply-demand balance process. However, a more subtle reason is the following: when *any* further resource is added to an existing set of capacity-providing resources which already contains storage, that pre-existing storage may *also* be rescheduled so as to enhance the usefulness of the additional contribution. This is particularly so when the additional contribution is firm capacity. A consequence of this, as we show formally at the end of Section 3 and demonstrate in the example of Section 6, is that the EFC of further storage added to an existing system is *less* (than it would otherwise be) in the case where that additional system already contains significant storage. An intuitive explanation is that the flexibility provided by further storage is less useful in the presence of existing, already flexible, storage.

In any theory of capacity adequacy, the two most commonly used risk metrics are *loss-of-load expectation* (LOLE) and *expected energy unserved* (EEU)—see [1, 11, 21]. LOLE is defined as the probabilistic expectation of the total length of time during a given year or peak season in which the system under study is undergoing *loss-of-load* or *shortfall*, i.e. the supply of energy is insufficient to meet demand, while EEU is the expectation of the total *volume* of demand which is unmet over all such periods of shortfall, i.e. the expectation of the total *unserved energy.* More formal definitions are given in Section 2. In the case where all capacity may be treated similarly to firm capacity—as is usually the case for appropriately de-rated conventional generation—an economic criterion based on the uniform valuation of unserved energy is equivalent to one based on either an LOLE or an EEU standard provided the various parameters are correctly aligned—see Section 5.

However, security-of-supply standards are presently more commonly defined via the use of the LOLE risk metric. We shall argue below that, in the presence of either variable generation or storage, LOLE is not be the best measure of system reliability. This is especially so in the case of storage. We discuss this in detail in Section 5, giving also the relevant mathematics, but for the moment a simple example suffices to indicate why the use of LOLE as a metric may be problematic. Suppose that a certain volume of stored energy is available to mitigate, but not wholly eliminate, a given period of shortfall (say of a few hours duration) in the supply-demand balance. The loss-of-load duration, and hence LOLE, will be minimised by using this additional stored energy to eliminate entirely the shortfall at those times within the above period at which this is most efficiently achieved, i.e. at those times at which the *depth* of the original shortfall is least—since this maximises the total length of time for which there is no longer any loss-of-load—see the left panel of Figure 1. However, it might reasonably be argued that in practice the available stored energy would be at least as well employed in instead minimising the *maximum* depth of shortfall in supply during the period concerned, a policy which would achieve the same reduction in unserved energy but quite possibly result in no reduction in loss-of-load whatsoever, as illustrated in the middle panel of Figure 1. Indeed, from an economic perspective, the latter policy would be clearly better (and indeed optimal within a deterministic environment) if the unit cost of unserved energy were an increasing function of shortfall depth. Further, regardless of how unserved energy is valued, the former policy of using stored energy to minimise loss-of-load during what would otherwise be the shortfall period does not make as effective a contribution to system reliability as would the use of firm capacity to achieve the same reduction in loss-of-load: unlike storage such firm capacity would continue to contribute even at those times when shortfall was not completely eliminated—see the right panel of Figure 1.

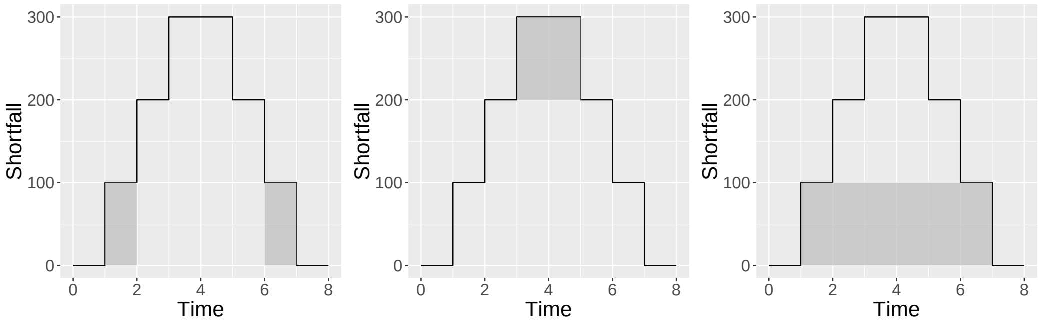

Figure 1: Use of a given volume of storage to mitigate a given period of shortfall. The left panel shows the reduction (shaded) in shortfall when the stored energy is used to minimise the loss-of-load duration; the middle panel shows the reduction when the same stored energy is used to minimise the maximum depth of shortfall; the right panel shows the reduction in shortfall (again shaded) which occurs when firm capacity is used to achieve the same loss-of-load duration as in the left panel.

The above difficulties in *measuring* the contribution of storage in particular arise on account of its time-limited duration combined with its flexibility of use. Notably, in the presence of storage, LOLE may be varied in a manner which (as illustrated in the above example) does not obviously relate to system adequacy as usually understood. To a lesser extent there may be analogous difficulties with variable generation—again see [2]; however, the times at which variable generation is available are not controllable in the same way as for storage. When electricity capacity is provided solely via conventional generation, and when the latter is capable of being modelled as if it were (appropriately de-rated) firm capacity, then *within a given system* the relation of LOLE to, for example, an economic measure of system adequacy such as uniformly-valued EEU is essentially predetermined, and the above difficulties do not then arise. We give a more precise mathematical treatment of these ideas in Sections 4 and 5.

Throughout the present paper we treat the process of electricity *demand* as given. However, demand management, or demand response, may also be used to assist in balancing electricity supply and demand. Demand management has many of the characteristics of storage—in particular any instance of such management can typically make a flexible contribution to the supply-demand balance process but only for a period of time of limited duration. Thus demand management may be viewed as a form of "virtual storage". Its contribution to electricity capacity, and its integration into capacity markets, may be analysed analogously to the treatment of storage given here.

Sections 2 and 3 of the paper study respectively some properties of risk metrics and of equivalent firm capacity. The latter is necessarily defined in terms of some given risk metric and is essential to the understanding of both capacity adequacy and the operation of a capacity market. The studied properties are implicit in the theory of present markets which are concerned primarily with conventional generation. However, so as to understand how to incorporate into such markets both variable generation and time-limited but flexible resources such as those provided by storage, it is necessary to make these properties and some associated assumptions more explicit. Risk metrics are treated in general; however, we pay particular attention to LOLE and EEU. Such metrics may be regarded as functions of the set of capacity-providing resources and, in order to obtain a tractable theory of capacity markets, we require *continuity* and *smoothness* assumptions about the behaviour of such risk metrics as the available capacity-providing resource is varied. Such assumptions are often implicit in other work, e.g. [2]. We formulate these assumptions explicitly, and argue that they are usually sufficiently satisfied in practice, including in present-day markets. The smoothness assumption yields an important *local additivity* property for equivalent firm capacities, which is essential for the correct operation of markets—even in the case where all resource is provided by firm capacity. In Section 3 we also show how to determine the EFC of marginal contributions of both variable generation and storage when the objective is the minimisation of EEU. In the case of storage this requires consideration of how it may be optimally scheduled.

Section 4 studies the operation of capacity markets when the objective is that of obtaining at minimum cost sufficient capacity to meet a given security-of-supply standard defined in terms of a risk metric. Section 5 studies the operation of such markets when, as discussed earlier, the objective is that of the minimisation of an overall economic cost. In particular we study the relation between economic and risk-based approaches to capacity market operation. It is known that, when all capacity-providing resource may be treated as firm capacity, an approach based on the minimisation of an overall economic cost in which the cost of unserved energy is proportional to EEU is equivalent to an risk-based approach using LOLE. We study the conditions required for this result to continue to hold in the presence of variable generation. Further, the theory requires substantial modification in the presence of storage, as discussed in Section 5.

The flexibility of storage scheduling has important consequences for the way in which a capacity market operates, and these are illustrated in the detailed example of Section 6. This example show the application of nearly all the above theory, and is chosen to be realistic in the context of a country such as GB. In particular it demonstrates the practical reasonableness of the assumptions required for a tractable theory.

2 Risk metrics

In the analysis of the adequacy of energy systems, the length of time over which system risk is to be assessed—typically either a year or a peak season—is usually divided into n time periods, each typically of an hour or a half-hour in length—see [1, 15]. Let the random variables D_t and X_t denote respectively the total energy demand and total energy supply in time period t. Then the *supply-demand balance* in the time period t is given by the random variable $Z_t = X_t - D_t$. In particular values of Z_t less than zero correspond to an energy *shortfall* or *loss-of-load* at time t, and then the level or *depth* of shortfall at that time is given by the random variable $\max(-Z_t, 0)$.

Any *risk metric* ρ is a function of the entire supply-demand balance process $(Z_t,\ t = 1, \dots, n)$. Risk metrics may either be used directly in the setting of appropriate reliability standards—as in the case of the present GB LOLE-based standard (see [15])—or may arise naturally in the context of *economic* approaches to determining security-of-supply (see Section 5). Commonly used risk metrics are LOLE and EEU given respectively by

$$\text{LOLE} = \sum_{t=1}^{n} \mathbf{P}(Z_t < 0), \tag{1}$$

$$\text{EEU} = \sum_{t=1}^{n} \mathbf{E}(\max(-Z_t, 0)) = \sum_{t=1}^{n} \int_{-\infty}^{0} \mathbf{P}(Z_t < z)\, dz, \tag{2}$$

where $\mathbf{P}$ denotes probability and where $\mathbf{E}$ denotes expectation—again see [1, 21]. Thus, in the present discrete-time setting, LOLE is the expected number of periods of shortfall during the season under study, while EEU is the expectation of the sum of the depths of shortfall during successive periods, i.e. the expectation of the total unserved energy.

In an economic context the use of EEU as a measure of economic cost corresponds to a uniform valuation of each unit of unserved energy, regardless of the overall depth of energy shortfall at any given time. The present paper mainly considers such a uniform valuation of unserved energy; however, in many systems modest levels (depths) of shortfall may be managed without significant ill effects by the use of emergency actions such as voltage reduction, while the avoidance of economic or other damage becomes increasingly difficult as the depth of shortfall increases. Thus it may be natural in such cases to value unserved energy more highly at higher levels of shortfall.

We take as given the demand process $(D_t,\ t = 1, \dots, n)$ over the successive time periods in which the system is to be studied. The energy supply process $(X_t,\ t = 1, \dots, n)$, and hence the value of any risk metric, is then determined by the set of capacity-providing resources (conventional generation, variable generation, and storage). We denote by R the set of such capacity-providing resources, and regard any risk metric ρ as a function $\rho(R)$ of the set R. We assume that the resources within the set R are optimally used for the minimisation of overall system risk—this is important in the context of flexible resources such as storage which may be used in different ways to support the system. We assume, as is the case with standard metrics such as LOLE and EEU, that any risk metric ρ is such that $\rho(R)$ is *decreasing* as the set of resources R is increased.

Typically, for a large system, the effect on the overall state of the system—as measured by the risk metric of interest—of the addition or subtraction of a single unit of resource is small. The optimality of the state of the system (by some criterion) is characterised by some appropriate form of *equilibrium* with respect to *marginal*, i.e. *relatively small*, variations in that overall state caused by the addition or subtraction of individual resources. So as to obtain a reasonable and tractable economic theory, we require continuity and smoothness assumptions with respect to such marginal variations. The *continuity assumption* is that of the reasonably continuous availability of capacity, i.e. that the state of the macroscopic system, as for example measured by the risk metric ρ, may be varied continuously—at least to a good approximation—by the addition or subtraction of individual units of capacity-providing resource. In large systems this is a reasonable approximation and one which is usually made in practice in the design of current capacity markets. An exception occurs where we are considering some very large individual resource, such as perhaps a large nuclear plant; however, as we discuss in Section 4, it is straightforward to deal with a small number of such large contributions. The *smoothness assumption* is that, for any set of resources R, the reduction in risk $\rho(R) - \rho(R \cup \{i\})$ resulting from the addition of some further marginal (i.e. small) resource i to the set R is, to a good approximation, unchanged by *small* variations in the set R. More formally this may be expressed by the requirement that, for any small resource j (representing the above variation in the set R) and for further small resource i, we have to a good approximation that

$$\rho(R \cup \{i, j\}) - \rho(R \cup \{j\}) = \rho(R \cup \{i\}) - \rho(R), \tag{3}$$

where by $R \cup \{i, j\}$ is meant the set of resources R supplemented by the further resources i and j, and where, as the contribution of the resources i and j tends to zero, the percentage error in the relation (3) becomes negligible. This smoothness assumption is a essentially a form of differentiability assumption—see the discussion on firm capacity below—and is generally well satisfied in most applications and for most risk metrics, including those discussed above.[1] In the extended example of Section 6, which is chosen to be reasonably representative of a system such as that in GB, we check the applicability of both the above assumptions.

The concept of *firm capacity*, defined as energy supply which is guaranteed to be available up to a given constant rate throughout the overall period under consideration, plays a particular role as a *reference measure* in assessing the usefulness of other forms of capacity providing resource—see Section 3. In the special case where we consider variations about the set of resources R given by variations in firm capacity, it is therefore helpful to write (in a mild abuse of notation) $R + y$ for the set of resources R supplemented by firm capacity able to supply energy at a further constant rate y, and to allow that y may be continuously varied. In particular this will be important when we consider *equivalent firm capacity* below. Then it is straightforward to show that, for such variation, the above smoothness assumption implies that there exists the *derivative* $\rho'(R)$ of $\rho(R)$ with respect to firm capacity;

[1] The smoothness assumption is not guaranteed: it is possible to imagine (in a very artificial situation) that two capacity-providing resources i and j might each make identical reductions in risk, as measured by the metric ρ, but be such that the use of both together achieved no further reduction in risk than the use of either singly, in which case (3) would fail.

the latter is such that, for small variations of the total resource R by firm capacity y,

$$\rho(R + y) = \rho(R) + \rho'(R)y, \tag{4}$$

(the relative error in this approximation tending to zero as y tends to zero). This derivative plays an important role in subsequent analysis.

3 Equivalent Firm Capacity

Throughout this section we take as given a suitable risk metric ρ. Then, given also any set of capacity-providing resources R, the contribution of any further resource i (conventional generation, variable generation, or storage) to be added to the set R may be measured by its *equivalent firm capacity* (EFC) $efc_R(i)$. This is the firm capacity which, if added to the set R in place of the additional resource i, would make the same contribution to security-of-supply, as measured by the risk metric ρ. Formally, the constant $efc_R(i)$ is the solution of

$$\rho(R \cup \{i\}) = \rho(R + efc_R(i)), \tag{5}$$

where, as defined in Section 2, the notation $R + efc_R(i)$ on the right side of (5) corresponds to the set of resources R supplemented by firm capacity equal to $efc_R(i)$—see, e.g., [11, 21, 15]. (Of course, when the resource i is itself firm capacity y_i for some $y_i > 0$, as may be the case with conventional generation, then we have $efc_R(i) = y_i$ regardless of the set R to which i is being added.) When the set R contains resources such as storage which may flexibly used, the addition of further resource to this set may result in a different pattern of usage of the resource within R itself: on both the left and the right side of (5) it is assumed that the total available resource is being optimally used. Note also that the EFC of a resource which is not itself firm capacity, for example, variable generation or storage, *depends also on the existing set of resources R to which it is being added*—see, e.g., [21, 5]. This will be particularly important in Sections 4 and 5 for developing satisfactory theories of capacity adequacy and capacity markets.[2]

It follows from (4) and from the definition (5) of EFC that, for the addition of any *marginal* (i.e. small) resource i to the set R,

$$\rho(R \cup \{i\}) = \rho(R) + \rho'(R)efc_R(i), \tag{6}$$

(where the derivative $\rho'(R)$ in (6) remains, as in (4), that defined with respect to variation in firm capacity). Thus, for any existing set of resources R, and with respect to the risk metric ρ, the contribution to risk reduction given by any further marginal resource i is proportional to its EFC $efc_R(i)$. It hence makes sense to

[2] It is also possible to define the *equivalent load carrying capacity* (ELCC) of any further resource i with respect to an existing set of resources R. This is the constant $elcc_R(i)$ given by the solution of $\rho(R \cup \{i\} - elcc_R(i)) = \rho(R)$. By writing $\rho(R)$ as $\rho(R - elcc_R(i) + elcc_R(i))$ we have that $elcc_R(i) = efc_{R - elcc_R(i)}(i)$, so that we shall find it sufficient in the present paper to work with EFCs. It is further an easy consequence of our earlier smoothness assumption that for any small additional resource i to be added to the set R, $efc_R(i)$ and $elcc_R(i)$ are approximately equal (again in terms of relative error).

value such additional resources proportionally to their EFCs—see Section 4. We further have *local additivity* of the EFCs of such marginal variations about the set of resources R, that is, for marginal additions i and j to the given set R,

$$efc_R(\{i,j\}) = efc_R(i) + efc_R(j), \tag{7}$$

where by the left side of (7) is meant the EFC of the combined resources i and j, and where the relative error in (7) again tends to zero as the additional resources i and j reduce in size. The property (7) is a straightforward consequence of the relation (6) and the further application of the smoothness assumption (3): for each term in the latter, the substitution of that term by the expression given by (6) yields (7) immediately. The local additivity result (7) does not hold in the case of larger variations about the set R, something which we discuss further in Section 4, and illustrate in Section 6.

Determination of EFCs. As discussed above, the EFC $efc_R(i)$ of any capacity-providing resource i added to an existing set of resources R is defined by the solution of equation (5). The solution of this latter equation may involve trial values of $efc_R(i)$ and may not always be straightforward. Provided that the additional resource i is marginal (i.e. small) in relation to the existing set of resources R, then it is generally more straightforward to obtain the EFC $efc_R(i)$ of the further resource i via the solution of the equation (6) above, i.e. as

$$efc_R(i) = \frac{\rho(R \cup \{i\}) - \rho(R)}{\rho'(R)}; \tag{8}$$

typically all the quantities on the right side of (8) may be readily estimated, e.g. via simulation. (Recall that the equation (8) is ultimately a consequence of the earlier smoothness assumption, the validity of which we check in our extended example of Section 6.)

In particular, as we argue in the Introduction and further in Section 5, it is often natural to take the underlying risk metric ρ to be given by EEU. Then, in order to use (8), we require an expression for the derivative $\mathrm{EEU}'(R)$ of $\mathrm{EEU}(R)$ with respect to firm capacity. For any set of capacity-providing resources R which consists entirely of generation, either conventional or variable, this is given by

$$\mathrm{EEU}'(R) = -\mathrm{LOLE}(R). \tag{9}$$

The result (9) is known in the case where all the resource within the set R is provided by firm capacity, for example conventional generation—see [6]; its proof in the present case—where R may also contain variable generation (but not storage)—is essentially the same and is given, for completeness, in the appendix. The results (8) and (9) therefore provide an efficient way to determine the EFCs (with respect to the risk metric given by EEU) of marginal contributions to capacity in an environment in which all capacity is provided by generation—a result which is also implicitly obtained by [2]. The result (9) also plays an important part in an economic theory of capacity markets—see Section 5.

When the set of resources R may also include storage, the result (9) is no longer valid and requires modification. It is necessary to understand how storage resources may be optimally scheduled for the minimisation of EEU, and indeed how they are rescheduled for this purpose as other capacity-providing resources are varied. Assume, for simplicity, that storage may be completely recharged between periods of what would, in the absence of such storage, be energy *shortfall*, but that storage may not usefully be charged during such periods of shortfall. (This is currently the case in many countries, such as GB, where periods of shortfall are generally well separated; typically there is at most a single period of shortfall—the period of evening peak demand—during any day, and storage may be fully recharged overnight.) Assume also that any process of what would otherwise be continuous energy shortfall is to be met as far as possible from available storage, with the aim of minimising any residual unserved energy, and that each store i delivers energy subject to a *rate* (power) constraint and a total *energy* constraint.[3] At each time t define the *residual lifetime* (or *to-go time*) of each store i as the residual energy in the store at time t divided by the maximum rate at which that energy can be served. Then the above minimisation of EEU is achieved by the use of the *greedy* algorithm in which, at each successive time t during any period of shortfall, that shortfall is reduced as far as possible from energy in storage and in which the use of the stores is prioritised in descending order of their residual lifetimes. The optimality of the above policy is proved in [4], and is also implicit in the results of [8]. Now let S_e be the set of stores which, under the above policy, are empty at the end of the day. (Note that, since the shortfall process is typically random, so also is the set S_e.) Then, in the presence of storage the result (9) is replaced by the more general result

$$\mathrm{EEU}'(R) = -\mathrm{LOLE}(R \setminus S_e), \tag{10}$$

where $R \setminus S_e$ is the set of resources in R other than those in the set S_e. This result is formally proved in [4]. However, in essence the reason for it is that when, under the above optimal scheduling algorithm, (marginal) firm capacity is added to the set of resources R, any store which is not in S_e (not empty at the end of the shortfall period) continues to contribute over that period to the reduction of EEU as if it were firm capacity with the same rate constraint, while those stores within the set S_e are rescheduled so that they continue to empty as before; the result (10) is then a straightforward extension of (9).

We further remark that comparison of the results (9) and (10) yields the important result that, when firm capacity is added to a set of resources R, it contributes more effectively to the reduction of EEU when the set R already contains storage resources which may then be rescheduled as above. A corollary is that the EFC of further storage resources is typically reduced by the presence of existing storage resources—see also Section 6 for an example of this.

[3]In other contexts the energy constraint might be referred to as the "capacity" of the store (as measured, for example, in MWh). However, in the context of capacity markets, "capacity" has units of power (as measured, for example, in MW), and we preserve this usage throughout this paper.

4 Capacity markets

In this section we consider the operation of a capacity market where the objective is that of obtaining at minimum cost a sufficient set of capacity-providing resources R so as to satisfy some condition

$$\rho(R) \leq k, \tag{11}$$

expressed in terms of some appropriate risk metric ρ. We assume throughout that there is a fixed process of net demand to be met by these resources. The level k may either be chosen so that the condition (11) defines some given security-of-supply standard, or it may be chosen according to some economic criterion (see Section 5). We allow that some resource may be provided by facilities other than firm capacity, for example variable generation or storage. The required theory is the same as that which is currently used in the presence of firm capacity (typically conventional generation) only, except that it is necessary to appropriately define the EFCs of other resources. Further some iterative and other additional calculation may be required in the operation of any auction associated with the capacity market—at least when resource other than firm capacity makes a substantial contribution.

An example is given by the capacity market which currently operates in GB [13], and which typically seeks to secure a level of capacity compatible with the GB LOLE-based security-of-supply standard (however, see also Section 5). The auction associated with this market determines a unit *clearing price* such that any successful bidder (capacity provider) in the auction receives this clearing price multiplied by its offered capacity. When the latter is other than firm capacity, e.g. when it is variable generation or storage, the bidder's EFC is used instead and, since the total such capacity is currently small, is reasonably estimated in advance of the auction—essentially using the methodology described in Section 3. (This EFC is usually represented by a *de-rating factor* which multiplies the nominal capacity of the resource.) The GB auction is then conducted as a *descending clock auction* [13]. An initially high unit price is set, and this is then gradually reduced; bidders may exit the auction at any point, and the auction stops at the point where there is just sufficient capacity remaining in the auction to meet the required reliability standard (11) for the given LOLE-based risk metric ρ; the unit price at this point then becomes the clearing price, and each successful bidder remaining in the auction is then paid as above.

A descending clock auction as described above relies on the EFC of any capacity other than firm capacity being clearly identified in advance of the auction. However, as discussed in Section 3, the EFC of any such capacity-providing resource depends also on the overall supply-demand balance process to which it is being added or of which it forms a part—which, in the case of an auction, is that determined by the overall set of resources R finally selected in the auction. It follows that, in the presence of a *substantial* number of resources other than firm capacity, the form of auction described above may require some adjustment—for example, as described below.

In order to provide a better basis for a coherent theory, instead of considering the minimum *unit* price (i.e. price per unit of EFC) which each bidder is prepared to accept for its capacity offering, we consider instead the minimum *total* price which each bidder is prepared to accept in return for its contribution. Thus, in analysing the operation of the capacity market we assume that each individual bidder or resource i is willing to make available some given capacity for a given (minimum) total price or cost c_i. For example, this capacity might be a given level of firm capacity for as long as might be required, or it might be storage which could be called upon as required and which could be used flexibly subject to specified rate and energy constraints. The societal problem is now to design an auction to choose at minimum cost a subset R of those resources competing in the market, such that the required reliability condition (11) is satisfied. We continue to assume that, in the interests of fairness, associated with the outcome of any such auction is a unit clearing price p such that, if R is the set of resources which are finally successful in the auction, then

$$\begin{aligned} c_i &\leq p \times efc_R(i), \qquad i \in R, \\ c_i &> p \times efc_R(i), \qquad i \notin R, \end{aligned} \tag{12}$$

and each successful resource $i \in R$ is paid in total $p \times efc_R(i)$. The relations (12) define a competitive equilibrium condition necessarily satisfied by the unit clearing price p and the required optimal set of resources R. That this is so depends implicitly on the continuity and smoothness assumptions of Section 2. The continuity assumption ensures that resource is reasonably continuously variable, and the smoothness assumption guarantees the local additivity property of Section 3: under these assumptions, were such a unit clearing price *not* to exist for the claimed optimal set of resources R, then resource from outside that set might be more cheaply substituted for resource from within it while continuing to satisfy the required reliability condition (11), contradicting the optimality of the set R. In general we might then expect the relations (11) and (12) to define the required set of resources R uniquely, though it is not clear that this is always guaranteed. (When, in the presence of one or more *large* capacity-providing resources, the above continuity assumption breaks down, there may be problems of "overshoot" in obtaining at minimum cost the required capacity. When the number of such large resources is small, these problems may usually be solved with a little experimentation; otherwise an integer optimisation approach is required.)

When the EFCs $efc_R(i)$ are known—or may be reasonably be estimated—in advance, the unit clearing price p may be identified by, for example, a descending clock auction as described above: any bidder or resource i exits the auction if the descending unit price falls below $c_i/efc_R(i)$. Alternatively, if bidders were willing to declare their minimum total prices c_i in advance of the auction, the auction could be conducted offline by ranking in ascending order the minimum unit prices $c_i/efc_R(i)$ (to give what is usually referred to in the context of energy markets as a *merit order stack* [17, 18]); the unit clearing price p would then be chosen so that the accepted set of resources R, such that (12) held, was just sufficient to satisfy the required reliability condition (11).

However, when there are substantial resources other than firm capacity participating in the market, then the final EFCs $efc_R(i)$ of these resources (estimated with respect to the finally accepted resource set R) may *not* be sufficiently well known in advance of any capacity auction. In order to identify a unit clearing price p and resource set R such that the required conditions (11) and (12) hold, it may be necessary to proceed iteratively: starting with initial estimates of the unknown EFCs, an initial clearing price p and initial accepted resource set R may be obtained—e.g. through the formation of a merit-order stack as above; on the basis of this set R, EFCs may be re-estimated and improved values of the clearing price p and accepted resource set R may then be obtained as before; since the set of capacity-providing resources participating in the market is finite, one might reasonably expect convergence to a solution of (11) and (12) within a fairly small number of iterations—see also the examples of Section 6.

Finally, recall also that, while the theory of this section depends on the local additivity of EFCs given by (7), such additivity does not hold over more extensive variation of a set of resources. In particular, when resource other than firm capacity is significant, *it is not in general true* that the sum $\sum_{i \in R} efc_R(i)$ of the EFCs of the individual resources in the set R required to meet the given reliability condition (11) is equal to the total firm capacity which would substitute for those resources in the set R. Hence, in running a capacity auction under such circumstances it would be necessary to periodically check, *via a full probabilistic recalculation of the value of the risk metric ρ for the set of resources so far obtained*, whether one had yet selected sufficient units from the merit-order stack as to satisfy the required condition (11). Indeed, such checking may be necessary even in markets—such as that currently operating in GB—in which most capacity is provided by conventional generation, as the latter is also not quite firm capacity, and so here too there is the above problem of nonadditivity of EFCs over significant numbers of resources. We give examples of these issues in Section 6.

5 Economic approaches

The previous section considered the determination of the optimal set R of capacity-providing resources needed to meet a given security-of-supply standard. However, it is also possible to consider explicitly economic approaches to the determination of electricity capacity. Thus one might choose the set of resources R so as to minimise an overall economic cost. Such a cost has two components: the first is the economic cost of the occasional failures of the set of resources R to provide such energy as is needed, while the second is the cost, within a capacity market, of providing the set of resources R itself. We consider the case where the first of these costs is given by $\text{VOLL} \times \text{EEU}(R)$ where the constant VOLL is a unit *value-of-lost-load* and, as usual, $\text{EEU}(R)$ is the *expected energy unserved* associated with the optimal use of the set of resources R. This measure is common in economic approaches to the determination of electricity capacity adequacy—see [1, 2, 6]. In particular we follow the approach which is sometimes used as a justification for the LOLE-based GB capacity adequacy standard (again see [6]) and which is there derived under the assumption that all generation participating in the capacity market is provided by firm capacity—typically conventional generation. (The GB capacity market does not attempt to take account of possible variation in *energy* costs according to the type of capacity selected; in practice this would be extremely difficult.) We examine the extent to which the above approach generalises to include variable generation and storage. Our particular interest is in the extent to which an economic criterion based on the valuation of EEU may continue to be reduced to a risk-based criterion expressible in terms of an LOLE-based security-of-supply standard. It turns out that this may or may not be possible in the case where the capacity-providing resources include variable generation, depending on the statistical characteristics of the latter, but is not readily possible when these resources include storage.

We thus consider the economic problem of choosing a set of resources R so as to minimise the overall economic cost

$$\text{VOLL} \times \text{EEU}(R) + c(R), \tag{13}$$

where VOLL is as above and where $c(R)$ is the cost of providing the set of resources R. We consider three cases of increasing generality.

All resource provided by firm capacity. We assume that all resource is provided by firm capacity, and that the latter is reasonably continuously available. Thus the set of resources R may be identified with the level of firm capacity it provides. The overall cost (13) is a convex function of that capacity, and it follows that this function is minimised at the level of capacity R such that

$$\text{VOLL} \times \text{EEU}'(R) + c'(R) = 0, \tag{14}$$

where, as usual, $\text{EEU}'(R)$ is the derivative of $\text{EEU}(R)$ with respect to firm capacity, evaluated at R, and where $c'(R)$ is the similarly the derivative with respect to firm capacity of the cost $c(R)$ of obtaining that capacity. In GB the derivative $c'(R)$, evaluated at the optimal level of capacity R as given by the solution of (14), is commonly referred to as the *cost of new entry* (CONE)—see [6]. It thus follows from (14) and from the earlier result (9) that the optimal level of capacity minimising (13) is given by the solution of

$$\text{LOLE}(R) = \frac{\text{CONE}}{\text{VOLL}}. \tag{15}$$

The cost function $c(R)$—and in particular the quantity CONE $= c'(R)$ evaluated at the optimal level of capacity R—is of course determined by the capacity market. However, $c'(R)$ usually varies slowly with R and may often be estimated in advance of any capacity auction (e.g. on the basis of earlier auctions—again see [6]). In this case the relation (15) re-expresses the economic criterion of the present section (that of minimising (13)) as a simple LOLE-based criterion, and the determination of the minimum-cost level of capacity R such that (15) holds is as described in Section 4.

The above theory is sometimes used as an economic justification for the present LOLE-based GB reliability standard. A description is given in [6], where the central values of VOLL (17/kWh) and CONE (49/kW per year) identified in that report correspond approximately, via (15), to the present GB reliability standard of a (maximum) LOLE of 3 hours per year. Nevertheless, the values of CONE identified in recent GB capacity auctions have varied very widely and are in all cases less than the value quoted above (see [13]), so that the above economic justification for the present GB standard remains a somewhat theoretical one. An alternative mechanism, now implemented within the most recent GB capacity auctions, is that of including provision for obtaining a total level of capacity which depends on the bidding taking place within the auction itself, so that the lower the level of bids the greater the capacity obtained. This may be done through the specification of a *demand curve* (see [14, 12, 22, 10]) which specifies the total level of capacity to be obtained as a function of the clearing price in the auction. This approach is straightforwardly implemented in, for example, the GB descending clock auction as described in Section 4: as the unit price is decreased in successive rounds of that auction, so the target capacity to be obtained is increased in line with the specified demand curve; the auction clears when the offered capacity remaining in the auction is equal to the current target capacity. While the GB demand curve is currently determined by government policy, it could of course be chosen so that the relation (15) was satisfied for the set of resources R obtained in the capacity auction, with CONE as the dynamically determined clearing price of the auction. The auction result would then satisfy the economic criterion (minimisation of (13)) of the present section.

All resource provided by generation of some form. Of interest now is the extent to which above theory generalises to the more complex situation in which all capacity is provided by either conventional or variable generation, but in which storage is not present. There is then no scalar measure of capacity which is sufficient to determine either EEU or LOLE. We do, however, have the following result.

Proposition 1. *Suppose that all capacity-providing resources are provided by some form of generation, and suppose further that, as the set R of such resources is varied, there is a one-one correspondence between values of* $\mathrm{EEU}(R)$ *and values of* $\mathrm{LOLE}(R)$*. Then the optimal set of resources R minimising the overall economic cost* (13) *is again that which minimises the cost $c(R)$ of providing them subject to the constraint* (15)*.*

A formal proof of Proposition 1 is given in the appendix. The essence of the argument is that the existence of above one-one correspondence ensures that minimisation of $c(R)$ subject to a constraint on $\mathrm{EEU}(R)$ is equivalent to minimisation of $c(R)$ subject to the corresponding constraint on $\mathrm{LOLE}(R)$, thereby making it possible to "pivot" a criterion based on the former to one based on the latter.

However, the above correspondence between values of $\mathrm{EEU}(R)$ and $\mathrm{LOLE}(R)$, while clearly trivial in the case where all resource is provided by firm capacity only (since both are decreasing functions of the level of that firm capacity), is not *guaranteed* in the case where resource is also provided by variable generation. It is possible that, in assessing its contribution to capacity adequacy, variable generation is capable of being treated *as if* it were firm capacity at a constant and appropriately "de-rated" level, so that the above one-one correspondence between EEU and LOLE is maintained. Provided the level of variable generation is not great, this will be close to being the case when, for example, the process of variable generation is statistically independent of that of demand, in which case the de-rated level of variable generation is close to its mean value—see [21]. However, it is also possible that the pattern of availability of variable generation—for example, throughout the day—is such that this correspondence is not maintained. For example, it is possible that on occasions periods of solar generation might be contained within periods of loss-of-load in such a way that this generation contributes to the reduction of unserved energy without reducing the duration of the loss-of-load periods. (This is particularly possible in countries where energy shortfalls tend to occur in the middle of the day.) Under such circumstances the determination of the optimal set of capacity-providing resources minimising the overall economic cost (13) becomes more complicated.

Resource also provided by storage. This above theory does not directly generalise to the case where the capacity-providing resources include storage, for then, for the reasons indicated in Section 3 and following from the flexibility of the way in which storage may be used, the result (9)—upon which the above theory depends—no longer holds. In this case the determination of the optimal set of capacity-providing resources minimising (13) again becomes more complicated.

Finally, we remark that the economic cost of unserved energy given by VOLL $\times$ EEU(R) may be replaced by other measures if the latter are more appropriate. Thus, for example, if unserved energy were valued more highly at higher levels of shortfall, one might replace EEU by some form of appropriately weighted risk metric. The above theory would then require only straightforward modifications.

6 Example

In this section we present a detailed example of energy storage and firm capacity competing in a capacity market auction, and designed to illustrate most of the theory of the present paper. In particular, we are concerned to show how to value correctly the contribution of individual stores within a market to which storage makes a significant contribution. The dimensions of the example correspond approximately to those of the current GB electricity system, except that we allow more storage than is currently present in that system. The objective within the example is that of choosing at minimum cost a set of resources R to meet a given EEU reliability standard. We take this to be 2746 MWh per year, as this corresponds to an LOLE of approximately 3 hours per year (the GB standard) for generation and demand broadly comparable to the current GB system.

We first create a credible *background* supply-demand balance process against which the capacity auction is to take place—it is almost always the case that in any capacity auction there is pre-existing capacity already committed, e.g. from multi-year auctions held in previous years, which therefore does not compete in the current auction. For this background process we assume a set of 230 conventional generators with a total of 61.36 GW of installed capacity. Capacities and outage probabilities for these conventional generators are taken from a 2015–16 National Grid scenario for GB, with random noise added to protect the sensitivity of the data. The availability of each generator is modelled as a two-state Markov process in which each generator is either completely available or completely unavailable, with a mean time to failure of 50 hours and a mean time to repair such that the equilibrium outage probability of the generator agrees with National Grid's scenario values (see [7] for more details). An empirical demand-net-of-wind process is created from paired hourly observations of GB electricity demand and wind generation for the winter season 2010–11 rescaled to 2015–16 levels of demand and an assumed installed total wind generation capacity of 14 GW (see [20] for more details of this). From this demand-net-of-wind process are subtracted 100 simulations of the conventional generation process to create 100 simulations of a residual demand process, which defines a sufficiently representative background process for the present example. The further firm capacity which would require to be subtracted from this background demand process in order to meet the target reliability standard of 2746 MWh per year is 1973 MW. This residual demand or background process is to be managed instead from the further generation and storage resources obtained in the capacity market. The volume of resource to be obtained in this market corresponds to that which might be required in a "top-up" capacity market, such as that held in GB one year ahead of real time.

Competition in the capacity market is provided by 120 stores and 30 units of firm capacity. The rate and energy constraints of the stores are chosen to be representative of those typically found in systems such as that of GB. However, in order to illustrate some of the concerns of the present paper, which are most relevant to future systems in which storage may play a larger part, we choose a relatively large number of such stores. We assume there are 60 stores with a rate constraint of 50 MW; 10 of these stores have an energy constraint of 12.5 MWh, 15 have an energy constraint of 25 MWh, 15 have an energy constraint of 50 MWh, and 20 have an energy constraint of 100 MWh. We further assume there are 60 stores with a rate constraint of 100 MW; 10 of these stores have an energy constraint of 25 MWh, 15 have an energy constraint of 50 MWh, 15 have an energy constraint of 100 MWh and 20 have an energy constraint of 200 MWh. The firm capacity units are assumed to have capacities between 10 MW and 100 MW. There are three units for each multiple of 10 MW capacity (i.e. three units with capacity of 10 MW, three with capacity of 20 MW, etc). The minimum total price c_i at which each store or unit of firm capacity i is prepared to make itself available (see Section 3) is set to be approximately proportional to the EFC of that resource with some statistical variation. (For a store this EFC is taken to be as estimated against an approximation of the set of resources R finally selected in the capacity market auction, while for a firm capacity resource its EFC is of course just that capacity.) This means that the stores and firm capacity units are generally competitive with each other in any capacity auction. The constant of proportionality is comparable with that which has held in recent GB capacity auctions. (The latter is very considerably lower than the reference value of 49/kW per year discussed in Section 5.)

The EFC $efc_R(i)$ of any storage unit i, relative to (the supply-demand balance process defined by) any set of resources R to which it is being added, is calculated as described in Section 3. In all cases storage is optimally scheduled for the minimisation of EEU as described in that section and on the assumption that all storage may be completely recharged overnight, but not at other times. Recall that, from the algorithm described in Section 3, the optimal scheduling of storage at any time t depends only on the shortfall process at time t and the charge states of the stores at time t. This scheduling may therefore be done separately for each simulation of the shortfall or background process to which the storage is being added. The EEU resulting from any configuration of resources in the capacity auction is thus estimated by the average remaining energy unserved calculated separately using the above algorithm for each of the 100 simulations of the initial background process.

The determination of the optimal set of resources R meeting the required EEU reliability standard at minimum cost is as described in Section 4. Thus it is necessary to determine a set of resources R—from amongst those participating in the capacity market—and a clearing price p such that the reliability condition (11) and the equilibrium conditions (12) are satisfied. (Recall, from Section 4, that the conditions (11) and (12) are necessarily satisfied at the optimal set R and may reasonably be expected to determine that set uniquely.) Again as discussed in Section 4, the solution of these equations would be straightforward in the case where the EFCs $efc_R(i)$ of the participating resources were known in advance: it would simply be a case of organising these resources in a merit-order stack and then choosing a minimum clearing price p such that, under (12), sufficient resources were selected so that (11) was also satisfied—the corresponding total price $p\sum_{i\in R} efc_R(i)$ would then be the minimum possible. However, for resources other than those capable of being modelled as firm capacity (at least to a good approximation, as is the case for suitably derated conventional generation), their EFCs $efc_R(i)$ depend on the finally accepted set R, and so are not known prior to the attempted solution of the conditions (11) and (12). Under those circumstances in which the nature of the final resource mix is reasonably predictable—as where there is relatively little resource which is not capable of being modelled as firm capacity, or where there is significant experience from previous auctions—it may be possible to make reasonable advance estimates of such EFCs. However, in the present example there is a relatively large volume of storage participating in the market, and the EFCs of the stores are not well known in advance of the solution of the conditions (11) and (12). We therefore solve these equations iteratively: we make initial estimates of the storage EFCs—as described below for the present GB capacity market—and use these estimated EFCs to obtain an initial clearing price p and set of resources R such that (11) and (12) are satisfied with these initial estimates substituting for the true EFCs (again this is done by forming an merit-order stack using these estimated EFCs); the set of resources R obtained is then used to form revised EFC estimates $efc_R(i)$ for all storage resources participating in the capacity market, and the process repeated until after several iterations R converges to a fixed set satisfying the conditions (11) and (12) with the EFCs now correctly calculated with respect to that set R. The solution obtained—the set of resources meeting the required reliability standard at *minimum cost* on a *pay-as-clear basis*—is a combination of 550 MW of firm capacity and a set of stores the sum of whose EFCs $efc_R(i)$ evaluated with respect to R is 1134 MW. These are the EFCs appropriate to the *marginal* contributions of the individual stores (as is appropriate to the optimal operation of the capacity market—see Section 4), and determine their payments received in the capacity market—the payment received by each individual store i being of course at least as great as the minimum price c_i which that store is prepared to accept. However, the EFC of the entire set of accepted stores—treated as a single unit—is 1423 MW, i.e. this is the amount of further firm capacity which would be required to substitute for the entire set of accepted stores in order to meet the required EEU reliability standard. This is a fairly extreme case, in the current example in which a large volume of storage is present, of the *nonadditivity* of EFCs over significant numbers of resources, as discussed at the end of Section 4. It is further a reflection of the observation, at the end of Section 3, that the EFC of further storage resources is typically reduced by the presence of existing storage resources. A similar phenomenon is well known in the case of certain variable generation such as wind power, although the reason for the latter is very different and results there from the temporal coincidence of wind resources at different locations.

In GB the capacity market is settled through a single-pass descending-clock auction which identifies the required clearing price. (Similar auctions are conducted in other countries.) Thus in GB the EFCs of storage facilities—which now participate in the GB capacity market—are estimated in advance of the capacity auction. The EFC $efc_R(i)$ of each store i is determined as described in the present paper, but is done so with respect to an "initial" set of resources R which is taken to be the amount of *firm capacity* which would be required in order to meet the GB reliability standard. In GB most resource currently participating in the market is indeed firm capacity, or at least conventional generation, and the storage EFCs estimated as above are close to their true values, i.e. to those calculated with respect to the set of resources R finally obtained in the market. However, the example of the present section is chosen so that storage plays a significant role—something which may very well also be the case in future real energy systems. If, in this example, the EFC $efc_R(i)$ of each store i is determined as is current practice in GB (i.e. as its contribution when added to an existing set of resources R consisting of firm capacity and just sufficient to meet the required reliability standard) then the storage EFCs so obtained prove to be considerable overestimates in comparison with their true values—those calculated with respect to the optimal set of resources determined as described above and meeting the reliability standard at minimum cost. The reason for this is again as discussed at the end of Section 3: storage added to existing storage is less valuable than when added to firm capacity providing the same level of reliability. A consequence, *in the present example*, of this overvaluation of the contribution of storage would be that—if uncorrected—*all* the resource obtained in the capacity market would consist of storage. Further, with the realistic resource costs chosen for this example, the cost of obtaining sufficient (all storage) resources to correctly meet the required reliability standard would be £58.1m, whereas the true cost of the optimal resource set (as identified earlier and consisting of a mixture of firm capacity and storage) to meet that reliability standard is £49.7m. The more careful evaluation of the contribution of storage in the present example therefore leads to a cost saving of 14.5% in the capacity market of this example.

Finally we test more carefully the extent to which the continuity and smoothness assumptions of Section 2 are applicable in the current example. Figure 2 shows the effect of starting with the background system of pre-existing capacity and gradually adding to it all the capacity-providing resources competing in the auction, taken in the order of the final merit-order stack. The figure plots residual EEU against cumulative EFC. At each point the latter is again the firm capacity which would substitute for the resources added so far while maintaining the same level of residual EEU (so that the underlying relationship defined by the plotted points is in fact independent of the order in which the resources are taken). We see that, as required for the continuity assumption, there are no large gaps between successive points. In particular these points are dense in the region corresponding to the target EEU for the auction—as represented by the horizontal line. We therefore conclude that the continuity assumption is sufficiently well satisfied.

Figure 3 examines the validity of the smoothness assumption (3). The set R is taken to be the set of resources identified by the capacity auction (as is appropriate in the application of the smoothness assumption.) Figure 3 shows, as a "heat plot", the percentage variation between the left and right sides of (3) where i and j both correspond to stores of 100 MWh energy and levels of power which in each case vary from 10 to 50 MW. The lowest powered stores contribute only modest additional capacity, and here the smoothness assumption (3) is seen to be very accurate. The highest powered stores contribute substantial additional capacity; nevertheless here the difference between the left and right sides of (3) is no more than 5%. Similar results would be obtained if the additional resources i and j corresponded to firm capacity or conventional generation, as in established current capacity markets. The smoothness assumption therefore also appears sufficiently well satisfied here.

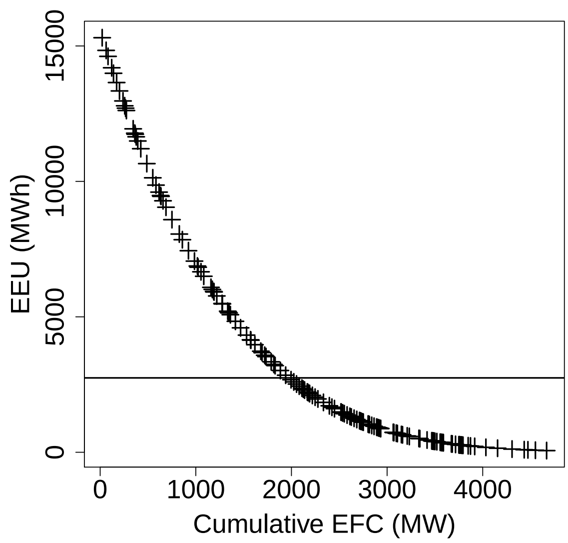

Figure 2: Plot of residual EEU again cumulative EFC (see text). The horizontal line corresponds to the target EEU standard.

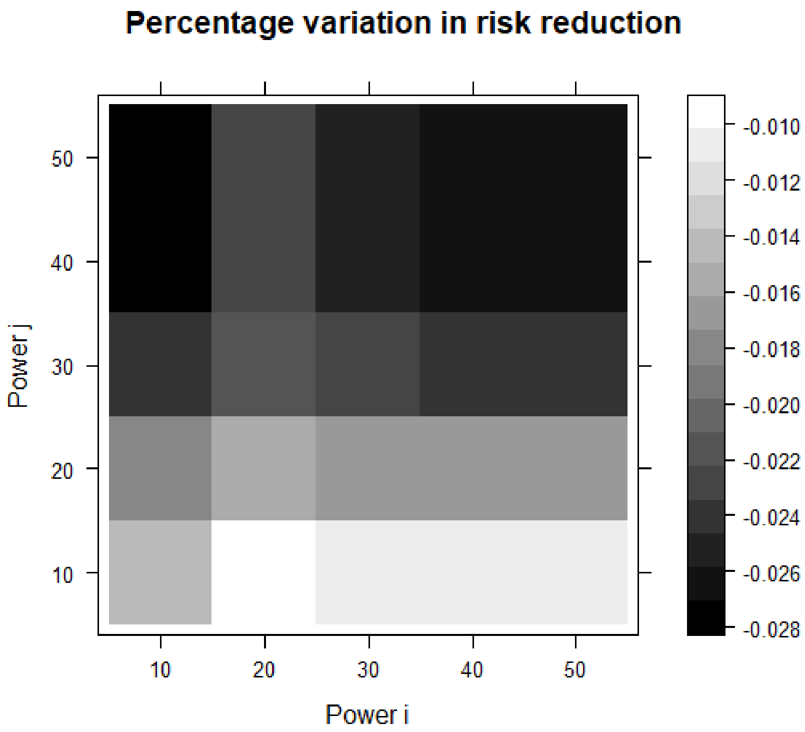

Figure 3: Heat plot to examine the accuracy of the smoothness assumption (3). The set R is as defined by the outcome of the capacity auction, the x- and y-axes give respectively the powers of the additional stores i and j, and the plot shows the percentage variation between the left and right sides of (3).

Acknowledgements

The authors would like to thank the Isaac Newton Institute for Mathematical Sciences for support and hospitality during the programme Mathematics of Energy Systems when work on this paper was undertaken. This work was supported by EPSRC grants numbers EP/I017054/1, EP/K002252/1, EP/K002252/1 and EP/R014604/1". The authors are grateful to many colleagues—in particular David Newbery, Benjamin Hobbs, Muireann Lynch and colleagues at National Grid ESO—for helpful comments and discussions.

References

[1] Roy Billinton and Ronald N. Allan. *Reliability Evaluation of Power Systems.* Springer, 2nd edition, 1996.

[2] C. Bothwell and B. F. Hobbs. Crediting wind and solar renewables in electricity capacity markets: The effects of alternative definitions upon market efficiency. *The Energy Journal*, 38(1):173–188, 2017.

[3] J. Bowring. Capacity Markets in PJM. *Economics of Energy and Environmental Policy*, 2(2):47–64, 2013.

[4] J. R. Cruise and S. Zachary. Optimal scheduling of energy storage resources. http://arxiv.org/abs/1808.05901, 2018.

[5] C. J. Dent and S. Zachary. Further results on the probability theory of capacity value of additional generation. In *2014 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, pages 1–6, July 2014.

[6] Department of Energy and Climate Change. EMR Consultation Annex C: Reliability Standard Methodology. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/223653/emr_consultation_annex_c.pdf, 2013. Accessed 3 August 2018.

[7] G. Edwards, S. Sheehy, C. J. Dent, and M. C. M. Troffaes. Assessing the contribution of nightly rechargeable grid-scale storage to generation capacity adequacy. *Sustainable Energy, Grids and Networks*, 12:69–81, 2017.

[8] M. Evans, D. Angeli, and S. H. Tindemans. Robustly maximal utilisation of energy-constrained distributed resources. https://arxiv.org/pdf/1710.06302.pdf, 2018.

[9] D. Helm. Cost of energy: independent review. https://www.gov.uk/government/publications/cost-of-energy-independent-review, 2017. Accessed 6 January 2019.

[10] B. Hobbs, M-C. Hu, J. G. Inon, S. Stoft, and M. P. Bhavaraju. A dynamic analysis of a demand curve-based capacity market proposal: The PJM reliability pricing model. *IEEE Transactions on Power Systems*, 22(1):3–14, 2007.

[11] A. Keane, M. Milligan, C. J. Dent, B. Hasche, C. DAnnunzio, K. Dragoon, H. Holttinen, N. Samaan, L. Sder, and M. OMalley. Capacity value of wind power. *IEEE Transactions on Power Systems*, 26(2):564–572, 2011.

[12] R. W. Moye and S. P. Meyn. Redesign of US electricity capacity markets. In S. Meyn, T. Samad, I. Hiskens, and J. Stoustrup, editors, *Energy Markets and Responsive Grids: Modeling, Control, and Optimization*, pages 73–103. Springer, 2018.

[13] National Grid plc. Capacity Market Auction Guidelines 2018. https://www.emrdeliverybody.com/Lists/Latest%20News/Attachments/197/Auction%20Guidelines%202018%20v2.0.pdf. Accessed 1 December 2018.

[14] National Grid plc. Capacity market: The demand curve. https://www.emrdeliverybody.com/CM/The-Demand-Curve.aspx. Accessed 1 July 2019.

[15] National Grid plc. Electricity Capacity Report 2018. https://www.emrdeliverybody.com/Lists/Latest%20News/Attachments/189/Electricity%20Capacity%20Report%202018_Final.pdf. Accessed 2 August 2018.

[16] D. M. G. Newbery. Missing money and missing markets: Reliability, capacity auctions and interconnectors. *Energy Policy*, 94:401–410, 2015.

[17] Ofgem. Wholesale Energy Markets in 2016. https://www.ofgem.gov.uk/system/files/docs/2016/08/wholesale_energy_markets_in_2016.pdf, 2016. Accessed 1 August 2018.

[18] I. Staffell and R. Green. Is there still merit in the merit order stack? the impact of dynamic constraints on optimal plant mix. *IEEE Transactions on Power Systems*, 31(1):43–53, 2016.

[19] S. Stoft. *Power System Economics: Designing Markets for Electricity*. IEEE/Wiley, 2002.

[20] A. L. Wilson, S. Zachary, and C. J. Dent. Use of meteorological data for improved estimation of risk in capacity adequacy studies. *PMAPS*, 2018.

[21] S. Zachary and C. J. Dent. Probability theory of capacity value of additional generation. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 226(1):33–43, 2011.

[22] F. Zhao, T. Zheng, and E. Litvinov. Constructing Demand Curves in Forward Capacity Market. *IEEE Trans. Power Syst.*, 33(1):525–535, 2018.

Appendix: technical results

In this appendix we formalise and prove two technical results in the present paper.

Proof of equation (9). We prove the result given by equation (9), namely that for any set of capacity-providing resources R which consists entirely of generation, either conventional or variable, we have that the derivative $\mathrm{EEU}'(R)$ of $\mathrm{EEU}(R)$ with respect to variation of firm capacity is given by $-\mathrm{LOLE}(R)$.

For each time t let the random variable Z_t be the supply-demand balance at time t corresponding to the use of the set of resources R. Then, from (2), for any addition to R of firm capacity equal to δ,

$$\mathrm{EEU}(R+\delta) = \sum_{t=1}^{n} \int_{-\infty}^{0} \mathbf{P}(Z_t + \delta < z)\, dz = \sum_{t=1}^{n} \int_{-\infty}^{-\delta} \mathbf{P}(Z_t < z)\, dz,$$

so that, differentiating with respect to δ and then setting $\delta = 0$, we have

$$\mathrm{EEU}'(R) = -\sum_{t=1}^{n} \mathbf{P}(Z_t < 0) = -\mathrm{LOLE}(R)$$

from (1) as required.

Proof of Proposition 1. For any possible *risk level* k of EEU, define R_k to be the set of resources R which minimises the cost $c(R)$ subject to the constraint $\mathrm{EEU}(R) = k$. Given the risk level k, the subproblem of determining R_k is the problem considered in Section 4. The additional problem in the minimisation of the overall economic cost (13) may therefore be viewed as that of determining the value of k such that R_k minimises

$$\mathrm{VOLL} \times \mathrm{EEU}(R_k) + c(R_k), \tag{16}$$

(with $\mathrm{EEU}(R_k) = k$) i.e. that of determining the optimal level of EEU to be then obtained at minimum cost. We may consider the effect of infinitesimal variation of the risk level k by considering the corresponding required infinitesimal variation in EFC, where the latter is defined with respect to R_k. At that value of k such that the overall economic cost (16) is minimised, we have stationarity with respect to such variation, so that at this value of k, analogously to (14),

$$\mathrm{VOLL} \times \mathrm{EEU}'(R_k) + c'(R_k) = 0, \tag{17}$$

where it follows from the definition of EFC in Section 3 that $\mathrm{EEU}'(R_k)$ may continue to be interpreted as defined in that section, i.e. as the derivative of $\mathrm{EEU}(R_k)$ with respect to firm capacity, and where $c'(R_k)$ may similarly continue to be interpreted as the cost of new entry (CONE) at the level of resource defined by R_k. Since it is assumed that all capacity-providing resource consists of some form of *generation*, the result (9) holds and so, analogously to (15), it follows that at the value of k such that the overall cost (13) is minimised,

$$\mathrm{LOLE}(R_k) = \frac{\mathrm{CONE}}{\mathrm{VOLL}}. \tag{18}$$

Now recall that, for each k, the set R_k is defined as minimising $c(R)$ subject to the constraint $\mathrm{EEU}(R) = k$. It follows from the assumed one-one correspondence between values of $\mathrm{EEU}(R)$ and those of $\mathrm{LOLE}(R)$ that R_k is also the set of resources R which minimises $c(R)$ subject to the corresponding constraint on $\mathrm{LOLE}(R)$. It thus follows from (18) that, exactly as in the case where all resource is provided by firm capacity only, the determination of the optimal set of capacity-providing resources minimising the overall economic cost (13) is again given by the minimisation of the cost $c(R)$ subject to the constraint (15).