Report for the Independent Market Operator

2014 Relevant Level Methodology Review Final Report

Dr Richard Tooth, Dr James Swansson

5 December 2014

sapere research group

About Sapere Research Group Limited

Sapere Research Group is one of the largest expert consulting firms in Australasia and a leader in provision of independent economic, forensic accounting and public policy services. Sapere provides independent expert testimony, strategic advisory services, data analytics and other advice to Australasia's private sector corporate clients, major law firms, government agencies, and regulatory bodies.

Sydney Level 14, 68 Pitt St GPO Box 220 NSW 2001 Ph: + 61 2 9234 0200 Fax: + 61 2 9234 0201	**Canberra** Unit 3, 97 Northbourne Ave Turner ACT 2612 GPO Box 252 Canberra City, ACT 2601 Ph: +61 2 6267 2700 Fax: +61 2 6267 2710	**Melbourne** Level 2, 65 Southbank Boulevard GPO Box 3179 Melbourne, VIC 3001 Ph: + 61 3 9626 4333 Fax: + 61 3 9626 4231
Wellington Level 9, 1 Willeston St PO Box 587 Wellington 6140 Ph: +64 4 915 7590 Fax: +64 4 915 7596	**Auckland** Level 17, 3-5 Albert St PO Box 2475 Auckland 1140 Ph: +64 9 913 6240 Fax: +64 9 913 6241	

For information on this report please contact:

Name: Dr Richard Tooth

Telephone: 02 9234 0216

Mobile: 0412 105 817

Email: rtooth@srgexpert.com

Contents

Glossary vii

Summary ix

1. Introduction **15**

1.1 Introduction 15
1.2 Scope of the review 15
1.3 Changes from the draft report 16
1.4 Outline of this paper 16

2. Background **17**

2.1 About Capacity Credits 17
2.2 Estimating capacity value 17
2.3 The current methodology 21
2.4 Criteria for evaluating the methodology 22

3. Performance of methodology **24**

3.1 Introduction and overview 24
3.2 The impact of the methodology 27
3.3 Performance of the methodology 28
3.3.1 Accuracy 28
3.3.2 Robustness 34
3.3.3 Volatility 35
3.3.4 Practical issues 36
3.4 Summary 37

4. Review of options **38**

4.1 Introduction 38
4.2 Method of determining IGF contribution 38
4.2.1 Penetration of intermittent generation and use of approximation based methods 38
4.2.2 Use of an alternative method 40
4.2.3 The use of LSG as a basis for selecting TIs and alternatives 41
4.2.4 How LSG is applied 41
4.2.5 Should existing IGFs be impacted by new IGFs? 43
4.3 K-factor adjustment 45
4.3.1 Revised estimate of the K parameter 45
4.4 The U-factor adjustment — Dealing with correlation between IGF output and peak demand 48
4.4.1 Introduction 48
4.4.2 Performance of IGFs at extreme temperature days 50
4.4.3 Overview of options 52
4.4.4 Using only Trading Intervals that are representative of the extreme peaks 53
4.4.5 The 'regression' approach 55

4.4.6 Update – U-factor 58
4.4.7 Summary 60
4.5 Other issues in the selection of Trading Intervals 60
4.5.1 The number of separate days (TIs) in a year 61
4.5.2 The number of years for selecting TIs 63
4.5.3 Weighting of Trading Intervals 65
4.5.4 Number of Trading Intervals per day 65

5. Conclusion and summary 67
5.1 Overview 67
5.2 Summary of recommendations 67

References 68

Appendices

Appendix 1 Theory and international experience 70
Appendix 2 Background information on the WEM 73
Appendix 3 Load for Scheduled Generation 74

Tables

Table 1: IGF categories used in this report 25
Table 2: Performance of IGFs at extreme peaks 51
Table 3: List of peak Trading Intervals on 40+ Degree Days 53

Figures

Figure 1: Map of renewable energy sources for the SWIS 26
Figure 2: Changes in Relevant Level 27
Figure 3: Value of Capacity Credits 28
Figure 4: Output of IGF facilities at peak TIs – five years to April 2014 31
Figure 5: Output of IGF operational since 2013 at peak TIs – two years to April 2014 32
Figure 6: An individual IGF results – example 1 33
Figure 7: An individual IGF results – example 2 33
Figure 8: RL method applied since 2013/14 without transition 36
Figure 9: Penetration of intermittent generation - capacity 39
Figure 10: Penetration of intermittent generation – energy supplied 40
Figure 11: Peak LSG vs maximum temperature in day by year 49
Figure 12: Air temperature and maximum temperature 50

Figure 13: Relationship between wind speed and air temperature at TIs between 4pm and 5pm on peak TI days 52

Figure 14: Output on peak days of IGF fleet that have been operational since 2007 54

Figure 15: Number of hot days available 55

Figure 16 Output of IGF established since 2007 on peak TIs (5 years to 2014) 56

Figure 17: Performance of IGF fleet on hot days 59

Figure 18: Time of the peak TIs by rank of day in year 62

Figure 19: Date of peak TIs by rank of day in year 62

Figure 20: Temperature at time of peak TIs by rank of day in year 63

Figure 21: Timing of MG peaks by year 64

Figure 22: Timing of LSG peaks by year 64

Figure 23: IGF output and timing of peak TIs 66

Glossary

COPT	Capacity outage probability table
DSP	Demand Side Programme
EFLSG	Existing Facility LSG (EFLSG)
ELCC	Effective Load Carrying Capability – a measure of the additional load that the system can supply with the particular generator of interest, with *no net change in reliability*
ELCC Risk Method	A method of calculating ELCC using iterative numerical calculations on data containing the probability distribution of output of each facility
IGF	Intermittent Generator Facility
IMO	Independent Market Operator
LSG	Load for Scheduled Generation
LOLE	Loss of load expectation, the sum of LOLP in a planning period e.g. year
LOLP	Loss of load probability, in an operation period, e.g. hour
MG	Market generation – the total of sent-out generation
MW	Megawatt
MWh	Megawatt hour
NFLSG	New Facility LSG
POE	Probability of Exceedance
2011 Report	The report by Tooth (2011) titled 'Capacity Value of intermittent generation: Public Report.'
RCM	Reserve Capacity Mechanism
RCT	Reserve Capacity Target
RL	Relevant Level
SWIS	South West Interconnected System
TI	Trading Interval
TD	Trading Day
WEM	Wholesale Electricity Market

Summary

Introduction

Generation facilities in the Wholesale Electricity Market (WEM), including Intermittent Generator Facilities (IGFs), are awarded Certified Reserve Capacity in recognition of their contribution in meeting the reliability criteria determined for the WEM. The Relevant Level (RL) is a quantity that sets the Certified Reserve Capacity for an IGF, subject to certain conditions.

In accordance with the WEM Rules (Market Rules) the Independent Market Operator (IMO) must undertake a review of RL methodology every three years. The RL methodology was changed in 2012 to more closely align the capacity awarded with the contribution to improving reliability.

As part of the IMO's review for 2015, Sapere Research Group (Sapere) was commissioned to review the performance of the RL methodology introduced in 2012, and any proposed amendments to the methodology, including possible amendments to the Market Rules.

A draft version of this report was published on 22 October 2014. This final report reflects some minor changes as a result of the public consultation which included a participant workshop and receipt of one submission to the draft report.

Background

The value of a Capacity Credit is the contribution that a generation facility makes towards the power system reliability in meeting demand. A widely accepted measure of capacity value is the Effective Load Carrying Capability (ELCC), which is the additional load that the system can supply with the particular generator, with *no net change in reliability*. In the WEM the most stringent planning criterion relates to the risk of not being able to meet the system peak and therefore it is appropriate to measure reliability in terms of the loss of load probability (LOLP).

The capacity value is not something that can be directly measured from observed data. However it is known that the capacity value will approximate the IGF's average output when there is no surplus capacity less an adjustment for the variance of its output. Additional adjustment may also be required to account for the risk that the output of IGFs during the periods examined differs to that when there is no surplus.

In summary a useful generic formula for simple methods to estimating capacity value is:

Capacity value =	1. Average IGF output in peak periods	***Less***	2. An adjustment for the variability of IGF output

The current method

The methodology (prescribed in Appendix 9 of the Market Rules) involves calculating the RL as:

RL for an IGF =

The mean IGF output during peak trading intervals (TIs), less

A K-factor adjustment, which is K times the variance of the IGF output during the peak TIs, where K is a parameter value (K is currently 0.003), less

A U-factor adjustment, that was introduced to reflect the issue of correlation between IGF output and peak demand. A cap is placed on the U-factor adjustment equal to one third of the mean facility output.

Peak periods are determined by the peaks in Load for Scheduled Generation (LSG), which is (with some minor qualification) the load not supplied by IGFs. These are the times when surplus capacity is lowest.

It is useful to decompose the RL methodology into two components being:

- the mean output at peak LSG less the K-factor adjustment, and
- the U-factor adjustment.

The first component is consistent with the probability theory mentioned above. The second component, the U-factor adjustment was added as pragmatic method given the limited available data at the time.

Performance of methodology and options

While the availability of data presents challenges in determining how well the RL methodology has performed, overall it appears to provide a reasonable estimate of the capacity value provided by IGFs.

Ultimately the methodology needs to be evaluated against the Wholesale Market Objectives; however, for the purposes of this report the criteria for evaluation have been distilled into consideration of accuracy, robustness, volatility and practicality.

Capacity value is in effect a forecast and cannot be directly observed from the data. As such, there is no clear benchmark for assessing the accuracy of the RL methodology in estimating the capacity value of IGFs.

The first component is an approximation method based on probability theory. Generally such an approximation method might be assessed by examining the performance against a more detailed risk calculation. However, this is not available and, regardless, may not be practical due to the lack of data at the very peak periods.

The U-adjustment was introduced to address concerns that IGF output would not be sufficient at peak TIs when it is very hot and the system would be under greatest stress. As such it is possible to assess this by examining the size of the U-factor adjustment and how IGF output varies between peak times and periods of extreme stress.

Given the lack of observations on days with very high temperatures, assessment is difficult and requires a degree of judgement. Our assessment is that the results of U-adjustment appear broadly consistent with the performance of the IGFs at very high temperatures. However, as we would expect, the methodology appears to apply more appropriately to some IGFs than others.

The results delivered by the methodology appear (again based on a subjective assessment of limited data) to be reasonable for solar and other new facilities. The methodology itself is very general; it is not dependent on any type of technology or any profile of demand.

The RL methodology does not appear to generate overly volatile results. While there has been some movement (excluding the impact of transition) in the RLs calculated over the last few years, the movement has been modest. The relative standard error (RSE) of these results by IGF is at most 15% and is under 10% in most cases.

While the procedure to determine RL has some complexity, there do not appear to be any material practical problems with its application.

Review of options to improve the methodology

For this review we considered a range of issues, largely driven by our scope of works, which we categorised as:

- the overall method (including the use of LSG)
- the K-value adjustment
- the U-value adjustment and the issue of IGF output being correlated with demand at extreme peaks, and
- other matters relating to the selection of the TIs.

Our overall assessment from this review is that most of the components of the methodology are appropriate. The greatest challenge is the lack of data at peak and determining the most appropriate form of the U-value adjustment.

The overall method (including the use of LSG)

The overall risk-based approximation approach appears appropriate given the alternative of a time-based approach and the impracticality of a detailed ELCC Risk Method.

There is a very clear logic to using LSG to identify peak TIs. Peak LSG is when surplus capacity is lowest and therefore (all else being equal) there is the greatest LOLP. A useful feature of using LSG to identify peak TIs is that it largely controls for the covariance between the output of IGFs; that is, it (appropriately) favours those IGFs who produce more at peak times when other IGFs are producing less.

There are some subtle possible improvements with how the contribution of IGF's is determined using LSG. However, in our opinion the potential benefits of any change are small relative to the costs of change.

The K-factor adjustment

The K-factor adjustment is based on recognised probability theory. The original K parameter was based on international benchmarks.

In updating the K parameter we undertook two key steps. Using information on the distribution of the forecast for peak demand and the probability theory we estimated a K parameter that is tailored for the WEM. In determining a final K value to be applied we have also taken into consideration how the IGF output is calculated. The results from these two steps offset each other leaving us to recommend a revised K parameter of zero.

The U-factor adjustment

Significant effort was undertaken to assess whether a viable alternative to the current U-factor adjustment was possible. The key constraint in considering alternatives is the lack of data at extreme peak TIs (i.e. TIs where the surplus capacity is close to zero). Consistent with the analysis in the 2011 review, we found a very strong relationship between peak demand (and peak LSG) and maximum daily temperature and a relationship between IGF output and maximum daily temperature.

Given the constraint of the available data, two alternatives were considered for detailed assessment:

- using only TIs from extreme temperature days, and
- using a regression approach to forecast IGF output on extreme peak days based on maximum daily temperature.

The TIs from extreme temperature days approach is not recommended due to the lack of sufficient TIs. The regression approach appeared to have some promise and address some theoretical issues; however, it did not produce results that were significantly superior to the current method. The results of the regression approach were also found to be volatile.

On the basis that no other valid alternative can be found we are left with the recommendation to continue with the current U-factor adjustment.

Review of TIs to be selected and other matters

In selecting TIs to be used in the methodology, the priority has been in selecting TIs that provide additional information (and therefore reduce volatility of results) but are most representative in determining the output of IGFs at an extreme peak. The current methodology is applied on 60 TIs. 12 TIs are selected from each of 5 years with each TI selected from a separate day and with each TI given equal weight. We considered whether any of these factors should be modified.

We did not see any justification to change the TIs selected. In summary:

- We assess there would be material risks to increasing the number of TIs per year and no justification for reducing the number of TIs selected per year.
- We found evidence that the profile of peak TIs is changing over time, but we assessed there is not sufficient cause to modify the number of years used.
- We considered the value of weighting TIs but assessed that the benefits of doing so would outweigh the costs.
- We assessed that continuing to use only 1 TI from each day is appropriate. This is because using additional TIs would add very little additional information (due to the high correlation of IGF output between adjacent TIs) and that using additional TIs in a day would risk selecting TIs at non-peak times.

Parameters

The K and U parameters were examined. Based on our analysis we recommend that:

- The revised K parameter value should be zero (i.e. 0.000).
- The revised U parameter value should be 0.635 (i.e. no change from the present)

No amendments to the Market Rules are required.

Conclusion

Our overall assessment from this review is that most of the components of the methodology are appropriate. We also note that no material concerns were raised in response to the draft report.

It is useful to consider the RL methodology as consisting of two key components. We have no material concerns with the 'mean output at peak LSG less K-factor adjustment' as this component is consistent with established approximation methods and is grounded in probability theory.

The results produced by applying the U-factor adjustment appear reasonable to date and, given the data constraints, we cannot find a better alternative at this stage. Prior to the next three-year review it would be appropriate to consider options as to how to improve the data available. This might, for example, involve attempts to forecast the distribution of weather conditions on extreme peak days.

1. Introduction

1.1 Introduction

The Independent Market Operator (IMO) is responsible for administering and operating the Wholesale Electricity Market (WEM) and, accordingly, it is responsible for maintaining and developing the rules and the market related procedures that govern the operation of WEM.

In accordance with clause 4.11.3C of the WEM Rules (Market Rules), the IMO must undertake a review of Relevant Level Methodology every three years. The methodology is used to determine the Relevant Level (RL), a quantity that, subject to certain conditions, sets the Certified Reserve Capacity for a Facility (typically an Intermittent Generator Facility, IGF) for a given Reserve Capacity Cycle. IGFs are awarded Certified Reserve Capacity in recognition of their contribution in meeting the reliability criteria determined for the WEM.

In 2012, the methodology was changed to attempt to more closely align the capacity awarded with the contribution to improving reliability. The revised methodology (prescribed in Appendix 9 of the Market Rules) involves calculating the RL, broadly as:

RL=

1. Average facility output during peak periods, less
2. A function of the variance of, and average of, facility output during peak periods and two parameters referred to as K and U.

Peak periods are determined by the peaks in Load for Scheduled Generation (LSG), which is (with some minor qualification) the load not supplied by IGFs.

The revised methodology was based on the recommendations contained in a report for the IMO Board (Tooth 2011, hereafter the 2011 Report) that included analysis of IGF output during extreme peaks in demand.

The IMO's initial review, for the three year period commencing on 1 January 2015, must be completed by 1 April 2015. The IMO proposes to complete the review by December 2014, to allow time for any recommended amendments to the methodology to be progressed through the rule change process in time for certification for the 2015 Reserve Capacity Cycle.

A draft version of this report was published on 22 October 2014.

1.2 Scope of the review

The IMO has specified a series of tasks that the review must include. A summary is that the review must include:

1. a review of developments in international best practice
2. analysis of the methodology since its implementation in 2012
3. consideration of the penetration of IGFs and whether there is a need to investigate alternative valuation methodologies prior to the next three-year review

4. consideration of whether any changes are warranted to how the peak Trading Intervals (TIs) are selected for the calculations
5. consideration of whether any changes to the methodology are warranted to account for the correlation of output between Intermittent Generators
6. consideration of the K and U factor adjustments
7. consideration of the effectiveness of the methodology in meeting the Wholesale Market Objectives and achieving an appropriate balance between simplicity and accuracy
8. recommended values for K and U to be applied, and
9. details of any proposed amendments to the methodology in the Market Rules.

1.3 Changes from the draft report

Following the release of the draft report there was a public consultation period involving a participant workshop and an invitation for submissions. Only one submission was received. Two substantive issues were raised in this submission. The first related to the cost of conducting the expert review. This is discussed in section 3.3.4 (Practical issues). The second issue raised was with regard to whether there is merit in integrating elements of the RL methodology into other areas of the market rules. The issue has been noted by the IMO, but is out of scope of this report and not discussed further.

1.4 Outline of this paper

The rest of this paper is organised as follows:

- Section 2 provides background to estimating the capacity value for IGFs.
- Section 3 reviews the impact of the methodology to date.
- Section 4 reviews aspects of the methodology and whether any modifications are appropriate.
- Section 5 concludes and provides a summary of the recommendations.

The appendices contain additional material:

- Appendix 1 provides additional information on theory and international experience.
- Appendix 2 provides additional relevant background on the WEM.
- Appendix 3 provides additional background on the calculation of LSG.

2. Background

2.1 About Capacity Credits

A Capacity Credit is a notional unit of Reserve Capacity provided by a facility (being a Scheduled Generator, Non-Scheduled Generator or a Demand Side Programme).

The IMO determines the Reserve Capacity Target that determines the Capacity Credits required in fulfilling two reliability criteria known as the Planning Criterion (see Appendix 2).[1] As noted by the IMO (2013, page 67) the 'Planning Criterion applies to the provision of generation and DSM capability. It does not specifically include transmission reliability planning or cover for a major fuel disruption such as a sudden and prolonged outage of gas supply.'

The first element of the Planning Criterion relates to the risk of not meeting forecast peak demand. The second relates to the expected energy shortfalls over the year. The most stringent criterion determines the Reserve Capacity Target. To date, and in the immediate future, this has been the peak demand criterion. As noted in the 2013 Statement of Opportunities (IMO 2013, page 68):

> *To date, load factors and plant availability have been such that the Reserve Capacity Target has been set by the first element of the Planning Criterion, relating to annual peak demand. For the 2015/16 Capacity Year, the peak demand-based capacity requirement exceeds the energy-based requirement by more than 700 MW. Based on this, it is expected that the peak demand forecast will continue to set the Reserve Capacity Target for the immediate future.*

Therefore, it is understood that the RL methodology should determine the contribution of facilities in meeting the peak demand criterion.

Of note, both system demand and the output of IGFs are random variables. Therefore, the highest risk of not meeting peak system demand (referred to as loss of load probability, LOLP) may not be when system demand is highest but rather when the system demand net of the IGF output is highest. This amount is LSG, which is a key concept in the current methodology.

2.2 Estimating capacity value

The challenge of estimating the capacity value of IGFs is not unique to the WEM. Individual IGFs are assigned capacity value in other capacity markets. In energy-only markets, the capacity value of individual facilities or the fleet of IGFs is calculated for the purpose of conducting adequacy assessments.

---

1 A useful discussion of the Reserve Capacity Targets and the Planning Criterion can be found in Section 6 of the 2013 Statement of Opportunities (IMO 2013).

Consequently, a great deal of work has been undertaken on methods of estimating capacity values. A useful summary of the theory and alternative methods is provided by Dent, Keane and Bialek (2010). Appendix 1 provides a brief summary of developments since the 2011 Report.[2]

Estimating capacity value is not a precise exercise. As summarised by Dent, Keane and Bialek (2010, page 2)

> *the 'capacity value' is not a quantity which can be calculated directly from observed data. Indeed, as there are a variety of possible definitions and calculation methods, there is not (even in principle) a single definitive value for the capacity value of a given generator; as a result we refer to simplified rather than approximate calculations. The capacity value should therefore be seen as an indicative quantity […], rather than something more precise.*

A starting point for valuing capacity is defining what is mean by capacity value. There are a number of theoretical definitions of the capacity value of an intermittent generator. The most common and preferred definition[3] is the Effective Load Carrying Capability (ELCC), which is the additional load that the system can supply with the particular generator, with no net change in reliability. A similar alternative measure is Equivalent Firm Capacity (EFC) which measures the capacity of a reliable[4] scheduled generator that would deliver the same reduction in risk.

There are a number of methods to attempt to measure capacity value of intermittent generators. The preferred approach, when there is no restriction on resources and there is adequate data, is a detailed risk calculation (hereafter, ELCC Risk Method) involving a series of iterative calculations with the probability distributions of the output of all generators (which are either estimated or taken from historical data) to estimate how the risk of outage changes with changes in demand. This method can be used to estimate capacity value when the goal is to minimise LOLP or Loss of Load Expectation (LOLE, which reflects the second reliability criterion). In addition to the data and computation resource requirements, a concern with such methods is that they lack transparency.

In capacity markets, where the capacity value of individual facilities is required, alternative approximation methods are more commonly used. The approximation methods can be categorised as:

- simple time-based methods, which are based on the average output of facilities during observed peak periods, and
- risk-based methods, which are based on the output of facilities during TIs in which LOLP is high.

A starting point for simple time-based methods and some risk-based methods is the average output of facilities at times when surplus capacity is lowest and therefore LOLP is highest.

2 See, for example, the application to solar in Madaeni et al (2012) and Lu et al (2012), and theoretical development by Dent and Zachary (2014 a & b).

3 This definition is preferred by Dent, Keane and Bialek (2010, page 2) and the International Energy Agency (IEA 2013, p.35).

4 An alternative definition is that of the Equivalent Conventional Power (ECP) which defines the capacity relative to a scheduled generator with some likelihood of forced outage.

This may seem surprising as the output of IGFs cannot be relied upon. However, as nicely summarised by Stoft (2008, page 1):[5]

> *'[…] reliability is a statistical property of the system, so the contribution of wind [or any IGF] to reliability must be based on its affect on the probability that net load will exceed available capacity'.*

As predicted by probability theory, the ELCC capacity value of an IGF will approximate its average output at these times less a small adjustment. More specifically (see Zachary and Dent, 2011), when the variance of the IGF output is small, the ELCC of an IGF can be estimated as:

$$\text{ELCC} \approx \bar{I} - K\sigma_I^2$$

Where: $\bar{I}$ and σ_I^2 are the mean and variance of IGF output when there is no surplus and K is a constant that reflects characteristics of the system.

All else being equal, the larger the system, the smaller the value of K and the closer the ELCC will approximate the average output of the IGF.

As it is extremely rare that surplus capacity is zero, it is necessary to use the output at times when surplus is lowest instead. Unfortunately, the output at these times may not be representative of the output when the surplus is zero. This may be because an IGF's output, at peak times, is correlated with demand and/or the output of other facilities. An implication is that further adjustments may be required.

A useful generic structure for the capacity value of small facilities using simple approximation methods is therefore:

Capacity value =	1. Average IGF output in peak periods	***Less***	2. An adjustment for the variability of IGF output

Most simple applications of time-based methods (which are based on the average output of facilities at particular times) do not include any adjustment.

The most simple of the risk-based approximation methods is known as the z-method. Consistent with the formula above, the z-method, involves using average output at the time of peak LOLP less an adjustment that is proportional to the variance of the facility output.[6]

Empirical work (e.g. Dent, Keane and Bialek 2010) and theory (e.g. Zachary and Dent 2011) provide evidence that the z-method is a reasonable approximation to the results of the more detailed ELCC Risk Method when the variance of the IGF output is small relative to that of the system. Dent, Keane and Bialek (2010, page 5) note the z-method 'is not of great use for

5 As pointed out by Stoft (2008, page 1), it can be helpful to think of IGF output as negative demand that reduces the load to be met by other generators.

6 The 'z' in the z-method refers to the z statistic (mean divided by standard deviation) for the system surplus, which is considered a reliability metric of the power system. The closed form solution of the z-method follows as the z-statistic is constant when new generation is added.

practical computation of the ELCC of entire wind portfolios; it may however give a reasonable result for the ELCC of individual wind farms […]'

An alternative approximation method that is commonly discussed but less rarely used is the Garver approximation method (Garver 1966). The method appears to perform better than the z-method; however the computational demand is still significant and the method is rarely used.[7]

A review of recent developments and methods used in other market is provided in Appendix 1. We understand that most capacity markets use simple approximation methods. We are aware of only one capacity market that includes the more complex ELCC Risk Method as part of determining individual facility capacity credits. The Midwest ISO (MISO) already used the ELCC Risk Method for adequacy planning before the introduction of a capacity auction in 2013 (Rogers and Porter 2012). MISO still calculates the system-wide wind ELCC by the probabilistic risk method. This capacity credit is then allocated to individual wind facilities deterministically using historic peak period data, based on the average capacity factor over the top 8 daily peak hours per year for the past 9 years (72 peak periods in total) (MISO 2014).

Of particular relevance to the WEM, is the issue that the output of IGFs appears to be correlated with the available surplus (i.e. capacity less system demand) at peak times. We have not found any methods being used in practice to deal with this issue; however, we are aware of some research being underway in the UK and the US.

7 See Dent, Keane and Bialek (2010) for a discussion of the Garver method.

2.3 The current methodology

The current method for determining the RL for IGFs is:[8]

RL = *1. Average facility output during peak TIs* **Less** *2. G × variance of facility output during peak TIs*

Where $G = K + U$, reflecting both known variability (reflected in *K)* and *uncertainty* of the distribution *of output* (reflected in U).

Facility output is measured in MW and the parameters K and U are measured in units of MW-1.

K = 0.003

U = 0.635/(average facility output during peaks).

A limit is place on the adjustment due to the U-factor of one third of the average facility output.

In the above formula, the peak TIs refer to the 60 TIs with the highest LSG selected from 12 separate days from each of the previous 5 years.

Peak LSG identifies the TIs when surplus capacity is lowest and therefore when the system is under greatest stress. LSG is calculated (in MWh) as:

- the Total Demand for energy, which is the sum of total sent out generation of all facilities plus the load that has been curtailed,[9] less
- the Total Intermittent Generator Output, which is generally just the sent-out generation of the IGF but includes adjustments for the impact of Consequential Outages and Dispatch Instructions on impacted IGFs.

It is useful to decompose the RL methodology into two components being:

- the 'mean output at peak LSG less the K-factor adjustment', and
- the U-factor adjustment.

The first component is a more general application of the z-method referred to in section 2.2 and is consistent with the probability theory in that section. The second component, the U-

[8] Further details are provided in Appendix 3.

[9] The Curtailed Load is generally zero. It is the sum of:

- Demand Side Programme (DSP) Reduction, the total quantity by which all DSPs reduced their consumption in response to Dispatch Instructions;
- Interruptible Reduction, the total quantity by which all Interruptible Loads reduced their consumption in accordance with the terms of an Ancillary Service Contract; and
- Involuntary Reduction, the total quantity of energy not served due to involuntary load shedding (manual and automatic).

factor adjustment, was added as pragmatic method given the limited available data at the time to reflect the issue of correlation between IGF output and peak demand.

2.4 Criteria for evaluating the methodology

Ultimately the methodology needs to be evaluated against the Wholesale Market Objectives. These are:

(a) to promote the economically efficient, safe and reliable production and supply of electricity and electricity related services in the South West interconnected system

(b) to encourage competition among generators and retailers in the South West interconnected system, including by facilitating efficient entry of new competitors

(c) to avoid discrimination in that market against particular energy options and technologies, including sustainable energy options and technologies such as those that make use of renewable resources or that reduce overall greenhouse gas emissions

(d) to minimise the long-term cost of electricity supplied to customers from the South West interconnected system, and

(e) to encourage the taking of measures to manage the amount of electricity used and when it is used.

However, to evaluate alternatives it is useful to develop a simplified and refined set of criteria that reflect the Wholesale Market Objectives and also the key issues and trade-offs in developing a RL methodology.

We have summarised these criteria as relating to:

- **Accuracy**

  As the capacity value of an IGF cannot be directly observed, it is not possible to measure precisely the accuracy of the RL methodology; rather accuracy may be thought of as the extent to which the RL methodology estimates the capacity value of an IGF.

  The more accurate the methodology the more likely the method will align with the Wholesale Market Objectives including those relating to reliability, technical neutrality and encouraging efficient entry. For example, if the methodology is inaccurate then:

  - there is a risk that the total capacity credits awarded to IGFs will be excessive and the reliability of the system is lower than is expected
  - there is a greater risk the methodology will favour particular technologies, and
  - there is a risk that inefficient entry will be encouraged and efficient entry will be discouraged.

- **Robustness**

  The methodology should be robust to changing circumstances (e.g. changes in the nature of demand and new technologies). This helps to reduce uncertainty, which may deter efficient entry, and to reduce the costs of managing the system. The costs of changing the methodology can also be significant, imposing a burden on existing participants and the IMO.

- **Volatility**

  It is important that the results of applying the formula are not too volatile; that is, sensitive to small changes. Higher volatility of results directly increases the volatility of returns to investors, which if unnecessary could deter efficient entry.

- **Practicality and simplicity**

  There is a range of practical considerations. It is important that the methodology is not overly complex to apply. Greater complexity leads to greater cost in terms of administration.

Accuracy is perhaps the primary objective. Accuracy is particularly important given the high value of Capacity Credits. A method that is overly generous may lead to inefficient investments and similarly, efficient investment may be deterred if insufficient Capacity Credits are awarded. The other objectives could be simply met using a time-based estimate. However, as discussed in this paper this would not be necessarily accurate.

The objectives are related. A formula that is inaccurate is also unlikely to be robust and may be more likely to provide volatile results.

3. Performance of methodology

3.1 Introduction and overview

This section considers the impact of the revised methodology since the methodology was developed and introduced in 2011.

The current methodology for determining the RL was developed in 2011 by Sapere Research Group (Sapere) in response to a request by the IMO Board to identify simple changes to two competing proposals to revise the prior method (see Box 1 below).

Box 1: Background to the current methodology

Prior to 2011, the methodology for assigning Capacity Credits to IGFs was recognised as inappropriate as it did not reward provision of output at peak periods when capacity is required. From the 2011 Report (page 1):

> *To examine the issue, the Independent Market Operator (IMO) established the Renewable Energy Generation Working Group (REGWG) to determine a new approach. A number of alternative proposals were considered but neither a consensus nor a compromise was achieved.*
>
> *Following the REGWG, two proposed rule changes were developed and submitted by the IMO (RC_2010_25) and Griffin Energy (Griffin) (RC_2010_37) in relation to the allocation of capacity credits to IGFs. This paper briefly assesses these two proposals and examines potential modifications which could be made to develop a new approach.*
>
> *Sapere Research Group (Sapere) was commissioned by the IMO Board to provide independent advice on the two proposals by Griffin and the IMO in relation to the allocation of capacity credits to IGFs. Sapere was asked to assess the two proposals and identify if there were modifications that could be made that would make them more robust and simpler.*
>
> *In particular, Sapere was asked whether simple changes could be made so as to allocate capacity credits based solely on individual performance while ensuring performance is during peak periods and significant volatility is not introduced. The report also examines the transition between the current capacity valuation and the proposed future method (i.e. provide a 'glide path').*

Source: The 2011 Report.

The methodology was designed and parameters determined based on analysis of the actual output of IGFs that were operational in 2011. These IGFs included a few wind farms and a few land fill gas (LFG) facilities. Since that time there have been some changes. In particular:

- a small number of LFG facilities are no longer operating as IGFs
- a number of IGFs (including several wind farms, several LFG facilities and one solar plant) have become fully operational, and

- a number of additional IGFs were assigned Certified Reserve Capacity for 2015/16 but have yet to become operational.

An implication of these changes is that there are varying amounts of data for different facilities by year available for the current review. For some facilities there is actual output data since the market start. For some there is a mix of actual data and estimated data and for others there is just estimated data to April 2013. A related complication is that the measure of LSG being used to identify peaks differs between existing facilities and new facilities.

A summary of facilities categorised by technology and status is provided in Table 1 below. For the purposes of assessing the methodology, naturally our preference is to rely on actual data from existing facilities that have been operational for some years. However, as this would exclude many new facilities that were operational in 2014, we have considered the broader set of facilities that were operational in 2014 for our main analysis.

In addition to the facilities listed in Table 1 there were a number of facilities that were not operational in the 2014 summer but for which the RL was calculated for the 2015/2016 capacity year. The results for these IGFs were examined but the estimated output data for these was not used for determining parameter values.

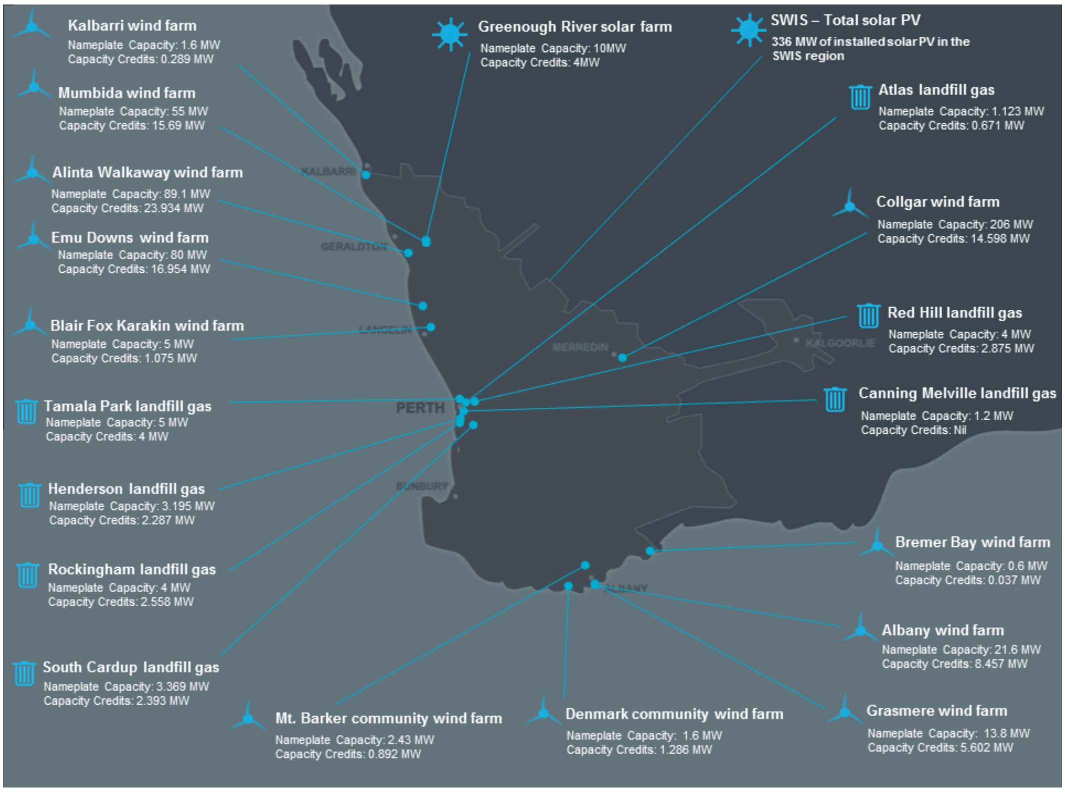
A map of the location of the renewable energy sources (in effect, IGFs) that were awarded capacity credits for the South West Interconnected System (SWIS) in the 2015/2016 capacity year is provided in Figure 1 below. The location of facilities is of interest as we would expect that the adjustments included in the RL methodology for the relationship between weather and peak load to be similar for wind facilities located near each other.

Table 1: IGF categories used in this report

Category	Facilities
Existing wind IGFs (operational since April 2007 or before)	ALBANY_WF1 ALINTA_WWF EDWFMAN_WF1
Existing LFG IGFs (operational since April 2007 or before)	RED_HILL TAMALA_PARK
Existing IGFs	Existing wind IGFs and Existing LFG IGFs
Ex-IGFs (were once registered as IGFs but no longer operational)	GOSNELLS CANNING_MELVILLE KALAMUNDA
New operational wind IGFs (operational in summer 2014 but are classed as a New Facility)	SKYFRM_MTBARKER_WF1 KALBARRI_WF1 BLAIRFOX_KARAKIN_WF1 BREMER_BAY_WF1 DCWL_DENMARK_WF1 GRASMERE_WF1 INVESTEC_COLLGAR_WF1 MWF_MUMBIDA_WF1

Category	Facilities
New operational Solar IGFs (operational in summer 2014 but are classed as a New Facility)	GREENOUGH_RIVER_PV1
New operational LFG IGFs (operational in summer 2014 but are classed as a New Facility)	ATLAS ROCKINGHAM SOUTH_CARDUP HENDERSON_RENEWABLE_IG1
New operational IGFs	New operational wind IGFs and New operational LFG IGFs
Operational IGFs	Existing IGFs and New operational IGFs

Figure 1: Map of renewable energy sources for the SWIS

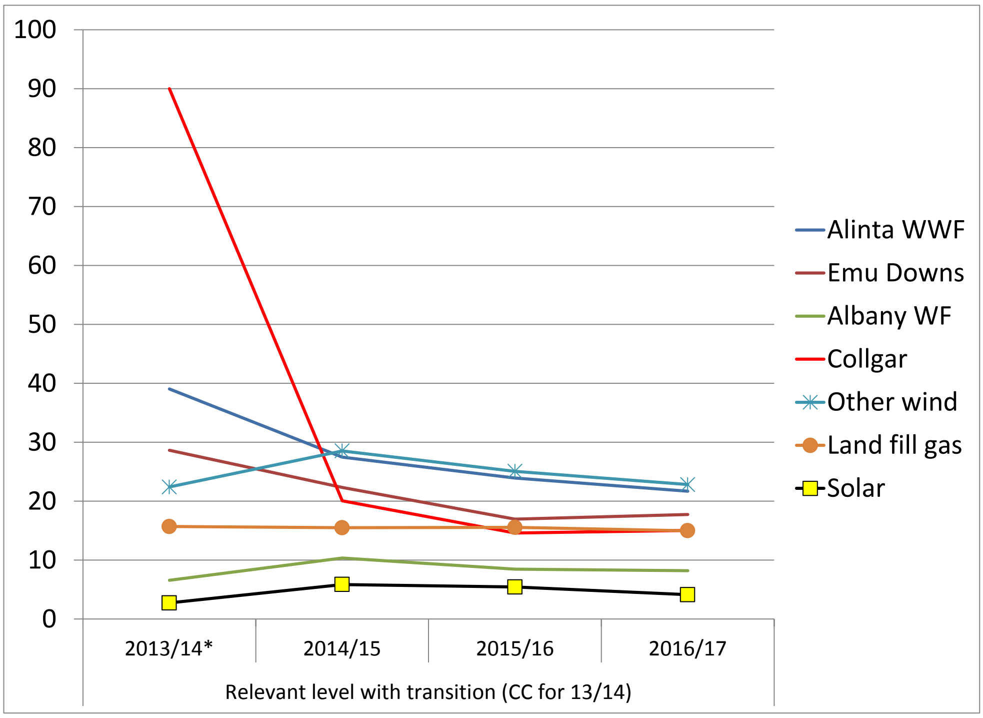


Source: SWIS Electricity Demand Outlook – June 2014. IMO (2014, page 46).

3.2 The impact of the methodology

The introduction of the methodology had a significant impact on the RL calculated for, and therefore the Capacity Credits awarded to, a number of IGFs. While for some (notably for solar and for a few wind farms), the methodology resulted in an increase in the RL, the methodology's most significant impact was to dramatically reduce the RL calculated for some wind farms. Figure 2 maps the RL (or, for 2013/14, the Capacity Credits awarded) for selected facilities and facility groups since the Capacity Year 2013/14, a year prior to when the new RL methodology came into effect. As shown in the figure, the introduction of the new method in 2014/15 resulted in a significant shift in the Capacity Credits for some facilities.

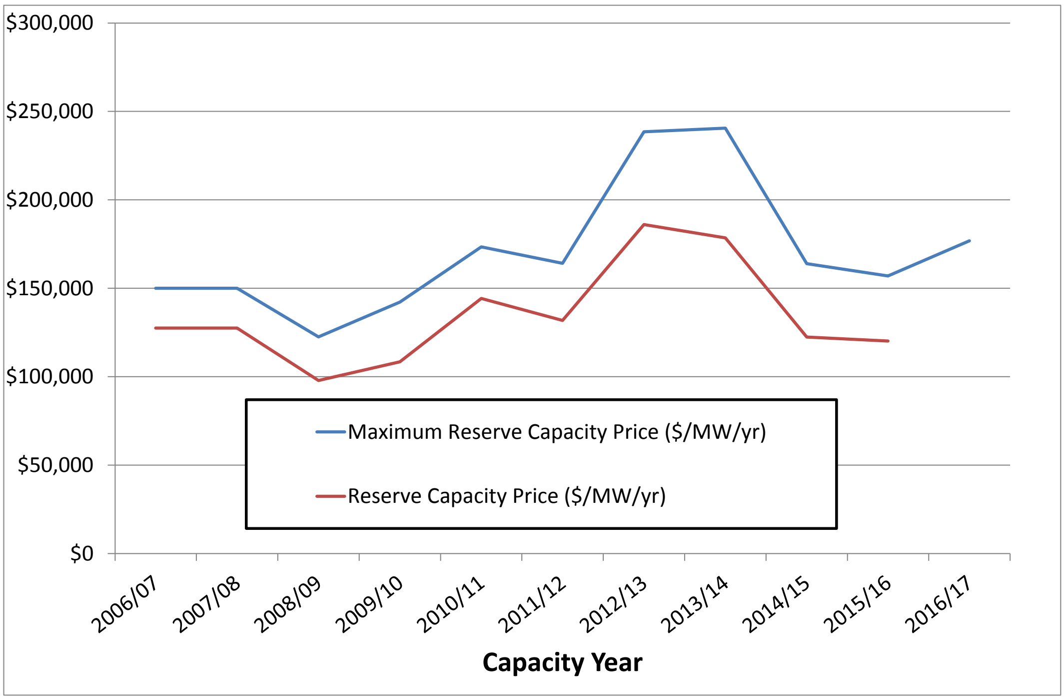
Figure 2: Changes in Relevant Level



Notes: For 2013/14* Capacity Credits (CC) are shown. For 2016/17, the RL is an estimate.

The introduction of the methodology coincided with a significant fall in the value of Capacity Credits (see Figure 3 below). The fall in the value of Capacity Credits in 2014/15 has reduced the financial impact of the RL methodology changes; however, it also has meant some facilities were doubly impacted by a fall in the capacity awarded and the price received for capacity.

Figure 3: Value of Capacity Credits



Source: http://www.imowa.com.au/reserve-capacity/maximum-reserve-capacity-price/maximum-reserve-capacity-price-overview.

3.3 Performance of the methodology

3.3.1 Accuracy

The RL methodology aims to estimate the contribution of IGFs in the event of an extreme peak in which all capacity is required. In effect, it involves forecasting outcomes during a future event, which typically does not have historical precedent and by design should be extremely rare. As such there is no actual benchmark[10] against which to compare the methodology.

To evaluate the accuracy it is useful to separate the current RL methodology into the two components being:

- an approximation method (mean output at peak LSG less the K-factor adjustment) that aims to measure ELCC (as if facility output accurately represented output when there is no surplus), and
- the U-factor adjustment.

10 This reflects the remark of Dent, Keane and Bialek (2010, page 1) that "the 'capacity value' is not a quantity which can be calculated directly from observed data."

Potentially, the first component could be compared with the results of an application of the ELCC Risk Method, which is generally considered to be the most accurate methodology. However, no such results for the WEM exist. Furthermore, even if there was, there would be the challenge of how to disentangle the U-factor adjustment affect.

There are, however, aspects of the first component that can be critically examined. These include matters like the value of the K parameter itself. The 2011 Report notes that the initial K parameter was based on international benchmarks and had not been tailored to the existing methodology and the WEM. There are other areas for potential improvements (such as the weighting of TIs) in the methodology. These and other issues are examined in Section 4 of this report.

It is possible to use the recent data to more critically analyse the performance of the U-factor adjustment. The U-factor adjustment was introduced to address the concern that — as it appeared from the available data — the output of IGFs would fall on very hot days when system demand would be highest and the system would be under greatest stress.

However, there are some risks that the U-factor adjustment will be inaccurate. These include that:

- the adjustment factor is based on the variance of output (divided by mean output) — it does not use any other controls to take into account whether a facility's output is lower during periods of extreme stress
- it was based on a limited amount of data that was available at the time, and
- it was designed before the introduction of solar and some other large facilities that have since become operational.

The next section more closely examines how the RL methodology, with its U-factor adjustment, compares with the output of facilities during periods of extreme stress.

The results during periods of extreme stress

To assess the performance of the methodology, and in particular the U-factor adjustment, we closely examined how the RL compared with the performance of IGFs at peak TIs and how this varied with temperature.

In conducting this analysis, we focussed on using the peak TIs that are currently used in the RL methodology. These are the peak 12 TIs taken from separate Trading Days in each year (year ending 31 March) where the peak is measured by LSG. These are used because, as discussed in Section 4, they are the TIs that are most likely to be representative of an extreme peak when LOLP is highest. Of note:

- A single TI in each day is used because (as discussed in section 4.5.4) additional TIs from the same day would add very little information (due to the high correlation of IGF output in adjacent TIs) and because, by construction, additional TIs would be selected at times that are not the peak in the day (and therefore less likely to be representative).
- Peak LSG and not peak market generation (MG) is used as peak LSG is when LOLP is highest.

We have, however, supplemented the data examined with output from additional days which have a very high maximum temperature but were not a peak LSG day because (typically) they are on weekends or public holidays. As discussed in section 4.4.4 these days provide some potentially useful information; however some caution is required in using the data as the peak TI on these days may occur at different times in the day to other peaks.

As discussed in section 4.4.2, we found the maximum daily temperature at Perth Airport to be the best single predictor of high system demand (and high LSG) days. We estimate that, based on historical weather data, the one in ten year peak that the system is designed to meet will occur when the maximum temperature is around 43.8 degrees.

Based on the above considerations, a useful method of examining the performance of the U-factor adjustment is to plot the output of facilities at peak TIs against maximum temperature in the day.

A summary example, for IGFs that have been operational since 2007, is provided in Figure 4 below. The figure shows the output of the facilities at the peak TIs and how this varies with daily maximum temperature.[11] All output data that is used in the RL methodology is included; however for the IGFs included in this chart there is only actual data. We have also included output on non-peak days (i.e. those not used in the RL methodology) when the maximum temperature is greater or equal to 39 degrees.

The figure also shows the RL, the mean output[12] and a line of best fit. These lines are all based on the output data for TIs that are used in the RL methodology; that is, they excluded the data on the non-peak days. One further line shown is the mean output of all data points shown (including non-peak days) when the maximum temperature is greater or equal to 41 degrees. This line ('Mean temp>=41 inc nonpeaks', hereafter also called 'Mean – hot days') provides a guide to the average performance of IGFs on very hot days.

At the far right of the figure is a small line (between 43.5 degrees and 44 degrees) that represents the RL excluding the K-factor adjustment ('RL excl K adj'). The line has been included as it is a more appropriate point of comparison when examining the average output at high temperatures. That is, this amount (being the mean less the U-factor adjustment) should approximate the average output at the one in ten year peak period (which we have assessed to most likely occur when the maximum temperature is around 43.8 degrees).

In this example, the 'RL excl K adj' line appears to coincide with what is predicted by the line of best fit. The 'RL excl K adj' line also appears consistent with the mean output at TIs where the temperature is at least 41 degrees.

Note, that the line of best fit in Figure 4 illustrates the negative relationship between IGF output at peak TIs and the daily maximum temperature.[13] However, some caution is required when examining the line of best fit as it assumes a linear relationship between maximum temperature and output and is sensitive to outliers.

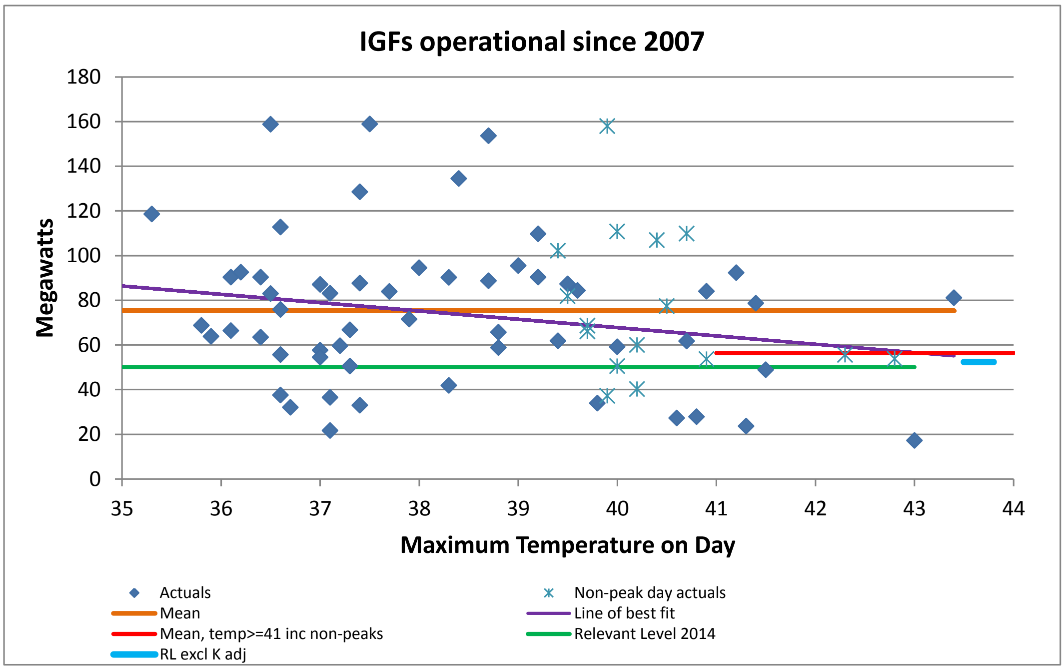
---

11 The peak TIs are those TIs used in the current methodology; that is, the 12 peak LSG TIs selected from separate days. The issue of how TIs are selected is discussed in Section 4.The discussion of temperature measure is also discussed in Section 4.

12 That is, the average of the points plotted.

13 This negative relationship is the rationale for the U-factor adjustment.

Figure 4: Output of IGF facilities at peak TIs – five years to April 2014



Charts plotting the performance of individual IGFs were examined but have not been included in the report for reasons of confidentiality. Given the lack of data at the very high temperatures, assessment is difficult and requires a substantial degree of judgement. Our assessment of the charts is that the RL appears broadly consistent with the performance of the IGFs at very high temperatures and on this basis appears reasonable. However, we acknowledge this assessment includes a degree of our own judgement.

Some elaboration and qualification of this assessment is warranted. As we would expect there is some variation and the RL methodology appears to apply more appropriately to some facilities than others. We expect that by applying local knowledge about weather conditions some patterns may be found that could be used in improving the assessment.

In our assessment (on the limited data available), the RL appears appropriate for solar. For solar there is very limited data — only 2 years of actual data and an additional 4 years of estimated data — and so assessment is difficult. In the case of solar the 'RL excl K adj' line is below what the line of best fit and the 'Mean – hot days' predicts; however:

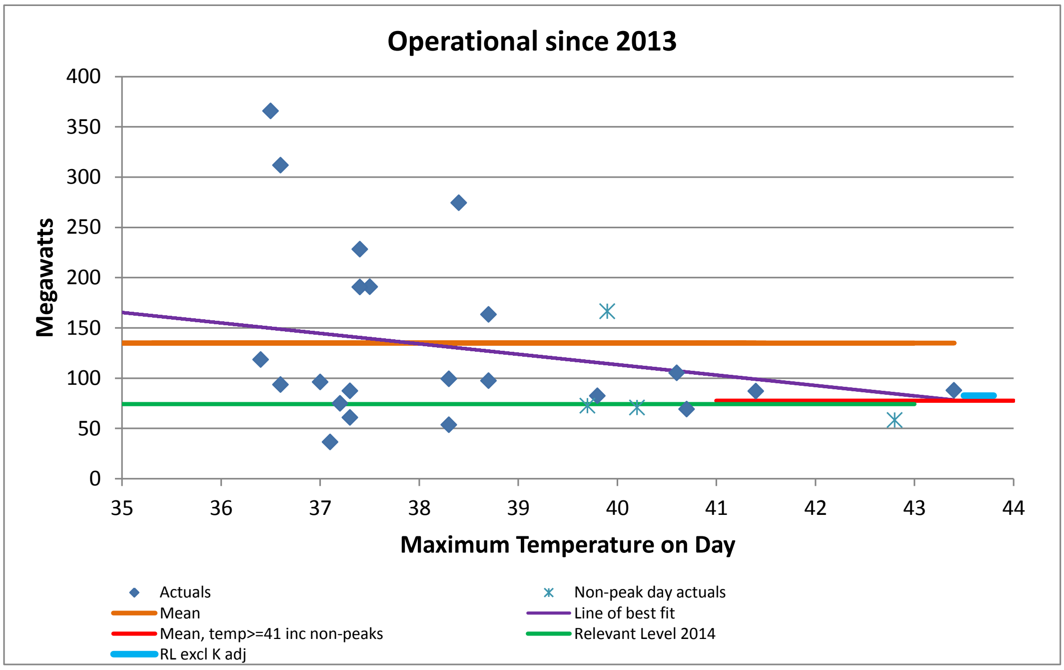
- the output on the highest temperature day for which actuals were recorded was very low
- the 'Mean- hot days' includes weekends etc when the time of the peak may not be representative of the peak
- there appears to be a material difference between the actuals and estimated data, where the peaks of the actual data are later in the day, and
- as discussed in section 4.4.2 the average actual output on days with a temperature of at least 40 degrees is very low.

Given that much of the new capacity proposed is solar it is of considerable interest as to how the methodology might perform in the future for solar facilities. Based on available data we can make some predictions. On hot days, the performance of solar follows a reasonably consistent pattern over the course of a day, and therefore the performance of solar largely depends on the timing of the peak TI during the day. As a result, if the timing of the peak is reasonably consistent then the variance of solar output will be small and the U-factor adjustment for solar will also be small; which appears appropriate. For example, based on the 2014 results, if the peak TI was always at 4:30pm then the U-factor adjustment for solar would be around 3% of the mean output (in contrast the difference between the mean output at 5pm and 4:30pm is around 30%). If however, the timing of the peak moves significantly then the variance of solar output and the U-factor adjustment will be large. This could be appropriate if (as the limited evidence available, discussed in section 4.4.2, suggests) some of the variance reflects that on hotter days the peak shifts to later in the day. The adjustment would be inappropriate if the reverse was true. As solar penetration increases, it will be important to continue to monitor the timing of the peak TIs.

An increase in the use of solar will shift the peak and thereby impact on the RL calculated for wind facilities. This may have a positive impact for wind facilities if peak TIs are pushed back later.[14]

In our assessment, the RL for newly operational facilities appears reasonable. Figure 5 includes data for facilities that have been operational for two years (again in this chart there is no estimated data to be included). Again in our judgement, the RL appears reasonable.

Figure 5: Output of IGF operational since 2013 at peak TIs – two years to April 2014



In some cases, the need for the U-factor adjustment appears less clear cut. Figure 6 below shows the anonymised results for a particular facility. In this case, while there is still great

[14] This would be appropriate as, if the peak were pushed later to when wind IGF output was greater, the capacity value of wind IGFs would increase.

uncertainty on how the facility would perform at an extreme peak, there does not appear to be as strong a justification for making the U-factor adjustment. A contrast is shown in Figure 7, which shows the anonymised results for another like facility. In this second example, there appears to be a much stronger justification for the U-factor adjustment.

Figure 6: An individual IGF results – example 1

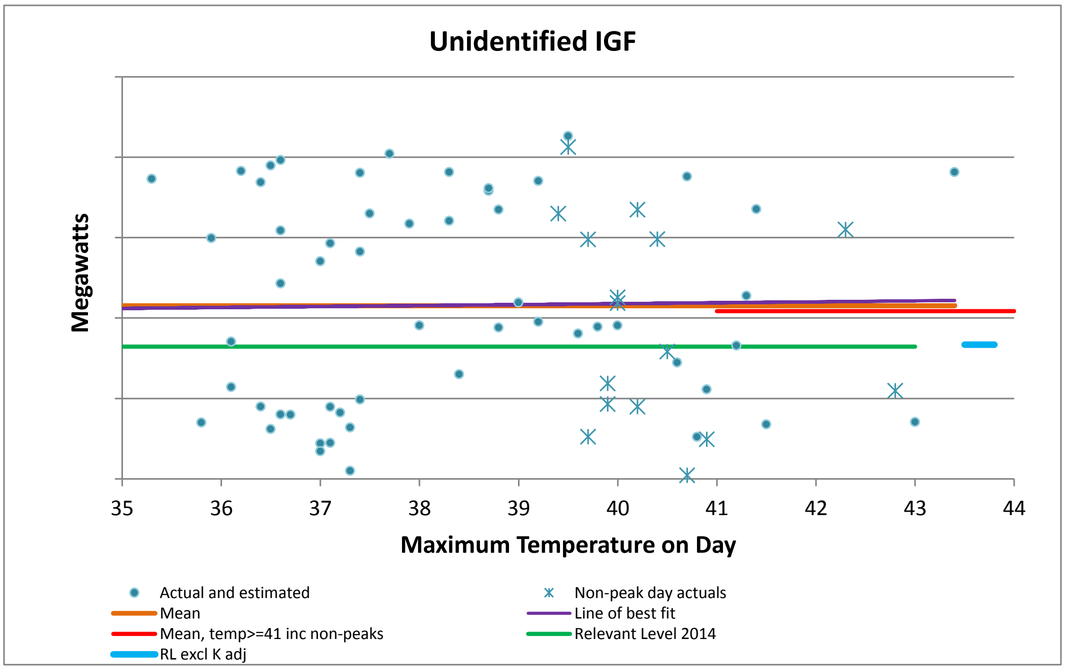
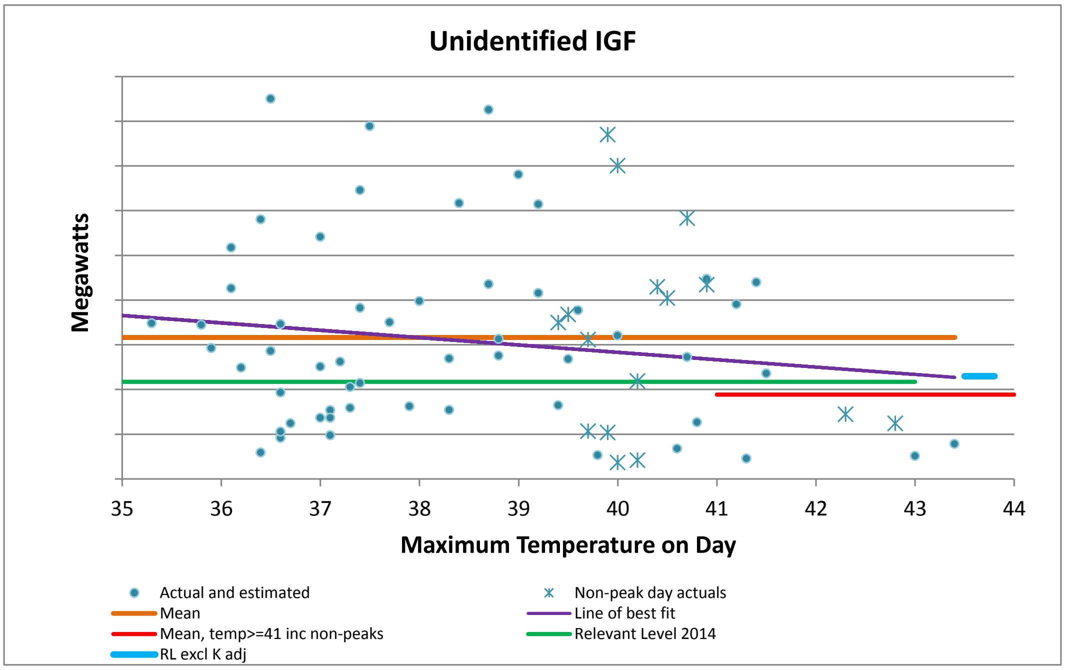


Figure 7: An individual IGF results – example 2



The accuracy of the estimated output of new facilities

New Facilities are those that have not been operational during the five-year period over which the RL methodology is applied. For these an estimate of the sent-out generation is used for TIs prior to the IGF being operational. The estimate is prepared by an expert consultant accredited by the IMO. Actual generation is used for TIs once the facility is operational.

A potential concern is that the expert estimates would not be reflective of actual performance and therefore the RL for New Facilities would be inaccurate.

For some New Facilities it is possible to compare the estimated sent-out generation during the Peak TIs before the operational date with actual sent-out generation in peak TIs once operational. It was found that relative to actual data there were situations where, for the estimated data:

- both the mean and variance of output were lower
- the mean of output was higher and the variance of output was lower, and
- both the mean and variance of output were higher.

In summary, there did not appear to be a systematic problem of bias (i.e. attempts to make the IGF appear more reliable) but the results suggest there is some potential for improvements in producing the estimates.

3.3.2 Robustness

In assessing the robustness of the methodology, it is again useful to consider separately the 'mean less K-factor adjustment' and the U-factor adjustment components.

We have no material concerns with the robustness of the 'mean less K-factor adjustment' component as this method is based on probability theory that applies to any intermittent facility whose variance of output is not overly large relative to the system. The selection of peak TIs based on highest LOLP mitigates the risk that the methodology will be impacted by changes in the profile of demand.[15]

Despite our concerns with the U-factor adjustment, it appears to have been reasonably applicable to new technologies and facilities. Since 2011, a number of new IGFs have become operational including the first solar facility. As discussed in the previous section, the adjustment appears to apply reasonably well to these new facilities. Of note, the U-factor adjustment methodology is not technology specific and does not rely on any particular feature of demand. Nevertheless, there has been limited data against which to assess the performance of the U-factor adjustment and we remained concerned that the U-factor adjustment will not be robust to updates using further data.

The methodology itself has been largely stable. There has been one rule change modification since the methodology was implemented in 2011. This rule change (RC_2013_17,

[15] As is noted in this report the profile of demand is changing, but this does not affect the application of the methodology. In contrast a shift in the profile of demand could impact on the accuracy of a time base approximation method.

implemented May 2014) related to revisions to the estimated output of an IGF when a Dispatch Instruction is issued. Of note; the rule change is not particular to the revised methodology — rather it might have been introduced for any methodology that relied on the output of intermittent generation to estimate the RL.

3.3.3 Volatility

While it is apparent from Figure 2 there has some movement in the RL since 2014/15, this change has been small (at least relative to the change following its introduction). Furthermore a portion of the change is due to the removal of transition arrangements over the three year period.

Nevertheless, there has been continued variation. Figure 8 below shows a summary of the RL as if the methodology had been applied from 2013/14 with no transition arrangements in place. As can be seen, there is some seemingly random movement and some more general trends.

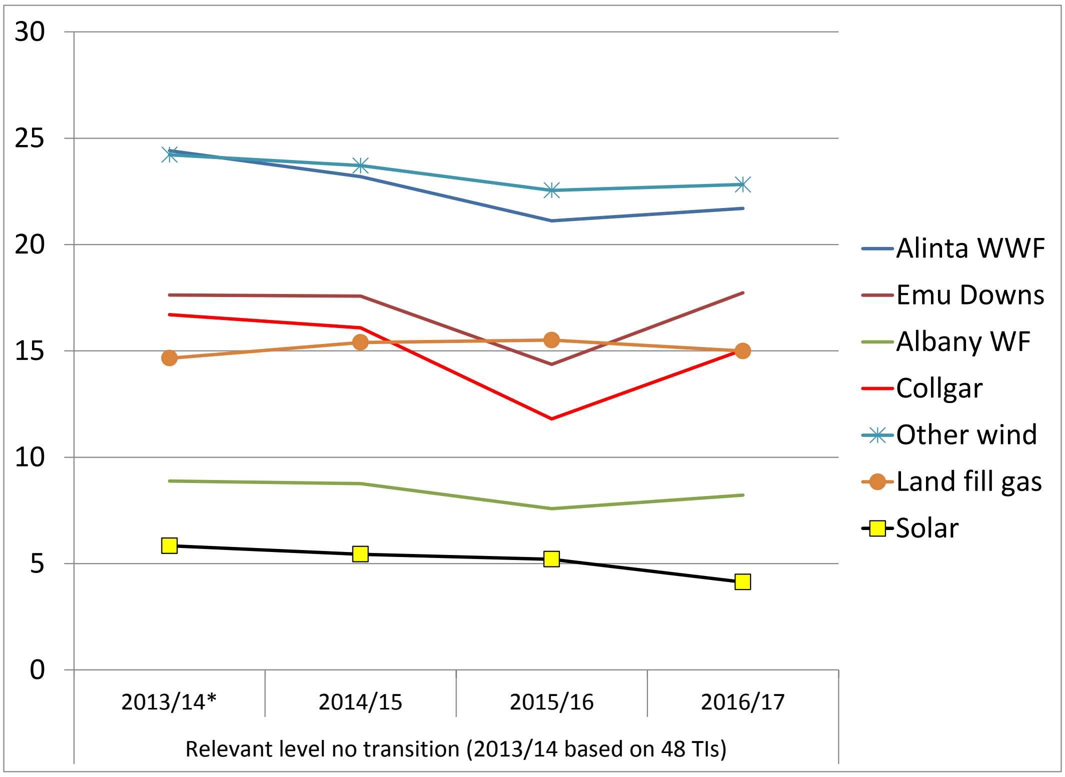
There are a number of reasons for the variation. Some of the volatility is due to changes in average output by year. The significant dip in the RL for two wind farms shown in Figure 8 can be attributed a significant fall in average output in the 2012/13 summer followed by a significant increase in the 2013/14 summer.

Some change over time is also expected due to the impact of other facilities. As noted above, the capacity value of an individual facility will be lower as new facilities come online which provide competing output at similar times.

We exampled the results for each IGF, including the minimum, maximum and the relative standard error.[16] Overall the degree of variation, as measured by the relative standard error, has been reasonable, varying between 3% and 15%. By way of comparison, as a rough rule of thumb, we would expect that 95% of the time the RL will be within a range of around plus or minus two standard errors; that is, in this case up to plus or minus 30%.

16 The relative standard error (calculated as the standard error divided by the mean) is a useful measure of the degree of volatility.

Figure 8: RL method applied since 2013/14 without transition



* For 2013/14 methodology has been calculated using 4 years' worth of data.

3.3.4 Practical issues

The process of applying the methodology involves collection of large data sets and performing a series of calculations applying a semi-automated procedure using Excel workbooks.

The process is reasonably complex. However, much of the complexity stems from dealing with issues that would apply to any methodology that relied on the output of IGFs during peak TIs. These include:

- making adjustments in determining total demand for curtailments
- making adjustments in determining the output of IGFs, and
- undertaking calculations for New Facilities.

Some of the complexity arises from having to work with LSG and in particular calculate LSG for all New Facilities. While it would be simpler if one measure of LSG was used, dealing with multiple measures of LSG is a relatively small inconvenience.

The cost of conducting the expert reviews was raised in in a submission in response to the draft report. In their submission[17] Community Electricity stated:

> *Given the complexity and data constraints, we would prefer simplification of the allocation of capacity credits to new facilities in order to reduce the cost of the expert review; in particular, the "experts" providing the estimate are subject to even more onerous constraints.*

This point was considered, however no further action was pursued at this time. An expert review would be a requirement for any methodology that relied upon the output of the IGFs. Furthermore, we would expect that investors in the IGFs would want such information to support their investment decisions and therefore that the additional cost of the expert review for the purposes of the RL methodology will generally be minor.

However, we recognise that the cost of expert reports may be an issue for very small generators. In such case it may be worthwhile in the future considering rule that accommodates such situations. For example, potentially a rule could be developed that allows new-generator owners to opt out of obtaining an expert report and returns a sufficiently conservative (i.e. low) RL so that they have the incentive to obtain an expert report when it is cost effective to do so.

In summary, the methodology appears to be reasonably practical to apply. There do not appear to be any material issues with the process or cost-effective opportunities to further streamline the process.

3.4 Summary

The introduction of the new methodology in 2011 had a significant impact on the Capacity Credits awarded to IGFs.

It is useful to consider the methodology as the sum of two components:

- a 'mean less K times variance' component that is consistent with probability theory, and
- the U-factor adjustment, which attempts to adjust for the concern that IGF output is correlated with demand at peak times.

There are reasons to have reservations with the U-factor adjustment. In particular, it could penalise facilities unnecessarily whose output is uncorrelated with the surplus at peak times. Nevertheless, the U-factor adjustment appears, *so far*, to have produced results that are consistent with our expectations.

---

17 Available at http://www.imowa.com.au/home/electricity/consultations/2014-relevant-level-methodology-review.

4. Review of options

4.1 Introduction

This section reviews potential modifications to the current methodology. The focus is on issues raised in the project scope. Nevertheless, a broad examination was undertaken.

Given the scope, we have organised the review into sections examining:

- the overall method of determining IGF contribution (including the use of LSG)
- the K-factor adjustment
- the U-factor adjustment and the issue of IGF output being correlated with demand at extreme temperature peaks, and
- other matters relating to the selection of the TIs.

As noted in the 2011 Report, the most significant issue for the WEM is the issue that IGF output appears to change with extreme temperatures. However, how this issue is addressed has implications for other aspects that are considered.

It should also be noted that there are costs to change and it would be inconsistent with the Wholesale Market Objectives to recommend a material change without significant benefit. The costs of change include both the administrative costs of change and the signal of uncertainty for investment decisions.

4.2 Method of determining IGF contribution

4.2.1 Penetration of intermittent generation and use of approximation based methods

As the penetration of intermittent generation increases, the complexity of calculating capacity value also increases. As noted in the 2011 Report, the approximation based methods (in particular the related z-based and K-adjustment methods) are recognised as being appropriate only for lower levels of intermittent generation penetration.

There is no simple definition of the level of intermittent penetration for which these approximation methods are valid – the level will reflect the various assumptions of each method that allow approximate calculation of ELCC. Furthermore these assumptions may still be valid for the calculation of the capacity value of individual generators in a system even when they are no longer valid for calculation of the capacity value of a fleet of generators.

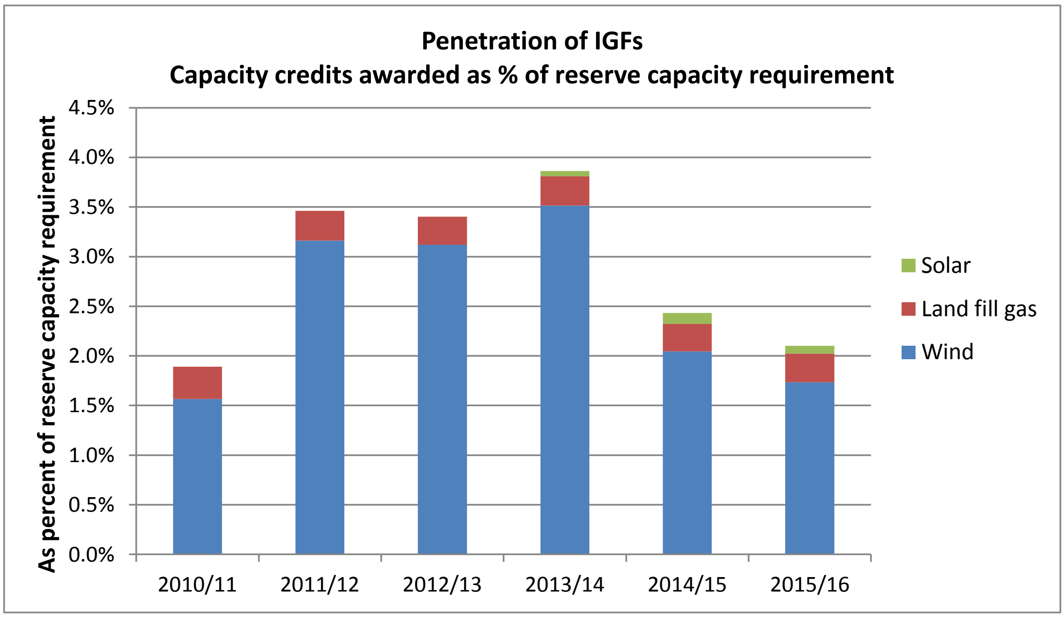
The primary assumption that underlies the 'mean less K-factor adjustment' approximation approach is that the additional generation does not change the shape of the probability distribution for available capacity. As a rough guide, Dent, Keane and Bialek (2010) calculate (for Great Britain) that the approximation begins to significantly deviate from the ELCC Risk Method calculation when the additional generation capacity at peak is approximately 0.7% of the total power system capacity. Above this level the approximation will tend to under estimate ELCC.

As a percentage of total energy supplied, the contribution of intermittent generation in the WEM is significant (see Figure 10); however, for the purposes of Capacity Credit valuation it is the contribution at peak times that is most relevant. In terms of Capacity Credits awarded (which with minor qualification[18] reflect the expected output of facilities when surplus is lowest) the percentage penetration of IGFs is around 2% of the Reserve Capacity Requirement (see Figure 9).[19]

Hence from Figure 9 it is not appropriate to apply the z-method to estimate the ELCC of the IGF fleet. However, it may still be appropriate for individual IGFs in the SWIS. The RL of any single facility in 2015/16 was no more than 0.5% of the Reserve Capacity Requirement, which is below 0.7% amount referred to above.

Of note, the accuracy of the K-factor adjustment approximation largely depends on the variability of the facility's output relative to the variability of surplus load. In one case, the measured variance of IGF output was particularly large (in total, greater than that of all the other IGFs combined). In this case the K-factor adjustment may not be an accurate approximation. Nevertheless, in this case it is apparent the observed output was correlated with demand and therefore the measured variance would not be representative of variance when the surplus was zero. Furthermore in this one case, the K-adjustment was small compared to the U-adjustment (which was so large it was bounded by the maximum adjustment) and, in our subjective assessment given the available data, the total adjustment for the particular IGF has been and continues to be reasonable.

Figure 9: Penetration of intermittent generation - capacity



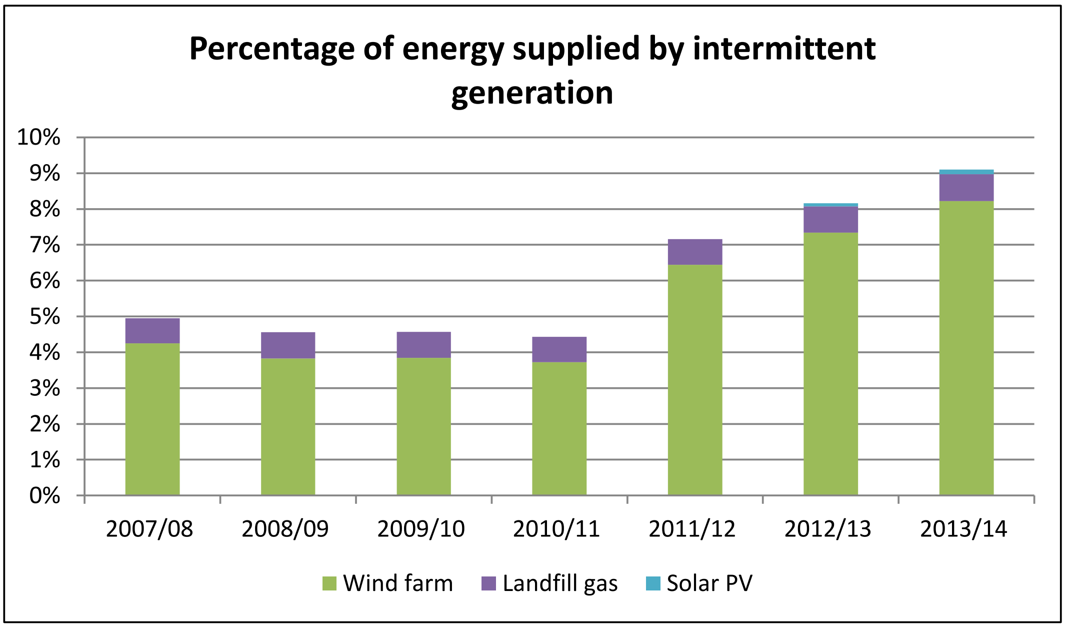
Source: http://www.imowa.com.au/reserve-capacity/capacity-credit-information, Statement of Opportunities reports 2008 to 2013.

---

18 Due to the K-factor adjustment that is made, the expected (i.e. average) output of an IGF when surplus is lowest is slightly more than the capacity credits awarded to it.

19 Of note, the fall shown in the figure is, in part, due to the revision of the RL methodology.

Figure 10: Penetration of intermittent generation – energy supplied



Source: Sent-out generation data provided by the IMO.

4.2.2 Use of an alternative method

Use of a time based approximation method

A common approximation method used in capacity markets elsewhere involves selecting TIs based on specific times of the day and year. While simple, this method runs the risk of selecting TIs that are not representative of extreme peak and therefore we would expect it to be less accurate.

A particular challenge of selecting a representative time period is that changes in system demand and changes in the output from IGFs may shift the times at which peaks occur. Evidence of the shifting system demand is shown in Figure 21 on page 64 below.

Using a time based approach would also be a significant change to the current methodology. Given the only apparent advantage of the approach is simplicity, we have deemed that is not an appropriate or realistic option.

ELCC Risk Method

An alternative to approximation methods is to undertake a detailed risk calculation using what is referred to in Section 2.2 as an ELCC Risk Method.[20] As noted in Section 2.2, while this is often a preferred method for estimating capacity value it relies on sufficient data being available.

[20] Dent, Keane and Bialek (2010) provide a description of the procedure.

We do not recommend this approach. The method is computationally costly and not transparent. In effect, the ELCC Risk Method would be an alternative to the 'mean less K times variance' component of the method. The method cannot be used to overcome the problem of insufficient data that the U-factor adjustment attempts to address.

4.2.3 The use of LSG as a basis for selecting TIs and alternatives

The current method is based on the output of IGFs during TIs as measured by peak LSG. The output of the IGFs at the time of peak LSG is typically not the same as the output at the time of peak system demand (as measured by MG). This is simply because the variability of IGF output shifts the timing of peak LSG.

There is a very clear rationale for using LSG as the basis for selecting TIs. All else being equal, the available surplus capacity is least, and therefore the LOLP highest, when LSG is highest. This is consistent with the ELCC theory of measuring output when surplus capacity is at its lowest level. The same cannot be said for MG. A facility that produces at peak MG but not at peak LSG does not improve the reliability of the system in meeting the peak.

The benefit of using peak LSG reflects that an IGF has less value if its output is correlated with the output of other IGFs. An IGF that produces (again all else being equal) when other IGFs do not produce has greater value.

If the output of IGFs at peak MG were used in the RL methodology instead of LSG, some alternative method would need to be used to adjust for the correlation of output between facilities. It is difficult to see how something could be introduced that would be simpler or more accurate than using LSG.

One potential alternative would be to introduce a covariance factor into the methodology, such that facilities with a lower covariance with other facilities are rewarded. However such an approach is problematic. A key issue is that the facility output at peak MG may not be representative of the facility output at peak LOLP (i.e. LSG) and therefore the covariance calculation based on peak MG data may not be appropriate. For example, peak MG tends to be later in the day than peak LSG (see Figure 21 and Figure 22 on page 64) and therefore the data selected at peak MG may not be representative of what IGFs (including both solar and wind) produce at peak LOLP. An additional issue is that a measure of covariance can be sensitive to outliers in the data. Finally, adding a covariance factor would add complexity to the method.

As, in addition to these issues, there does not appear to be any benefit of moving to the use of a covariance factor, we conclude that it is appropriate to continue using LSG in determining peak TIs.

4.2.4 How LSG is applied

There are some subtle issues that stem from the current application of LSG.

Peak reduction versus output at peak

To apply the RL methodology we collect data on IGF output. However, there is a challenge in measuring IGF output. Ultimately we are interested in the contribution of the output of

IGFs in reducing the peak LSG. Because an IGF's output can shift the timing of peak LSG, the size of the peak reduction will be between the output of the IGF at peak LSG calculated with and without the IGF's output. For the RL methodology LSG is calculated *including* the IGF's own output and therefore may underestimate the marginal contribution of the IGF. Similarly the output of the IGF at the peak LSG calculated *excluding* the IGF's output (hereafter, LSG Ex-The IGF) may overestimate the marginal contribution of the IGF.

A hypothetical example to illustrate this issue is provided in Box 2 below.[21]

Box 2: Peak reduction vs output at peaks

The table below provides a hypothetical example of the effect of two intermittent generators (whose output is labelled IG1 and IG2) on the peak to be met by scheduled generation.

In the absence of the two intermittent generators, the peak (calculated as the peak of MG) would be at 16:00. Peak LSG (which is calculated as MG – IG1 – IG2) is at 15:30.

Trading interval	MG	IG1	IG2	LSG	LSG excl IG1	LSG excl IG2
15:00	2120	20	50	2050	2070	2100
15:30	2190	35	40	2115	2150	2155
16:00	2200	40	60	2100	2140	2160
Maximum (over the TIs)	2200			2115	2150	2160
	a			***b***	***c***	***d***

<u>Calculations</u>

1	Fleet output at peak MG	100	*= IG1 + IG2 at 16:00*
2	Peak reduction due to the fleet	85	*= a - b*
3	*Marginal peak reduction of IG1*	*35*	*= c – b*
4	*Marginal peak reduction of IG2*	*45*	*= d – b*
5	Sum of individual peak reductions	80	
6	Fleet output at peak LSG	75	*= IG1 + IG2 at 15:30*
	<u>Marginal benefit of IGs considered in order</u>		
7	*Peak reduction of IG1 (assuming no IG2)*	*40*	*= a - d*
8	*Peak reduction of IG2*	*45*	*= d - b*
9	Sum of individual peak reductions (considered in order)	85	*= a- b*

In the hypothetical example provided there are just two IGFs labelled IG1 and IG2. The marginal contribution of IG2 in reducing the peak is 45 MW (see line 4), which is between

21 The 2011 Report (Box 3, p.14) also illustrated this point.

its output at peak LSG (40 MW at 15:30) and output at peak LSG Ex-The IGF (60 MW at 16:00).[22]

A potentially viable alternative to using output at peak LSG is to use a measure of an IGF's contribution in reducing the peak LSG on the peak Trading Days (i.e. the Trading Days that are currently used in the methodology). The contribution in reducing the peak LSG for any Trading Day could be simply estimated as:

Peak day reduction = Maximum LSG in day <u>without</u> the facility output, less
Maximum LSG in day <u>with</u> the facility output

This is a straightforward calculation to perform. Of note, the peak reduction in a day may still underestimate the marginal contribution of an IGF in reducing the peak as the peak may shift across days; that is, the true peak reduction due to an IGF within a day may be less than the peak reduction within a broader period. However measuring the peak reduction within a year is computationally much more complex.

The advantage of using peak reduction within a day as the 'output' measure is that it is a more accurate measure of an IGF's marginal contribution. The key disadvantage is the cost of changing procedures.

However, the benefits of using peak day reduction over output at peak LSG appear to be minor because of the related K-factor adjustment. The use of output at peak instead of peak reduction has a very similar effect on an individual IGF to that of an increase in the K parameter value; both having the effect of penalising IGFs whose output is more variable. Accordingly, if the method was changed to using peak day reduction, there would need to be an offsetting increase in the size of the K parameter. Therefore the benefits of a switch to using peak day reduction would be reduced as any change would have implications for the calculation of the K parameter.

Given the lack of substantial benefit and the cost of change we do not recommend using the peak reduction measure for individual IGFs. Nevertheless, this issue of peak reduction is a consideration in the calculation of the K-factor adjustment and discussed in section 4.3.

4.2.5 Should existing IGFs be impacted by new IGFs?

The example shown in Box 2 also serves to illustrate another set of issues. The sum of the marginal contributions of each IGF will typically be less the total contribution of the fleet of facilities. In terms of the Box 2 example, line 2 'Peak reduction due to the fleet' is greater

22 The example also helps to illustrate why the output at peak MG (i.e. system demand) is not an accurate measure of the contribution of a facility in reducing the peak and using an LSG based measure is preferable. In the example, the output of the IGFs at peak MG is 100 MW, which is significantly more than the contributions of the individual facilities in reducing the peak. The difference is, in part, attributable to the output of the two IGFs being correlated. Using an LSG based measure is preferable in that it helps to control for correlation in the output of IGFs. All else being equal, an IGF has lesser value (to reliability) if its output is highly correlated with that of other facilities as this means its output is more likely to be low when other IGFs is low and therefore when LSG and LOLP are high.

than line 5 'Sum of individual peak reductions'. The effect of the IGFs is to reduce the peak in the day by 85 MW; however the sum of their marginal contributions (line 5) is 80 MW. [23]

This outcome simply reflects that the marginal benefit of a facility is less if there are like facilities. It is analogous to the effect in the electricity market of higher competition at particular times of the day reducing the marginal value of electricity at those times.

An important implication is that the marginal contribution of an existing facility — and similarly the IGF output measured at peak LSG — tends to fall as new 'like' facilities become operational. In effect, the current methodology compensates IGFs for their marginal contribution because new facilities (once operational) affect the calculation of Existing Facility LSG.

There are two 'apparent' issues with awarding IGF's Capacity Credits simply based on their marginal contribution.

1. The marginal contribution and therefore the existing facilities' Capacity Credits awarded are affected by new entrants.

2. There may be a gap in the Capacity Credits that are awarded and the total capacity that is provided by facilities.

However these 'issues' may not affect efficiency in terms of signals for investment. These issues are analogous to a competitive market where new competition affects the price received by existing providers and the total value provided to consumers exceeds the amount paid to providers. In theory, it is efficient that facilities are compensated based on their marginal contribution. Doing so provides investors with the appropriate investment signals.

The effect of how new facilities impact on existing facilities is gradual as Existing Facility LSG is based only on actual output data (i.e. it does not use the estimated data that is used to calculate New Facility LSG) that is collected from a period of 5 years. Therefore the full impact of a new facility on other facilities will not be felt until it has been operational for 5 years.

Eventually, there may be a gap between the Capacity Credits that are awarded and the incremental capacity value of the IGF fleet. While this may seem an anomaly, it is not inefficient as it is the marginal contribution of facilities that is important for encouraging efficient investment decisions. There are, however, other considerations. The uncertainty over the impact of new entrants may be an unreasonable deterrent for efficient investment. Furthermore, the Reserve Capacity Mechanism, in effect, assumes that the Capacity Credits awarded are equal to the total capacity provided.

To ensure that the Capacity Credits awarded match the incremental capacity provided by the IGF fleet, there appear to be two practical options. One is to determine an IGF's RL based on its marginal contribution calculated in the absence of newer facilities. This could be

[23] It can be more if the output of the facilities is negatively correlated and the individual facilities are complementary in their contribution to reducing the peak. This can be seen by using the example in Box 2 and setting the output of IG1 at 16:00 to 10 MW, in which case the sum of individual peak reductions will exceed the peak reduction of the fleet.

simply achieved by, for each facility, selecting peak periods using LSG whose calculation does not include newer facilities. There are, however, some issues with this option. It could lead to a race to be registered and reduce the efficient incentive by existing facility owners for expansions and modifications.

The second alternative is to share the additional capacity contribution of the fleet among the existing fleet. This second option is, in effect, what is done under the current methodology, as the U-factor parameter was selected so as to align the RL of the fleet with — what is estimated to be — its contribution to capacity.

Given the U factor adjustment is again being proposed, we recommend the second option is maintained.

4.3 K-factor adjustment

4.3.1 Revised estimate of the K parameter

The K-factor adjustment is used to address the issue that the IGF output will add to the volatility of the output to be met by other generators. The K-factor adjustment is based on probability theory that is described in Box 3 below.

The original K value was approximated from international experience. It was noted at the time, that the K-factor adjustment was very small relative to the U-factor adjustment. Furthermore, as the U-factor was set to meet the gap between the K-factor adjustment and the total adjustment, the selection of the K-factor did not affect the size of the adjustment for the fleet of IGFs.

For this review, further exploration of the K-factor adjustment has been possible. Two key steps were undertaken to:

- estimate a K value relevant to the SWIS, and
- modify the K value to reflect how LSG is used in determining peak IGF output.

Box 3: The basis for the K value

It can be shown[24] that where the variance of IGF output is small compared to the variance of the surplus load (equal to the surplus of available capacity over load) then:

$$\text{ELCC} \approx \bar{I} - K\sigma_I^2 \qquad (1)$$

Where: $\bar{I}$ and σ_I^2 are the mean and variance of peak load IGF output and K is a constant.

The formula holds when IGF output is independent of surplus load. Where output and surplus load are dependent the formula still holds as long as the mean and variance of the IGF output are representative of when there is no surplus.

The value of K can be estimated with knowledge of the distribution of surplus. If the surplus load follows a Normal distribution, then K may be simply estimated as:

$$K = \frac{\bar{S}}{2\sigma_S^2} \qquad (2)$$

where $\bar{S}$ and σ_S^2 are the mean and variance of surplus load.

More generally K can be estimated as:[25]

$$K = \frac{f_M'(0)}{2f_M(0)} \qquad (3)$$

where $f_M(0)$ is the density function of the distribution of surplus load when surplus load is zero.

When the variance of IGF output is large the formula will not be accurate, and appears to underestimate the ELCC value (see Dent, Keane and Bialek 2010).

Source: Adapted from Zachary and Dent (2011).

A K value appropriate for the SWIS

As noted in Box 3 above, K can be estimated using information on the probability distribution of surplus load at the peak. As surplus load is the available capacity less system demand, the distribution of surplus load can be estimated from measures of capacity and the distribution of peak demand.

Information on the probability distribution of peak demand can be found in the 2014 SWIS Electricity Demand Outlook (SEDO).[26] Unfortunately detail on the full probability distribution is not available; however, this report (Fig 30) includes data for peak demand for the 10%, 50% and 90% probability of exceedance (POE).

---

24 Zachary and Dent (2011) discus the theory. Stoft (2008) provides numerical examples.

25 Zachary and Dent (2013).

26 IMOWA (2014).

For the measure of capacity we used the Reserve Capacity Target (RCT) net of Minimum Frequency Keeping Capacity. This is calculated as the 10% POE peak demand plus the reserve margin of 7.6 per cent. We use the RCT and not the capacity awarded as the RCT is the desired level of capacity and therefore should reflect the required capacity contribution of a facility. Furthermore, the RL is being determined for future years beyond the date for which capacity has been awarded.

Both the RCT and the peak demand distribution change over time and therefore the estimated value of K will change over time.

The SEDO forecasts indicate a skewed distribution of peak demand, so that the peak demand mean is higher than the median (given by the 50% POE). As a result we used equation (3) referred to in Box 3 above to estimate K. To estimate the distribution of the surplus we fitted the POE data points to a three parameter skewed probability distribution function.[27] This enabled us to calculate the values required to estimate K. Over the capacity years 2017/18 to 2019/20 we estimated K to range from 0.0024 to 0.0022 with an average of 0.0023.

Our initial expectation was that, as the SWIS is a relatively small electricity market, the K value would be relatively large compared with international benchmarks. This expectation reflected that all else being equal, an identical IGF will have a more significant impact in a small system than in a large system. The relatively low value of K reflects that the forecast distribution of peak demand is skewed; which has the effect of stretching the tail of the distribution to look more like a bigger system.[28]

Other adjustments to the K value

The K-factor adjustment from statistical theory is an adjustment to the contribution of a facility to account for an increase in the variance of the remaining LSG.

However, as discussed in section 4.2.4 an IGF's output measured at peak LSG may underestimate its marginal contribution in reducing the peak, particularly for those facilities with larger variance of output. Therefore the use of IGF output at peak LSG may reduce the need for the K-factor adjustment.

To estimate the size of this effect, we calculated for each operational IGF, on peak TI days, the average difference (using all actual data) between the peak day reduction and the actual output at peak LSG. We found this amount across all IGFs to average 5 MW. In total we found this to be similar in effect to a K parameter value of around 0.00165. We found the size of the effect was also similar to the effect of a K-factor adjustment for individual

27 A shifted log-logistic distribution function was used. See https://en.wikipedia.org/wiki/Shifted_log-logistic_distribution#Shifted_log-logistic_distribution. This uses three parameters being: mu, the median; sigma, a scale parameter, and eta a shape parameter. The median values are given. The remaining two parameters were estimated using a numerical data solve procedure. For 2018-19 the parameters were mu=4,470, sigma = 131.8 and eta = 0.15809.

28 We also estimated the K value using a log-normal distribution (a 2 parameter skewed distribution) with just the 50% and 90% POEs. This gave similar results. However if a symmetrical distribution is used very different results are obtained. For example, If the 50% and 90% POE were unchanged and the distribution of the peak was normal (as is often assumed) then the K value would be in the order of 0.005; that is, 0.002 greater than the current K parameter.

facilities; that is, as expected, the difference in the peak day reduction and actual output at peak LSG is greatest for the facilities with the higher variance of output at peak; that is, those facilities which incur a higher K-factor adjustment.[29]

The above estimate reflects the peak reduction of facilities over a day. There will be some additional peak reduction that occurs from movement between peaks across days. We estimated the average peak reduction in 2014 across all operational facilities for the top 12 Trading Days was 6.9 MW above the fleet output;[30] (that is an additional 1.9 MW). This total peak reduction is equivalent to a K-factor adjustment of around 0.0023, which is coincidentally our estimate of the unadjusted K value.

There are a few other minor considerations. These include that some of the measured variance in facility output will be due to the change in output with higher temperatures and that facility output will vary within a TI. For these reasons it is not possible to be too precise.

Nevertheless, in light of the above considerations, we recommend that the appropriate K parameter to be modified to be zero to reflect:

- an unadjusted K parameter of 0.0023 reflecting the forecast distribution of surplus peak capacity, less
- 0.0023, an adjustment for how IGF output is determined.

4.4 The U-factor adjustment — Dealing with correlation between IGF output and peak demand

4.4.1 Introduction

The 2011 Report examined the relationship between demand, temperature and IGF output. The results of the analysis highlighted a potentially significant issue that the contribution to reducing peak demand by IGFs fell at very high temperatures. The U-factor adjustment was introduced to address this issue.

To assess the extent of the underlying issue and whether the U-factor adjustment could be removed or replaced, we more closely examined the IGF output during periods that have weather characteristics of the highest peak. To support this analysis we obtained weather data from a number of weather stations in the Perth region.

We examined how peak demand and peak LSG change with different weather measures. We found that of the measures available (see Box 4), the temperature as captured at the Perth Airport weather station was the best predictor of peak demand.

---

29 There were some differences but we did not assess these to be significant.

30 The average peak reduction was calculated as the average MG in the 12 peak TIs (1 TI per day) as determined by MG less the average LSG in 12 peak TIs (1 TI per day) as measured by LSG.

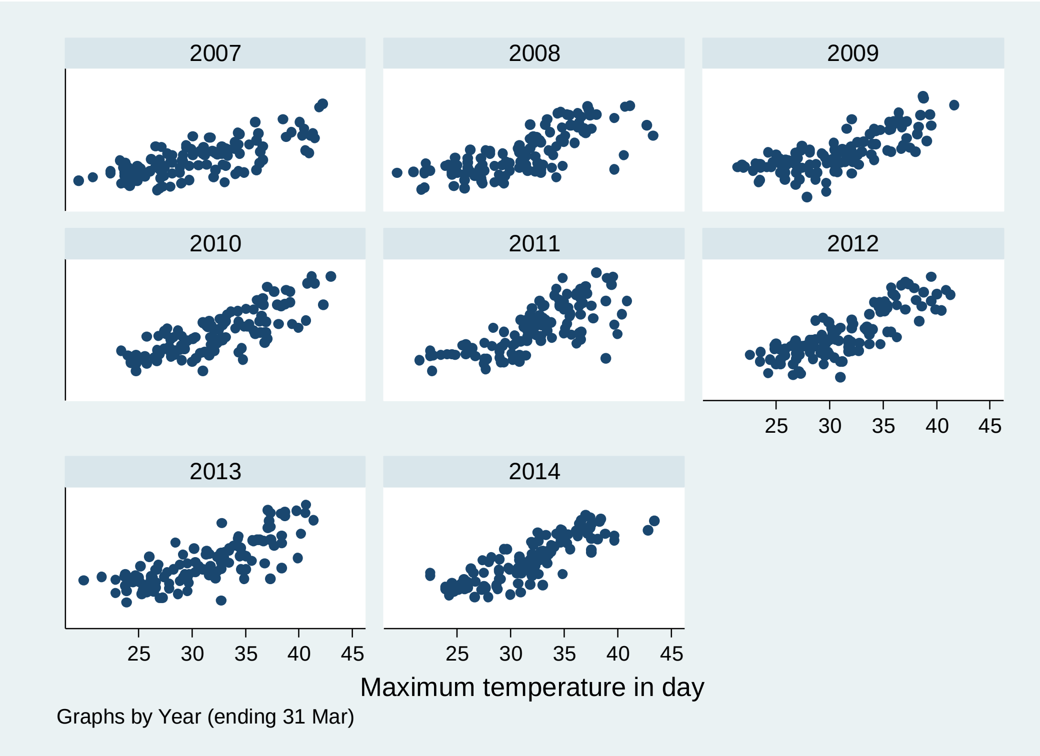
Box 4: Weather measures used in this study

All references to temperature in this report are to the air temperature as measured from Perth Airport weather station.

For the purposes of this review weather data was obtained from the Bureau of Meteorology (BOM). Data from a number of Perth weather stations was examined. The weather station whose temperature data correlated most strongly with demand was Perth Airport. From this station half-hourly temperature data, wind-speeds and cloud cover were obtained.

An example of the relationship between demand and maximum temperature is shown in Figure 11 below. It clearly demonstrates that the peak LSG is most likely to occur on days with high maximum temperatures.[31] For peak TIs the relationship between peak demand and maximum temperature appears reasonably linear.

Figure 11: Peak LSG vs maximum temperature in day by year



Note: Based on days December to March only

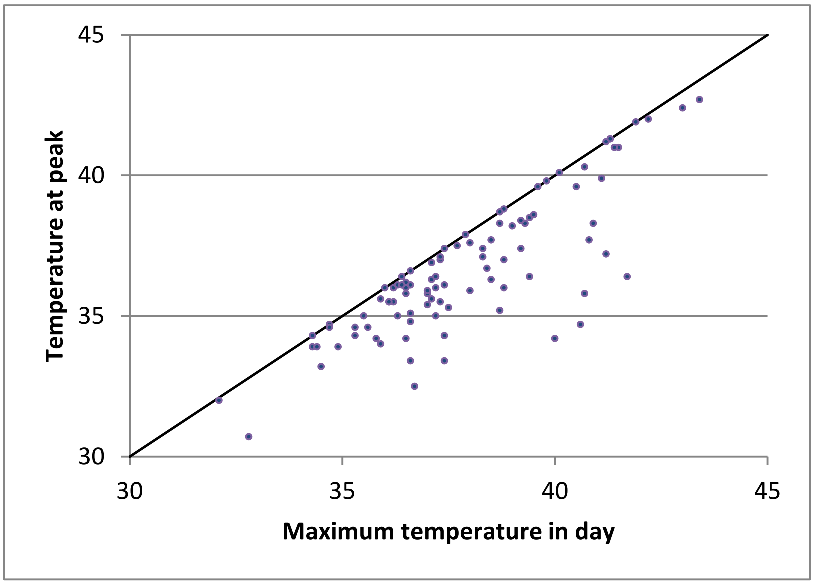
Of the weather variables available, we found that the maximum temperature in the day was the best single predictor of total demand. Based on a simple linear regression the peak load increases by around 70 MW for each additional degree of maximum temperature.

31 The picture is very similar for MG.

To estimate the likely maximum temperature on the peak day for which capacity may be required, we examined the historical daily maximum temperature data for Perth Airport. This data is available for 70 years from 1944 to 2014. Based on this data we estimate the 10% POE maximum temperature to be around 43.8 degrees. This suggests that the one in ten year peak demand day will occur when the temperature is around 43.8 degrees.

Of note, the peak TI (in terms of LSG or demand) does not tend to occur at the time of peak temperature and therefore the air temperature at the peak TI tends to be less than that of the maximum temperature in the day (see Figure 12 below). The air temperature at the peak is on average around 1 degree less than the maximum temperature in the day.

Figure 12: Air temperature and maximum temperature



Note: Based on peak TIs from 2007 to 2014.

4.4.2 Performance of IGFs at extreme temperature days

The performance of IGFs at extreme peaks differs from that on other peaks. Table 2 below highlights the issue. The table summarises the impact of IGFs (from wind, LFG and solar) in terms of peak reduction[32] in the 12 peak Trading Days for each of the 8 years from 2008 through to 2014. When it is very hot,[33] the output of IGFs appears to be significantly lower.

[32] The impact is measured in terms of peak reduction in a day. As discussed in the previous section, peak reduction a more accurate measure of the contribution of IGFs than output at peak LSG TIs.

[33] For this analysis and following sections we have focussed on days where the temperature exceeded 40 degrees (40+ Degree Days). The selection of 40 degrees as the cut-off is for presentational purposes. As noted above, we expect that the maximum temperature at the one in ten year peak to be around 43.8 degrees.

Table 2: Performance of IGFs at extreme peaks

	Peak (MW) reduction by			
	Wind	LFG	Solar	Count of days
Max temp<40	97.21	14.18	4.60	77
Max temp>=40	75.16	13.75	1.79	19
All	92.85	14.10	4.13	96
% reduction on 40+ Degree Days	-23%	-3%	-61%	

1. Data taken from years 2008 through to 2014. When a facility is not operational the data is treated as missing
2. Output shown is the average peak reduction in the day by type of facility.

It is of interest to understand why IGFs do not perform as well on very hot days. Potentially this could be because the timing of the peaks is different or because of a relationship between temperature and the output of IGFs.

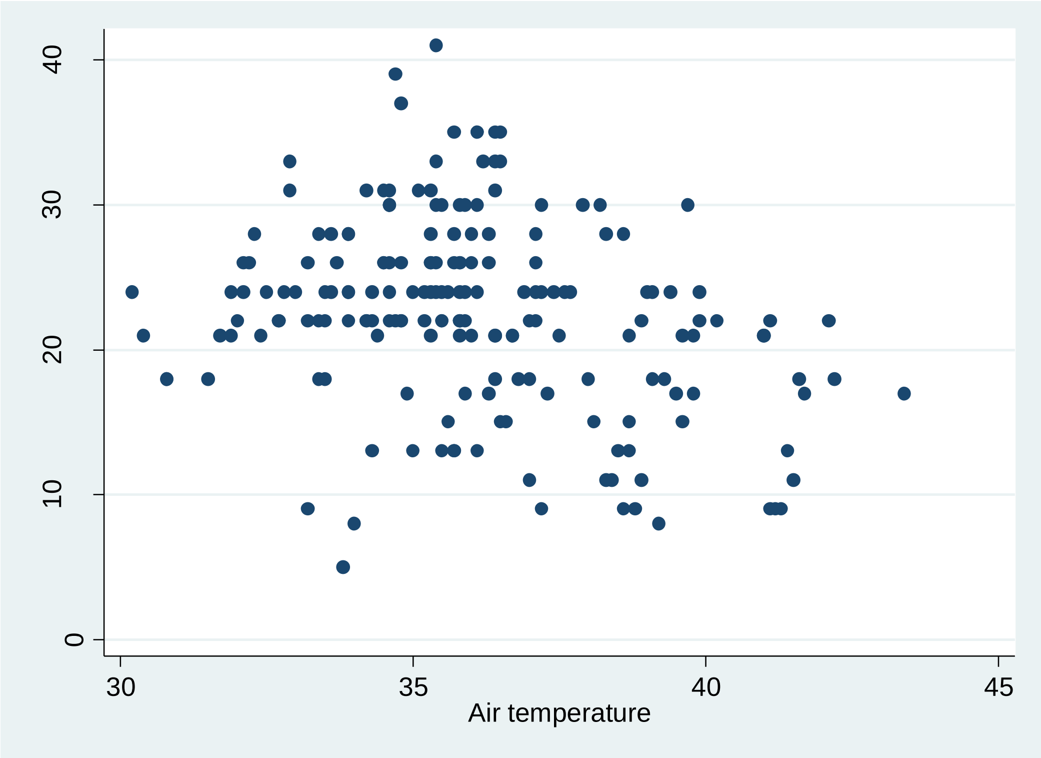
The issue with wind appears not to be that wind shifts the peak in the day, but rather that the wind may drop on very hot days.

Figure 13 maps the relationship between wind-speed and air temperature in a constant one hour period (the 2 TIs between 4 and 5 pm) on peak TIs. There a statistically significant negative relationship between temperature and wind speed.[34] While the weather data is based only on Perth Airport weather station, it nevertheless may be indicative of the patterns elsewhere in the SWIS.

For solar, the issue appears to be that on very hot days, the peak TI is more likely to be later in the day; however, there is little actual data at this stage that can be used. The poor result for solar on 40+ Degree Days in the table above is based on 4 data points.

[34] Based on a simple ordinary least squares regression of the data points shown the wind speed falls by approximately 0.75 knots (95% confidence interval of 0.4 to 1.0) with every 1 degree increase in air temperature.

Figure 13: Relationship between wind speed and air temperature at TIs between 4pm and 5pm on peak TI days



Note: Points represent the air temperature and wind speed on peak LSG days (i.e. top 12 days per year) taken from the Perth Airport weather station.

4.4.3 Overview of options

There appears to be no publically documented approaches to address the issue of correlated intermittent generation and demand.[35]

Given the limitations of the data available, we have assessed that there are three pragmatic options to deal with the issue identified above. These are:

1. Attempt to use only TIs that are representative of the extreme peaks.
2. Use the broader set of TIs but make an adjustment to reflect the negative relationship between demand and output at the extreme peaks. Within this there are two options:
   (a) the current approach, whereby an adjustment is based on the variance (and mean) of output at the peaks, and

[35] We understand that there is some experimental work being undertaken in the UK and US on what seem to be similar issues (Appendix 1 makes reference to the US work).

(b) a 'regression' approach, whereby the adjustment is based on correlation between output and temperature (or other factors that drive peak demand).

These options are considered below.

4.4.4 Using only Trading Intervals that are representative of the extreme peaks

The key challenge with the first approach is the small number of TIs that may be selected. As shown in Table 2 above, there were only 19 40+ Degree Days categorised as peak days over the last 8 years. There were no peak days on 40+ Degree Days in 2010/11 and only one in 2009/10.

Furthermore, while the high temperature days are generally ranked higher in terms of peak demand, this is not always the case. Thus there is a question as to whether the lower-ranked high temperature days are representative and if so what method can be used to distinguish the days that are likely to be extreme peaks and those that are not.

We explored this latter concern by examining why some of the very hottest days were not also the highest ranked days in terms of demand. Table 3 lists the peak TIs for 40+ Degree Days. The days which were not highly ranked are highlighted. As can be seen by the table these lower ranked very hot days tend to be weekends or holidays when demand is dampened.

Table 3: List of peak Trading Intervals on 40+ Degree Days

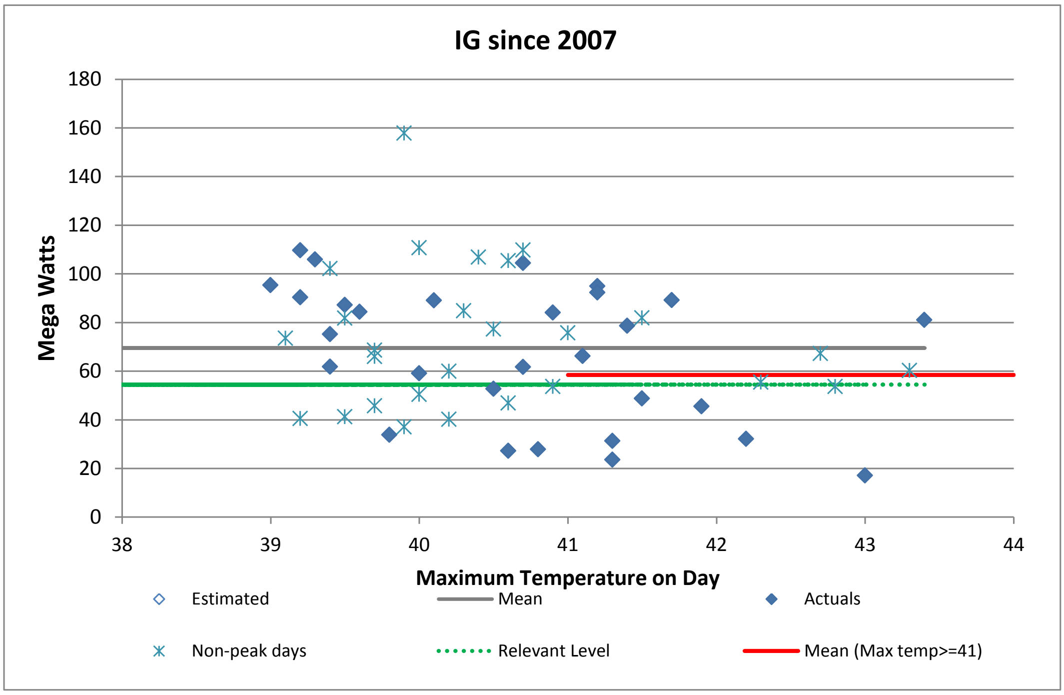
Year	Trading Interval	Rank of day in year	Max temperature on day	Comment
2006/07	Wed 7/3/2007 15:30	1	42.2	
2006/07	Tue 6/3/2007 15:30	2	41.9	
2006/07	Fri 2/2/2007 15:30	4	40.1	
2006/07	Sat 3/2/2007 14:00	6	40.5	Weekend
2006/07	Sun 28/1/2007 16:00	9	41.3	Weekend
2006/07	Sat 27/1/2007 16:30	12	41.1	Weekend
2007/08	Thu 28/2/2008 15:30	1	41.2	
2007/08	Thu 17/1/2008 15:00	3	40.7	
2008/09	Fri 16/1/2009 14:30	3	41.7	
2009/10	Thu 25/2/2010 16:00	1	41.2	
2009/10	Mon 18/1/2010 14:00	2	43	
2009/10	Fri 12/3/2010 16:00	3	40.8	
2009/10	Tue 19/1/2010 15:00	4	41.5	
2011/12	Sat 28/1/2012 16:30	7	40.9	Weekend
2011/12	Thu 26/1/2012 15:00	11	41.3	Holiday
2012/13	Tue 12/2/2013 16:30	1	40.7	
2012/13	Thu 21/2/2013 17:30	5	40.6	
2012/13	Mon 31/12/2012 16:30	9	41.4	Christmas eve
2013/14	Sat 11/1/2014 17:30	5	43.4	Weekend

These results suggest that an appropriate approach may instead be to select days based on extreme weather conditions rather than just peak LSG. For example a weekend may be an appropriate day to use even though demand is dampened on the weekend because IGF output is not influenced by the day of the week.

Including other very hot days significantly increases the number of data points that can be used at any given temperature. Figure 14 provides an illustration. The figure shows the output of the fleet of IGF facilities that have been operational since 2007 (the 2007 IGF Fleet) on peak days and any other day on which the maximum temperature reached 39 degrees. Also shown is the mean output, the combined RL and the mean output on days when the temperature reached 41 degrees.

The figure also helps to highlight the issue of IGF performance on high temperature days. The output of the fleet is noticeably lower for higher temperature days. The RL appears to be a reasonable estimate of the likely output of the fleet. Of note, the mean output of the fleet on the 41+ degree days is similar to the RL. This suggests that potentially TIs from very hot days on weekends and holidays could be used to increase the number of TIs and that just using the very high temperature days might make the U-adjustment factor redundant.

Figure 14: Output on peak days of IGF fleet that have been operational since 2007



There are, however, some challenges with using these extreme temperature days. These include:

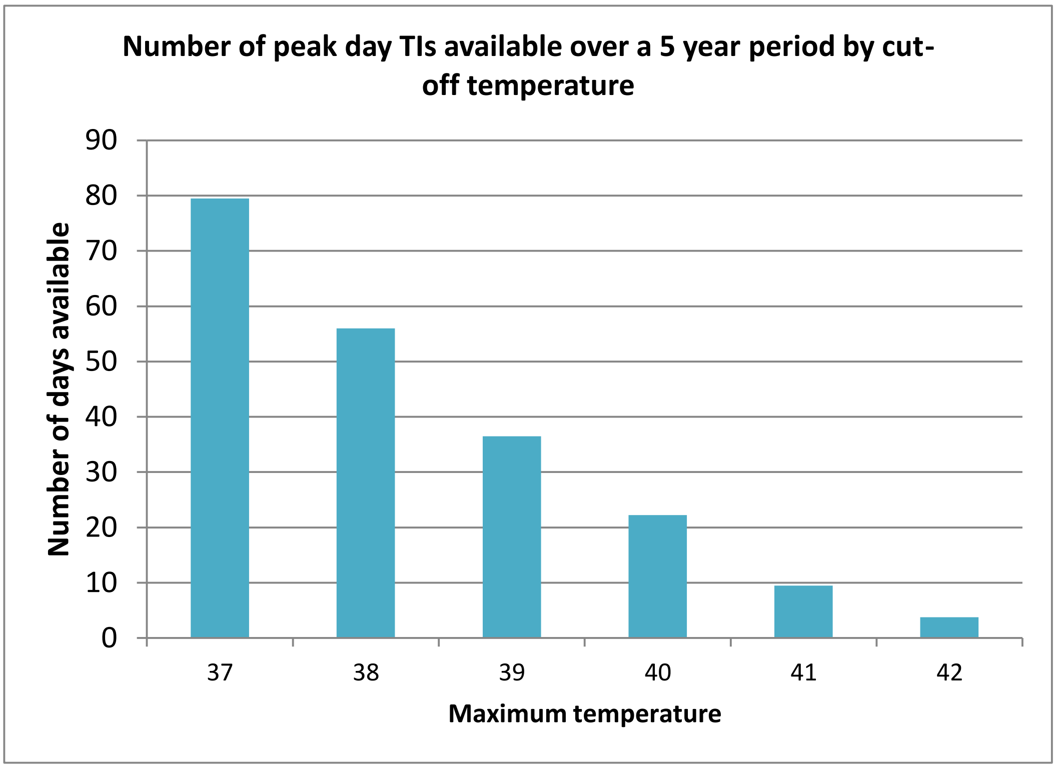
- There are still few very high temperature days. Figure 15 below shows the number of potential available days over a 5-year period by minimum maximum temperature. On average there are around twenty 40-degree days and less than ten 41-degree days. A risk is that the small number of TIs will lead to significant volatility in results.
- The seemingly negative relationship between temperature and outputs suggest that some additional adjustment might still be required.

- Peak TIs on very hot non-peak days (which are generally weekends and public holidays) tend to occur at different times to those on peak days. On non-peak days the peak TIs tend to be more spread-out.

We analysed similar charts to Figure 14 above for each individual facility. Our assessment, upon reviewing the charts, is that just using the output from very hot days would give results that are too volatile and not always representative of what would appear to be an appropriate RL.

Given the considerations above we have rejected the alternative of using just the peak days with very high maximum temperatures.

Figure 15: Number of hot days available



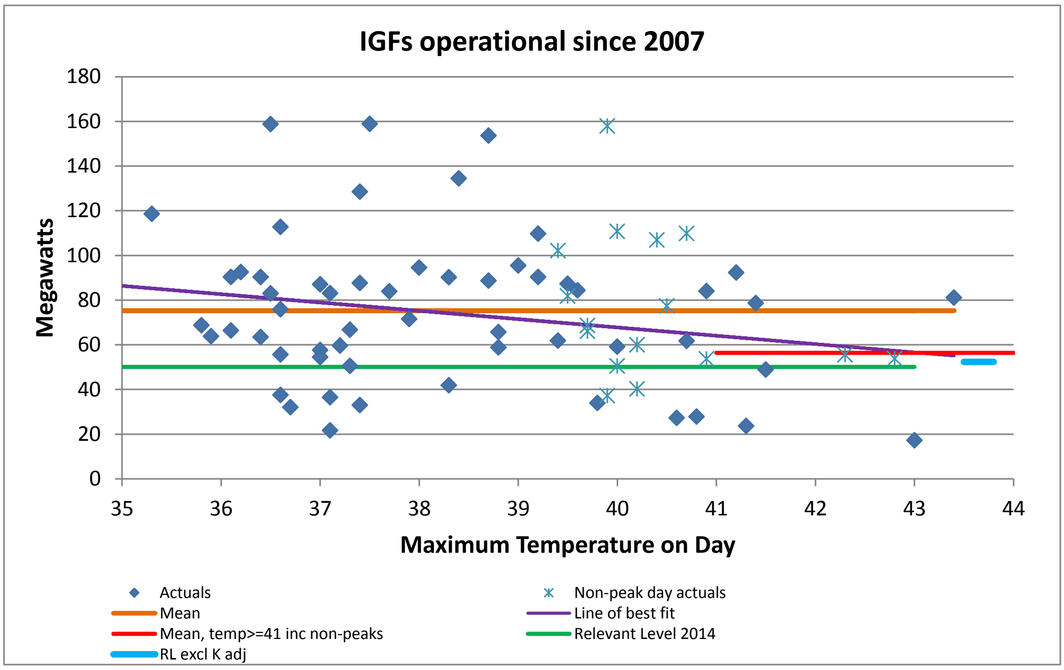
4.4.5 The 'regression' approach

Another potentially viable alternative is to make an adjustment that reflects how an IGF's output at peak TIs varies with maximum temperature.[36] In effect, this involves forecasting what IGF output will be at very high temperatures that are representative of the one in ten year peak.

[36] A potential alternative would be to use peak demand (or conversely surplus) as the independent variable in the regression. We have assessed this as not being a viable alternative. One challenge with this alternative is that it would not be simple to combine data from multiple years. Furthermore, our analysis suggests there is significant additional noise in the peak demand data and there this approach leads to more volatile results.

This option is illustrated in the Figure 16 below. The figure includes a regression line ('Line of best fit') that plots the relationship between the output of the IGF fleet at peak TIs and maximum daily temperature. The approach would then involve determining the RL based on a point on that regression line.

Figure 16 Output of IGF established since 2007 on peak TIs (5 years to 2014)



The calculation of the line of best fit uses a fairly simple formula. In the example, the slope of the line is calculated as:

$$B = \frac{Covariance(IGF\ output\ at\ peak\ TIs, \max temperature)}{Variance\ (\max temperature)}$$

The predicted output at a particular maximum temperature, *T* is then:

$$Y(T) = \bar{Y} + B(T - \bar{T})$$

where:

$\bar{Y}$ is the average output

$\bar{\mathrm{T}}$ is the average maximum temperature

The U-factor adjustment would then simply be replaced by an adjustment equal to $B(T - \bar{T})$. The maximum temperature target *(T)* could be the daily maximum temperature at which the 10% POE peak demand occurs. We estimated this (as discussed in the beginning of section 4.4.2) to be around 43.8 degrees for a one in ten year peak.

The approach is attractive as it would not penalise those IGFs whose output does not decrease on hotter days and more appropriately penalises those whose output is more likely to fall at the extreme peaks.

One might question whether the forecast could be improved by including additional variables in the formula above. For example, the time of the peak may be of interest if (as noted in section 4.4.2) on very hot days the time of the peak varies to other days. Such a modification is not advised. Ultimately, we are attempting to forecast what will occur on the very peak days. We have good reason to believe that maximum daily temperature is a driver of demand and thus a good variable to use in forecasting. This is not the case with the time of day. Changes in the time of the daily peak are more likely to reflect changes in demand behaviour on very hot days and therefore be correlated with the temperature variable.[37] Furthermore, inclusion of additional variables would result in a significantly more complex formula to estimate output.[38]

The key issues and challenges with the approach are that:

- the relationship between maximum temperature and output is not necessarily linear, and
- the result is sensitive to outliers.

Potentially, statistical techniques could be used to address such issues; however, such techniques (such as transforming the variables and limiting the selection of variables) are reasonably complex.

Furthermore, the approach would appear to be an innovation for a capacity market. We are aware that some similar work exploring the relationship between temperature and wind output is being undertaken for adequacy assessment in Europe and the US; however, we have not found any practical applications of the approach.

We analysed the 'line of best fit' for each facility. Of note:

- In many cases the results are very similar to the current U-factor adjustment.
- In some cases there are differences, whereby the regression approach appears to provide more realistic results, in particular where the IGF output does not appear to be correlated with temperature.
- There are cases which appear unrealistic, for example, where the regression line is upward sloping despite the average output on the very hot days being below the average output overall all peak days.

While the approach appeared to have some merit we were concerned about the volatility of the results. Our analysis of the results indicated that the standard error of the RL using this regression approach would be (at the median) double that of the U-factor adjustment. That is, for example, the confidence interval range for estimating the RL would most typically double.

This coupled with our concerns that the relationship between temperature and IGF output is not linear leads us to reject this approach.

---

37 In effect, the changes due to time of day effects will be reflected in the changes that are correlated with temperature.

38 It would require the use of matrix algebra or many additional terms to the formula listed.

4.4.6 Update – U-factor

In the 2011 Report the U-factor was determined in combination with the K-factor to ensure that the RL for the IGF fleet was consistent with the peak reduction contribution[39] of the IGF fleet at the extreme peaks.

For this report, a similar process was undertaken to determine a revised U parameter value. As with the previous report, we identified a U parameter value that would deliver results that are consistent with the level of peak reduction by the IGF fleet at extreme peaks. However, for this report, we estimated the U-factor separately to that of the K-factor. This was done by (as with the analysis in section 3.3.1) comparing the value of peak reduction with the RL excluding the K-factor adjustment.

The process of estimating U was complicated (compared to the process for the 2011 Report) by the large number of additional facilities that have recently become operational, with the implication that the size of the operational IGF fleet has changed by year. To draw some comparison over different years we examined for the IGF fleet in each year the ratio of peak reduction on the peak Trading Days to the mean output of the fleet for the peak TIs used in the RL methodology and how this varied with temperature.

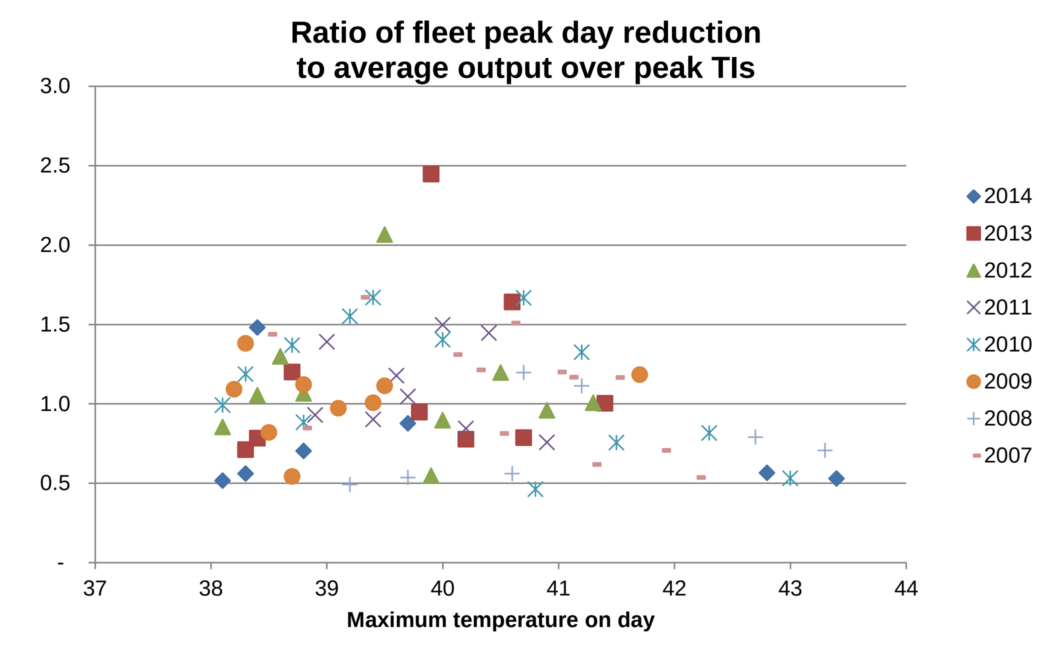
A summary of the results for days which contained peak TIs (used in the methodology) and where the temperature reached at least 38 degrees is shown in Figure 17 below. Each point represents a single day. A ratio of 1 indicates that the reduction in peak LSG on the day was equal to the mean output of the fleet over all peak TIs in that year. For example, the 2014 point in the far right corner of the chart is 0.53 which equals the peak reduction on that day (114 MW) divided by the mean output of the IGF fleet in the 12 peak TIs for 2014 (215 MW).[40]

The figure again highlights the problem of the performance of IGFs on very hot days. The mean output of the fleet of operational IGFs at peak TIs in 2014 was 215 MW whereas the peak reduction by the fleet in 2014 on the 4 hottest days was 149 MW (a 33% fall from the mean output) and on the days with a maximum temperature at least 37 degrees was 162 MW (a 25% fall from the mean output).

[39] That is, the amount (in MW) by which LSG was reduced by the IGF fleet.

[40] Of note, the output during the peak TIs varies substantially by individual year. The average output of the currently operational fleet over 5 years to 2014 (including estimated data) was substantially less than the 2014 output. Therefore the 2014 points in graph are abnormally low.

Figure 17: Performance of IGF fleet on hot days



Ultimately, selecting a U-factor parameter requires a degree of judgement. In forming our judgement we considered the following.

First, as discussed in section 3.3, the results of the RL methodology using the current U parameter appear to be, in general, appropriate.

Second, the most recent results are consistent with the current U parameter. The peak reduction of the fleet of IGFs that was operational in 2014 on the hottest day (the far right hand point in the graph) was 114 MW[41], almost precisely what is predicted by the current RL methodology (excluding the K-factor adjustment).[42] The other points in the graph suggest that the far right point is not an anomaly but rather a reasonable representation of the performance of the IGF fleet on very hot days.

In summary, the most recent results are consistent with the current parameter values for the U-factor adjustment. We therefore have no justification to change the current parameter value.

The maximum adjustment

The current methodology includes a maximum level of adjustment attributable to the U-factor of one third of the mean IGF output. Like the U-factor adjustment itself, the maximum adjustment is not based on theory; rather it is a pragmatic measure that prevents the U-factor adjustment being excessive, thereby providing some protection to IGF owners

41 The mean output of the current operational fleet over the last 5 years was 151 MW.

42 More specifically, for IGF fleet that is operational in 2014 we calculate the RL excluding the K-factor adjustment to be 113.7 MW. The peak reduction on the highest temperature day in 2014 was 113.5 MW.

from excess variations. The maximum adjustment would not be required if a reasonable alternative to the U-factor adjustment could be found.

The size of the maximum adjustment can be assessed by comparing the effect of the adjustment with the actual results. Based on the 5 years to 2014, it would appear to be binding in 3 of 18 cases. In two of the cases the maximum level is just binding. In one case the maximum adjustment prevents the RL methodology returning a negative result.

This current maximum level of adjustment appears reasonable. In the one case that the maximum adjustment has a significant impact the resulting RL (excluding the K-factor adjustment) closely matches the IGF's output on the peak TI for the hottest day for which data is recorded.

4.4.7 Summary

Making an adjustment for poor IGF performance at very high temperature days is a substantial issue.

The current approach appears to have been reasonably effective in that the U-factor adjustments appear to largely match the output we would expect at peak TIs. Nevertheless, there is a risk that the approach will penalise an IGF whose output at peak TIs is not correlated with maximum temperature.

Viable alternatives were examined, however, we have assessed:

- it is not feasible to select just a small number of TIs, and
- a regression approach appears attractive but results are not significantly different and it may introduce additional volatility.

On the basis that there is no better alternative identified, we recommend continuing with the current approach. We have also concluded that the current U-factor parameter is appropriate.

4.5 Other issues in the selection of Trading Intervals

The current methodology is applied on sixty TIs. Twelve TIs are selected from each of five years with each TI selected from a separate day and with each TI given equal weight. This section considers whether any of these factors should be modified.

In selecting the TIs, there are general trade-offs between selecting the TIs that are representative[43] and selecting a sufficient number so as to ensure that the methodology is not too sensitive to the selection of a few TIs.

[43] Of note, we are concerned with finding TIs that are *representative* of the output of IGFs. The TIs themselves need not represent the highest demand periods. For example, demand on a particular day may be much lower simply because it is a holiday or weekend.

As discussed in the previous section, there is clearly a concern that the current sixty TIs are not representative of what will occur during an extreme peak. Nevertheless, for applying the methodology some data is required and therefore the approach might be considered as attempting to find the *most* representative TIs in the data that is available.

If volatility was not a concern, a single highest peak TI would be used to determine the 'average' output. Additional TIs are selected to provide additional information so as to reduce the volatility of the results. These are selected on the basis that that they provide additional information and are representative of what could be the one in ten year peak. They need not be the TIs with the next highest LSG (LOLP); for example, on the day with the single highest peak, the LSG at 6:00pm may be higher than other TIs used, but it would be inappropriate to use this TI because, based on history, 6:00pm is very unlikely to be the time of the extreme peak in LOLP.

4.5.1 The number of separate days (TIs) in a year

The current method uses TIs selected from 12 separate days. While using more days provides more information, it also increases the risk that the days, and the peak TIs selected from those days, are unrepresentative of the 1 in 10 year peak.

To assess the number of separate days to use, we undertook some analysis to understand how many days could be selected from a year such that the peak TI of each day would still be potentially representative of a peak. Using data from the 7 years from 2007/08 through to 2013/14, we examined how the key characteristics of the peak TIs used in the RL methodology varied with each extra day; that is, for example, from moving from the 5th highest ranked day to the 6th highest ranked day in a year.

For example, Figure 18 shows how the timing of the peak TI in the day varies between the highest ranked day (rank=1) through to the 30th highest ranked day in the year. The figure suggests that the timing of the peak does not materially change for 12 highest ranked days but gets more unpredictable for lower ranked days. For example, the latest time of the peak TI of the 16th highest ranked day is later than any TI selected from the higher ranked days. Based on the this figure alone we would be concerned that the peaks TIs on days ranked much lower than around 12 would be unrepresentative.

Figure 19 shows the date range of the peak TIs. From this figure we can see that in the years 2007/08 through to 2013/14 the top 15 ranked days all fell between 15 December and 25 March but that lower ranked days fell outside this range.

Figure 20 shows how the air temperature at the time of the peak TI varies by rank of day in the year. Consistent with the finding that temperature is strongly correlated with peak demand we see a slight downward trend between air temperature and the rank of the day. The difference in temperature by rank of day becomes relatively large for days outside the top 15 ranked days.

Based on these results we conclude that it would be inappropriate to increase the days (and therefore TIs) used per year beyond the current number of 12. There also appears no compelling reason to reduce the number of TIs. We therefore recommend that the methodology continue to use 12 TIs per year.

sapere research group

Figure 18: Time of the peak TIs by rank of day in year

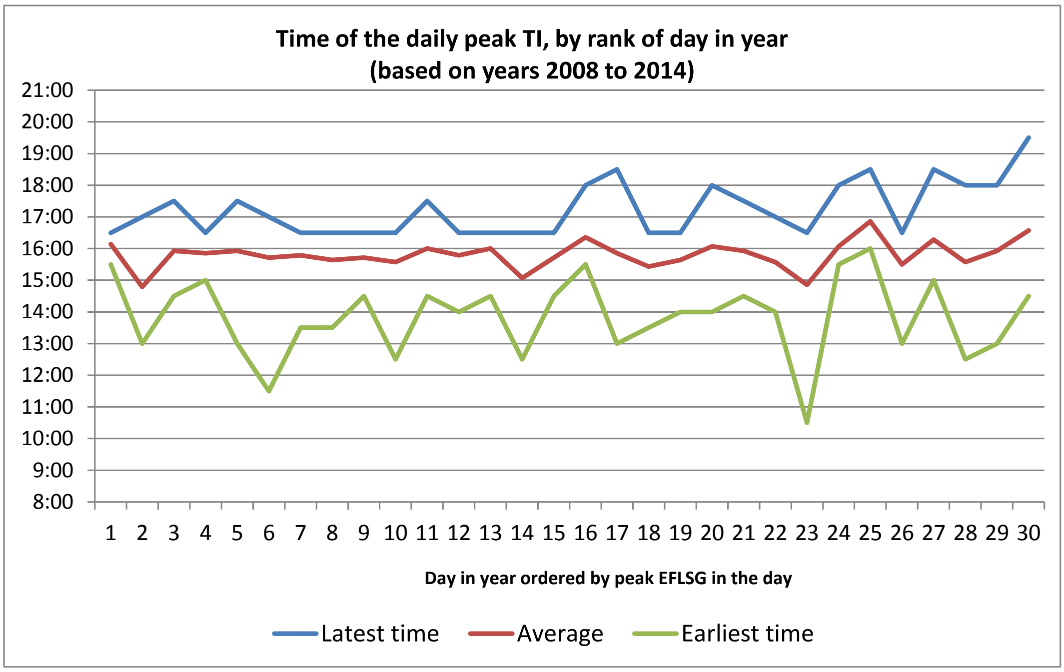


Figure 19: Date of peak TIs by rank of day in year

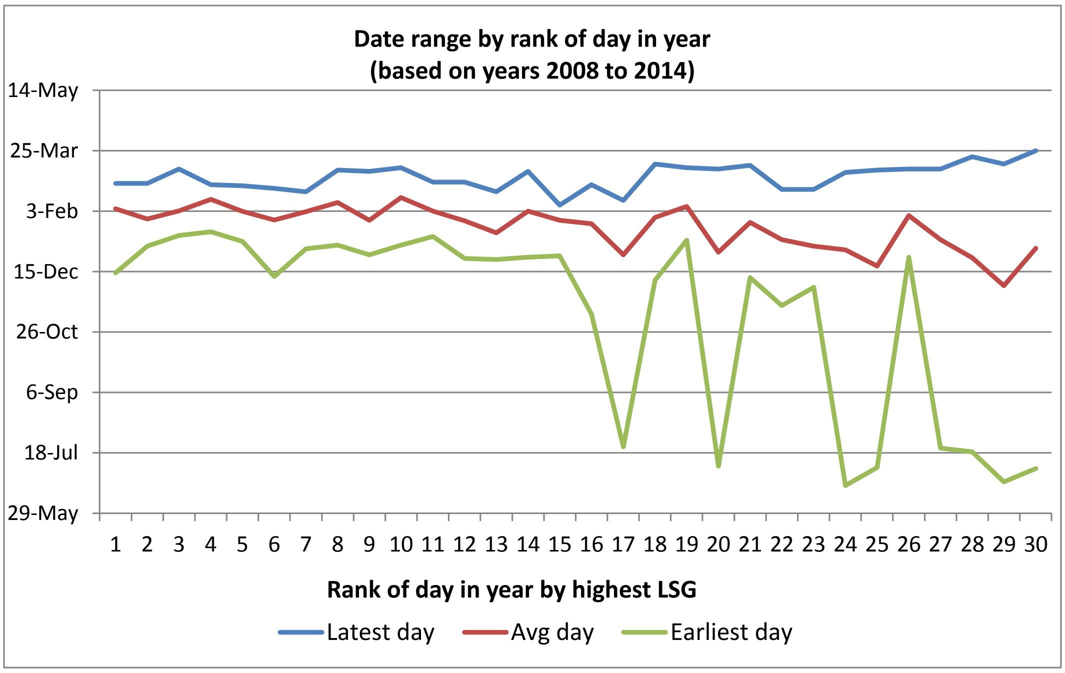
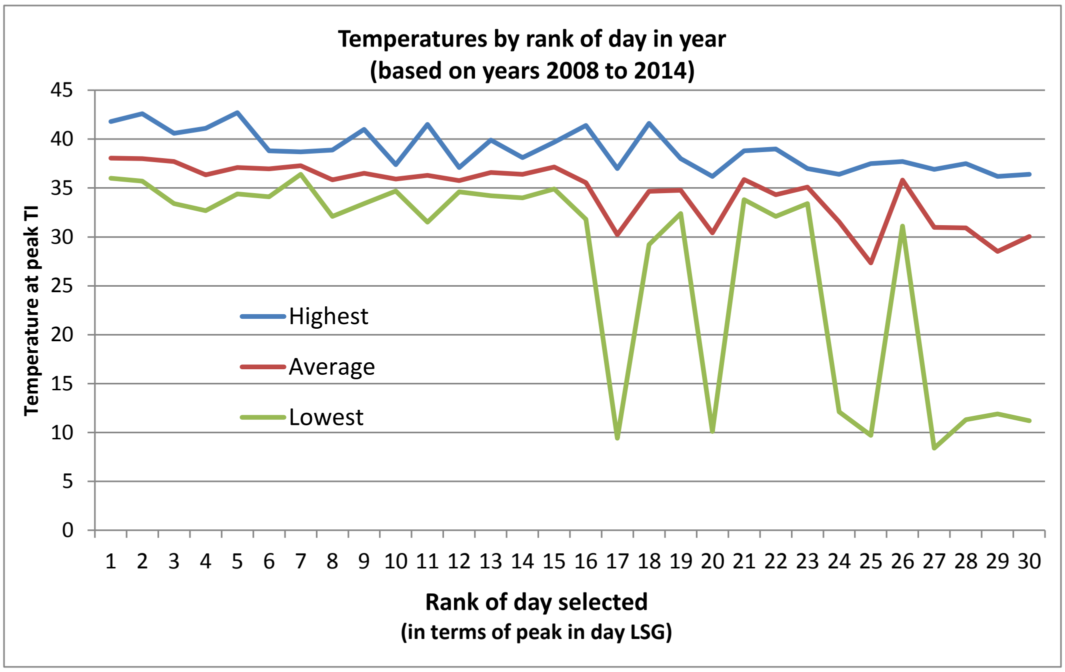


Figure 20: Temperature at time of peak TIs by rank of day in year



4.5.2 The number of years for selecting TIs

The current method selects TIs from 5 years. The choice of 5 years was based on the availability of data at the time (2011 Report, pages 12-13).[44] Since the 2011 Report an additional 3 years of data has become available to be used.

In general, using more years of data is desirable in that it provides greater reliability of results. However, there are downsides to using more years of data. These are:

- The older the data that is used the less likely that it will be representative. This is particularly important if the profile of demand changes over time.
- The more years of data that is used, the more that data from expert reports will need to be relied upon for estimating the RL for new facilities.
- There are administrative costs to changing the number of years.

Figure 21 below is helpful in examining the first of these issues. It shows the timing of the peak demand TI (i.e. MG not LSG)[45] over the last 8 years. There is a clear trend of the peak shifting to later in the day.[46] Thus the results suggest that the time of peak demand is changing and therefore a time based method that is appropriate in one year may not be appropriate in a following year. These results also suggest that there is a large risk in using a large number of years.

---

44 Of note, the original IMO proposal in 2011 was for using 8 years' worth of data.

45 Because the IG fleet changes by year, it is useful in this instance to examine results based on MG.

46 We presume the increase in installed solar PV has contributed to this trend. There may be other factors. The map in Figure 1 on page 26 includes an estimate of the installed solar PV system capacity of 366 MW.

The timing of peak LSG is shown in Figure 22. This figure shows a similar trend. We note that since 2010/11 the timing of the peak TIs has been stable but that the average time of the top 12 TIs is getting later. These results suggest that there is some risk that prior years will not be representative of the current peak. However, in our opinion there is not enough movement to warrant a reduction in the number of years selected.

Figure 21: Timing of MG peaks by year

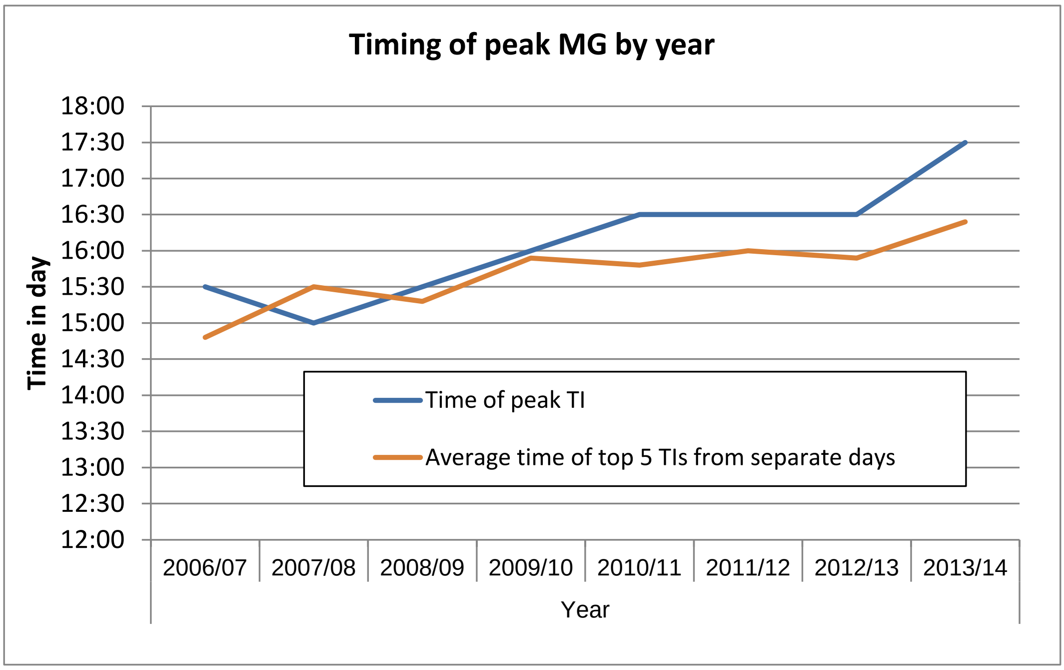
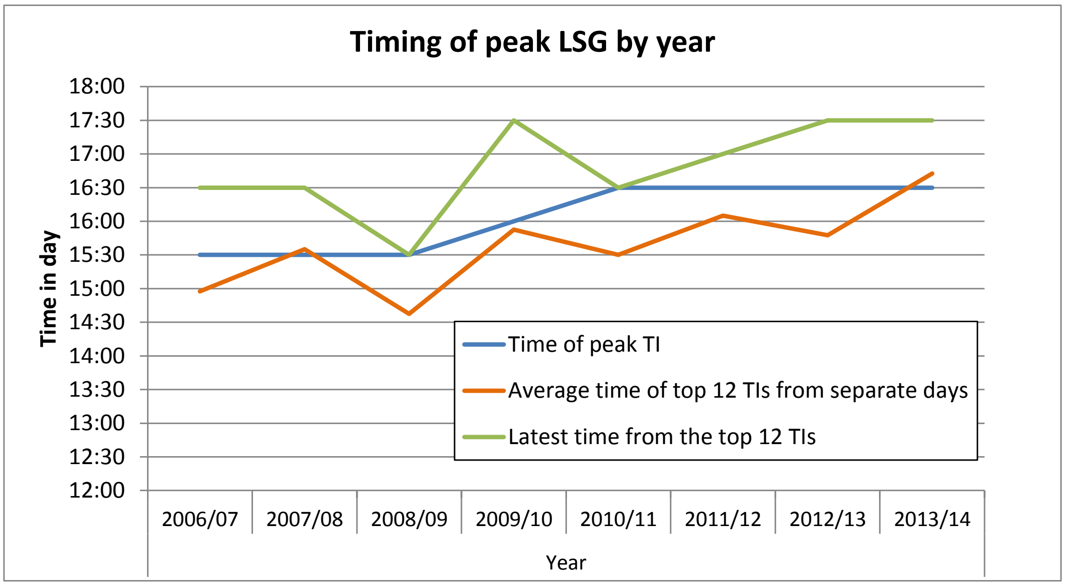


Figure 22: Timing of LSG peaks by year



4.5.3 Weighting of Trading Intervals

The current method gives equal weight to all TIs selected. An option that may improve accuracy is to weight the TIs used such that the TIs that are more representative of the peaks will be given greater weight.

There is no limit to the weighting options. In theory the greatest weight should be given to the TIs that are most likely to represent the extreme peaks. In this case, we might place greater weight on TIs with the highest LOLP or from days with the highest maximum temperature. Simple weighting procedures could for example, place double the weight on days when the peak temperature exceeds 41 degrees.

We experimented with a few different types of weights based on LSG and maximum temperature in the day. Consistent with the findings in section 4.4 that the IGF fleet performed worse on extreme peaks, applying a weighting based on LOLP or temperature resulted in a lower average output. However, the modification was not significant and did not make a material difference to the need for another adjustment. For example, using a weighting of the square of LSG resulted in a fall in the mean output of the Existing IGF fleet of about 1 MW; which is very small relative to the U-factor adjustment of above 18 MW for this group.

The downside of applying weights is that it can increase the volatility as the results will be more greatly influenced by the highest weighted TIs. A further downside is that applying weights adds complexity to the formula used.

Given the limited additional benefit for the added complexity, we rejected using weights.

4.5.4 Number of Trading Intervals per day

The 2011 Report recommended selecting a single TI from each day. The arguments against using additional TIs from a single day are that it:

- has little benefit, and
- has unwanted effects.

Using additional TIs from a single day has little benefit because the output of IGFs in a TI is highly correlated with that of adjacent TIs.[47] A simple regression analysis[48] highlights that over 97% of the variation in the IGF fleet's output is explained by the output of the preceding TI. The implication of this is that using more TIs from the same day provides little additional information.

A seemingly possible benefit of using additional TIs from the same day is that, if it is combined with a rule selecting the top TIs in a year, then the effect is to place additional weight to the days that contain the highest peak TIs. However, as noted in the 2011 Report (page 33) and discussed in Section 4.5.3 there are simpler and more accurate ways to apply a weighting to the highest peak TIs.

---

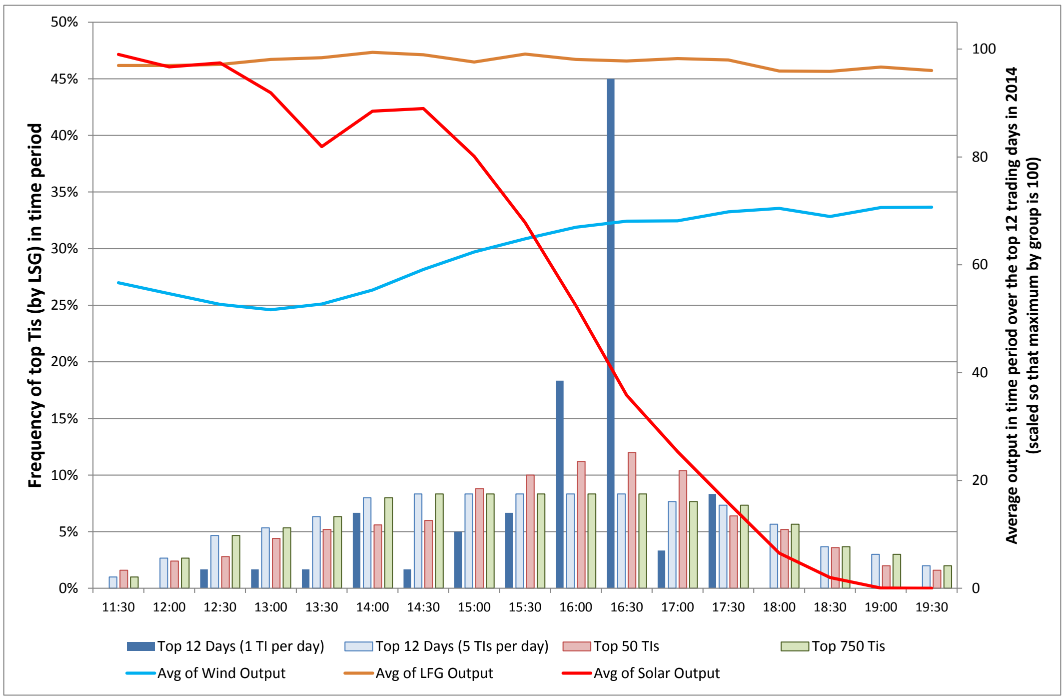
47 As described in the 2011 Report it is similar to conduct a phone survey and repeatedly calling the same household.

48 Regression of fleet IGF output on the previous period's fleet IGF output.

There are unwanted effects of using additional TIs from the same day. By construction the second, third, fourth etc TIs are not representative of the peaks. This is of particular concern as the outputs of many IGFs follow a similar pattern throughout the day. This is highlighted in Figure 23, which shows the frequency of when peak TIs occur and how the average output of IGFs changes over the course of the day. The figure shows a variety of selections of TIs including the current selection (that is the Top 12 days with 1 TI per day), and selections which involve more than 1 TI from the same day. As can be seen in the figure using more TIs in a day results in the distribution of TIs used being spread over the day away from the period (16:30) that by far the most likely time for the peak to occur. The figure also shows (for the peak Trading Days in 2014) the pattern of output for IGFs for different technology groups over the course of the day. It is clear from the figure that the output of solar and wind is related to the time of day.

The key risk is that a biased result is obtained meaning that the TIs used would be unrepresentative of a potential one in ten year peak. For example, the lower ranked TIs in a day may be later in the evening (which would bias against solar) or possibly earlier in the day (which would bias against wind).

Figure 23: IGF output and timing of peak TIs



Notes: Frequency data from 5 years to April 2014. Output data was limited to actual output in the year ending April 2014. Top 50 and top 750 refer to the top 50 and 750 TIs in each year. In all cases the top TIs are by LSG.

5. Conclusion and summary

5.1 Overview

The RL methodology introduced in 2011 can be thought of as containing two components:

- the mean output at peak LSG less the K-factor adjustment, and
- the U-factor adjustment.

The first component is consistent with established probability theory on capacity valuation and an established approximation method. Some possible minor improvements were identified in applying the method; however, we assess the effect of these to be minor.

The second component, the U-factor adjustment was added as pragmatic method given limited available data to address the issue that IGFs appeared to perform worse at very high temperature days when LOLP was likely to be highest.

The U-factor adjustment generally provides (in our subjective assessment given the limited data available) reasonable results at a fleet and individual facility level.

However, there are theoretical issues with the U-factor adjustment and therefore it may not prove robust in the future. Alternative options were explored; however, given the limited data available, the current structure appears to be the best viable option in the immediate future. Improving on it will require overcoming the problem of a lack of information about how IGFs will perform in extreme temperature periods.

Prior to the next three-year review it would be appropriate to consider options as to how to improve the data available to forecast the conditions and from this the performance of IGFs at extreme peaks. In effect, a forecast of the probability distribution of IGF output on the very hot days that represent the extreme peaks is required. This might, for example, involve forecasting the distribution of weather conditions that affect demand on extreme peaks days and also the related weather conditions where IGFs are located.

Other aspects of the current methodology were examined. There was no compelling reason to make any other adjustments. We also note that no objections to this conclusion were raised (and similarly no material issues raised or modifications proposed) during the public consultation period following the release of the draft version of this report.

5.2 Summary of recommendations

We recommend that no changes are made in the short-term to the current methodology.

The K and U parameters were examined. Based on our analysis we recommend that:

- The revised K parameter value should be set to zero, and.
- The revised U parameter value should not be changed; that is it should remain at 0.635.

References

Agency for the Cooperation of Energy Regulators (ACER) (2013). Capacity Remuneration Mechanisms and the internal market for electricity.

Belgian Regulatory Commission for Electricity and Gas (2012). Study on capacity remuneration mechanisms.

Calpine, CPS Energy, LCRA, Luminant and NRG Energy, 2011, 'Winter Weather Readiness for Texas Generators', mimeo.

Dent, C. J., A. Keane, and J. W. Bialek, 2010, "Simplified Methods for Renewable Generation Capacity Credit Calculation: A Critical Review," in IEEE PES General Meeting, 2010.

Dent, C. J., & Zachary, S. (2014). Further results on the probability theory of capacity value of additional generation, International Conference on Probabilistic Methods Applied to Power Systems (PMAPS) Durham, England, 7-10 July 2014..

Dent, C. J., Keane, A., & Bialek, J. W. (2010, July). Simplified methods for renewable generation capacity credit calculation: A critical review. In Power and Energy Society General Meeting, 2010 IEEE (pp. 1-8). IEEE.

Dent, C., & Zachary, S. (2013). Capacity Value of Additional Generation: Probability Theory and Sampling Uncertainty. arXiv preprint arXiv:1305.6479.

Dragoon K, Dvortsov V., 2006, Z-Method for Power System Resource Adequacy Applications. IEEE Trans Power Syst. 2006 May;21(2)

Garver, L.L. (1966). Effective Load Carrying Capability of Generating Units., IEEE Trans. on PAS (PAS-85); pp. 910–919.

Independent Market Operator WA (IMOWA) (2013). Electricity Statement of Opportunities – June 2013. Available at http://www.imowa.com.au/docs/default-source/Reserve-Capacity/soo_2013_rev1.pdf?sfvrsn=2.

Independent Market Operator WA (IMOWA) (2014). SWIS Electricity Demand Outlook — June 2014. Available at http://www.imowa.com.au/docs/default-source/Reserve-Capacity/soo/imo_2014-swis-electricity-demand-outlook5EA7F337C272.pdf?sfvrsn=0.

International Energy Agency (2013). Expert Group Report on Recommended Practices - 16. Wind Integration Studies, 1st Edition

Keane, A., Milligan, M., Dent, C.J., Hasche, B., D'Annunzio, C., Dragoon, K., Holttinen, H., Samaan, N., Soder, L., O'Malley, M., 'Capacity Value of Wind Power', May 2011, IEEE Transactions on Power Systems, vol. 26(2) pp564-572.

Lu, S., Diao, R., Samaan, N., & Etingov, P.,(2012). Capacity Value of PV and Wind Gneration in the NV Energy System, Pacific Northwest National Laboratory

Madaeni, S. H., Sioshansi, R., & Denholm, P. (2011). Capacity Value of Concentrating Solar Power Plants. National Renewable Energy Laboratory, June 2011..

Madaeni, S. H., Sioshansi, R., & Denholm, P. (2012). Comparison of Capacity Value Methods for Photovoltaics in the Western United States, National Renewable Energy Laboratory, Technical Report NREL/TP-6A20-54704 July 2012

Milligan, M., & Porter, K. (2008, June). Determining the capacity value of wind: An updated survey of methods and implementation. In Wind Power.

Midwest ISO (2013). Planning Year 2014-2015 Wind Capacity Credit

Milligan, Porter, 2008, Determining the Capacity Value of Wind: An Updated Survey of Methods and Implementation, Conference Paper NREL/CP-500-43433 June 2008.

North American Electric Reliability Corporation (NERC), Methods to Model and Calculated Capacity Contributions of Variable generation for Resource Adequacy Planning, March 2011

PJM (2014). Manual 21: Rules and Procedures for Determination of Generating Capability

Rogers, J., & Porter, K. (2012). Summary of Time Period-Based and Other Approximation Methods for Determining the Capacity Value of Wind and Solar in the United States: September 2010-February 2012 (No. NREL/SR-5500-54338). National Renewable Energy Laboratory (NREL), Golden, CO.

Stoft, Steven E., The Surprising Value of Wind Farms as Generating Capacity (August 23, 2008). Available at SSRN: http://ssrn.com/abstract=1250187.

Tooth, Richard. (2011). Capacity Value of intermittent generation: Public Report. 18 August 2011. Available at: http://www.imowa.com.au/docs/default-source/rules/rejected/imo_-_board_report_tweaks_final_v5_001.pdf?sfvrsn=2.

Zachary, S. & Dent, C.J. (2012). 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. Draft version available at http://www.supergen-networks.org.uk/publications.

Zhang, N., Kang, C., Kirschen, D. S., & Xia, Q. (2013). Rigorous model for evaluating wind power capacity credit. Renewable Power Generation, IET, 7(5).

Zachary, S., & Dent, C. J. (2014). Estimation of joint distribution of demand and available renewables for generation adequacy assessment. IEEE transactions on power systems.

Appendix 1 Theory and international experience

Introduction

A review of the literature and practices was conducted for the 2011 Report. Since that time there has been some additional relevant work undertaken. In general the then emerging preference for detailed ELCC Risk Method analysis has been verified (Keane et al 2011, NERC 2011, IEA 2013). Some analysis refines existing theories (Dent and Zachary 2014, Zachary and Dent 2014), while other analysis extends the application of approaches examined with wind generation to solar generation (Madaeni et al 2012).

Similarly some additional work has been undertaken surveying practices, particularly in North America. These indicate minimal change, although a new capacity market has commenced in the Midwest ISO.

Overview

The value of a capacity credit is the contribution that a generation facility makes towards the power system reliably meeting demand. A reserve margin must exist in a power system because all generators have non-zero probability of failure, and loads cannot be known with certainty in advance. The methods and data requirements of capacity value are well established in traditional resource adequacy planning, and have been adapted to variable generators including wind and solar.

In the past decade a number of markets in Europe and the USA have responded to increasing uncertainty in the investment in generation capacity by introducing mechanisms to remunerate generators for available capacity as well as produced energy. Hence the capacity credit has acquired a financial value to generators reflecting their contribution to power system reliability.

There is a variety of such mechanisms that provide generators revenues complementary to their sales of electricity in wholesale markets. The details of such mechanisms reflect the specific circumstances of each market, however there are three main types of market-wide mechanisms. The simplest type is a direct capacity payment that is defined and controlled by an independent body. Market mechanisms include, inter alia, decentralised 'capacity obligations' or 'resource adequacy requirements' that place future reserve requirements (set in part by an independent body) on large consumers (i.e. retailers) and suppliers creating a primary market for bilateral agreements, and capacity auctions where the role for capacity procurement on behalf of total demand is centralised through an independent body.

EU power systems

The European Union (EU) is dominated by energy-only markets. However as conventional generators are retired and replaced by renewable generators, some member states have implemented or are investigating capacity remuneration mechanisms to ensure sufficient capacity is available. Five markets (Ireland and Northern Ireland, Greece, Italy, Portugal and Spain) have existing capacity remuneration systems.[49] Currently Italy and Spain are revising their capacity payment mechanisms and other states are considering introducing capacity markets including France, Germany and Great Britain (ACER 2013).

However little information is available regarding the detailed calculation of capacity credits for renewable generators, and in the absence of an overarching regulatory framework most European studies are focused on the effectiveness of capacity remuneration mechanisms or energy-only markets within states meeting the policy objectives of those states (e.g. Belgian RCEG 2012).

North American power systems

Five of America's seven cross-jurisdictional Regional Transmission Organizations (RTO)/Independent System Operators (ISO) operate capacity markets. Four ISOs operate forward capacity auctions where retailers can obtain capacity for future periods (PJM, MISO, NYISO and ISO-NE). The California Public Utilities Commission resource adequacy requirements require retailers to have capacity to meet their annual peak including regulatory reserve margins.

There being an overarching regulatory framework administered by the Federal Energy Regulatory Commission, there is a stronger focus of studies in North America on how these regulatory objectives are implemented by RTO/ISOs and local utilities to fulfil adequacy planning, including the operation of capacity markets. A summary of wind generation capacity values and methods employed in US power systems was given in 2008 (Milligan and Porter 2008) and updated in 2012 (Rogers and Porter 2012). As noted therein, broadly approximation methods are risk-based or time-based, where the latter attempt to capture risk indirectly by assuming a high correlation between LOLP and demand, and assume all hours are weighted evenly.

The following summarises time-based methods used to determine ELCC for intermittent generators by the five RTO/ISOs with capacity markets.

California Public Utilities Commission (CPUC)/California ISO

CPUC sets the capacity value of wind generators monthly through a capacity factor approximation using a 70% exceedance factor; meaning the output achieved in 70% or more of the sample intervals each month. The sample intervals are the previous 3 years of generation between 4 and 9 pm for November to March and 1 and 6 pm for April to October. This initial capacity value so determined is modified to account for the benefit of system diversity. This involves allocating the difference between the 70% exceedance of the generation fleet as a whole and the sum of individual generators in proportion to an individual generator's contribution to fleet production.

Independent System Operator New England (ISO-NE)

ISO-NE assigns variable energy facilities capacity credits using a capacity factor approximation based on a rolling average of the median net output of the variable renewable energy from the previous 5 years during designated periods – 1 through 6 pm

[49] Capacity payments targeted to maintain strategic reserves also operate in Sweden, Finland and Poland.

from June through September for the summer capacity credit, and 5 through 7 pm between October and May for winter capacity credit. Additional hours are included to adjust for periods of system-wide shortages and/or performance during local shortages in import-constrained zones. New facilities are assigned a capacity credit based on 1 year of onsite data (Rogers and Porter 2012).

Midwest ISO (MISO)
MISO introduced a capacity market in 2013. MISO performs a full ELCC study to determine the capacity value of the wind fleet, using actual historic load, generation and wind values for the previous year. This capacity credit is then allocated to individual wind facilities deterministically using historic peak period data, based on the average capacity factor over the top 8 daily peak hours per year for the past 9 years (72 peak periods in total) (MISO 2014).

New York Independent System Operator (NYISO)
NYISO determines wind facility credits based on the capacity factor from the previous year during designated periods – between 2 and 6 pm June through August for the summer capacity credit and 4 and 8 pm December through February for the winter capacity credit. New wind facilities are assigned credit as a fixed proportion of their nameplate capacity – 10% for summer and 30% for winter for onshore facilities and 38% for offshore facilities (Rogers and Porter 2012).

PJM
PJM determines capacity values for renewable generators based on the capacity factor using a rolling 3-year average of net capacity between 2 and 6 pm local prevailing time from June through August. Hours when PJM directed the wind generator to reduce its output are excluded from the calculation of the capacity factor so as not to penalize the wind generator for following PJM directives. For new wind projects with insufficient wind generation data, PJM applies a "class average" capacity value (currently 13% for wind, 38% for solar), to be replaced by the wind generator's actual capacity value once the project is in operation for at least a year (PJM 2014).

Developing temperature correlated data
The North West Resource Adequacy Advisory Committee (formerly ad hoc Forum) currently assesses wind credits (for planning purposes) using 30% of installed capacity for the annual value and 5% for the sustained-peak value, the percentages being anecdotally derived from a historic review of wind generation data from one member. The forum would like larger data resources to assess wind ELCC and is exploring the development of synthetic wind generation data that is temperature-correlated.

Appendix 2 Background information on the WEM

The Planning Criterion

The current design of the WEM derives the required amount of capacity (**Reserve Capacity Target**) from the Planning Criterion. The Planning Criterion (clause 4.5.9 of the Market Rules) sets a minimum standard for the acceptable level of generating capacity and has 2 parts:

- A "defined event scenario" that sets out the requirement for reserve generating capacity which must be available during system peak as the greater of:
  - 7.6 percent of the forecast peak demand (including transmission losses and allowing for Intermittent Loads); and
  - the maximum capacity, measured at 41 degrees Celsius, of the largest generating unit,

  while maintaining Minimum Frequency Keeping Capacity for normal frequency control. The forecast peak demand should be calculated to a probability level that the forecast would not be expected to be exceeded in more than <u>one year out of 10</u>; and
- A requirement that there be sufficient reserve to ensure that expected energy shortfalls are restricted to 0.002 percent of annual energy consumption.

The Reserve Capacity Target is set annually based on the <u>most stringent element</u> of the Planning Criterion.[50]

[50] Note that the Planning Criterion applies to the provision of generation and Demand Side Management capability and does not include transmission reliability planning.

Appendix 3 Load for Scheduled Generation

Note: This appendix is an abridged version of a public guidance note of LSG.[51]

The formula for determining the number of Capacity Credits for Intermittent Generation facilities that apply from the 2012 year is:[52]

Capacity credits =	*1. Average facility output during peak Trading Intervals*	***Less***	*2. G × variance of facility output during peak Trading Intervals*

Where $G = K + U$, reflecting both known variability (reflected in *K)* and *uncertainty* of the distribution *of output* (reflected in U).[53]

LSG is the measure used to select the peak Trading Intervals. Peak LSG identifies the Trading Intervals where surplus capacity is lowest[54] and thus when the system is under greatest stress. LSG is calculated (in MWh) as:

- the Total Demand for energy less
- the Total Intermittent Generator Output.

Using LSG, a total of sixty peak Trading Intervals are selected from a five year period. Twelve Trading Intervals are selected from each of the five years with the requirement that each Trading Interval occurs on a separate Trading Day.

LSG is determined for a five-year period, being the five years up to and including the most recent Hot Season (which ends on 31 March). In each year of the five-year period, LSG is used to determine the peak 12 Trading Intervals.

- The Trading Intervals are sorted in order of highest LSG and the top 12 Trading Intervals are selected with the requirement that these are selected from separate Trading Days.
- The output[55] of the candidate facilities during the 60 selected Trading Intervals over the relevant five-year period is then used in determining the Relevant Level.

---

51 Availabile at http://www.imowa.com.au/docs/default-source/Reserve-Capacity/lsg_help_guide_10feb2012.pdf?sfvrsn=2.

52 For further details of the methodology for determining the capacity credits for Intermittent Generators see the rule change report and related documents for Rule Change RC_2010_25 available at www.imowa.com.au/RC_2010_25.

53 Facility output is measured in MWs and the parameters K and U are measured in units of MW^{-1}. K is initially set at K = 0.003 and U is initially set at U=0.635/(average facility output during peaks).

54 Note: Surplus capacity = Capacity (Scheduled generation + intermittent generation) *less* demand,
= Scheduled generation *less* LSG.

So as to capture all demand, the Total Demand used in the calculation of LSG is the sum of total sent out generation of all facilities (Operational Load) plus the load that has been curtailed. When there is no Curtailed Load, LSG measures (as the name implies) the load that will be met by Scheduled Generators.

The Curtailed Load is generally zero but may be significant at peak times. This is the sum of:

- ***DSP Reduction,*** the total quantity by which all Demand Side Programmes reduced their consumption in response to a Dispatch Instruction;
- ***Interruptible Reduction***, the total quantity by which all Interruptible Loads reduced their consumption in accordance with the terms of an Ancillary Service Contract; and
- ***Involuntary Reduction***, the total quantity of energy not served due to involuntary load shedding (manual and automatic).

Total Intermittent Generator Output

The Total Intermittent Generator output is generally just the total metered sent-out generation of intermittent generation facilities; however there are some circumstances where this will not be the case.

First, adjustments are made to sent-out generation for the impact of Consequential Outages and Dispatch Instructions on impacted Facilities. When Consequential Outages occur or a facility's output is reduced in response to a Dispatch Instruction, an estimate of the output of the facility is used so as to ensure that the true ability of the facility to produce output during that particular Trading Interval is reflected in the Capacity Credit calculation.

Second, as new Facilities will typically have not been operational during the five-year period, an estimate of the sent-out generation for these Facilities, as prepared by an expert consultant accredited by the IMO, is used prior to the facility being operational.

The Capacity Credit formula is designed so that the estimated data for new facilities does not affect the LSG calculation relevant for other Intermittent Generator Facilities. To achieve this, the Market Rules (refer to Appendix 9 of the Market Rules for further details) make use of two LSG measures based on when a facility becomes fully operational:[56]

- An Existing Facility LSG (EFLSG) used for all candidate facilities that have been fully operational since the start of the period for which LSG is being calculated; and
- A New Facility LSG (NFLSG) that is calculated separately for each new or upgraded candidate facility that does not have five years' worth of actual metered output for the facility under the configuration for which the facility is being certified (known as new candidate facilities).

The New Facility LSG calculated for a Trading Interval is identical to the Existing Facility LSG from the time the new candidate facility is in full operation. Prior to the full operational

---

[55] This will be actual output from the operational date and estimated output prior to the full operational date. An adjustment is made to actual output for consequential outages.

[56] The full operation date for a Facility is a date provided by the facility in its Reserve Capacity certification.

date the two measures differ simply in that Existing Facility LSG is based on actual output of all Intermittent Generators applying for certification, whereas New Facility LSG also uses the estimated output of the new candidate facility for Trading Intervals prior to its full operation date.[57]

Existing Facility LSG will be used in most circumstances. In effect, Existing Facility LSG will be used for all Trading Intervals from when a facility is in full operation.

For each new candidate facility, a New Facility LSG will be calculated based on Existing Facility LSG and the actual and estimated data up until the facility's operational date.

At times a facility may be upgraded to expand capacity. As the Capacity Credit formula uses historical metered output, the additional output of the upgraded facility will, over time, result in an increase in the Capacity Credit valuation. To enable a facility to receive the benefit of an upgrade once operational, a facility will apply for certification of the upgraded amount through the IMO's normal certification processes.

For the purposes of determining the level of Capacity Credits to be assigned to an upgraded facility, the upgraded facility will be treated as a new candidate facility so that the estimated output of the upgrade does not impact on the valuations assigned to existing Facilities. This is achieved, in effect, by the relevant Market Participant advising the new date on which the facility became fully operational under the configuration for which certification was sought (i.e. the upgraded configuration). Where a full operation date is not provided, the IMO will assume that the facility is not yet operational and therefore treat the facility as a new candidate facility by default.

As Existing Facility LSG is based on actual metered output, it is not affected by the decision of a facility to undertake an upgrade (and seek certification for the upgraded amount). As always, the New Facility LSG will equal the Existing Facility LSG adjusted for the difference between estimated and actual output of the facility.

As with other new facilities, the estimated output of the upgraded facility will be based on an assessment of an accredited expert. Although there is no explicit requirement for such, the IMO expects that the estimated output would consist of the actual output of the facility's existing installation generation plus an estimate of the output of the upgraded generation.

57 For a new candidate facility:

NFLSG = EFLSG less actual output of the new facility + estimated output of the new facility.

From the full operational date, NFLSG will equal EFLSG because from this time actual output will equal estimated output.