(DRAFT) SEMI-SCHEDULED GENERATION SELF-FORECAST – ASSESSMENT PROCEDURE

Published: **July 2018**

AEMO
AUSTRALIAN ENERGY MARKET OPERATOR

IMPORTANT NOTICE

Purpose

AEMO has prepared this document to provide information about the process for AEMO's initial and ongoing assessment of participant-provided forecasts for semi-scheduled generation.

Disclaimer

This document is a draft only and 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.

1. INTRODUCTION

The Australian Wind Energy Forecasting System (AWEFS) and Australian Solar Energy Forecasting System (ASEFS) provide unconstrained intermittent generation forecasts (UIGF) for solar and wind generating units, respectively, covering timeframes from 5 minutes to 2 years ahead.

AWEFS and ASEFS use a combination of statistical and numerical weather prediction-based forecasting models fed by the following inputs:

- Supervisory Control and Data Acquisition (SCADA) measurements from the power station.
- Numerical weather prediction data from multiple weather data providers.
- Standing data from the power station as defined in the Energy Conversion Model.
- Additional information provided by the power station, including inverters/ turbines under maintenance and upper MW limit on the facility.

The 5 minute ahead UIGFs from AWEFS and ASEFS are used by AEMO's National Electricity Market Dispatch Engine (NEMDE) to determine the dispatch of semi-scheduled wind and solar generating units.

In 2018, the **Market Participant Self-Forecasting** trial was initiated as a collaboration between AEMO, ARENA and industry (forecasting service providers, existing and new wind/solar generators) to demonstrate the benefits of self-forecasting wind and solar generation. The initial focus of the trial is on the benefit of using the participant 5 minute ahead self-forecast (SF) in dispatch, in preference to the forecast from AWEFS or ASEFS.

This document explains the process and criteria that AEMO intends to apply when assessing the suitability of participant self-forecasts for use as UIGFs in dispatch.

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
AEMO	Australian Energy Market Operator
ASEFS	Australian Solar Energy Forecasting System
AWEFS	Australian Wind Energy Forecasting System
AWEFS_ASEFS	AWEFS/ASEFS forecasting systems
DUID	Dispatchable Unit Identifier
ECM	Energy Conversion Model
MAE	Mean Absolute Error
MT PASA	Medium Term Projected Assessment of Supply Adequacy
NEMDE	National Electricity Market Dispatch Engine
PP	Possible Power
RMSE	Root Mean Squared Error
SCADA	Supervisory Control and Data Acquisition
SF	Self-Forecasts
UIGF	Unconstrained Intermittent Generation Forecast
Web API	Web Application Programming Interface

2. INITIAL ASSESSMENT

2.1 Process

Consistent with its power system security responsibilities, AEMO seeks to use the most accurate forecasts reasonably available.

Before a participant self-forecast (SF) is first used in dispatch, AEMO will assess its performance against the forecasts produced by AWEFS or ASEFS as applicable (AWEFS_ASEFS). This assessment is performed to provide a reasonable level of assurance that the SF will not provide materially worse inputs to dispatch than the current forecasts from AWEFS_ASEFS.

The overall stages in the initial assessment process are:

1. Participant works with AEMO to implement their SF into AEMO's web API pre-production and then production environments. During this stage, the participant is able to submit SF to AEMO's production environment for assessment purposes, but AEMO suppresses its use in dispatch.

2. Participant advises when AEMO can start to use its SF for assessment purposes.

3. AEMO conducts the initial assessment of SF performance, using a 12 week sample period[1] of SF data collected within its production environment. As a resilience measure, an unsuppressed SF must be received for at least 99% of all intervals before it can be assessed. The assessment will not differentiate between different model forecasts[2].

4. AEMO provides its initial assessment of SF performance to the participant:

    a. If the SF passes the initial performance assessment, AEMO enables the SF for use in dispatch by unsuppressing the SF[3], and AEMO commences ongoing performance assessment of the SF.

        AEMO will issue an automatic report to the relevant participant when it has unsuppressed the SF.

    b. If the SF does not pass the initial performance assessment, or there are insufficient SF received to perform the assessment, steps 2 to 4 are repeated on a weekly basis.

[1] AEMO's current wind and solar forecast vendor requires 3 months' of data to tune their Dispatch forecast model.
[2] If a participant submits multiple model forecasts, AEMO may, on the participant request, assess the performance of each model separately and provide feedback to the participant, but AEMO will not use this as the basis for acceptance. The participant must uniquely identify each model using the confidential Model field in their web API submission.
[3] Even after AEMO approves a SF for use in dispatch, the relevant participant may suppress that SF for any dispatch interval via its Web API submission.

2.2 Initial Performance Assessment

The initial assessment involves comparing the relative performance of the SF and AWEFS_ASEFS against measures of Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) over the sample period. These measures are defined below.

The threshold for passing the MAE assessment is that, for the 12 week sample period, MAE_{SF} must be less than or equal to MAE_{AWEFS_ASEFS}, that is:

$$\boldsymbol{MAE_{SF} \leq MAE_{AWEFS\_ASEFS}}$$

The threshold for passing the RMSE assessment is that, for the 12 week sample period, $RMSE_{SF}$ must be less than or equal to $RMSE_{AWEFS\_ASEFS}$, that is:

$$\boldsymbol{RMSE_{SF} \leq RMSE_{AWEFS\_ASEFS}}$$

The SF must pass both the MAE and RMSE initial performance assessments before AEMO will enable the SF for use in dispatch.

Mean Absolute Error

For i = 1…n 5-minute intervals in the 12 week sample period, the calculations are:

$$MAE_{SF} = \frac{1}{n}\sum_{i=1}^{n} Absolute(SF_i - Actual_i)$$

$$MAE_{\text{AWEFS\_ASEFS}} = \frac{1}{n}\sum_{i=1}^{n} Absolute(\text{AWEFS\_ASEFS}_i - Actual_i)$$

Where;

SF_i = 5-minute ahead MW forecast from the participant for time i, which is the highest priority forecast with the latest Offer_DateTime (prior to the MP5F gate closure time at i-70 seconds) that is not suppressed by the participant

$AWEFS\_ASEFS_i$ = 5-minute ahead MW forecast from AWEFS/ASEFS for time i, or where there is no AWEFS/ASEFS forecast, the SCADA Initial MW used as the default forecast

$Actual_i$ = IF semi-dispatch cap does not apply at time i ($SEMIDISPATCHCAP_i = 0$)

$$Actual_i = MAX(0, SCADA\ Initial\ MW_i)$$

ELSE semi-dispatch cap applies at time i ($SEMIDISPATCHCAP_i = 1$ [4])

[4] The SEMIDISPATCHCAP flag is set = 1 when the DUID is capped within a binding or violated constraint or has a dispatch target less than its UIGF.

$$Actual_i = MAX(0, SCADA\ Initial\ MW_i\ , PP_i)$$

PP_i = SCADA Possible Power (MW) at time i, if available and good quality[5], else ignore

Root Mean Squared Error

For i = 1…n 5-minute intervals in the 12 week sample period, the calculations are:

$$RMSE_{SF} = \sqrt{\frac{\sum_{i=1}^{n}(SF_i - Actual_i)^2}{n}}$$

$$RMSE_{AWEFS\_ASEFS} = \sqrt{\frac{\sum_{i=1}^{n}(AWEFS\_ASEFS_i - Actual_i)^2}{n}}$$

Calculation Notes

- For a DUID, it is assumed that the SCADA MW and the 5-minute ahead MW forecasts for AWEFS_ASEFS and SF are all referenced to the agreed dispatch point as defined in the ECM.
- For a solar DUID, sample intervals for assessment are between dispatch intervals ending 0405 and 2100 AEST inclusive, to minimise biasing the assessment during night-time periods.

[5] Note that the approach of using a higher SCADA Possible Power rather than SCADA MW during down-regulated periods will tend to reduce the SF error.

3. ONGOING ASSESSMENT

3.1 Process

AEMO expects to assess ongoing[6] performance on a weekly basis, to verify that the SF remains no less accurate than AWEFS_ASEFS over the relevant sample period.

The ongoing assessments use similar MAE and RMSE measures to the initial assessment, but differ in the shorter sample period:

- If the SF is not suppressed by AEMO at the time of the assessment, the sample period is up to four previous consecutive weeks where the SF passed the ongoing performance criteria.
- If the SF is suppressed by AEMO at the time of the assessment (see below), the sample period is the previous one week. This allows the performance measure to reflect more recent, potentially large improvements in SF performance.

In cases where the SF does not pass the ongoing performance assessment for the relevant sample period:

1. AEMO will suppress the SF for a period of one week (until the next weekly assessment) and issue an automatic report to the relevant participant.
2. The following week, AEMO will reassess the SF performance.
3. If the SF passes the ongoing performance assessment, AEMO will unsuppress the SF and issue an automatic report to the relevant participant.

3.2 Ongoing Performance Assessment

The threshold for passing the MAE assessment is that, for the relevant sample period, MAE_{SF} must be **X %** less than or equal to MAE_{AWEFS_ASEFS}, that is:

$$\boldsymbol{MAE_{SF} \leq MAE_{AWEFS_{ASEFS}}\, x\ [1 - \frac{X_{\text{ongoing}}}{100}]}$$

The threshold for passing the RMSE assessment is that, for the relevant sample period, $RMSE_{SF}$ must be **Y %** less than or equal to $RMSE_{AWEFS\_ASEFS}$, that is:

$$\boldsymbol{RMSE_{SF} \leq RMSE_{AWEFS_{ASEFS}}\, x\ [1 - \frac{Y_{\text{ongoing}}}{100}]}$$

Where;
$X_{ongoing}$, $Y_{ongoing}$ = 0 % initially, but is configurable

[6] AEMO may decide in future that ongoing assessment of SF is not required.

The SF must pass both the MAE and RMSE ongoing performance assessments to remain unsuppressed. Failing either or both will result in AEMO suppressing the SF for a further one week.

4. ONGOING MONITORING FOR SYSTEM SECURITY

In cases where the AEMO control room considers that a SF is causing, or could cause, a threat to system security:

1. AEMO may suppress that SF and default to using AWEFS_ASEFS in dispatch.

   AEMO will automatically report to the relevant participant when it has suppressed that SF. AEMO may not contact the participant directly prior to suppressing an SF.

2. AEMO will review the SF at the next ongoing weekly assessment, and if it passes the assessment AEMO will unsuppress the SF and automatically report that to the relevant participant.

AEMO is seeking stakeholder feedback on the following:

Question 1 – What other validations should AEMO perform on the sample data?

Question 2 – 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?

Question 3 – 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?

Question 4 – Should we 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?

Question 5 – 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?