AEMO
AUSTRALIAN ENERGY MARKET OPERATOR

Semi-Scheduled Generation Dispatch Self-Forecast - Assessment Procedure Consultation

November 2018

Draft Determination

Important notice

PURPOSE

On 11 July 2018, AEMO invited written submissions from NEM Semi-Scheduled Generators and Interested Parties on the draft procedure "Semi-Scheduled Generation Self-Forecast – Assessment Procedure". The procedure describes how AEMO proposes to assess the suitability of self-forecasts for use in dispatch.

AEMO has considered all feedback received and revised its draft procedure.

DISCLAIMER

This document or the information in it may be subsequently updated or amended. This document does not constitute legal or business advice, and should not be relied on as a substitute for obtaining detailed advice about the National Electricity Law, the National Electricity Rules, or any other applicable laws, procedures or policies. AEMO has made every effort to ensure the quality of the information in this document but cannot guarantee its accuracy or completeness.

Accordingly, to the maximum extent permitted by law, AEMO and its officers, employees and consultants involved in the preparation of this document:

- make no representation or warranty, express or implied, as to the currency, accuracy, reliability or completeness of the information in this document; and
- are not liable (whether by reason of negligence or otherwise) for any statements or representations in this document, or any omissions from it, or for any use or reliance on the information in it.

Contents

Definitions, acronyms, and abbreviations

Terms defined in the National Electricity Rules have the same meanings when used in this document. Acronyms used in this document are explained in the table below.

Term	Definition
ASEFS	Australian Solar Energy Forecasting System
AWEFS	Australian Wind Energy Forecasting System
DUID	Dispatchable Unit Identifier
ECM	Energy Conversion Model for AWEFS and ASEFS
MAE	Mean Absolute Error
NEMDE	National Electricity Market Dispatch Engine
NER	National Electricity Rules
POE	Probability of Exceedance
RMSE	Root Mean Squared Error
SCADA	Supervisory Control and Data Acquisition
SF	Participant's 5-minute ahead Dispatch Self-Forecast of unconstrained intermittent generation from a semi-scheduled generating unit
UIGF	Unconstrained Intermittent Generation Forecast

1. Introduction

Currently, AEMO uses the Australian Wind Energy Forecasting System (AWEFS) and Australian Solar Energy Forecasting System (ASEFS) to provide 5-minute ahead unconstrained intermittent generation forecasts (UIGF) for solar and wind generating units, respectively, for use in 5-minute dispatch.

From January 2019, AEMO may use the 5-minute ahead UIGF provided by the participant (self-forecast) as an alternative to the AWEFS/ASEFS UIGF, subject to satisfactory performance assessed by AEMO on an initial and ongoing basis.

On 11 July 2018, AEMO released a draft "Semi-Scheduled Generation Self-Forecast – Assessment Procedure"[1] (the draft procedure). The procedure describes how AEMO proposes to assess the suitability of self-forecasts for use in dispatch.

In preparing the draft procedure, AEMO considered feedback received following the self-forecasting technologies workshop on Thursday 15 February 2018 (the previous workshop)[2].

[1] See https://www.aemo.com.au/-/media/Files/Stakeholder_Consultation/Working_Groups/Other_Meetings/5-Minute-Forecast/Semi-Scheduled-Generation-Self-Forecast---Assessment-Procedure.pdf.

[2] See https://www.aemo.com.au/-/media/Files/Stakeholder_Consultation/Working_Groups/Other_Meetings/5-Minute-Forecast/Feedback---Market-Participant-Self-Forecasting-Workshop.pdf.

2. Summary of consultation

On 11 July 2018, AEMO invited written submissions from NEM Semi-Scheduled Generators and Interested Parties on the draft procedure.

Submissions closed on Friday 10 August 2018, with five respondents:

- Infigen Energy (Infigen).
- Meridian Energy (Meridian).
- Neoen Australia (NEOEN).
- Proa Analytics (PROA).
- Tilt Renewables (Tilt).

AEMO also received verbal feedback from attendees to the AWEFS/ASEFS Working Group forum on Friday 20 July 2018. AEMO subsequently held telephone meetings with each party to clarify their feedback.

AEMO wishes to thank all respondents for providing their valued feedback. AEMO has considered all feedback received and revised its draft procedure.

3. Next steps

This document summarises stakeholder feedback on the draft procedure (including responses to specific questions posed by AEMO within the draft procedure) and AEMO's response to that feedback.

This has resulted in changes to the initial draft procedure, herein referred to as the "revised procedure".

AEMO has published the revised procedure here: https://www.aemo.com.au/Stakeholder-Consultation/Industry-forums-and-working-groups/Other-meetings/Market-Participant-5-Minute-Self-Forecast

AEMO is now seeking written feedback on the revised procedure from NEM Scheduled and Semi-Scheduled Generators and Interested Parties, by Wednesday 28 November 2018.

Based on feedback, AEMO aims to publish the final Semi-Scheduled Generation Dispatch Self-Forecast Assessment Procedure in mid-December 2018.

4. Stakeholder feedback

4.1 Assessment process

4.1.1 Draft proposal

The draft procedure proposed a two-stage process to assess SFs for use in dispatch:

1. Initial assessment process:
   - AEMO assesses the reliability of the SF process.
   - AEMO assesses the relative accuracy of the SF compared to AWEFS/ASEFS forecast over the initial assessment window.
2. Ongoing assessment process:
   - AEMO assesses each week the relative accuracy of the SF compared to AWEFS/ASEFS forecast over the ongoing assessment window
   - AEMO control room monitors the SF for gross errors to the extent they cause, or could cause, a threat to system security (as currently done for AWEFS/ASEFS)

AEMO repeats the initial assessment process weekly until the SF passes both the reliability and accuracy tests, whereby the SF is eligible for AEMO acceptance by unsuppressing the SF for use in dispatch. If the SF fails the weekly assessment, or the control room determines the SF is grossly in error causing system security issues, AEMO rejects (suppresses) the SF and it is not used in dispatch until it passes the next weekly assessment.

Note that a participant can choose to submit multiple SFs for a farm[3] that reflect different forecasting models, however only the unsuppressed SF with the highest forecast priority[4] prior to gate closure may be used in dispatch. A participant does not need to suppress all the SFs from existing models before submitting a SF from a new model for assessment and use in dispatch. Instead, they can choose to demote the forecast priorities of all existing SF models below that of the new SF model.

4.1.2 Feedback

Infigen supported the draft proposal for a rolling 12-week initial assessment and an ongoing assessment comprising a 4-week assessment window to suppress a SF and a 1-week assessment window to unsuppress a SF, adding the process should be automated in the long term.

However, Infigen offered a preferred approach – an automated rolling 60-minute evaluation:

- Automatic switching between the SF and AWEFS/ASEFS forecasts depending on which has been producing more accurate results.
- A switching hysteresis, to prevent forecasts switching too frequently.

Infigen noted their preferred approach gave less weighting to which forecast was more accurate over the longer term, because the purpose of 5-minute self-forecasting was to get the most accurate forecast at any point in time.

[3] In this document, "farm" refers to the relevant semi-scheduled generating unit

[4] For example, in a SF submission an unsuppressed SF assigned a "forecast priority" value = 2 is used in dispatch ahead of an unsuppressed SF with "forecast priority" value = 1

Infigen was concerned that AEMO would treat a grossly inaccurate SF differently if it was proactively suppressed by the participant rather than by the AEMO control room. Infigen added that AEMO should be encouraging self-awareness of an inaccurate SF.

NEOEN asked whether AEMO would assess any SFs that were not submitted to Pre-Production[5], considering that the 12-week initial assessment would likely only capture the performance of the early forecasting models, with other forecasting models expected to be progressively implemented over time.

PROA considered that AEMO should accelerate the initial assessment process for a new SF from a proven third party forecast provider if it is based on the same forecasting model as a SF for an existing farm that AEMO has already accepted for use in dispatch.

PROA noted the AWEFS/ASEFS forecast for a new farm only requires four weeks to implement and suggested the initial assessment window for a SF should also be four weeks, rather than the proposed 12 weeks.

4.1.3 AEMO response

AEMO disagrees with Infigen's alternative proposal to replace the proposed ongoing weekly assessments with an automated 60-minute rolling assessment, because this would result in forecasts switching too frequently depending on prevailing conditions. AEMO notes that Infigen did not suggest a design to avoid this occurring.

AEMO will consider Infigen's suggestion of an automated rolling assessment as part of a future review of the process. As part of this review, AEMO will consider the value of implementing a very short-term online assessment to assist the AEMO control room in detecting grossly inaccurate forecasts.

As NEOEN observes, AEMO's initial assessment would only be based on the overall performance of the SF over that period, rather than the underlying performance of the individual forecasting models. However, AEMO expects that the forecast provider would perform their own assessment of each new forecasting model within their own systems (and potentially in AEMO's Pre-Production environment) before deciding to unsuppress their model forecast within AEMO's Production environment for AEMO's assessment and potential use in dispatch. AEMO has revised the draft procedure to clarify this expectation.

AEMO supports Infigen's suggestion that the participant should be encouraged to pro-actively suppress a grossly inaccurate SF before the AEMO control room suppresses it. Under the draft procedure, if the AEMO control room suppresses a grossly inaccurate SF it remains suppressed until AEMO's next weekly assessment, whereas if the participant suppresses that SF (rather than AEMO) and this resolves the issue then the participant can choose when to unsuppress that SF once they have addressed the underlying issue.

Note that AEMO proposes to revise the ongoing monitoring process in the draft procedure, to make it clear that the AEMO control room can suppress a Dispatch forecast (either SF or AWEFS/ASEFS forecast) if it is grossly inaccurate and is causing, or could cause, market or power system security issues, rather than only power system security issues. This aligns with AEMO's current practice for suppressing an AWEFS/ASEFS forecast.

AEMO does not support PROA's view that AEMO should accelerate the initial assessment of a SF for a new farm that uses a "type-tested" forecasting model, because:

- AEMO is not aware of the confidential details of each individual forecasting model, its inputs and the supporting algorithms. Hence AEMO must treat each as a unique "black box" for assessment purposes.
- A SF is potentially based on multiple forecasting models developed and managed by potentially multiple third-party providers, whereas the AWEFS/ASEFS forecasting models are developed and managed by a single vendor contracted to AEMO. Hence AEMO is more familiar with the underlying design of the AWEFS/ASEFS forecasting model and its inputs and has greater confidence in its expected performance.
- AEMO is not only assessing the overall performance of the forecasting model and supporting algorithms, but also the quality and reliability of the unique inputs to that forecast model and their impact on

[5] AEMO interprets this as referring to Production.

performance. Note that AEMO does not request its AWEFS/ASEFS vendor to commence development of a forecasting model for a new farm until the critical inputs to that model are good quality and reliable, a process that takes up to four weeks post-energisation.

AEMO has revised the draft procedure to add a new "Participant Dispatch Self-Forecast" section.

This section will provide a definition for the Participant Dispatch Self-Forecast (SF) and clarify the roles and responsibilities of AEMO and the participant, as follows:

- AEMO may use the SF in dispatch if satisfied that it meets the reliability and accuracy requirements described in this procedure.
- AEMO expects the participant would have already tested any new forecasting model underlying their SF before submitting it to AEMO for assessment and use in dispatch.
- AEMO expects each SF submission to include the unique Model ID assigned by AEMO that is associated with each underlying forecasting model. However, AEMO will only assess the overall SF performance for use in dispatch, and not differentiate between different model forecasts.
- AEMO only intends to interact with the relevant participant in respect of its SF, and not via any of their third-party forecast providers.
- The relevant participant is responsible for coordinating the provision of SF from their forecast providers.

AEMO has revised the assessment process in the draft procedure, as follows:

- AEMO will only assess SF data collected within its Production environment.
- Participant will advise the earliest dispatch interval after which AEMO should include SF in its initial assessment.

AEMO has revised section 6. "Ongoing Monitoring for System Security" in the draft procedure, to:

- Rename section to "Ongoing Monitoring for Gross Errors" and insert the following paragraph: "The AEMO control room monitors the SF and AWEFS/ASEFS Dispatch forecasts for gross errors and may suppress the relevant forecast if their use in dispatch is causing, or could cause, market or power system security issues."
- Make it clear that AEMO may suppress a SF until the next weekly assessment, unless the relevant participant has already chosen to do so.

4.2 Assessment pre-requisites

4.2.1 Draft proposal

The draft procedure proposed that, before AEMO could undertake an initial assessment, an unsuppressed SF must be received for at least 99% of all intervals over the 12-week assessment window.

4.2.2 Feedback

Infigen and Meridian disagreed with the draft proposal and considered the 99% minimum availability requirement prior to undertaking an initial assessment was too stringent.

Infigen considered the minimum availability requirement was too stringent because it:

- Potentially disadvantaged participants who actively suppress their SFs over periods where AWEFS/ASEFS was likely to produce a more accurate forecast.
- Did not make a reasonable allowance for planned forecasting system downtime (a single days' worth of outage would preclude an initial assessment).
- Infigen suggested two alternatives:
  - Lower the minimum availability requirement to 95% or less, or
  - Re-define the initial assessment window as 24,000 DI's with unsuppressed SF (excluding semi-dispatch intervals), regardless of how long this takes to accumulate.

Meridian considered the 99% minimum availability requirement was unnecessarily limiting and should not require a continuous period of 12 weeks, but "a sufficient number of data periods that are representative of plant operation". For example, during commissioning a participant might suppress its SF for more than 20% of the time.

Meridian proposed that the initial assessment window cover a minimum of eight weeks of "representative operation" and apply the minimum availability requirement to any continuous four-week period within that window.

4.2.3 AEMO response

Minimum DIs for SF reliability

AEMO agrees that a forecast provider is likely to initially suppress the SF from a new forecasting model, or actively suppress the SF for an existing model under some conditions where they consider the AWEFS/ASEFS forecast performs better.

The minimum availability requirement is intended to be a measure of how reliably the participant forecasting system delivers a valid SF submission every dispatch interval, regardless of whether the participant has suppressed that SF.

AEMO agrees with Infigen and Meridian that the 99% minimum availability criterion is too stringent and should be relaxed to 95% as proposed by Infigen.

AEMO agrees with Infigen that the purpose is to assess the overall reliability of the SF submission process, not whether those SFs are also unsuppressed. AEMO will clarify that, although a new SF is initially suppressed by AEMO, if the SF remains unsuppressed by the participant (and meets the gate closure requirement) then the SF is counted towards meeting the minimum availability requirement and is used in the initial assessment.

AEMO disagrees with Meridian's proposal to apply the minimum availability requirement over any continuous four-week period within that window, which equates to the provision of a SF for only 50% of the time.

AEMO agrees with Meridian's proposal to reduce the initial window for the initial assessment process from 12 to eight weeks, although AEMO disagrees that this should reflect "representative operation", because this is open to interpretation and would likely require its own guidelines to determine. If the initial performance

assessment fails, AEMO also proposes to allow this initial eight-week window to extend each week by a further week up to a maximum of 16 rolling weeks. Beyond that, the SF assessment window would remain at 16 rolling weeks until the SF passes the initial performance assessment. AEMO considers that the increasing window would ensure there were enough valid samples for a meaningful assessment despite periods of substantial curtailment[6] over the initial eight-week period, while the maximum 16-week window would avoid using a SF that was too old to be relevant to the performance assessment.

Minimum DIs for performance assessment

AEMO clarifies that, if a participant has suppressed its SF, AEMO would not use it in dispatch and hence we would not include it in any forecast performance assessment.

AEMO considers that, before it can undertake any performance assessment of a SF (whether initial or ongoing), there must be a minimum number of dispatch intervals over the assessment window for which:

- There is a valid SF received within 70 seconds of gate closure that is not suppressed by the participant, and
- The relevant farm is not dispatched below its UIGF, unless AEMO received a good quality SCADA Possible Power for use as the performance benchmark.

AEMO would include all dispatch intervals that satisfy this test in its performance assessment.

AEMO proposed the second condition in response to Tilt's concerns with using SCADA Actual MW as a performance benchmark during semi-dispatch intervals, particularly where the farm's output was heavily curtailed.

Some respondents proposed that AEMO should conduct an assessment excluding all semi-dispatch intervals. However, if this were taken to the extreme, this could result in AEMO assessing forecast performance and potentially suppressing or unsuppressing the SF based on only a handful of DIs over the assessment window.

Instead, AEMO proposes that if the "Minimum DIs for Performance Assessment" test fails then AEMO would not conduct a SF performance assessment (either initial or ongoing) and would not change the status of the SF.

AEMO has revised the draft procedure, to introduce preliminary tests that must be satisfied before AEMO can undertake a SF performance assessment:

- Minimum DIs for Reliable SF test (initial assessments only):
  - For at least 95% of dispatch intervals over the current assessment window, AEMO received a SF at least 70 seconds prior to the gate closure time for the dispatch interval
- Minimum DIs for SF Performance Assessment test (initial and ongoing assessments):
  - For at least 80% of dispatch intervals over the assessment window the following criteria are satisfied
    - AEMO received an unsuppressed SF at least 70 seconds prior to gate closure time
    - Unit's energy target was greater than or equal to its UIGF (that is, the unit is not constrained off), unless the participant submitted a good quality SCADA Possible Power for the dispatch interval

As part of these changes, AEMO has revised the draft procedure to:

- Relax the minimum availability requirement from 99% to 95%, and
- Include all SFs that meet the gate closure requirement, whether suppressed or unsuppressed

[6] Curtailment refers to semi-dispatch intervals where the farm's energy target (dispatch level) is below its dispatch UIGF

4.3 Performance metric

4.3.1 Previous workshop

In the previous workshop, there was broad agreement to using Mean Absolute Error (MAE) as the performance metric, although some suggested using MAE as a normalised percentage of installed capacity. There were also suggestions to use the Root Mean Squared Error (RMSE).

4.3.2 Draft proposal

The draft procedure proposed to measure the accuracy of the SF and AWEFS/ASEFS forecast using the MAE and RMSE of each forecast against either SCADA Actual MW or SCADA Possible Power.

For each dispatch interval, the assessment is based on the latest, highest priority unsuppressed SF received at least 70 seconds before the gate closure for that interval. The AWEFS/ASEFS forecasts already meet that gate closure requirement.

For dispatch intervals where the AWEFS/ASEFS forecast is unavailable or suppressed by AEMO[7], SF accuracy is measured against the accuracy of the default persistence forecast which is the SCADA Actual MW.

AEMO posed these specific questions in this area[8]:

1. What other validations should AEMO perform on the sample data?
2. Should AEMO use a probability of exceedance (POE) error measure (in place of the RMSE measure) to indicate forecast accuracy for large ramping events? If so, what POE would be appropriate? The 99% POE error[9] ?
3. Are there other performance metrics that would be appropriate? Specifically, what metrics are useful for quantifying ramping events? Should there be a time-of-day specific error measure?

4.3.3 Feedback

Infigen, NEOEN, and PROA supported the draft proposal to use the MAE and RMSE performance metrics. NEOEN added that the RMSE metric could be complemented by an error bias check, and that a POE error metric might be more useful for solar farms.

Meridian and Tilt considered the performance metric should place greater weight on large forecast errors. Meridian proposed either using the RMSE metric alone, or to combine it with a MAE metric with a higher error exceedance threshold. Tilt proposed to combine the MAE metric (with a small error exceedance threshold) with a POE error metric rather than a RMSE metric.

Meridian and NEOEN considered all metrics should be reviewed and refined over time with more self-forecasting experience.

Infigen considered the MAE and RMSE metrics were sufficient because the assessment process should not be too complicated at its inception.

Meridian suggested to only use the RMSE metric because it is more sensitive to larger (and therefore more relevant) errors, adding the MAE metric only makes the test less relevant and does not improve the ability to assess relevant forecasting accuracy. Further, if AEMO were to include the MAE metric, then it must have a wider error exceedance threshold than the RMSE otherwise the MAE metric would overwhelm the benefits of using the RMSE metric.

Meridian added there was insufficient evidence that the RMSE metric was better than a POE error metric and suggested that AEMO monitor this to see how the different measures perform once the systems have been in place for some time. They suggested there may be value in AEMO reporting those periods where NEMDE

[7] Typically, this occurs during the initial commissioning of a new farm, or if the forecast is grossly inaccurate during weather events.

[8] Appearing as questions 1, 4, and 5 respectively in the draft procedure.

[9] This is a typo. The question should have read: "The 1% POE error".

produced dispatch outcomes likely to impact forecasting performance (for example, semi-dispatch caps with large variations, or significant price changes likely to induce responses).

NEOEN suggested that AEMO treat the proposed metrics as guidelines not rules which are subject to refinement over time once SF data starts to flow. They added that an additional bias check might be useful, with a slightly higher RMSE but low bias preferred over a lower RMSE but high bias.

NEOEN considered the POE error metric might be more valuable for solar farms with its larger ramps.

PROA considered the draft proposal of using MAE and RMSE metrics was sufficient, and that a POE error metric approach was not required.

Tilt considered that, although a comparison on average performance (MAE/RMSE) indicates the overall accuracy of a SF under "ordinary conditions", it did not completely assess the power system security benefits that arise from improved forecasting of outlier wind events such as:

- High wind speed cut-out.
- High temperature cut-out / de-rating.
- Wind sector management.
- Fast moving weather patterns (clouds, winds, temperature).

Tilt added that using a simple average performance metric across all periods would "swamp" forecast performance during relatively infrequent outlier events when forecast improvements provide greatest benefit to power system security. In many cases, outlier events such as wind direction changes can affect multiple sites, further increasing the impact of poor forecasts on power system security.

Tilt suggested a separate assessment of error outliers to accurately assess and reward the benefits to power system security, with an error threshold selected using a histogram of output volatility during unconstrained dispatch intervals. Tilt did not suggest how to select the error threshold from that histogram.

4.3.4 AEMO response

AEMO agrees with Infigen, NEOEN, and PROA that the MAE and RMSE performance metrics should be used as the basis for accepting SF for use in dispatch.

AEMO disagrees with Tilt that the RMSE error metric should be replaced by a POE error metric, because there is insufficient evidence that the RMSE metric would under-weight the larger errors.

AEMO disagrees with Meridian and Tilt that a SF should be considered as performing better than the AWEFS/ASEF forecast if its average error exceeds AWEFS/ASEF by a small threshold when assessed over a sufficiently long window, because this means that AEMO would potentially be using a worse forecast in dispatch.

However, as suggested by Meridian and NEOEN, AEMO will work with stakeholders to review the effectiveness of the SF assessment process (including metrics, benchmarks, thresholds, and exclusions) after gaining sufficient experience with self-forecasting.

As part of that review, AEMO would consider whether to integrate bias and POE error into the performance metrics, as suggested by NEOEN and Tilt.

AEMO will use the proposed MAE and RMSE metrics, and not revise its draft procedure.

However, given the different views on the appropriate metric, AEMO will review the effectiveness of the SF assessment process (including metrics, benchmarks, thresholds and exclusions) after gaining sufficient experience with the use of self-forecasts in dispatch.

4.4 Performance benchmark

4.4.1 Previous workshop

In the previous workshop, stakeholders held different views on the appropriate performance benchmark, with some suggesting comparing forecasts against the persistence (do nothing) forecast based on SCADA Actual MW.

Several stakeholders disagreed with the use of SCADA Possible Power as a benchmark during semi-dispatch intervals, noting that it is defined at the vendor/OEM's discretion making it inconsistent between facilities and potentially an unreliable benchmark. Some of those stakeholders suggested excluding all semi-dispatch intervals from performance assessment

AEMO disagreed with excluding all semi-dispatch intervals from the performance assessment, as the absence of such monitoring for unconstrained forecasts would unduly expose AEMO to the risk of scheduling errors.

4.4.2 Draft proposal

The draft procedure proposed to measure the accuracy of the SF and AWEFS/ASEFS forecast using the MAE and RMSE of each forecast compared to:

- Maximum of (0, SCADA Actual MW) outside semi-dispatch intervals.
- Maximum of (0, SCADA Actual MW, SCADA Possible Power) during semi-dispatch intervals.

If SCADA Possible Power is unavailable or not good quality, forecasts are compared to SCADA Actual MW.

If SCADA Actual MW was unavailable or not good quality, forecasts are compared to the energy target from the previous dispatch interval, which is always good quality.

AEMO will define the optional SCADA Possible Power in the Energy Conversion Model (ECM) as the real-time now-cast estimate of unconstrained intermittent generation.

4.4.3 Feedback

Infigen and Tilt opposed the draft proposal to use SCADA Possible Power as a forecast performance benchmark during semi-dispatch intervals.

Infigen also opposed using SCADA Actual MW as a forecast performance benchmark during semi-dispatch intervals, whereas Tilt were opposed to using SCADA Actual MW only during those semi-dispatch intervals where the farm output was deeply curtailed.

Infigen considered all assessments should exclude all semi-dispatch intervals

Infigen considered that using SCADA Possible Power as a forecast benchmark would encourage inaccurate forecasting particularly when farm output is heavily curtailed by more than, say, 10%. This is because Possible Power is assumed to lie between the AWEFS/ASEFS forecast and Actual MW, hence there is an incentive to provide a SF marginally lower than the AWEFS/ASEFS forecast to ensure the SF is considered more "accurate".

Infigen asked whether the SCADA Possible Power, if used as a forecast benchmark, would be taken at around the time that the SF applied.

Infigen further considered that using SCADA Actual MW as a forecast benchmark would encourage extremely inaccurate forecasting because it would incentivize the generator to provide a SF as close to SCADA Actual MW as possible to be deemed more 'accurate' than the AWEFS/ASEF forecast.

Tilt strongly opposed using SCADA Possible Power as a forecast performance benchmark during semi-dispatch intervals, because it is:

- An algorithmically-derived estimate with no guarantee on its accuracy, particularly during outlier events, and suffers the same modelling issues as any other forecast.
- Manufacturer- and model-specific, and not within the easy reach of the participant to improve.

- Not guaranteed to account for high-wind cut-out, wind-direction effects, or wind-sector management, all the outliers a SF would strive to improve.

Tilt added that it did not make sense to dedicate manufacturers' resources in developing a Possible Power algorithm for use as a performance benchmark against a similar algorithm that would already be developed for producing the SF. Tilt instead suggested that the now-cast estimate from the SF algorithm itself would be a better performance benchmark than the Possible Power estimate developed by the manufacturer.

Tilt was also concerned with the draft proposal using SCADA Actual MW as the benchmark during semi-dispatch intervals if the SCADA Possible Power was unavailable or bad quality, because in a deep-curtailment situation this would effectively compare an unconstrained forecast to a constrained forecast, a comparison of no value.

Tilt proposed an alternative approach for assessing forecast performance against SCADA Actual MW during semi-dispatch intervals which depends on the level of curtailment:

- In pre-operation and operation with deep curtailment [10]:
  - AEMO visually compares the SF versus AWEFS, to look for general agreement and outliers
- In pre-operation (AWEFS used in dispatch), but farm not curtailed:
  - AEMO only includes outlier errors in assessment:
    - Bad outliers (SF error is far higher than AWEFS error)
    - Good outliers (SF error is far lower than AWEFS error) – for example, when wind direction has an impact
- In operation (SF used in dispatch), but farm not curtailed:
  - AEMO assesses SF error against a threshold using an average and outlier assessment, with thresholds derived from a histogram of output volatility during unconstrained dispatch intervals.
  - This would be a one-sided test (SF error is far higher than AWEFS error), because a SF that is too low is hidden by the farm's output being limited to the incorrectly low SF. It would be instructive in quantifying assessment thresholds to look at the existing gap between the actual MW and the AWEFS forecast in this situation.

4.4.4 AEMO response

AEMO agrees with Tilt that the SCADA Actual MW should not be used as a forecast performance benchmark during semi-dispatch intervals when the farm output is curtailed.

However, AEMO still sees value in using the SCADA Possible Power as the performance benchmark during such intervals and disagrees with Tilt and Infigen that those intervals should be excluded from the assessment.

AEMO does not accept Infigen's assumption that Possible Power "lies between actual park MW and the forecasts", or how the generator could reliably create a SF that is marginally lower than the AWEFS/ASEFS forecast during semi-dispatch intervals. If the generator were to systematically under-forecast to ensure continued use in dispatch this would result in over-curtailment, an undesirable outcome and not sustainable.

AEMO notes that a generator might be incentivised to create a SF marginally higher than its Possible Power and then over-estimate both during semi-dispatch intervals so that the SF is always assessed as more "accurate" than the AWEFS/ASEFS forecast. However, this would result in AEMO using a SF that is worse than AWEFS/ASEFS during semi-dispatch intervals, particularly where the farm is not curtailed and is generating well below its SF, potentially causing market or power system security issues and resulting in the AEMO control room suppressing the SF.

[10] AEMO understands "pre-operation" and "operation" to mean intervals where AWEFS/ASEFS and SF is used as dispatch UIGF, respectively.

To Infigen's question, AEMO confirms that the SCADA Possible Power (and the SCADA Actual MW) used in forecast performance assessments are taken at the end of each 5-minute dispatch interval, the same time as the SF (and AWEFS/ASEF forecast) applies.

AEMO accepts Tilt's concerns with the accuracy of the Possible Power algorithm developed by the manufacturer and agrees with Tilt that a viable alternative might be to develop the now-cast estimate from the SF algorithm itself. AEMO has no preference for who implements the Possible Power algorithm or how it is implemented.

AEMO proposes to only exclude curtailed dispatch intervals[11] from its performance assessments where the relevant farm's SCADA Possible Power is either not provided or not good quality in that dispatch interval.

AEMO has revised the performance benchmarks in initial and ongoing assessments of the draft procedure, as follows:

- Use SCADA Actual MW as a performance benchmark for dispatch intervals where the unit energy target is not less than the dispatch UIGF (that is, not curtailed)
- For all other dispatch intervals, either:
  - Use the SCADA Possible Power as the benchmark in the performance assessment if it is available and good quality; or
  - Exclude the dispatch interval from the performance assessment if the SCADA Possible Power is not available or not good quality.

    If more than 80% of dispatch intervals are excluded over an assessment window, then the "Minimum DIs for Performance Assessment" test fails, and the assessment is not conducted

[11] Semi-dispatch intervals where the farm's energy target (dispatch level) is below its dispatch UIGF, whether the UIGF is set by the SF or the AWEFS/ASEFS forecast

4.5 Performance threshold

4.5.1 Draft proposal

The draft procedure proposed that AEMO accept a SF for use in dispatch if the SF error does not exceed the AWEFS/ASEFS forecast error, on both an MAE and RMSE basis. This applied to the initial assessment and (at least initially) for the ongoing assessment.

For the ongoing assessment, the threshold by which the SF error must be lower than the AWEFS/ASEFS forecast error was initially zero, but configurable. AEMO's intention was to increase this threshold over time as SF accuracy improved relative to the AWEFS/ASEFS forecast.

AEMO posed two specific questions in this area[12]:

1. What relative performance thresholds for the ongoing MAE/RMSE assessments are reasonable? Should the SF performance be better than the AWEFS/ASEFS performance, and how much better?
2. Should there be a maximum acceptable MAE/RMSE for the SF performance, even if the SF has a relatively lower MAE/RMSE than AWEFS/ASEFS?

4.5.2 Feedback

Infigen and PROA supported the draft proposal that AEMO accept a SF if its error did not exceed the AWEFS/ASEFS forecast error, on both an MAE and RMSE basis. Meridian and Tilt considered the SF should still be accepted if its error exceeded the AWEFS/ASEFS forecast error by less than a pre-defined threshold. NEON and Tilt suggested a further outlier error check before rejecting the SF.

Infigen and Tilt disagreed with the implied requirement in the ongoing assessment that AEMO reject a SF if it does not outperform the AWEFS/ASEFS forecast by more than a pre-defined threshold.

Infigen considered that introducing a maximum MAE and RMSE would needlessly penalise participants with a less predictable location even though these were sites with the most to gain. Infigen added that some relative improvement in SF over AWEFS/ASEFS should always be preferable to no improvement in absolute terms.

Meridian considered that the relative performance threshold should balance the ability to pursue novel forecasting approaches while meeting the requirements of power system security and an efficient market. For example, if the SF for a solar farm performed slightly worse than ASEFS on many clear-sky days and was much better on fewer overcast days so on average the SF was slightly worse, this would be the wrong outcome.

To this end, Meridian suggested that AEMO should assess the SF over a long-, medium-, and short-term window and accept the SF if its error does not exceed the AWEFS/ASEFS forecast error by more than a pre-defined threshold[13], with increasing thresholds for shorter windows.

Meridian proposed:

- A 12-week SF assessment with zero relative error exceedance.
- A four-week SF assessment with a small relative error exceedance.
- A one-week SF assessment with a large relative error exceedance.

NEOEN considered that the SF MAE could exceed the AWEFS/ASEFS forecast MAE in extreme circumstances, such as when a farm trips and its SF is 0.1 MW higher than actual. However, consistent large errors, or occasional extreme errors were potentially damaging and should result in the SF being suppressed.

NEOEN suggested that if the SF MAE check failed, then this should be qualified by further checks for consistent large errors or occasional extreme errors rather than immediate SF suppression – at least until a rule can be determined that worked under all scenarios. NEOEN suggested further checks, such as the 90th

[12] Appearing as questions 2 and 3 respectively in the draft procedure.

[13] Meridian did not suggest values for these thresholds or how they would be defined

percentile error (if SF error is higher it suggests consistent large errors) or the ratio of SF to AWEFS/ASEFS MAE (1.01 is not so bad, 1.8 is probably bad).

NEOEN considered the RMSE is unlikely to be affected by a few extremes, so the suppression of SF based on failing the RMSE check is probably fine.

Tilt had strong concerns about using a simple "less than or equal to" comparison of SF versus AWEFS/ASEFS forecast because this was inconsistent with the SF error being "materially worse" and it risks rejecting a SF that would deliver greater power system security benefits overall. They added that a participant has a clear incentive to ensure the SF is as accurate as possible through the causer-pays process.

Tilt instead suggested that AEMO assess a SF as more accurate than the AWEFS/ASEFS forecast if the SF error does not exceed the AWEFS/ASEFS forecast error by more than 1 MW, which is "well within the noise in the current dispatch process, where generators aim to be within 6 MW of their dispatch levels".

In other words, the SF passes the forecast performance assessments (initial and ongoing) if:

$$\boldsymbol{MAE_{SF} \leq MAE_{AWEFS_{ASEFS}} + 1}$$

and

$$\boldsymbol{RMSE_{SF} \leq RMSE_{AWEFS_{ASEFS}} + 1}$$

Tilt was also concerned that section 3.2 of the draft proposal appears to say that the SF must outperform the AWEFS/ASEFS forecast by some configurable threshold to remain unsuppressed and continue to be used in dispatch. Tilt added there would be a detriment to power system security in reverting to a lower-performing AWEFS/ASEFS forecast if the SF did not continue to improve its performance over time relative to AWEFS/ASEFS.

4.5.3 AEMO response

AEMO agrees with Infigen and Tilt that a SF does not need to outperform the AWEFS/ASEFS forecast in the ongoing assessments to remain unsuppressed and continue to be used in dispatch. AEMO revised the draft procedure to remove the configurable threshold ($X_{ongoing}$, $Y_{ongoing}$) from the ongoing performance assessment.

AEMO disagrees with Meridian and Tilt that a SF should be considered as performing better than the AWEFS/ASEF forecast if its error on average exceeds AWEFS/ASEF by a small threshold because this means that AEMO would potentially be using a worse forecast in dispatch.

AEMO accepts Tilt's view that performance assessments should place greater weight on outlier errors because forecasting improvements would deliver greater benefits power system security. However, AEMO considers this approach risks ignoring other periods where forecasting improvements can deliver more efficient market outcomes.

As noted in section 4.3, AEMO will work with stakeholders to review the effectiveness of the SF assessment process (including metrics, benchmarks, thresholds and exclusions) after gaining sufficient experience with self-forecasting.

AEMO has revised the draft procedure to:

- Remove the configurable threshold (Xongoing, Yongoing) from the ongoing performance assessment test.

AEMO will review of the effectiveness of the SF assessment process (including metrics, benchmarks, thresholds and exclusions) after gaining sufficient experience with the use of self-forecasts in dispatch.

4.6 Assessment window

4.6.1 Previous workshop

In the previous workshop, AEMO considered publishing a 15-minute rolling minute MAE for each of the SF, AWEFS/ASEFS, and "persistence" forecasts.

Some participants considered that such monitoring might be suitable for capturing gross forecast errors. However, participants generally cautioned against assessing relative forecast performance over less than one week, because this would not adequately capture the variations in solar and wind conditions.

Draft Proposal

The draft procedure proposed:

- For the initial assessment, use a 12-week window.
- For the ongoing weekly assessment, use a different window depending on whether the SF is currently suppressed by AEMO:
  - If the SF is not suppressed: use all previous consecutive weeks where the SF passed the ongoing performance criteria, up to a maximum of 4 weeks
  - If the SF is suppressed: use the previous week only, to allow the performance measure to reflect more recent, potentially large improvements in SF performance

4.6.2 Feedback

Infigen considered that the SF would not normally be worse than the AWEFS/ASEFS forecast if assessed over a reasonable time frame, suggesting a minimum of four weeks.

Meridian disagreed with the draft proposal to apply a different window for the ongoing assessments depending on whether the SF was currently suppressed (one week) or unsuppressed (one to four weeks).

Meridian noted this could result in the following perverse outcome:

- Start of Week 1 – SF initially unsuppressed, but becomes substantially inaccurate over next week
- Start of Week 2 – AEMO suppresses SF for failing the previous four-week test (poor week 1).
- Start of Week 3 – AEMO suppresses SF for passing the previous week test (good week 2).
- Start of Week 4 – AEMO suppresses SF for still failing the previous four-week test (poor week 1), even though there was no change in SF accuracy in weeks 2 and 3.

Instead, Meridian proposed that AEMO assess performance over a long (12 weeks), medium (four weeks), and short (one week) window, based on relative RMSE only and excluding periods of substantial forecast similarity. The SF would pass the 12-week assessment if its error did not exceed the AWEFS/ASEFS forecast error (zero error exceedance threshold), with higher thresholds for the shorter windows. AEMO would suppress the SF if the assessment failed over ANY of these windows.

Meridian did not suggest how to determine the thresholds for "substantial forecast similarity" or "relative error exceedance".

Meridian argued its proposal was better because:

- Decisions are based only on periods likely to more appropriately assess forecast accuracy.
- The same approach is used to suppress and unsuppress the SF.
- Both long-term and short-term inaccuracies are captured.
- It should be relatively easy for AEMO to automate.

4.6.3 AEMO response

AEMO agrees with Infigen that the SF would not normally be worse than the AWEFS/ASEFS forecast if assessed over a sufficiently long timeframe.

For the ongoing assessments, AEMO agrees with Meridian that using different assessment windows depending on whether the SF is initially suppressed or unsuppressed would result in unnecessary swapping between the SF and AWEFS/ASEFS forecasts. AEMO also agrees with Meridian that there should be different assessment windows to capture short and long-term inaccuracies.

However, AEMO disagrees with Meridian that a SF should be assessed as performing better than the AWEFS/ASEFS forecast if it has a marginally higher forecast error on average over the medium- or short-term windows. AEMO accepts there can be prolonged periods where the SF and AWEFS/ASEFS forecasts can be substantially similar. However, AEMO considers that the factors that create "noise" in the SF would equally apply to, and affect, the AWEFS/ASEFS forecasts.

AEMO reasonably expects that the accuracy of the SF should be no worse on average than the equivalent AWEFS/ASEFS forecast under most conditions before it is used in dispatch.

For the initial performance assessment, AEMO proposes to change the SF assessment window from a fixed eight-week rolling window to an initial eight-week window which (if the initial performance assessment fails) extends each week by a further week up to a maximum of 16 weeks. Beyond that, the SF assessment window would remain at 16 weeks until the SF passes the initial performance assessment.

An increasing window ensures there will be enough valid samples for an assessment to occur despite where there are periods of substantial curtailment over the initial eight-week period, while the maximum 16-week window avoids using a SF that is too old to be relevant to the performance assessment.

For the ongoing performance assessment, AEMO proposes the following:

- AEMO assesses SF performance over a long- (eight weeks), medium- (four weeks), and short-term (one week) window.
- AEMO will not exclude periods of "substantial forecast similarity" from an assessment.
- The SF passes the performance test if the SF error does not exceed the AWEFS/ASEFS forecast error over the assessment window.
- AEMO accepts the SF for use in dispatch if the SF passes the ongoing performance test for ANY of the three assessment windows.

The benefits of this approach are:

- The SF does not need to pass the ongoing performance test for ALL three assessment windows, as proposed by Meridian.
- AEMO does not need to determine thresholds for "substantial forecast similarity" or "relative error exceedance".
- AEMO can ignore substantial poor SF performance of earlier weeks if SF performance is better in more recent weeks (for example, due to improved SF models or SF suppression techniques).

AEMO has revised the initial assessment window in the draft procedure as follows:

- If the SF fails the preliminary assessments, AEMO does not conduct a SF performance assessment and repeats the preliminary assessments next week extending the current rolling assessment window by one week, up to a maximum of 16 rolling weeks.
- This process repeats each week until the SF passes the initial assessment.

AEMO has revised the ongoing assessment windows in the draft procedure as follows:

- AEMO will assess ongoing performance of the SF over a long (8 weeks), medium (4 weeks) and short-term (1 week) window.
- AEMO will not use different windows depending on whether the SF is initially suppressed or unsuppressed.
- AEMO intends to accept the SF if the SF passes the ongoing performance test for ANY of the three assessment windows.

4.7 Assessment intervals

4.7.1 Draft proposal

The draft procedure proposed to exclude the following dispatch intervals from a performance assessment:

- Dispatch intervals where there is no SF received 70 seconds prior to gate closure that is not suppressed by the participant.
- For solar units only, dispatch interval ending 2105 to 0400 AEST inclusive, to minimise biasing the assessment during night-time periods.

AEMO did not propose to exclude semi-dispatch intervals, although it was noted during those intervals the use of SCADA Actual MW as the benchmark would tend to increase the SF error.

4.7.2 Feedback

Infigen considered all assessments should exclude all semi-dispatch intervals, whereas Tilt suggested to exclude intervals of deep curtailment and Meridian suggested to exclude intervals with small differences between the SF and AWEFS/ASEFS forecasts.

Infigen considered all assessments should exclude all semi-dispatch intervals, because:

- It was impossible to reasonably calculate which forecast would theoretically have been more accurate, and
- There are issues with using SCADA Possible Power or SCADA Actual MW as benchmarks at those times, and to do so would encourage inaccurate forecasting because:
  - If SCADA Actual MW were used, there is an incentive to provide a SF as close to this figure as possible to be deemed 'accurate'. During times when parks are being capped well below their actual possible generation this could encourage extremely inaccurate forecasts.
  - If SCADA Possible Power were used, there is an incentive to provide a SF marginally lower than the AWEFS/ASEFS forecast because it is assumed to lie between Actual MW and those forecasts.

Meridian considered all assessments should exclude periods where forecasts only marginally differ due to measurement or forecasting noise and focus only on periods of significant forecast differences between the SF and AWEFS/ASEFS forecasts. They suggested excluding periods from the assessment where the forecast differences were below (say) 25% of the plant's dispatch non-conformance small error trigger limit STRIGLM[14].

Tilt considered it was unreasonable to bundle capped and uncapped dispatch intervals into one assessment, because the impact on power system security and the ability to assess the forecast accuracy for a capped situation is very different from an uncapped situation.

Tilt considered all assessments should exclude semi-dispatch intervals involving deep output curtailment. For example, the SA system strength constraint routinely curtails wind farms by 30%. When this occurs, the accuracy of the 5-minute ahead forecast was irrelevant because the farm was dispatched far below its forecast and improving forecasts in those periods provided no benefit to power system security. Further, when the system strength constraint bound the forecast makes no difference to the dispatch outcome because dispatch split curtailment between wind farms relative to their offered capacity not their forecast availability.

Conversely, there are other semi-dispatch intervals where the farm is dispatched close to its forecast and the farm is essentially operating uncurtailed. In these situations, the 5-minute ahead forecast is very important to power system security. For example, constraints on Victoria's South Morang F2 transformer involve numerous generators and typically "constrain" low-priced generation such as wind farms at their forecast.

In this situation, if the forecast is too high due to say, high wind speed turbine cut-out then this results in a potentially large dispatch error and risk to power system security risk. The SF would aim to identify and model

[14] The Small Error Trigger threshold is defined in Appendix A "Non-Compliance Calculations and Process Overview" in AEMO's Dispatch System Operating Procedure: https://www.aemo.com.au/-/media/Files/Electricity/NEM/Security_and_Reliability/Power_System_Ops/Procedures/SO_OP_3705---Dispatch.pdf.

these situations, and assessments should give sufficient weighting to performance at these times. However, if the SF is too low, the farm will be able to meet the dispatch level, so there is no dispatch error and no power system security risk but would directly cost the participant's production.

4.7.3 AEMO response

AEMO agrees with Infigen that if SCADA Actual MW is used as the forecast performance benchmark during semi-dispatch intervals, then the forecast that is closest to the actual curtailed output would be deemed the most "accurate". AEMO accepts this might perversely incentivise the generator to provide a constrained SF, although the counter to this is that the generator would deliberately lose dispatch volume and risk running its dispatch down to zero.

As noted in section 4.4.4, AEMO does not accept Infigen's assumption that Possible Power "lies between actual park MW and the forecasts". Further, AEMO does not accept that a generator would be incentivised to create a SF systematically lower than the AWEFS/ASEFS forecast during semi-dispatch intervals simply to ensure its continued use in dispatch because this would result in over-curtailment, an undesirable outcome for the generator.

AEMO disagrees with Meridian that intervals with small differences between the SF and AWEFS/ASEFS forecast performance should be excluded from assessment. As the assessment of SF and AWEFS/ASEFS forecasts are performed on the same sample set of intervals, and the MAE/RMSE metrics are compared between the two, intervals of similar performance are effectively nullified.

AEMO accepts Tilt's view that it is unreasonable to bundle capped and uncapped dispatch intervals into one assessment. However, separating out those intervals into two samples may result in one of those samples being too small for statistical significance.

AEMO notes Tilt's view that different intervals be assessed with different weightings based on the significance of those intervals to system security concerns, to incentivise participants to improve forecasts at those times. However, AEMO considers that participants are already sufficiently incentivised to improve forecasts for all intervals, irrespective of dispatch caps being applied which may not anticipated or correlated with weather.

AEMO has revised the assessment process in the draft procedure to:

- Use SCADA Actual MW as a performance benchmark for dispatch intervals where the unit energy target is not less than the dispatch UIGF (that is, not curtailed).
- For all other dispatch intervals, either:
  - Use the SCADA Possible Power as the benchmark in the performance assessment if it is available and good quality; or
  - Exclude the dispatch interval from the performance assessment if the SCADA Possible Power is not available or not good quality.

    If more than 80% of dispatch intervals are excluded over an assessment window, then the "Minimum DIs for Performance Assessment" test fails, and the assessment is not conducted

4.8 Other feedback

4.8.1 Assessment principles

Meridian suggested the draft procedure should include a section after the Introduction that defines the underlying principles that inform the procedure's purpose, that guides AEMO on the application of the procedure, and assists in assessing potential solutions and forecasting approaches.

Meridian proposed the following:

> AEMO will attempt to use forecasts for unconstrained intermittent generation that are both accurate and consistent. In particular, forecasts used in NEMDE in the short term (e.g. 5 min ahead) should seek to utilize best available information about the generator and the factors likely to impact levels of generation.
>
> Where generators are in possession of information or understandings about the performance of their plant that is likely to be more accurate and more consistent than industry wide models (e.g. AWEFS/ASEFS) AEMO should seek to utilize that information where feasible.
>
> There should be clear guidelines and processes that set out, for the benefit of the market, how AEMO will exercise its decision making in this area so that both generators and the wider market can have confidence that the forecast used are likely to be accurate and consistent.
>
> Any procedures adopted should not be overly complex and be capable of ready adoption and maintenance by both AEMO and generators.

AEMO response

AEMO agrees with the intent of the proposed principles, and considers those principles are already embodied in the self-forecast assessment procedure.

4.8.2 Assessment reporting

Meridian suggested that, as part of performance assessment reporting, AEMO identify dispatch intervals with high market or system impact, such as semi-dispatch intervals with large forecast variations or intervals with large negative prices.

AEMO response

AEMO considers that it should provide its weekly assessment reports to the participant at the summary level to keep its process manageable while providing sufficient information to assist the participant in reconciling the assessments against their own data accessed via market systems and their own SCADA database.

AEMO will not separately identify dispatch intervals with high market or system impact as suggested by Meridian, because a participant should be able to readily access this information from the market systems.

AEMO will provide the following summary level information in the SF assessment reports made available to the participant each week in respect of their semi-scheduled generating units:

1. Initial SF Assessment Report
   - DUID
   - Assessment Period
   - Count of Total DIs
   - Minimum DIs for Reliable SF Test
     - Count of Minimum Required DIs

- Count of Valid DIs
- Test Pass/Fail

– Minimum DIs for SF Performance Assessment Test
  - Count of Minimum Required DIs
  - Count of Included DIs
  - Count of Excluded DIs:
    - Constrained-off DIs with no SCADA Possible Power
    - DIs where no unsuppressed SF meeting gate closure
  - Test Pass/Fail

– SF Performance Assessment
  - SF MAE (MW)
  - SF RMSE (MW)
  - AWEFS/ASEFS forecast MAE (MW)
  - AWEFS/ASEFS forecast RMSE (MW)
  - Assessment Pass/Fail

2. Ongoing SF Assessment Report
   – DUID
   – For each assessment window:
     - Assessment Period
     - Count of Total DIs
   – Minimum DIs for SF Performance Assessment Test
     - Same as for Initial SF Assessment Reports
   – SF Performance Assessment
     - Same as for Initial SF Assessment Reports.

4.8.3 Changes to Energy Conversion Models

Infigen suggested that AEMO's proposed changes to the Wind and Solar Energy Conversion Models to include the optional SCADA Possible Power should be published as soon as possible to allow participants to provide thoughtful feedback on whether it is a reasonable benchmark.

Tilt considered that the proposed changes to the Energy Conversion Model should be treated as material rather than administrative because of the new definition of SCADA Possible Power and its use in determining whether to use a SF in dispatch.

AEMO response

AEMO agrees with Tilt that we should seek feedback on the definition of new ECM signals (such as SCADA Possible Power).

AEMO published the proposed changes to the Energy Conversion Model (including the addition of the optional SCADA Possible Power) in November 2018 for further feedback before finalising.

4.8.4 Changes to EMMS Portal – Intermittent Generation interface

Tilt suggested that the usability of the Intermittent Generation availability interface of the EMMS portal would be greatly improved by offering a similar interface to that for energy and FCAS offers, whereby the user can

enter changes down a column by entering the value and pressing enter. Currently the availability interface requires mouse-clicks to enter each value, which makes it very time consuming to enter an outage that covers a 48-period day, or multiple days.

AEMO response

AEMO will consider the changes proposed by Tilt as part of a future review of the Intermittent Generation Availability interface of the EMMS portal.