Ref: EPR0053

5 May 2017

Attention: Mr Sebastien Henry
AEMC
ELECTRONICALLY SUBMITTED

Dear Mr Henry

Directions Paper: System Security Market Frameworks Review

CS Energy welcomes the opportunity to respond to the Directions Paper on the System Security Market Frameworks Review.

Our initial assessment was existing policymakers see system security as a technical problem not to be resolved by the power market auction. AEMO's work, and the Consultation and Direction Papers have confirmed this assessment.

This premise is leading to the drafting of Rules for:

- powers to AEMO in defining inertia standards, specifying services, managing specific events and approving contracts;
- obligations on network monopolies for the provision of inertia and local system strength services;
- technical standards for generators for the capability to provide faster frequency response;
- non-market contracts for a faster ancillary service from a particular type of nascent technology; and
- possibly a future market for an ancillary service for a highly specified service separate to the existing power market auction.

Instead of seeing distinctions between security and reliability, or inertia and faster frequency response for that matter, CS Energy simply observes a scarcity of a resource best allocated through a market's marginal price incentives.

We recognise there is a market failure with the current auction because it fails to price the scarcity of system security services.

In our opinion, the correct approach is to upgrade the power market auction to incentivise market participants, including new entrants, to keep the system reliable and secure. This should be superior to allocating the responsibility to regulated networks or prescribing by law, the provision of the service.

We have the following vision for the power market auction:

Brisbane Office
PO Box 2227
Fortitude Valley BC Qld 4006
Phone 07 3854 7777
Fax 07 3854 7300

Callide Power Station
PO Box 392
Biloela Qld 4715
Phone 07 4992 9329
Fax 07 4992 9328

Kogan Creek Power Station
PO Box 41
Brigalow Qld 4412
Phone 07 4665 2500
Fax 07 4665 2599

Wivenhoe Power Station
PO Box 38
Fernvale Qld 4306
Phone 07 5427 1100
Fax 07 5426 7800

CS Energy Ltd ABN 54 078 848 745 **www.csenergy.com.au**

The Rules should look to allocate the responsibility for reliability and system security on Market Participants as this will reveal an efficient cost through competition.

Once the power market auction's deficiencies are resolved, such as using real time calculations and more sophisticated computation, security limits could be incorporated into dispatch calculations and scarcity reflected in the price(s) when the auction clears.

With efficient marginal price signals, in the longer run, new and old technologies compete in the energy-only market without heavy regulation, specifications or compliance obligations. This will include consumers revealing their price elasticity, removing the missing money problem that occurs in energy only markets when consumer demand is rationed without reference to price (load shedding).

This vision can be achieved through improvements to the auction.

You will find attached a report prepared by Intelligent Energy Systems (IES) that details a comprehensive package of improvements to the NEM power market auction. The centrepiece of the package is the adjustment to system marginal prices to include system frequency.

We hope the Commission will find the report useful for the numerous concurrent processes[i] that pertain to the auction's operation.

Yours sincerely,

David Warman

EGM Energy Markets (Acting)

Enquiries: David Scott
Telephone 07 3854 7440

[i] 5 minute settlement, Non-scheduled generation, Scheduling Load as well as anticipated changes to marginal loss factors and regulation frequency control "Causer Pays" allocations.

A PACKAGE OF IMPROVEMENTS FOR THE NEM AUCTION:

A Report Prepared by Intelligent Energy Systems for CS Energy

18 April 2017

Intelligent Energy Systems

Executive Summary

CS Energy has commissioned Intelligent Energy Systems (IES) to prepare a report which, in summary:

"documents a credible and coherent package of NEM auction reforms that implements marginal pricing principles and allows more autonomous operation by participants"

The motivation is a perception that the National Electricity Market (NEM) relies more than is necessary on performance standards, regulation, participant direction and action by non-competitive entities such as monopoly networks, as well as on mechanisms that have market elements but which do not implement marginal pricing as well as they could or should.

The report covers three closely related areas where arrangements can be improved to gain the benefits of marginal pricing, namely:

- arrangements that operate within the half hour trading interval;
- the dispatch and pricing engine (NEMDE); and
- fees and charges

Each of these broad areas is described and elements identified where outcomes could be improved. Specific proposals for improvement are then presented and discussed in each of these areas. They are then brought together into a coherent improvement package with an associated implementation strategy.

Arrangements within the half hour trading fail to implement marginal pricing principles in many ways, including:

- the process of averaging 5 minute prices for settlement;
- the artificial step changes in price that occur between trading interval and, potentially, between dispatch intervals;
- the completely different treatment of Frequency Control Ancillary Services (FCAS) which are based on enablement rather than performance;
- the cost allocation process for FCAS which, at best, is distorted (as with causer pays) and at worst ineffective (as with contingency FCAS cost allocation); and
- the lack of any current mechanism to value and encourage inertia and fast frequency response.

We propose an upgrade package which takes the form of an additional service. It would have the effect of converting the current distorted energy and FCAS pricing into a smooth price trajectory which dynamically adjusts to promote frequency and Time error stability under a range of disturbances. Participants responding to these marginal price signals will help keep the system secure and reliable. The package would also would support the emerging need to promote and support inertia and fast frequency response. It could be implemented in stages, but relatively quickly if the commitment were made.

The second component of the package is a set of possible upgrades the to the National Electricity Market Dispatch Engine (NEMDE). The current NEMDE has served the market well but its underlying modelling approach, while it has been improved over time, is now about 20 years old and could benefit from an upgrade. Areas that could be considered for improvement, along with associated Rule changes, include:

- determination of Marginal Loss Factors (MLFs);
- treatment of security constraints;
- contingency FCAS cost allocation;
- modelling of offer profiles and regional demand; and
- valuing reactive power, voltage and system strength.

In some but not all cases an upgrade may involve inclusion of a full network model using either DC load flow and, for advanced prototyping, an AC power flow model. These options in turn may require, or could benefit from, use of a non-linear solver. The advantages of these upgrades are that the dispatch process could be made more transparent, more stable and able to make fuller use of the network. Rule changes may be required in some cases e.g. to move to dynamic MLFs.

The upgrades to the intra- half hour processes and the NEMDE would in themselves promote better cost allocation in key areas, specifically in the allocation of FCAS costs. However, we recommend that the allocation of all fees and charges be reviewed to ensure that price sensitive participants do not make inappropriate decisions based on how those fees and charges are levied.

One important area for review is how market development is to be funded. Pressure from participants can lead to underfunding when market improvements are most needed. There is a role for governments to support market research, development and demonstration activity in a more focussed way than it does at present.

Finally, we present a co-ordinated improvement package and implementation strategy that draws all these elements together. Key features of the implementation strategy are:

- a focus on fast prototyping to uncover development issues and to give participants an insight into the new arrangements before going live;
- even when going live, a focus on phasing in the arrangement wherever possible, so participants can learn how to adjust their operations if they need to;
- minimal disturbance of existing systems on implementation;
- low cost and development pathways initially; and
- scope to upgrade to more advanced implementations over time, in order to support system requirements for services such as inertia and fast frequency response.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AC	Alternating Current
AEMC	Australian Energy Market Commission
AEMO	Australian Energy Market Operator
AER	Australian Energy Regulator
DC	Direct Current
FCAS	Frequency Control Ancillary Service
IES	Intelligent Energy Systems
MDP	Metering Data Provider
MLF	Marginal Loss Factor
NECA	National Electricity Code Administrator
NEMDE	National Electricity Market Dispatch Engine
NEMMCO	National Electricity Market Management Company
NEO	National Electricity Objective
NER	National Electricity Rules
NMI	National Meter Identifier
R&D	Research and Development
RAS	Ramping Ancillary Service
SCADA	System Control and Data Acquisition

1 Introduction

1.1 Background

CS Energy has commissioned Intelligent Energy Systems (IES) to prepare a report which, in summary:

"documents a credible and coherent package of NEM auction reforms that implements marginal pricing principles and allows more autonomous operation by participants"

The motivation is a perception that the National Electricity Market (NEM) relies more than is necessary on performance standards, regulation, participant direction and action by non-competitive entities such as monopoly networks, as well as on mechanisms that have market elements but which do not implement marginal pricing as well as they could or should.

Marginal pricing is associated with competitive markets. Promotion of competition in electricity markets is a central tenet of the Australian Energy Market Commission (AEMC). In a 2016 speech on storage technologies, the AEMC Chairman, John Pierce, said:[1]

"In short, new technologies and business models, coupled with consumer-led development of the energy sector, mean the biggest gains will come from network tariff reform and redrawing the line between what is subject to economic regulation and what is competitive".

As this report will demonstrate, re-drawing the boundary between what is regulated or poorly priced and what is competitive and priced according to marginal pricing principles can be applied to the NEM more broadly, not just to network pricing and regulation.

1.2 What this Report is About

The focus of this report is improving the operational and pricing machinery of the wholesale market of the NEM and, specifically, the operations and pricing close to real time. We see progress in this area as a necessary condition for dealing with the raft of challenges facing the electricity sector in Australia in 2017 and beyond.

Maintaining system frequency is a requirement for the system to remain in a secure state. The auction marginal price signal can be improved by incorporating a real time price component into settlements based on deviations from system base frequency, which are driven by very short term fluctuations in supply and demand. By including system frequency error (and time error) into prices, through adjustments to the marginal price, participants will keep the system secure and reliable and be paid for doing so.

[1] "New technologies: re-drawing the line between what is subject to economic regulation and what is competitive", John Pierce, Chairman, Australian Energy Market Commission, a paper presented at the Electricity Energy Storage Future Forum, 23 February 2016, Sydney Australia, p1.

Upgrades to the dispatch engine may allow for more stable and predictable dispatch price outcomes, more effective optimisation of security constraints, better treatment of marginal losses and, potentially, reactive power, reducing costs by optimising asset utilisation. Improved calculations will improve marginal decision making by participants, reducing costs. Potentially also, the opportunities and requirements for improving system strength could become more transparent.

Following this approach, the NEM will need to rely less on interventions, regulations, technical performance standards, separately defined ancillary services, network monopoly provision, obligations, price caps and other heavy handed mechanisms.

With improvements in technology and measurement, financial incentives will provide the boundaries for market participants to compete, with a reduced need for specifications, central systems and rigid compliance requirements. With these developments market participants will increasingly include consumers, storage as well as producers, with no requirement for direct control by networks or dominant retailers.

In a transparent market with such participation, there won't be the anxiety over failure of Energy Only Markets that exists today, (with regulated price caps and non-price rationing of demand). Consumers will more easily reveal their value of reliability though their consumption decisions and the market will clear.

1.3 What this Report is Not About

Current policy challenges in the NEM include gas supply and pricing, market concentration, competition in retail markets, the approach to Australia's emissions obligations in the sector, aspects of governance and a pervading sense of policy uncertainty inhibiting investment. Important as they are, these aspects are not the subject of this report.

More innovative approaches to marginal pricing can also substitute for heavier-handed approaches within distribution networks. For example, it is not hard to envisage a pricing regime for distribution network constraints which could render superfluous centralised control or even ownership and/or control by regulated networks of technologies such as batteries. However this report focuses on wholesale market improvements only.

1.4 Opportunities for Improving the NEM Auction

The National Electricity Objective (NEO), as stated in the National Electricity Law, is:

"to promote efficient investment in, and efficient operation and use of, electricity services for the long term interests of consumers of electricity with respect to – price, quality, safety, reliability, and security of supply of electricity; and the reliability, safety and security of the national electricity system."

To these can be added supplementary principles that support the NEO, including:

- effective support for government environmental objectives; and

- The market design principles set out on Chapter, Clause 3.1.4 of the National Electricity Rules (NER) which, in summary, includes the following requirements (See Appendix A for the full text):
    1. allow Market Participants the greatest amount of commercial freedom to decide how they will operate in the market;
    2. maximise the level of market transparency to allow for responses that reflect underlying conditions of supply and demand;
    3. avoid any special treatment of different technologies used by Market Participants;
    4. consistency between central dispatch and pricing;
    5. equal access to the market for existing and prospective Market Participants;
    6. as far practicable, market ancillary services should be acquired through competitive market arrangements determined on a dynamic basis;
    7. the relevant action under section 116 of the *National Electricity Law* or direction under Clause 4.8.9 must not be affected by competitive market arrangements;
    8. ancillary services, charges should be allocated to provide incentives to lower overall costs of the NEM; and
    9. where arrangements provide for AEMO to acquire an ancillary service, AEMO should be responsible for settlement of the service.

The trade-offs between price, reliability, security and environmental objectives (specifically, accommodating renewable energy targets) appear particularly stark at the time of writing. One way of improving this trade-off is through innovation – doing things better than they were done before. Our review of opportunities to improve pricing in the wholesale market has focused on three related areas. These will together make up a coherent package that can be implemented quickly and then developed and improved over time. These areas are:

- Market operations within the half-hour trading interval

  Within each half-hour, generators and market loads are given energy dispatch targets each 5 minutes and ancillary services are also deployed to keep the system secure. The design is intended to dispatch efficiently, but the confused pricing of energy and ancillary services within the half hour is inadequate to deal with current security challenges and emerging changes in technology. <u>The failure to include security services in wholesale pricing is a market failure that can be rectified.</u> Current rule change proposals under consideration by AEMC do not fully address this issue.

- National Electricity Market Dispatch Engine (NEMDE)

  This core system was designed and implemented 20 years ago and has undergone only incremental improvements since. Although it has served the NEM well until recently, its capabilities are now well behind those of state-of-the-art systems now being deployed in other markets. NEMDE is due for significant upgrade, not only to bring it to state-of-the-art, but to better deal with the specific challenges now facing the NEM.

- Allocation of costs and fees

  The allocation of costs and fees in the NEM are generally highly smeared and do not provide clear marginal cost signals for participants to minimise the costs. Even in cases where some attempt has been made to allocate costs efficiently, such as Causer Pays in FCAS regulation, implementation has smeared the price signal by removing the application of the price from real time. The structure of market fees has also put a brake on market development as AEMO has come under pressure from participants to reduce costs. The other recommendations of this report open up opportunities to revisit cost allocation arrangements in the NEM.

1.5 Structure of the Report

As noted above, this report covers three closely related areas where arrangements can be improved to gain the benefits of marginal pricing, namely:

- arrangements that operate within the half hour trading interval;
- the dispatch and pricing engine (NEMDE); and
- fees and charges

Sections 2, 3 and 4 review the current status of these arrangements. They also offer a critique of where these arrangements could better meet the National Electricity Objective and the principles for the dispatch process laid out in the National Electricity Rules.

Sections 5, 6 and 7 present specific proposals for improving these arrangements.

In Section 8, these proposals are brought together into a market auction upgrade package and implementation strategy.

Section 9 contains the Conclusions of the report.

2 Current Arrangements within the Trading Interval

2.1 Energy Market Dispatch and Pricing

While the trading interval for settlement of the spot market and, by extension, settlement of most energy contracts is half an hour, the market is dispatched and a price of energy determined each 5 minutes. Separate prices are determined for each of the 5 NEM regions, and settlement amounts are adjusted by fixed MLFs whose values are periodically revised by AEMO.

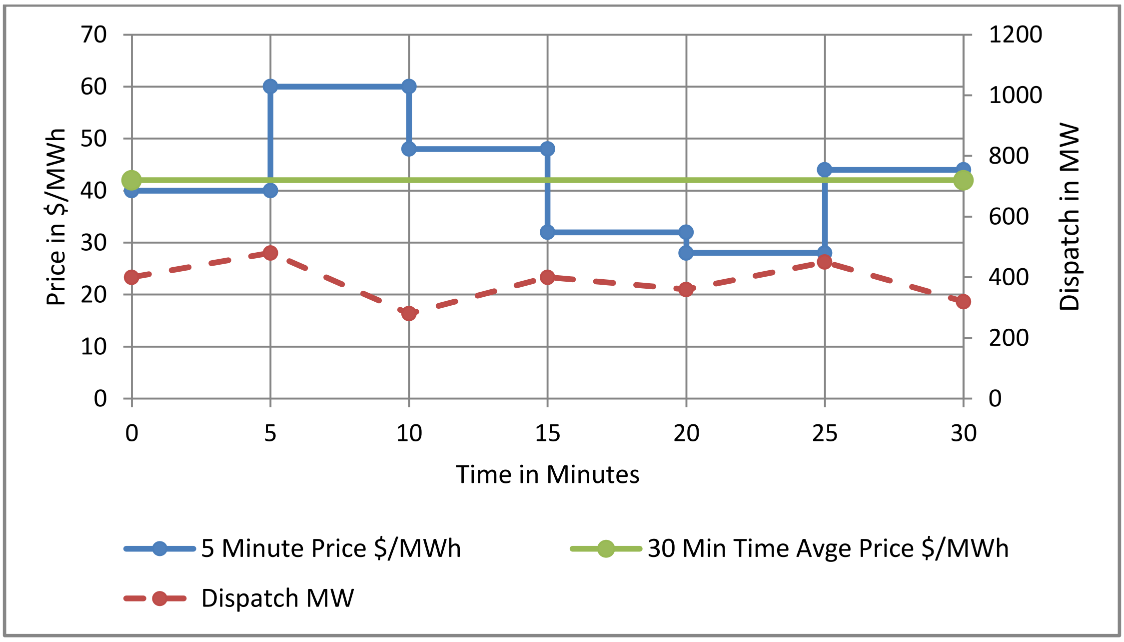
The logic for determining the 5-minute dispatch schedule and spot price is implemented within NEMDE. This is considered further in Section 5. This Section is concerned with how the 5-minute price is used for settlement. The detailed process is as follows:

- Just before each 5-minute boundary, AEMO gathers data on the current state of the system through its SCADA system and, using this as a starting point along with participant offers and other data, runs NEMDE to determine the operating schedule to meet the regional loads projected 5 minutes ahead.
- NEMDE assumes that participant plant ramps linearly to the scheduled operating level 5 minutes in the future. AEMO requires participants to conform to this dispatch schedule as closely as possible.
- The regional prices that NEMDE determines are actually the marginal value of energy *at a time 5 minutes in the future – at the next 5-minute dispatch boundary*. Because this price is calculated ahead of time, it is called an ex ante price. 5 minute dispatch with ex ante pricing is a feature of the Australian NEM.
- To determine a settlement price:
  - settlement is based on the regional price in each region, adjusted by MLFs;
  - the 5-minute ex ante price is assumed to apply over the whole of the 5 minute period prior to its application;
  - the six 5 minute prices are arithmetically averaged to get the half hourly regional price for the region, which is adjusted by fixed Marginal Loss Factors (MLFs) for participant settlement;
  - the half hourly energy used for settlement is that measured by half hourly participant revenue meters, even though the system is managed in real time using SCADA metering.

We can illustrate this process in the diagram of Figure 1 below. A Time = 0 (minutes) NEMDE determines a price of $40/MWh and a schedule for the illustrated participant requires the unit ramp from its current level to around 440 MW in 5 minutes, as shown in dashed red. The price of $40/MWh is assumed to apply over this whole 5-minute period. At the 5 minute boundary a new price a further 5 minutes ahead is determined at $60/MWh, together with a trajectory for the unit, to ramp down in this case (the unit must have increased its offer price). That $60/MWh price applies for the whole interval. This process continues to the end of the half hour, as shown.

The settlement price for the half hour is calculated as the arithmetic (time) average, as shown by the green horizontal line. This is used for settlement (with MLF adjustment), notionally of the energy under the red dotted trajectory. However, this physical trajectory is not used, for two reasons. First, the unit will never follow this trajectory exactly. Second, the real time operations are based on SCADA real time readings, which are less accurate than revenue meters. Therefore, settlement is based on the more accurate revenue meter half hour readings, and the averaged half hour price.

Figure 1: Example of 5 Minute Dispatch and Half-hour Pricing



It is immediately evident that, within the half hour, this logic violates marginal cost pricing principles. Nowhere does the time average price equal a calculated marginal price from NEMDE, which is the efficient price.

Why is the price calculated in this way? At the start of the market, the half hour trading interval was agreed early on. Later in the design phase of the market, the advantages of 5-minute dispatch were recognised (and remain valid today); 5 minute dispatch greatly reduces the need for ancillary services to balance supply and demand services and for system operator intervention to keep it secure.

This change in philosophy raised the question of how the half hour trading interval for settlement was to be reconciled with the 5-minute dispatch price. The simple solution was to arithmetically (time) average the six 5 minute prices in each half hour. The effect of that decision is the average of the 5 minute piecewise constant ex ante prices as described previously.

This was recognised as a compromise and an approximation, but acceptable because no-one imagined that loads in particular could or would do much within each half hour. Furthermore, generators could live with the approximation involved.

Interestingly, and in contrast, the FCAS markets implemented a few years later (in around 2001), have been settled on a 5 minute basis from the start. But this approach for FCAS did not present the same revenue meter challenges as 5-minute settlement in the energy market could have done, because they are not settled the basis of measured energy.

At the time of writing the AEMC is considering a rule change proposal from Sun Metals which would aim to implement 5 minute settlement for energy. In essence, the aim is to settle on the stepped price profile in Figure 1, rather than the averaged flat price profile. The aim is to provide a sharper price signal to participants, and specifically fast response options. It would be implemented by settling each dispatchable element on its own generation or load weighted price. Non-dispatchable loads and generation would be treated as a residual. A range of variations on this are being considered by AEMC. Submissions suggest little current participant support for the change, based on a range of arguments including the extent of system changes that would be required, difficulties for gas fired generation with start-up times and the disruption that could occur to the contract market. These issues will be discussed further later in this report.

Another factor that sometimes results in a departure from marginal pricing principles for some participants is the market design element that settles all participants in a region at the regional spot price adjusted by a constant marginal loss factor. This is usually a reasonable approximation to marginal pricing, except when losses vary widely or when a unit is constrained on or off by an intra-regional security constraint. When a constraint is binding, there will be a discrepancy between pricing and dispatch. This issue has been considered in the past by the AEMC in a rule change proposal called “Optional Firm Access”. An alternative approach is to implement full nodal pricing with firm transmission rights. Nodal pricing is technically possible but can shrink the size of regional markets.

2.2 Energy Spot Price Step Transitions

The concept of 5 minute settlement appears sound on the basis for first principles because the variations in marginal prices that can occur within a half hour would be shown to participants and their generation or load settled on that basis. This is important for options such as batteries that can respond quickly. Further, with increasing penetration of renewables into the system, system security will require that more fast response options be available as time goes on, including a minimum level of inertia.

However, the stepped price profile of Figure 1 cannot represent a true marginal cost, as cost curves of generation are, taken as a whole, continuous and smooth.[2] Even when costs are modelled as piecewise constant marginal cost curves (as they are in the NEM), the price would not step suddenly to the price 5 minutes ahead. This is not just a theoretical proposition of no practical interest. Large scale penetration of batteries could respond rapidly to this step price change (if exposed to it) and destabilise the system. Generators would also be incentivised to respond, yet this is presently prohibited under the Rule 4.9.8 (which requires absolute compliance with dispatch instructions). It would be against the

[2] Generators and loads can have step changes in their cost curves and also “forbidden regions”, but these occasional local discontinuities do not affect the continuous and smooth nature of an aggregated cost curve.

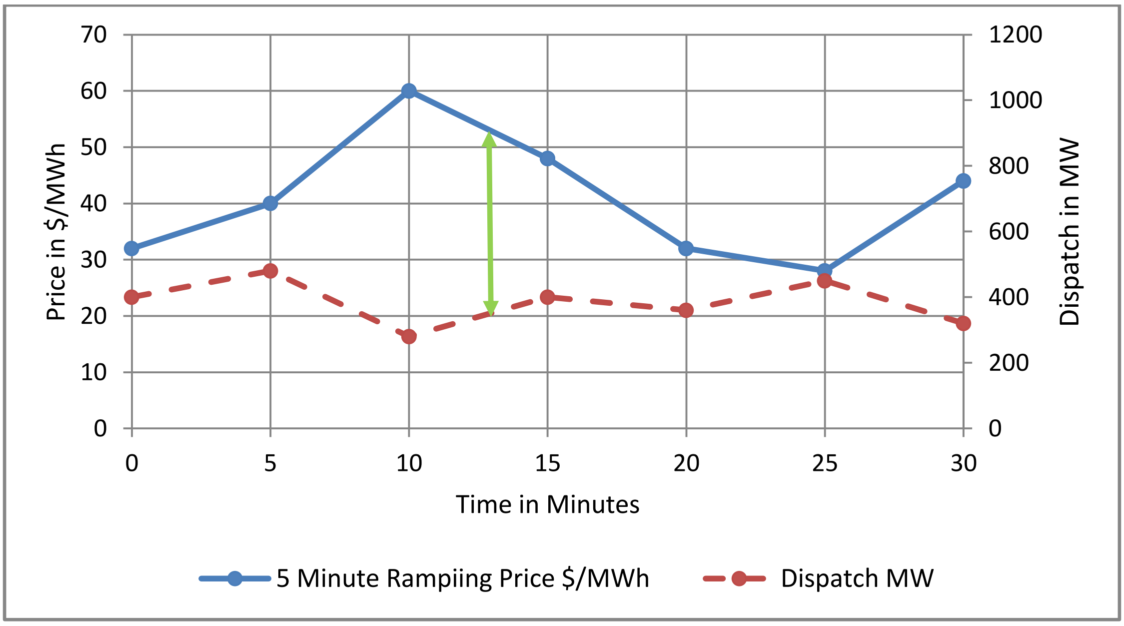
intent of the market to use rules and regulations to manage this problem; one seeks a market approach so marginal prices must be improved.

When the NEMDE is run at a specific 5-minute boundary, it is targeting a time 5 minutes ahead. The dispatch targets and the prices it calculates apply to that time. The scheduled generators are modelled to ramp linearly from the current MW level (as measured by SCADA) to the target level. Use of the calculated price over the whole prior interval is a convenient artifice for simple settlement, but not reflective of marginal costs. A much better approximation is illustrated in Figure 2 below; just as scheduled units ramp from one MW level to another, so does the price. The ramped price is the solid line and the ramped schedule is shown dashed. To simplify the diagram, we have assumed that participants follow their dispatch trajectories.

Under a ramped price regime, participants would not be encouraged to switch at 5 minute boundaries; they would wait for a price that suits them. This would encourage switching diversity and greater price and operational stability.

To calculate settlement under this arrangement, one needs to accumulate the product of MW over a small period (to measure energy over the interval) with the corresponding price as indicated by the green double-headed arrow. This is not a difficult calculation with the appropriate metering on place, either in real time through SCADA or a with a suitably programmed revenue meter.

Figure 2: Ramping Price and Dispatch Trajectories



The trajectories shown can be interpreted as scheduled trajectories for dispatch and price[3]. In practice, generators never follow their linear trajectories; nor is there a compelling

[3] AEMO actually schedules generators from their SCADA-measured starting limits from each 5-minute boundary. Thus AEMO would view dispatch trajectories as discontinuous. For our proposed development options, we prefer to think of these trajectories as moving from one target to another to avoid price discontinuities.

reason to require them to do so if frequency and other security elements are managed within bounds. Also, we can conceive of a mechanism that varies the price seen by a market element, so that the payment adjustment (positive or negative) depends on its performance in helping to stabilise the system. This is normally the job of Frequency Control Ancillary Services (FCAS) as described in the following sub-section.

2.3 Frequency Control Ancillary Services (FCAS)

Five minute schedules, even if diligently followed, are not sufficient to manage the energy balance in the system. Load Forecast errors, generators either non-scheduled or simply not following their schedules, short term load variations, and the differences between model outputs and reality all combine to require a real time adjustment of power inputs and outputs to keep the system within prescribed frequency bounds. This is the role of FCAS.

The NEM defines two classes of FCAS; regulation and contingency. Regulation deals with the small but cumulative variations that occur all the time. Contingency services deal with larger events such as large generator, load or interconnector (single line) trips. Why the difference? The main difference is that contingency services need to respond to a frequency deviation event quickly and autonomously, whereas slower moving regulation is currently managed centrally though real time SCADA measurements sent to AEMO, which sends lower and raise signals to generators depending on high or low frequency.

Because some supply FCAS options are unidirectional, FCAS services are also classed as raise or lower. Within each such class, there is one regulation service and there are three contingency services requiring responses at 6 seconds, 60 seconds and 300 seconds respectively, so there are 8 FCAS services in total. It could be possible to define additional services, covering faster response and inertia, but that has not been done thus far.

Unlike energy, FCAS is procured on an MW “enablement” basis. That means the providing unit is essentially contracted each 5 minutes to stand by to provide the service when needed. In the case of regulation, they will be used all the time, but not necessarily to the full MW enabled. In the case of contingency, they will be called upon much less often, perhaps doing nothing for weeks at a time.

Each FCAS is procured competitively each 5 minutes though a bidding process integrated with the energy dispatch process, managed and optimised within the NEMDE. The requirement is generally set by AEMO, so AEMO pays for the amount enabled at the market price determined by NEMDE each 5 minutes. So some key features of FCAS are:

- AEMO pays for enablement, not actual performance. Each supplier is responsible for its performance or it faces enforcement action by the Australian Energy Regulator (AER).
- AEMO must recover the cost of each service. Generally, raise services are recovered from generators pro-rated by energy, with lower services recovered from loads in a similar way. Networks are excluded from the cost recovery process.

- The exception to the above method of cost recovery is with regulation, where a process called Causer Pays is applied. Under this process, SCADA measurements of the extent of "cause" of frequency variation are accumulated over a settlement period and expressed as a percentage factor. Regulation raise and lower costs from the following settlement period are then allocated according to those pre-calculated factors. This process appears to be moderately effective but is watered down by not being applied in real time. Further, the approach is one sided; it charges (indirectly) for causation but does not reward service provision, whether it be enabled or not. Performance with the provision of FCAS when enabled has to be managed administratively by enforcing regulatory Rules.

2.4 Other Security Requirements

Within the half hour, security requirements other than frequency control also need to be managed. These are typically requirements to protect the network and stable operation in case of network outages, implemented as security constraints within NEMDE as outlined in Section 3.2. The effect of these constraints is to adjust the dispatch targets to ensure that the system remains stable after a network outage. Generally, the only requirement here is that dispatch targets be followed to a reasonable degree; there is no real time requirement except in special cases where a protection scheme might be implemented.

There are also other mechanisms that operate within the half hour, such as the Fast Start Inflexibility Profile used to commit so-called fast start generators.

2.5 What's Wrong with the Current Arrangements?

Areas where marginal pricing principles are violated in these processes are summarised below.

- The time averaging of 5 minute dispatch energy prices to get a half-hourly energy price for settlement destroys the price signals operating within the half hour, most notably when prices are changing rapidly for some reason, such as the binding of security constraints.
- Even if this problem were to be fixed with some form of 5 minute settlement as currently under consideration by the AEMC, step changes in price every 5 minutes, made available ex ante, could ultimately be exploited by fast acting technologies such as batteries, or by generators deliberately deviating from dispatch instructions. While a controlled response of such technologies is desirable, if driven by artificial step changes in price their responses could be detrimental to the efficiency of the market and even system stability. While AEMO could attempt to correct for such response by adjusting its load forecasting, relying so much on AEMO's forecasting to determine market outcomes is not a satisfactory approach. AEMC has not so far mentioned this issue in its evaluation of the 5-minute settlement rule change proposal.
- The regulation FCAS markets are entirely driven by enablement rather than actual performance in provision. Poor performance in provision does not have a direct financial consequence. The regulation FCAS requirements (raise and lower) are relatively arbitrarily set, based on experience. Regulation FCAS cost allocation is by

"causer pays" factors, which is a marginal cost concept. However, its effect is diluted by having them apply to settlement in the month after their values were determined. The enablement of regulation FCAS from each participant can undergo a step change at each 5-minute interval, possibly inhibiting smooth and efficient operation.

- As with regulation FCAS, contingency FCAS markets are driven by enablement rather than provision. This is more understandable than it is for regulation because the need to provide contingency FCAS is relatively rare; contingency FCAS is like insurance. Costs are allocated to generators (for raise) or customers (for lower) as marginal costs in $/MWh volumetric charges, yet cannot in practice be avoided by changing such volume. The actual causers of contingency FCAS incidents (including networks) are immune from any financial penalty from that incident. The requirements are set by the largest credible contingency affecting a region or set of regions, but these are generally not optimised within NEMDE (interconnector contingencies being an exception) and costs are not allocated according to that logic.
- FCAS services in practice can only be provided by direct market participants and much of it under centralised control by AEMO. Each market is highly specified to create homogeneity (a prerequisite in commodity markets), yet this creates boundary issues and barriers to entry for incumbent generators and entrants that don't quite fit in. This philosophy restricts the range and volume of possible providers. A recent "aggregator" rule change may help but is only one component of a solution.
- While energy provision and FCAS enablement are optimised within NEMDE, there is no evident relationship between energy and FCAS prices. Philosophically, there should be a smooth transition between the provision of FCAS within the 5 minutes and the energy market.
- FCAS services are generally system wide except under specific circumstances where a credible contingency may cause system separation. There may be merit in considering more regionally-based FCAS provision, especially for regulation FCAS. Amongst other things, this could support tighter control of interconnector flows.
- While AEMO and AEMC have recognised a need for inertia and fast frequency response services, there are as yet no serious proposals under consideration for how such markets might work or, more sensibly, how existing market signals could be amended to incentivise these services. The default approach is to hand the task to regulated entities–TNSPs. AEMC's current approach is to consider 5-minute settlement as a standalone issue very much focussed on a particular model put forward by Sun Metals as a rule change proposal. Treating every issue as distinct or separable makes no sense.
- The NEM is an Energy Only Market that operates with a price cap and a separate mechanism to procure supplies with costs in excess of the cap (Reliability and Emergency Reserve Trader). When the system operator sheds load, this rationing is not based on price; in these circumstances the market hasn't cleared. Preferably, an energy only market would not have a price cap, but instead have sufficient demand elasticity so that consumers are happy to respond by rationing their own consumption when price escalates. The non-price rationing of demand, when consumers have differing levels of utility, is a market failure.

2.6 Assessment against Market Principles

Current arrangements can also be assessed against the market principles of Appendix A. The current context needs to be kept in mind. For example:

- the system today may need more options that can respond quickly within the half hour;
- newer technologies such as batteries and improved communications and control can deliver that capability more effectively than 20 years ago;
- high electricity costs have ignited more interest in managing loads;
- even though a service may be procured competitively at present, there may be alternative mechanisms that better reflect marginal pricing principles to deliver a more efficient outcome.

Table 1 following summarises the assessments of the previous sub-section against these principles, taking account of the current context.

Table 1: NEM Dispatch Process Conformance to Market Principles

No.	Principle (summarised)	Comment on NEM Approach as of March 2017	Assessment
1	Commercial freedom for participants	• Scheduled participants have freedom to participate or not. • Freedom within the 5-minute dispatch interval is circumscribed by conformance rules, sometimes unnecessarily.	• Could be made more flexible and opened up to a wider set of participants.
2	Market transparency	• Generally good in the wholesale spot market part of the NEM.	• Maintain transparency with any new development.
3	Technology Neutrality	• Current rules appear to favour large, slow moving schedulable technologies; albeit circumscribed by compliance obligations. • New technologies such as batteries and other fast but limited duration options may be restricted by price distortions within the half hour, even though the system would value their services, especially with increased renewable penetration. • Specifically, inertia and faster frequency response is not supported in market arrangements.	• The current rules were formulated when most participants were slow moving and could be scheduled. The requirement for large, fast responses was less than it is now or will be in future.

No.	Principle (summarised)	Comment on NEM Approach as of March 2017	Assessment
4	Consistency between central dispatch and pricing	• Violated within the half hour due to current pricing rules. This distortion greatly affects fast response options but less for slow moving options. • Some distortion also from the use of regional rather than local nodal prices for settlement. • Non-scheduled market participants not circumscribed by compliance obligations.	• The current distortions within the half hour can and should be corrected. • Regional pricing should remain.
5	Equal access for existing and prospective Market Participants	• Non-scheduled fast response potential participants (such as battery operators) have limited opportunities to participate in energy and FCAS markets. • Definition of numerous FCAS markets with systems, standards, registration, etc. are more difficult to access than frequency control costs integrated into system marginal prices.	• Rules can be developed which are more neutral and sympathetic to new market participants who do not wish to be scheduled.
6	Ancillary Services market to be competitive and dynamic	• FCAS enablement markets are competitive and dynamic, yet require complex rules, standards and systems to function. • For scheduled participants, there are no incentives (restrictions as well) to compete in real time within the 5 minute dispatch interval. For non-scheduled participants there is the freedom to respond as they wish.	• Improvements are possible with a real time usage market in FCAS.

No.	Principle (summarised)	Comment on NEM Approach as of March 2017	Assessment
7	Section 116 of the National Electricity Law not affected	• Not applicable to this project.	• Not applicable
8	Ancillary Service Costs allocated so as to be avoidable if possible	• Contingency FCAS is smeared across broad classes such as generators & loads as a marginal cost associated with the consumption or supply of energy, yet changing consumption or supply (unless the largest contingency) does nothing to avoid these costs. • Networks do not receive any cost allocation for contingency FCAS, even though they can often affect FCAS costs • Regulation FCAS costs are imperfectly allocated under causer pays due to lagged application of Causer Pays factors, netting rules and other compromises in the calculations.	• If a FCAS usage market is implemented, a lesser burden would be placed on FCAS enablement markets. Costs could be allocated so as to be avoidable and therefore more efficiently. This is not done at present apparently for fear that the burden would be too heavy on some participants.
9	AEMO to settle Ancillary Services market	• Done at the spot market level • Any bilateral contracting for these services should not be settled by AEMO	• Any improved arrangements would continue to be settled by AEMO.

3 The Current Pricing and Dispatch Engine

The NEM Dispatch Engine was designed 20 years ago and has only been maintained and tweaked at the margin since then. In the intervening period there have been major advances in computer and optimisation technology, bringing a far more capable dispatch engine within reach. Other comparable markets have or are in the process of refreshing their systems to take advantage of this improved capability.

The NEM dispatch engine implements the dispatch process as defined by the Rules, but many details of that process are not defined in the Rules; nor should they be. The following sub-sections define areas of the NEMDE which could be improved to meet the emerging challenges facing the NEM.in the light of the analytical technology now available.

3.1 Determination of Marginal Loss Factors

The NER specified that Marginal Loss Factors (MLFs) should be constant over the day, week and year but updated from time to time based on a forward-looking analysis done by AEMO. Only losses between regions are modelled dynamically. Under the current approach, AEMO carries out forward-looking off line studies to estimate future weighted average of MLFs within each region, as well as equations that adequately model (hopefully) the inter-regional loss relationships for use in the dispatch model. The requirement for intra-regional MLFs to be fixed is a particular constraint was justified initially (pre 1998) to make contracting simpler.

This loss treatment has proved to be problematic. While many MLFs are small or workably constant, in more remote areas the network flows can be highly variable and even tidal, especially where intermittent energy sources are involved. MLFs should vary accordingly. AEMO has implemented a number of fixes to partially deal with these problems where they occur (for example, by implementing directional MLFs) but with more remote power supplies coming on stream, the problem can only get worse.

Participants understand that hedging can never be perfect and could work with dynamic MLFs. They would quickly learn the pattern of behaviour of their MLFs and contract accordingly. Contracting in an environment where AEMO updates MLFs periodically also carries risks. One major advantage of introducing dynamic MLFs would be elimination of a whole layer of AEMO activity that in many ways pre-empts market outcomes.

Problems with loss modelling also arise within NEMDE when prices become negative, as NEMDE then uses the high loss segments without using the low loss ones first – physically impossible. This could be fixed at some unknown cost to performance with a mixed integer formulation, but AEMO chooses to detect such situations and do a second run, with the possibility of a non-physical solution locked out with a heuristic. This is an inefficient process, especially when re-runs are required for other purposes as well. The treatment of this issue could be improved with more suitable optimiser technology.[4]

[4] The issue is so-called “non-convexity” in the model when prices become negative. While non-convexity cannot be made to go away, an upgraded, non-linear modelling technology offers a simpler and more elegant solution.

3.2 Treatment of Security Constraints

To operate securely, the system must remain on or within so-called “security envelope”. Invariably this envelope is defined so that the system remains stable (but not necessarily reliable) after any credible contingency such as a line or generator outage.

The mechanism used in the NEM to define this security envelope for network contingencies is to apply a set of so-called generic constraints to the dispatch model. A generic constraint is a linear inequality (a weighted sum of injections and offtakes) applied to the dispatch model that must be respected when reaching a solution, if possible (sometimes it’s not possible!).

There are many possible contingencies that could occur at any one time and each contingency could affect several critical transmission elements. One cannot know which ones might be critical in advance, so sets of constraints are developed which are wheeled in when required to deal with credible scenarios. These sets cover broad categories of contingency issues such as:

- keeping the thermal loading of network elements within desirable tolerances after a line outage;
- keeping voltage levels stable after a line outage; and
- avoiding oscillations on the system after a disturbance from any source, specifically over the longer interconnectors.

Implementation of security-constrained dispatch is a large and important task for AEMO. Over time, some improvements have been made to make these constraints more dynamic in nature (such as implementing so called “feedback” constraints for thermal contingencies) but the operation remains substantially offline.

As with constant MLFs, one consequence of formulating security constraints off line and in advance is that one cannot fully account for the current state of the system. So “worst case” constraints are formulated which inevitably contain a degree of slack. This means that on many occasions the network is more limited in capability than it really needs to be if current conditions could be better accounted for.

Many other markets have moved to the dynamic generation of most security constraints, to be imposed on a dispatch model “on the fly”. Each solve actually operates in a tight loop; a trial solution is tested for security and a new constraint generated. If “hot starts” are used and other tricks implemented to reduce the requirement to test different cases, this process can be made very efficient. Advantages of generating security constraints dynamically are:

- the task of generating these constraints manually in advance would be substantially reduced, with a corresponding reduction in the risk of error;
- opportunities would open up to define constraints with less safety margin based on current conditions, thereby extracting more capability from the existing network without jeopardising security;

- potentially, a constraint formulated dynamically could be relieved by parties or technologies not yet considered for that purpose;

AEMO has gone some way down this path with the implementation of so called "feedback" thermal constraints, but more could be done.

3.3 Contingency FCAS Cost Allocation

The current NEMDE has a detailed arrangement to co-optimise the dispatch of FCAS services and energy, based on offers made into all of those markets. The requirement for each is managed with a series of generic constraints, but in many cases the requirement is set externally by looking to the largest generator online and setting the requirement at that level, after allowing for load relief.

The biggest problem with the current setup is the allocation of contingency FCAS costs. Regulation costs are allocated according to causer pays logic which, although garbled on implementation, attempts to allocate marginal costs to parties who can act to avoid them in their marginal decision making (how well they comply with dispatch instructions). But contingency raise services are smeared across generators and lower services are smeared across loads. A decision to change consumption or supply does not change the requirement unless the decision is made in advance by the contingency that sets the requirement.

The formulation to set the FCAS requirement optimally within the dispatch engine is straightforward in principle. The requirement (raise or lower) should be set to be no less than any credible contingency, from a generation, load or network element. In practice, the dispatch process would likely equalise members of a set of equal-sized contingencies to achieve a least cost outcome.

As with the energy balance equation, a dispatch solution would also yield a shadow price for an increment of the critical contingency, allocated efficiently across the affected parties. This could and should be used to allocate costs, as participants that represent these contingencies (loads, generators, networks) can, in principle and in practice, do something to reduce these costs. However, this solution has been considered and rejected in the past, apparently because of a fear that high costs would be assigned to a single participant who happened to get dispatched at the highest level.

This fear is largely unfounded because there will normally be a set of critical contingencies, not just one, as a natural outcome of the dispatch optimisation. Each load, generator or network participant owning a critical contingency will be exposed to a marginal cost of this contingency. In any case, FCAS prices are generally quite low. However, in unusual situations (e.g. when a region could separate), this cost burden could be high.

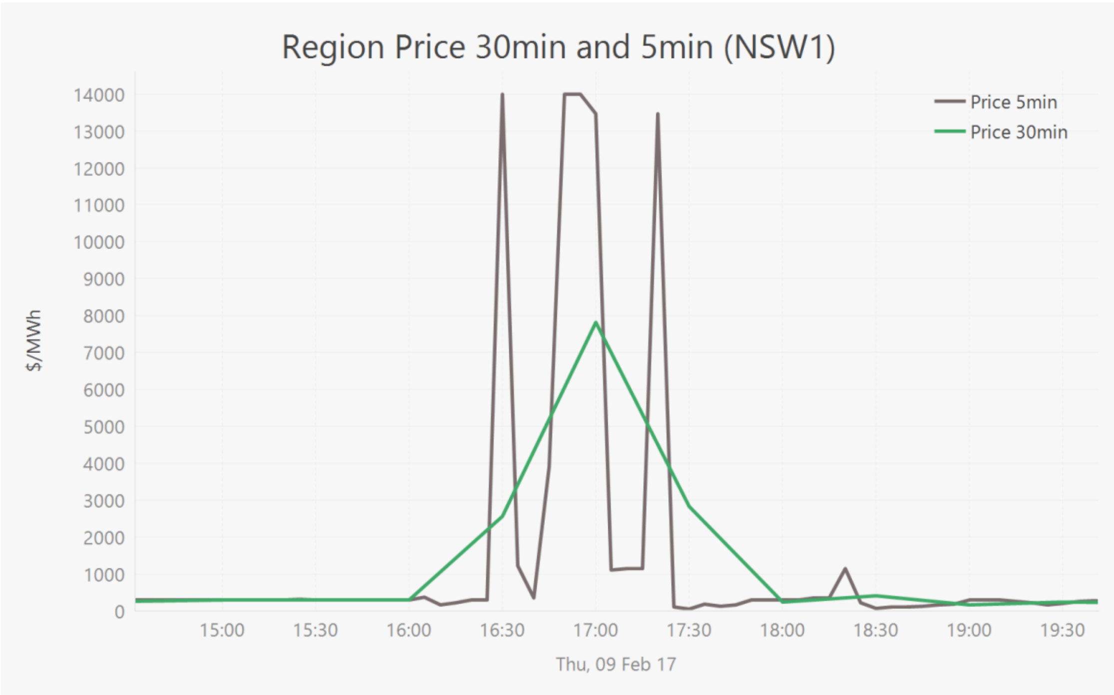
If a means could be found to reduce the cost of the FCAS enablement markets, by implementing a real time market, for example, this more efficient approach to contingency FCAS cost allocation could be revisited.

3.4 Modelling of Offer Profiles and Regional Demand

For practical reasons driven by available technology 20 years ago, NEM MW offer profiles are constrained to 10 bands. Further, the price of each band over a day is required to be constant. Generally, participants can construct offer profiles within these constraints that meet their requirements. However, at the high price end of their offer stacks, the ability to fine tune offer price necessarily becomes very coarse.

This is an irritant for participants because dispatch becomes more of a lottery than it should be. It's also unhelpful for the market as a whole, as prices at the high end can and do oscillate each 5 minutes. To some extent these oscillations are smeared out by the price averaging process previously described. However, if the NEM moves to some form of 5-minute pricing for settlement, this issue could come to the fore as the market becomes more volatile due to increasing renewable penetration. Figure 3 below illustrates one example of this 5 minute price instability during a recent market episode n NSW.

Figure 3: Example of likely artificial 5-minute price volatility



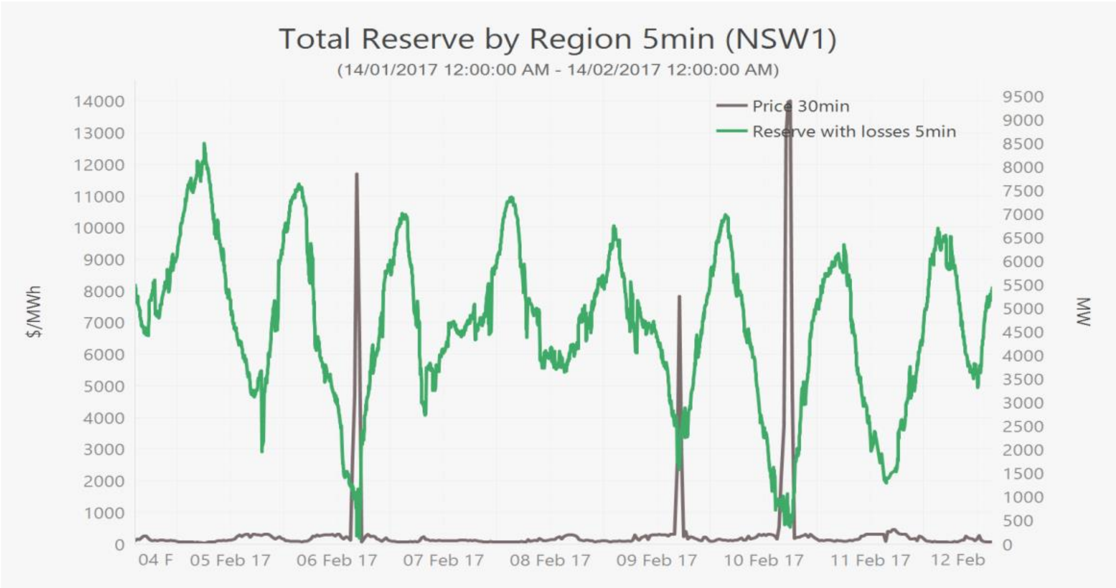
There are several factors at work here. One is the inability to fine tune the offer profiles at the high end, as just described. The other is possible load response to the high ex ante 5 minute price, which leads to a following price collapse, and then a later spike. Likely these two factors are interacting. Other technical matters, such as the Aggregate Dispatch Error; Dynamic Regulation Requirements; unit commitment under the Fast Start Inflexibility Profile commitment run; changes in the ratings of security limits, SCADA set-points (mill limits, FCAS trapeziums) and generating unit ramp rates all contribute to significant changes in price upon scarcity.

What could be done to improve matters? One approach to improve the flexibility of offer profiles would be to increase the number of offer bands available, although this could only ever be a partial solution. A more elegant and computationally efficient approach would be to allow offer band prices to vary continuously between bid bands. This could be made optional right down to the level of individual bands. This approach will be developed further later in this report.

Unscheduled load and small participant response to pricing will likely become more important over time, so this should be recognised in the dispatch process as well.[5] However, AEMO is inhibited in its ability to accommodate price sensitivity of demand as it could be seen as directly influencing the market if it did so.

However, there are some physical indicators that could be included in AEMO's forecasting algorithm that could take some account of load responsiveness. One indicator that IES maintains is total regional reserve, which is the margin of available generation over demand in a region, taking account of potential support over interconnectors. There is a high correlation between low levels of regional reserve and high regional prices, as illustrated by the recent example in Figure 4 below.

Figure 4: Regional Reserve and Regional Price



To implement, AEMO would simply need to incorporate regional reserve as a variable in its 5-minute load forecasting algorithm. A more sophisticated approach would be, in addition, to forecast price elasticity as a function of the current state, including the current level of total regional reserve. To implement an elastic demand, a non-linear NEMDE would be required, as will be discussed later[6]. If estimated demand and demand elasticity are based

[5] There is a current Rule change proposal under consideration by AEMC to lower the MW threshold for loads to be become scheduled. However, requiring load response to be scheduled could inhibit load flexibility for facilities that are not dedicated to servicing the electricity market.

[6] A stepped demand function could also be used, but that would both be more approximate and more difficult to estimate.

on measureable physical variables both past and present, AEMO cannot be considered to be active in the market.

3.5 Valuing Reactive Power, Voltage and System Strength

The objective of the dispatch process in the NEM is cost minimisation to meet the demand based on participant offers and subject to real power flow laws (highly aggregated and approximated), network security constraints and various other requirements such as the efficient dispatch of FCAS enablement.

Like most other markets, the NEM dispatch process is based in real power flows; it does not deal explicitly with voltage and reactive power or even flows over individual lines. Reactive power and voltages are not incorporated into the dispatch process for several reasons. One is that the solver technology and practical computer hardware of 20 years ago was probably inadequate for the task. Even without that constraint, which is probably not present today, there are some challenging modelling issues to be addressed.

Another major reason is the nature of reactive power and voltage. The effects of reactive power on voltage are local and the effect on real power flow relatively weak, at least within normal voltage bounds. The same applies to system strength, which is generally a local issue.[7] It is therefore difficult to imagine a competitive market in reactive provision and voltage control. For this reason, the general approach to reactive power provision and voltage control has been to impose technical standards.

Despite these challenges, a dispatch process that frees up the mix of real and reactive power provision offers potential benefits to the market. Where reactive power limits active power flow or unit dispatch there may be opportunities to reduce costs. How this could be approached in practice will be discussed further in a later section.

3.6 Assessment Against Market Principles

Table 2 following contains an assessment of the selected features of the current dispatch and pricing engine against the market principles in Appendix A. The clearest failures are the failures to price losses and contingency FCAS. The requirement for contingency services is generally not co-optimised with energy dispatch and in any case the cost of the requirement is smeared rather than allocated to the parties who could act to avoid those costs. Other areas could be improved although there are likely to be divergent views on some of them.

One matter likely to become more pressing in future is the ability to define offers more flexibly, and for AEMO's forecast demand to better reflect demand price responsiveness in some way, without AEMO becoming an active market participant. We would expect demand elasticity under the NEM upgrade package (that is proposed in this paper) to be more predictable and measurable.

[7] For AEMO's description of the system strength issue, see https://www.aemo.com.au/-/media/Files/Electricity/NEM/Security_and_Reliability/Reports/AEMO-Fact-Sheet-System-Strength-Final-20.pdf

Table 2: Pricing and Dispatch Engine Conformance to Market Principles

No.	Principle (summarised)	Comment on NEM Approach as of March 2017	Assessment
1	Commercial freedom for participants	• Commercial flexibility limited by allowed structure of offers. • Onerous obligations to comply with dispatch instructions irrespective of the need or cost of doing so. • Technical standard approach to reactive and voltage control may be too restrictive and inefficient.	• While the NEM generally rates well on an international scale, commercial freedom in some areas could be improved to allow costs to be minimised.
2	Market transparency	• Generally good, but processes for generating MLFs and security constraints can be opaque and somewhat subjective.	• MLFs and constraint generation could be made dynamic and reflective of current market conditions.
3	Technology Neutrality	• There are proposals, such as the Demand Response Mechanism, to attempt to force more demand-side into the scheduling regime.	• Improved pricing rather than enforcing scheduling and imposing overly-rigid technical standards could make the NEM more open to newer technologies and give more flexibility to existing technologies.
4	Consistency between central dispatch and pricing	• Generally well managed at the 5 minute boundaries but distorted by pricing rules and lack of a real time price.	• Greatest improvement is from revamping the settlement arrangements.

No.	Principle (summarised)	Comment on NEM Approach as of March 2017	Assessment
5	Equal access for existing and prospective Market Participants	• This should be interpreted to include non-scheduled, small scale participation.	• Small scale participants responding to market signals should not be required to be scheduled.
6	Ancillary Services market to be competitive and dynamic	• Current enablement markets meet this requirement, but definitions, standards and systems create artificial barriers to participation.	• FCAS Enablement markets should be supplemented with a real time FCAS usage market to improve performance and competition.
7	Section 116 of the National Electricity Law not affected	• Not applicable	• Not applicable
8	Ancillary Service Costs allocated so as to be avoidable if possible	• This requirement is not met at all for FCAS contingency markets and only partially for FCAS regulation markets.	• Both these elements could be greatly improved as recommended in this report.
9	AEMO to settle Ancillary Services market	• Done at present • Any bilateral contracting for these services should not be settled by AEMO	• Any improved arrangements would continue to be settled by AEMO. • AEMO could allow for the “reallocation” of FCAS settlement amounts or exposures between Participants.

4 Fees and Charges

The efficient way to allocate the cost of ancillary services is most clearly specified in the market design principles set out in Chapter 3 of the Market Rules and reproduced in Appendix A. The part on ancillary service cost allocation is set out below. The same principle can be applied to the allocation of fees generally.

3.1.4 (a) (8) where arrangements require participants to pay a proportion of AEMO costs for ancillary services, charges should where possible be allocated to provide incentives to lower overall costs of the NEM. Costs unable to be reasonably allocated this way should be apportioned as broadly as possible whilst minimising distortions to production, consumption and investment decisions;…

In the previous section we identified FCAS services contingency as being too broadly allocated when it is possible to allocate them to be more avoidable. To a lesser extent FCAS regulation cost allocation could also be improved. These possibilities will be reviewed later in this report.

For fees and charges more broadly, the current and developing context is relevant. Established generation has come under pressure from renewable penetration and older coal units are being shut down. Right now, tight gas supply means that these plants are not being replaced with more flexible plants. Therefore, it is worthwhile re-visiting these cost allocations to see whether some fees and charges, other than the FCAS charges referred to above, might also be re-allocated to lower the burden on generation or to improve market outcomes more generally.

One area which could be improved is the allocation of market fees. These are allocated broadly to registered market participants and, understandably, these participants apply pressure on AEMO to keep those fees to a minimum. While pressure to raise efficiency is a good thing, AEMO also has an important role in improving market systems to meet emerging needs. In fact, there is and has always been a specific reference in the Rules to market improvement activity, specifically on the workings of the dispatch and pricing engine. The current wording of this rule compared with an earlier version is reproduced in Appendix B.

In the earlier NECA version of the rule, NEMMCO (predecessor to AEMO) was required to (“must”) investigate scope for the development of the dispatch algorithm. Somewhere along the line this was changed to “may”, which is where the wording remains today. While the dispatch algorithm has been tweaked here and there in the interim, no major upgrades have occurred for a long time. NEMMCO merged with the transmission planning entity VENCORP, diffusing its focus from solely market and system operations to transmission network planning and gas market and network development.

The AEMC, despite its rhetoric, is appearing to focus on regulated outcomes, such as the recent approach in the realm of system security, rather than competitive approaches based on marginal pricing principles. In any case, if the market operator, the specialist in market systems, has a propensity to focus on regulated, network and central planning options it is likely that the Commission will simply follow.

In light of this, there is a case that the market development task should be funded by consumers or governments, and that AEMO and AEMC be given a fresh mandate to be far more pro-active in researching opportunities for improvement than they are at present.

In any case, existing participants will need to invest in their own systems upon any revisions to NEMDE and the settlement approaches. These internal costs could be offset by some measures to remove the inefficient fees participants are presently exposed to: AEMO's costs of market operation presently charged as fees.

If we expect the NEM to be efficient, in that participants can make efficient decisions at the margin to produce, consume or invest in electricity services, it is only acceptable to allocate costs to participants when participants can change behaviour to avoid them. Clearly this is not the case for AEMO's routine fees, which should therefore be allocated on a fixed fee basis to the most inelastic demand – residential consumers, by connection point. Pricing as a fixed fee per customer connection point is the least distorting approach. In addition other fixed costs, such as System Restart Ancillary Services, should be allocated in the same way.

5 Improving the NEM Auction within the Half Hour

5.1 Overview

As noted Section 2.5, the current pricing arrangements within the half hour in the NEM are the outcome of a set of compromises involving energy and FCAS activities priced in quite different and internally inconsistent ways,. Specifically, they offer no coherent marginal pricing signals within that timescale and are therefore poorly set up to support an efficient market-oriented approach to the current and future needs of the electricity system.

The AEMC is currently considering a range of proposals that relate to this field, including 5 minute settlement, the power system security review and the scheduling of loads and non-scheduled generation. AEMO is also reviewing the FCAS regulation "Causer Pays" procedure. While AEMC recognises there are relationships between these activities, to date there seems to be no integrated approach to address them as a package.

The package to improve marginal pricing within the half hour, to be described in this section, has the following useful properties. It:

- deals with energy and frequency control in a coherent package;
- supplements but does not replace AEMO systems, at least initially;
- is quickly and relatively easily implemented within AEMO operations;
- initially involves few changes to participant systems, other than those they might wish to change to improve their operations;
- can be phased in gradually to allow market participants to adjust their operations progressively as may be desirable;
- has good error properties when SCADA is used in a first stage; and
- supports extension in a further stage to faster response services, including FFR and inertia, by using specially programmed revenue metering.

For the purpose of presentation, each element of the intra-half-hour part of the package will be introduced progressively in the following sub-sections, before describing the integrated package and the scope for future development.

5.2 A Ramping Ancillary Service to Fix the 5 Minute/Half Hour Problem

This is a well worked-over issue which is currently undergoing further AEMC review under the name of "5-minute Settlement". This rule change proposal, by Sun Metals, is currently meeting strong resistance from most incumbent participants.

The general proposal is to use SCADA data available for scheduled participants to calculate a half hourly a dispatch volume weighted price (different for each scheduled participant) which is then used for half hour settlement. Non-scheduled participants would be treated as residuals for the time being although this is a point of contention. The effect is to convert 5-minute dispatch prices into the actual prices used for settling energy in that 5 minutes.

One problem for participants could be the label "5 minute settlement", which implies a major upheaval of existing systems and practices, especially with settlement systems and contracting. We propose an alternative implementation which deals with some of these concerns. Under this approach, we:

- leave current half hourly settlement logic alone;
- introduce a Ramping Ancillary Service (RAS), which, conceptually, is settled at the difference between 5 minute prices and volumes and half hourly prices and prices and volumes *as measured by SCADA*; settlement for a participant is then:
  - The settlement amount as currently determined, plus
  - The settlement amount for the RAS ; and
- measure MW with SCADA or, in future, with suitably programmed revenue metering.

The RAS settlement can be expected to be relatively small compared with the half-hour settlement. This approach confines any errors in SCADA measurements to the relatively small RAS component only. Further, we can show that any error in the RAS settlement amounts is proportional to meter scaling error rather than to any offset error, as we can show that the RAS settlement is based simply on measured MW differences [8]. Also, any lags of a second or two in measurements are not significant at the 5-minute level.

While the RAS settlement carries these errors when real time SCADA measurements are used, current practice would ignore that RAS component altogether. Clearly, use of real time metering from SCADA is acceptable for an initial implementation to improve marginal pricing within the half hour. While a RAS approach and a dispatch weighted price approach would deliver a similar outcome, the RAS approach has the following advantages:

- it leaves existing AEMO and user systems largely untouched, although an additional facility would need to be handled;
- implementation by AEMO is of a similar order of complexity as Causer Pays and therefore relatively straightforward;
- being separately settled, metering errors affecting settlement would be confined to the RAS. The effect of these errors is well defined;
- the RAS easily supports enhancement to smooth out prices and to take account of real time pricing of FCAS usage and precision, as will be described in following sub-sections; and
- a separate RAS can be phased in transparently at levels of, say, 0%, 20%, 40%, 60%, 80% and 100%, to allow participants time to learn how to adjust their operations.

5.3 Smoothing Price Shocks

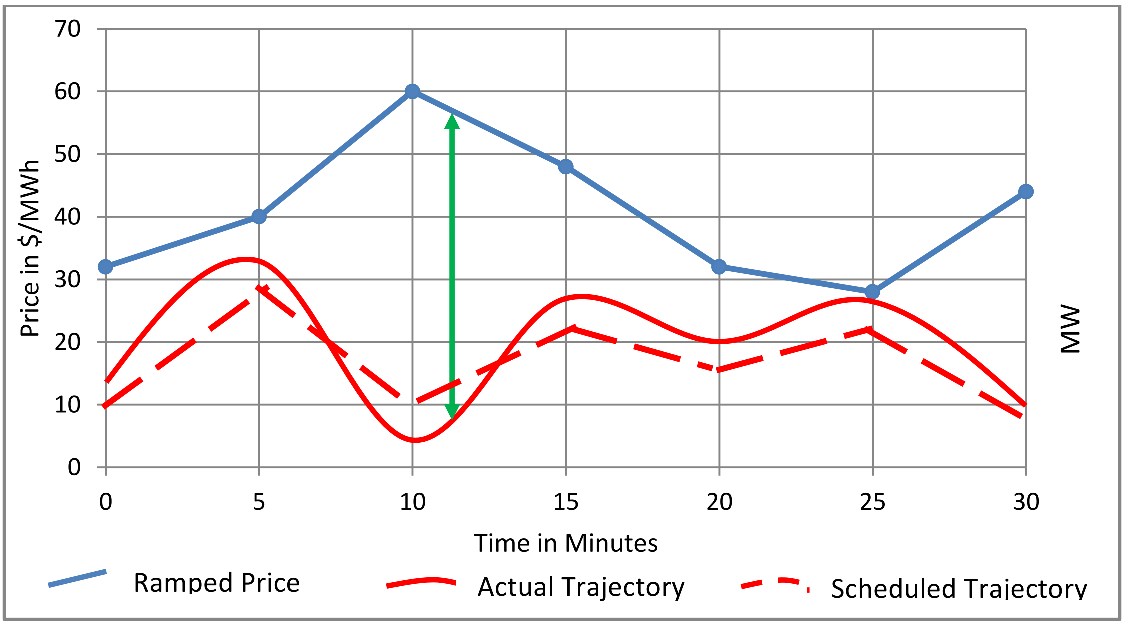
As argued in Section 2.2, moving to an effective 5 minute settlement arrangement actually magnifies and sharpens the price transitions at settlement boundaries. In time, one can

[8] Analysis of settlement errors when a dispatch weighted price is used shows a similar result, but the settlement error is based on the meter offset error rather than the scaling error. As a result, the error under that approach increases as the measured MW become small relative to the range of the real time meter and, presumably, the size of the facility.

expect large responses from unscheduled generators with potential for some price or even system instability if not addressed.

We deal with this problem by ramping price from one 5 minute dispatch boundary to the next, as shown in Figure 5. Assuming a linear price ramping model as proposed in Section 2.2, we need to deal with the ramped price trajectory (in blue) and real time measured load or generation trajectory (in red) to calculate a settlement amount. The scheduled trajectory is shown dotted in red.

Figure 5: Load or Generation Trajectory with 5 Minute Ramped Pricing



We can carry out the core calculations in real time (either using SCADA or within a specially programmed revenue meter) by accumulating certain measured values for each 5 minutes for later settlement using 5-minute dispatch price data.

The essence of the task is to simply calculate, for each small time interval (say 4 seconds if SCADA is used), the product of price, MW quantity and interval duration (delt = 4 seconds), and to accumulate those values over each 5 minutes. The green two-sided vertical arrow in the diagram indicates the price and MW quantity to be multiplied. To avoid the need to use 5 minute prices and the risk of temporary communication failures (if implemented in a local revenue meter), we can accumulate two quantities, Q1 and Q2, each 5 minutes:

Q1 = sum(x(alpha)*delt)

Q2 = sum(alpha*x(alpha))*delt)

Where:

Q1 and Q2 are the accumulated quantities

alpha is the fractional time along the 5-minute dispatch interval x(alpha) is the measured load/generation at the point alpha along the interval

delt is the chosen measurement interval. This would be 4 seconds for SCADA and as low as, say, 100 milliseconds in a specially programmed revenue meter [9]

* is multiplication and sum(.) is the summation of values over the interval

This is a trivial addition to the simpler task of simply accumulating MW values (the quantity Q1 above) if stepped pricing logic is used. Then, at the time of settlement it is easy to show that the gross settlement amount within the 5 minutes, S5gross, is:

S5gross = pstart*Q1 + (pend – pstart)*Q2

where pstart and pend are dispatch prices at the start and end of the dispatch interval.

To get the half hour RAS settlement over the half hour we:

sum the gross 5-minute settlement values for each 5 minutes

subtract the half hour average 5-minute ex ante price multiplied by real time measured energy summed over the half hour.

When we add back the existing energy settlement we get the gross settlement again, but with any real time metering error restricted only to the RAS component of the settlement.

The benefit of the ramped price approach is an incentive on market participants to respond to prices at a range of times and extents rather than to focus on dispatch interval boundaries.

However, to make price responsiveness truly robust, we need to send a price signal which feeds back to participants the frequency and time error state of the system, in order to encourage them to stabilise those errors.

5.4 Maintaining Frequency Stability: Stage 1 – Regulation

Maintaining frequency within narrow bounds under normal operation and after contingencies is a core requirement for system operations. As discussed in Section 2.4, any upgrade to marginal pricing principles within the half hour must be able to integrate the treatment of FCAS smoothly into the new arrangements.

Consider the task of controlling a system to achieve a balance between frequency stability and cost, given an initial indicative trajectory and assuming ongoing small disturbances. There is a standard theory (linear quadratic control combined with Kalman filtering) that fits our requirement to control frequency and time error with resources that have energy and ramping costs, or indeed, any other form of linear dynamics and quadratic costs. The approach also delivers marginal prices for energy. Further, the control strategy can be implemented locally with limited and locally measured quantities (frequency error, time

[9] The value of such a short measurement interval is more evident with the real time FCAS usage market discussed in the following sub-section.

error and the state of one's own plant). It can account for and reward inertia and load relief. The essence of the approach is as follows:

- we model the system as states; states include frequency error, time error and the level of generation or load by different units or "virtual" units, as we also assume that it takes time and expense to move from one level of generation or load to another;
- the system moves from one set of states to another over a given time interval, driven by physical dynamics and controls applied (e.g. ramping a generating unit up or down);
- the system is subject to random noise that disturbs the operating trajectory;
- we have limited measurements to estimate the current state of the system, in this case we assume frequency and time error only, to which we could add one's own state;
- all generation and ramping levels are relative to some set point (e.g. for frequency and time error) and incur costs per unit time which increase quadratically around a mean of zero. A linear cost term can be included but can be added later; and
- the aim is to find a control strategy (controls as a function of measurable state) that minimises costs, which are the sum of the costs of deviations from set points (frequency, time error and control costs) over some long time horizon.

While this may seem like a centralised control system[10], it can be decentralised to some degree through the prices that emerge from the system model. A centralised implementation with a focus on regulation - a first stage - could operate as follows:

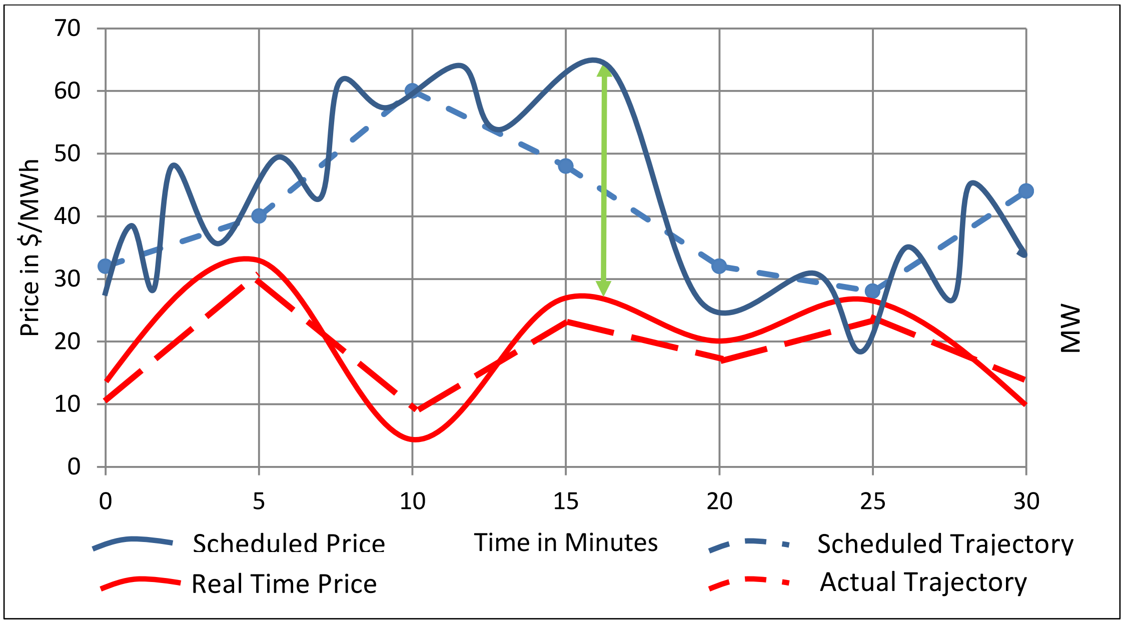
- AEMO models the system as described, noting that only underlying dynamic responses need to be captured, not necessarily individual units. This will lead to a control strategy based on the recent history of frequency and time error. It should be very similar or identical to the control strategy currently used by AEMO for regulation FCAS;
- the dynamic price from the control strategy is then used to reward and charge participants for performance. This is the essence of a two sided market as if we treat currently non-SCADA metered units or loads as residuals; and
- while the units are centrally controlled, the pricing and settlement based on performance would reward good response.

Figure 6 Illustrates settlement logic with the additional real time price component included. Instead of settling on the straight line ramped price, we settle on the ramped price with an additional real time price superimposed. The real time price component works to stabilise frequency and time error. Settlement is based on the product of price (solid blue line) and MW volume (solid red line) as indicated by the green double-headed arrow. These values are summed over the dispatch interval and, eventually, over the trading interval. The dotted lines show the corresponding scheduled prices and trajectories.

The logic is very similar to that described in Section 5.3, but would now include a real time price variation. The precise details would depend on some implementation decisions, such as whether real time price variations are scaled by dispatch prices or not.

[10] AEMO essentially uses it already for controlling enabled regulation.

Figure 6: Load or Generation Trajectory with Real Time Pricing



To illustrate how all this could work, we have built a simple demonstration system based on toy data. The system consists of:

- a virtual flat load that is subject to random noise with a mean of zero;
- a virtual generator that can ramp and operate away from set points to control frequency, but at (an additional) cost to ramp and operate away from a set point; and
- an objective designed to minimise the expected value of a weighted combination of frequency and time error costs and the costs of ramping and operating of the generation away from set points.

Figure 7 shows traces for frequency error and time error over 15 minutes. Each run will give a different trajectory because of the random noise in the load. We assume that deviations at the end of each 5 minutes are ramped back into the energy market, so they do not grow and so that control is smooth. As deviations are ramped back into the energy market, new deviations occur and are controlled.

The figure shows frequency being controlled around a set point of zero (corresponding to 50Hz). We have included a correction term for time error. This lags somewhat behind the frequency. Note that time error reaches a local extreme value when the frequency error passes though zero, as expected.

Figure 7: Frequency and Time Error Traces for Toy System

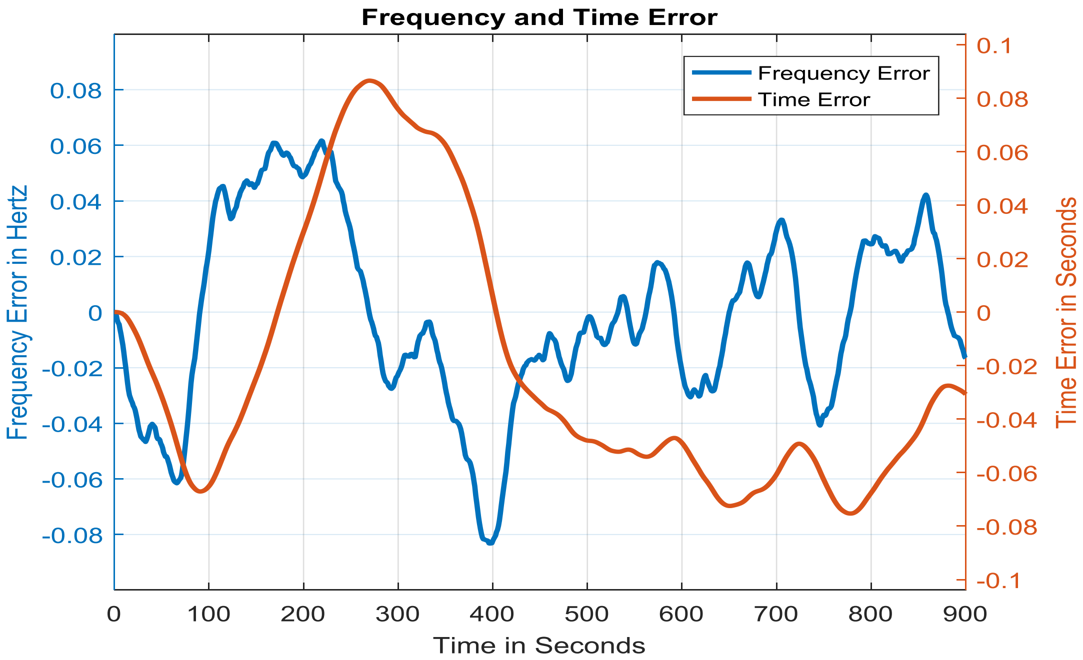


Figure 8 illustrates a similar run over a period of two hours. Success in controlling both frequency and time error over that longer period is evident.

Figure 8: Frequency and Time Error Traces for Toy System

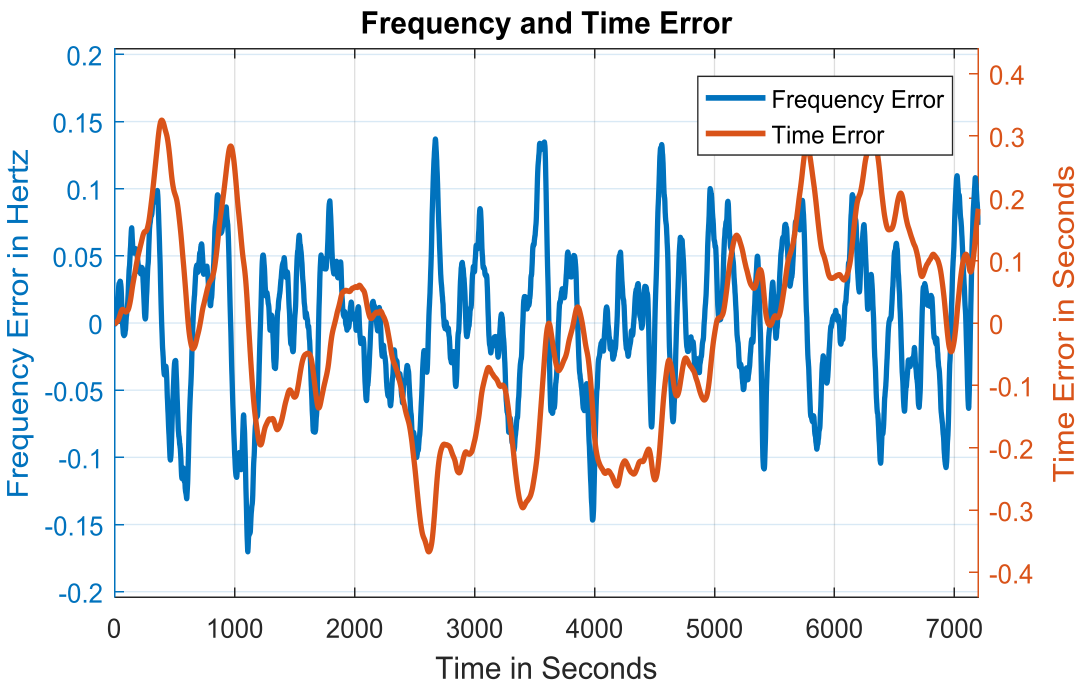
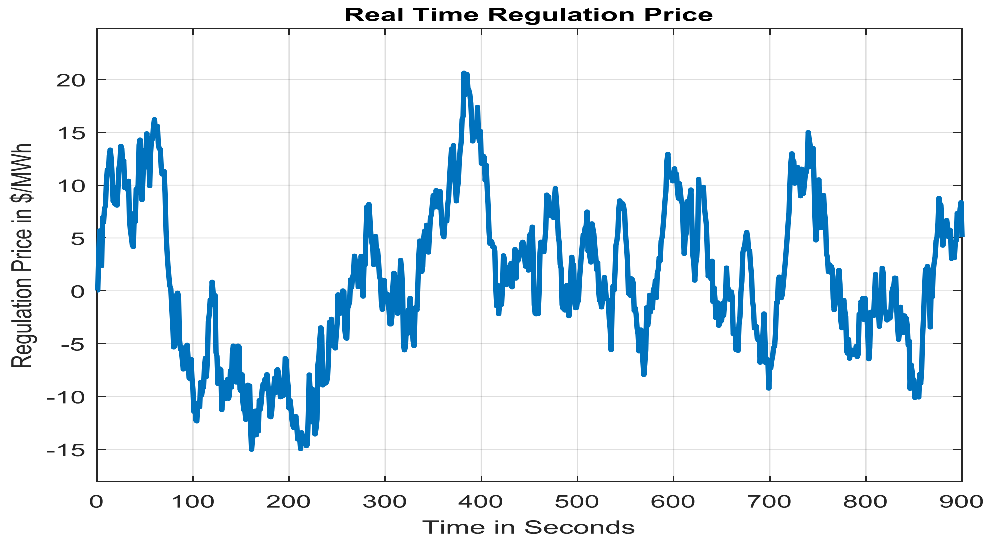


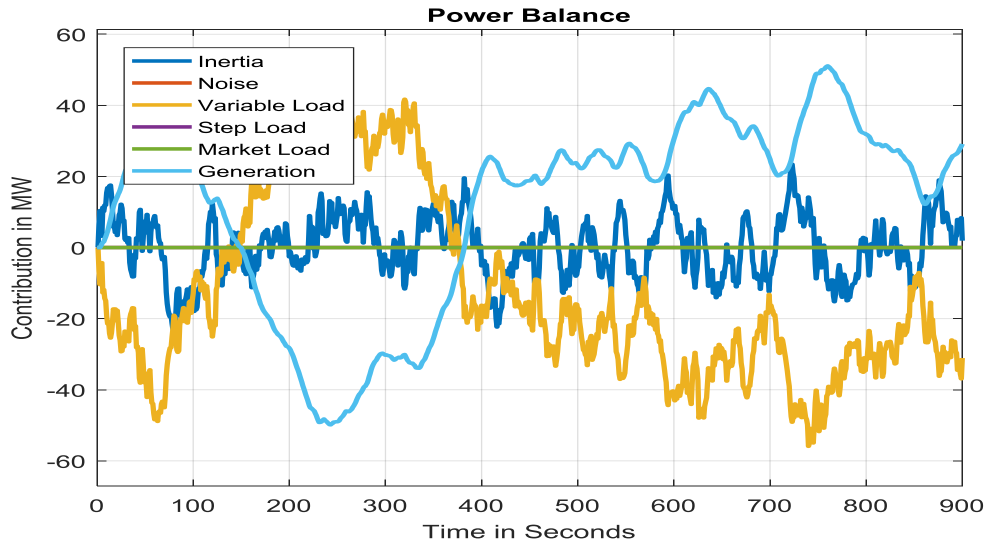
Figure 9 shows the real time price for the data assumed. Note that the real time price discussed here is a variation around the price trajectory established in the 5-minute dispatch process. As would be expected, there is an approximately inverse relationship between price and frequency deviation. For example, if frequency increases, the price tends to dip to encourage the generation to back off. However, this relationship is not rigid as there are ramping costs in play as well. In a more complex system with different speeds and costs of response, the response patterns would likely be more complex again.

Figure 9: Real Time Price in Toy System



To illustrate how the frequency is being controlled, Figure 10 shows the contributions of power from different system elements. Elements marked as “Noise”, “Market Load” and “Step Load” in the legend are not part of this run.

Figure 10: Power Balance in Toy System



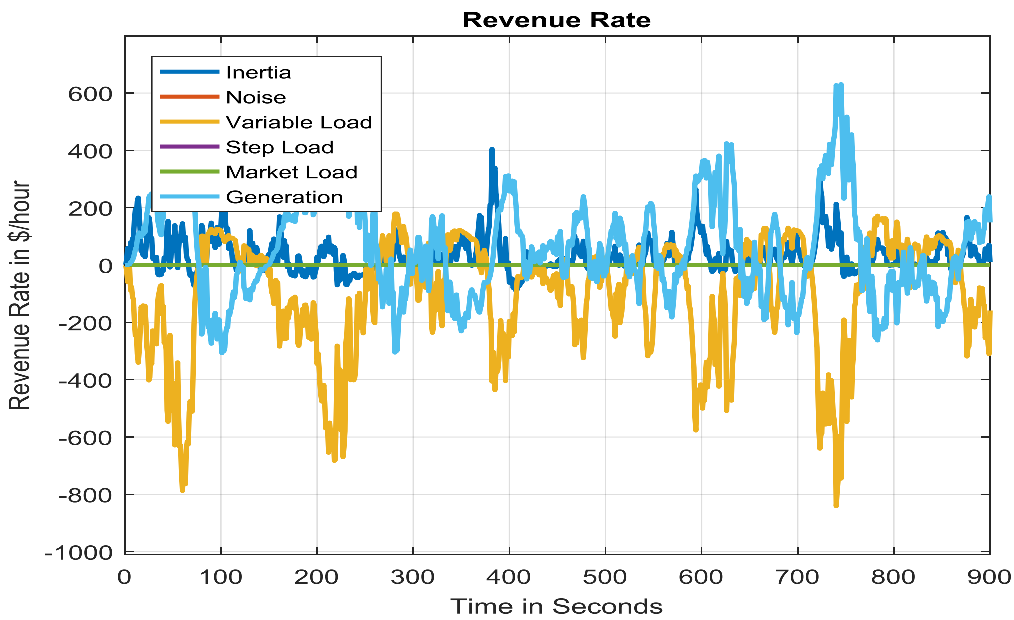
In this run, the load (in yellow) is essentially autonomous and is subject to random noise as is evident. Load deviations from the noise injection do not grow indefinitely because, at the end of each dispatch interval, any deviation is ramped back into the energy market over the next 5 minutes[11] . More deviations occur during that process but they stay within manageable bounds. The regulation process works along similar lines at present, with some differences in detail.

In the chart, we see two elements responding; controlled generation (in light blue) and inertia (in dark blue). In fact, these two components are part of the same unit and electrical measurement would not separate them. They only appear separately in the dynamic modelling. In essence the light blue represents the mechanical energy input, while the dark blue inertia represents the energy stored in rotation, taking energy from the system (when accelerating) and contributing energy (when decelerating). In this simple lossless system, the sum of the mechanical generation and inertia power contributions precisely balance the load[12], in all cases relative to a reference trajectory.

In this run the generator is undertaking a regulation function. We note that the generator's inertia tends to absorb the very short term load fluctuations, while the generator's mechanical power input tracks at a slower pace, exhibiting a smoother trajectory.

The settlement outcome when the real time price is applied to the generation (including inertia) and load is illustrated in Figure 11. The settlement amounts are expressed as rates ($/hour) for illustrative purposes.

Figure 11: Rate of Revenue Accumulation in Toy System



[11] The load itself is also modelled as having mean reverting behaviour, but this behaviour may not be evident over a 5-minute dispatch interval.
[12] This is a statement of Newton's second law of motion applied to a rotating system, as modelled in the system.

The chart shows that the load is consistently on the negative side of the ledger; it is disturbing the system rather than controlling, and is a net "causer" as in the causer pays mechanism. The inertia (dark blue) and mechanical generation (light blue) are generally on the positive side of the ledger, because inertia (passively) and mechanical generation (actively) are working to maintain frequency. As the power variations including inertia contributions are balanced, the settlement amounts are balanced also.

Valuing system frequency and time errors in real time will encourage participants to control frequency when it is cost effective for them to do so, and not otherwise.

5.5 FCAS Contingency Services under Stage 1

In the in the previous sub-section, the workings of a real time control and pricing system (which could evolve into a market) were discussed in the context of an example where the task was to control second by second and minute by minute small variations in the system; regulation function. However, such a system would also respond to the frequency fluctuations caused by larger contingencies. As we have not included in our toy system any very fast response options we would not expect ideal response patterns, but the system should nevertheless do what it can to keep frequency and time errors stable.

Figure 12 shows the response pattern of frequency and time error after a 600 MW loss of generation contingency occurs in the system. The broad pattern of behaviour is stable, although the frequency droop is larger and the recovery time longer than might be desirable because our modelled response options are limited. If the system and the corresponding control and pricing model were to include fast response options, and if metering were adequate to capture those responses accurately, a much improved outcome could be expected.

Figure 12: Frequency and Time Error after a Contingency in Toy System

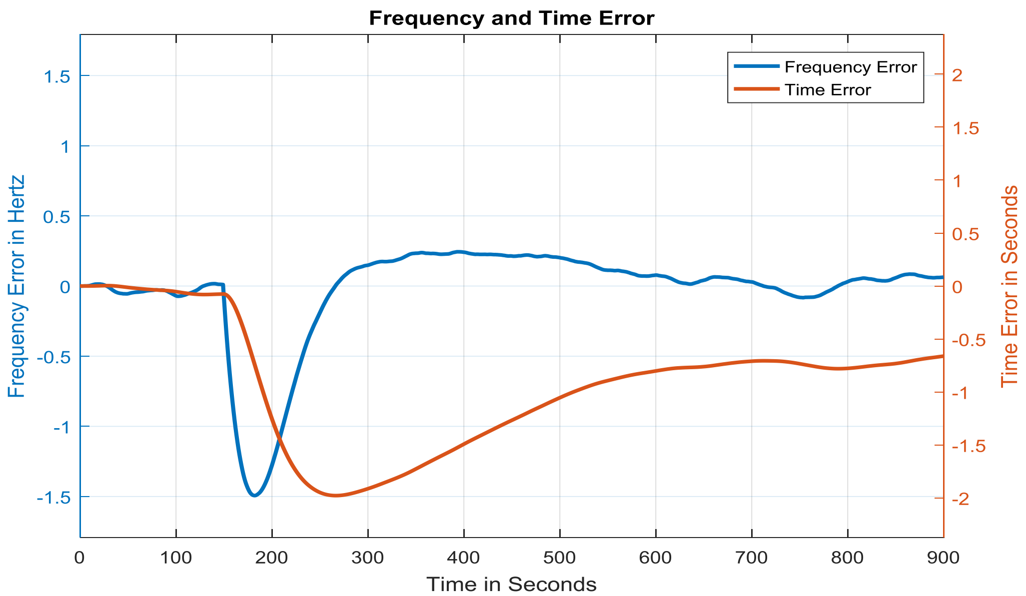


Figure 13 shows the resulting real time price pattern. As would be expected, there is a sharp peak corresponding to the contingency, which falls away as the frequency error is corrected and the new generation pattern is ramped back into the energy market as illustrated the next figure. Note that the pattern of prices in this toy example may differ markedly in a real system.

Figure 13: Real Time Price after a Contingency in Toy System

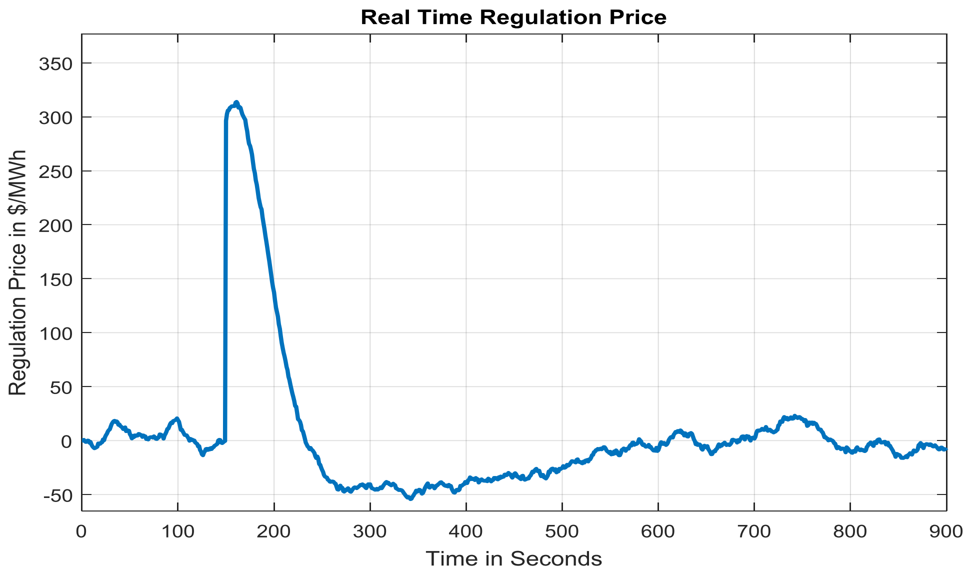
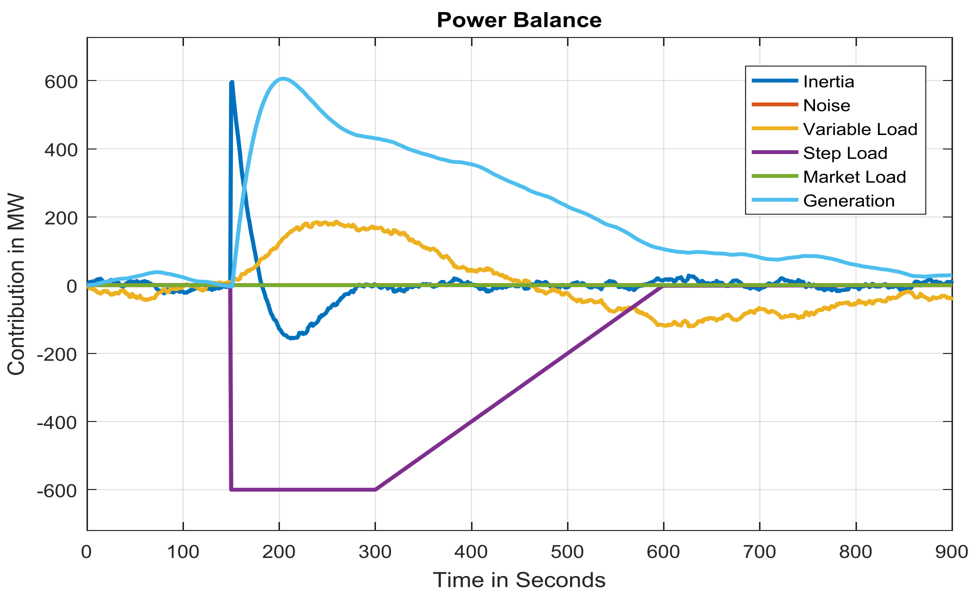


Figure 14 shows the power balance in the system from the contingency. The contingency is shown in purple, including how its effect is ramped back into the energy market over the following dispatch interval.

Figure 14: Power Balance after a Contingency in Toy System

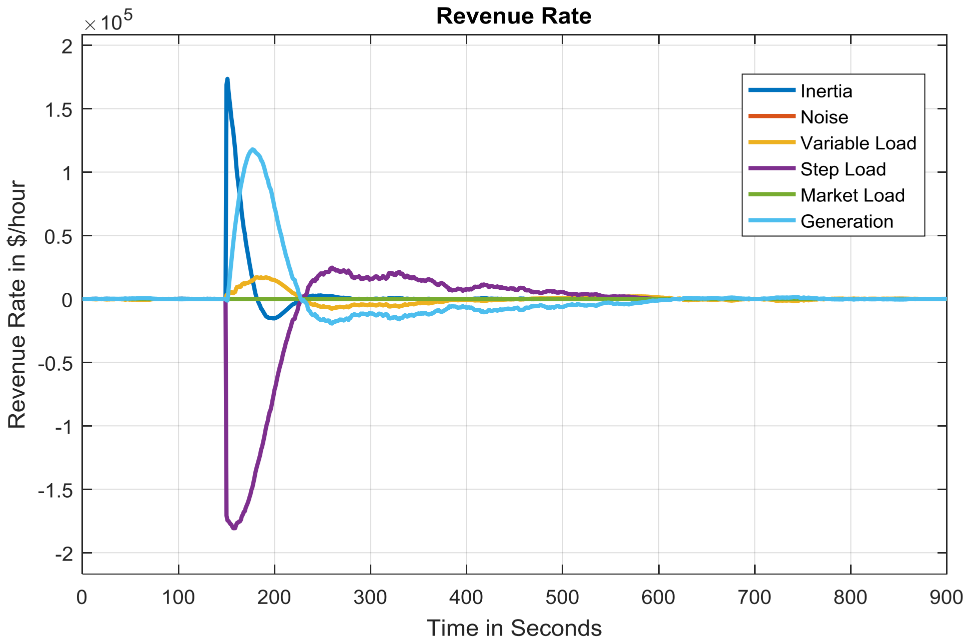


Several interesting features can be observed from this chart. First, the lost power input is entirely absorbed by the system inertia (belonging to a single modelled generator in this case) in the immediate aftermath of the contingency, as shown by the dark blue line. Next, we can see some frequency response in the variable load line (in yellow), as we have modelled load relief in the control and pricing model. The generator is also responding fairly quickly in this case. Everything gradually returns to a new equilibrium over the following few dispatch intervals.

Finally, we show the rate of revenue accumulation in Figure 15. Clear features are:

- the unit suffering the contingency absorbs the cost of the contingency;
- The inertia component of generation earns a significant part of the revenue in the immediate aftermath of the contingency;
- The (mechanical) generation earns a positive amount after the inertia falls away; and
- The positive and negative revenues are balanced.

Figure 15: Rate of Revenue Growth in Toy System after a Contingency



A key observation is that revenue and costs are allocated where they should be and that inertia is a significant contributor, both physically and financially. It follows that this pricing mechanism offers a pathway to encourage inertia into the market. Whether it is a complete solution is another matter, but it can certainly contribute and may well be sufficient in many cases.

There are design questions to be resolved on implementation of the proposed real time market. These and other issues should be investigated in a prototyping environment.

- Should the real time price be scaled by the current regional energy price, or operate autonomously across the whole NEM (except when there is islanding)?
- How should units be controlled (is AGC relevant) and would compliance obligations (such as Rule 4.9.8) become redundant in time?
- Is there an implied contract when the 5-minute dispatch target is set?
  - ✓ Our view is "no", because that approach places too much market weight on the accuracy of AEMO's load forecast. Our settlement logic takes the approach that the dispatch process sets a trajectory for a reference price around which there can be real time variation, but there is no implied contract.

5.6 Scope for Further Development: Stage 2

In a second stage we can remove the assumption of central control and reward and charge for performance even if not under central control. Further we aim to support a real time market in fast frequency response and inertia.

To implement this stage, we use specially programmed revenue meters that can take local measurements of frequency and power down to sub-second (programmable) time intervals[13]. Such meters could support markets in faster acting responses such as inertia and fast frequency responses because:

- they will pick up and measure sub-second responses from the inertia of units as well as fast frequency responses;
- measurements matching the frequency based price formula with MW will be valid for very short time intervals because SCADA lags will be eliminated; and
- real time logic (using local frequency) will be programmed into the meter for later use in settlements.

The pricing algorithm used will need to reflect the presence and relevance of fast frequency response options. The role of Inertia is already inherent in the price modelling. Apart from the sub-second measurement interval, the pricing logic is otherwise very similar to that used in the SCADA implementation of Stage 1.

A Stage 1 implementation would look very similar to the current causer pays logic operated by AEMO which is, by AEMO standards, a simple system. A Stage 2 decentralised implementation is logically similar but requires communication infrastructure to make it work. An indicative configuration for a participant operating in a decentralised mode is shown in Figure 16. The figure shows the following elements:

- a plant which is subject to control;
- a (participant operated) controller;
- a specially programmed meter;
- a Meter Data Provider (MDP) responsible for the plant's metering; and
- AEMO as the market operator.

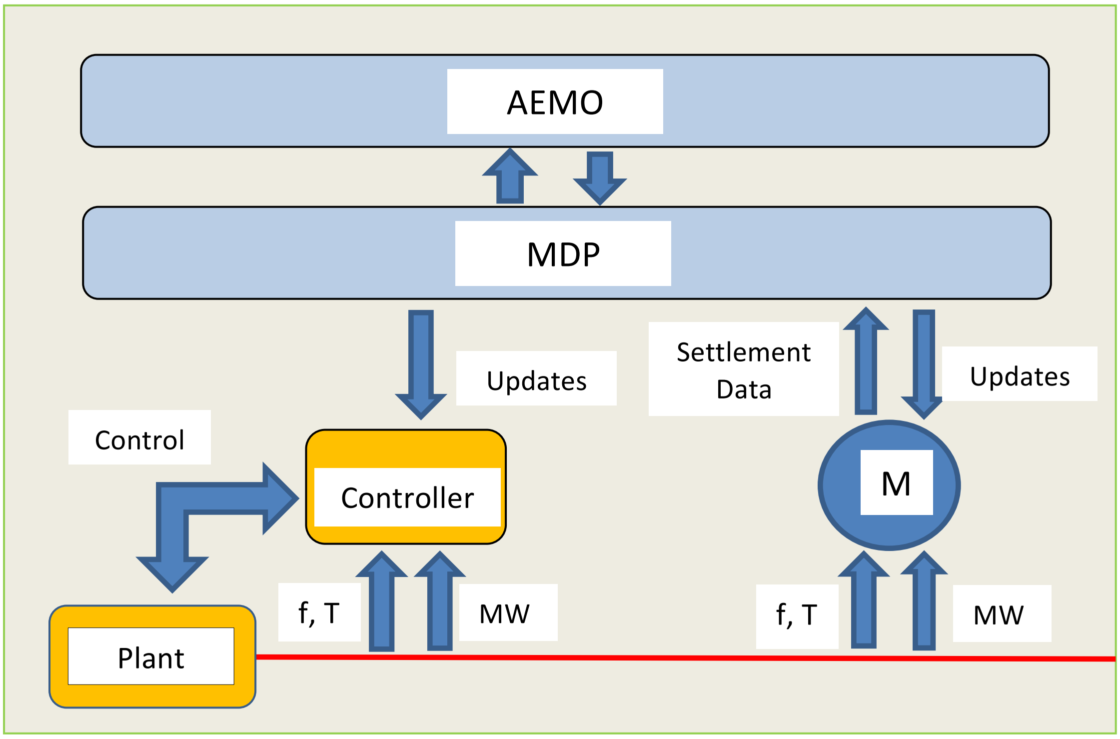
[13] Electronic Revenue meters already use software to calculate and accumulate instantaneous real power and other electrical values at intervals in the order of a millisecond or less. A built-in real time pricing algorithm is an extension of that concept.

The specially programmed revenue meter measures (or accesses a measurement of) local frequency error (f), time error (T) and MWs as well as other electrical values. It processes these data in real time to produce 5 minute accumulated values for settlement. These values are periodically uploaded to the MDP and then to AEMO for settlement. Periodically, AEMO downloads updates to the MDP and then to the participant meter to:

- update the local pricing algorithm as required;
- update the meter clock; and
- update the current recorded meter time error to correct for drift.

The participant's controller receives similar updates and also regular dispatch prices and other standard AEMO data such as ST PASA (not required for the revenue meter). It also measures or accesses MW, frequency and time error (not necessarily to the same accuracy as the revenue meter). From these data it can reproduce the pricing in real time and determine a control for plant operations. The electrical outcome is ultimately recorded by the revenue meter in a form suitable for settlement.

Figure 16: Configuration for Participant in a Decentralised Real Time Market



Note that the data going to and from AEMO and the MDP are not required in real time. The strictly real time activity including measurement and control is all handled locally. In this way, very fast responses including inertia can be properly handled.

6 Improving the NEM Dispatch Engine (NEMDE)

6.1 Optimisation Technology

The potential improvements to the dispatch and pricing process outlined and discussed in Section 3 all require changes to the NEMDE software system. NEMDE is an optimisation model based on the technique of linear programming. In broad terms it maximises benefits (equivalent to minimising cots costs) at the 5-minute time horizon, subject to technical, security and commercial constraints (e.g. on offer bands) of the system. It delivers a dispatch schedule as well as 5-minute market clearing energy and FCAS prices.

In linear programming models, all the relationships in the problem must be linear, but variables can and typically do have upper and lower bounds. The economist Harold Hotelling was an early critic of the potential value of George Dantzig's original Simplex (linear) solution algorithm, noting that the real world is non-linear. Despite this self-evident truth, since its invention in 1949 linear programming has proved to be a remarkably potent and widely applicable tool in many fields, not least in electricity markets. Non-linear elements in electricity markets, such as the modelling of losses, can be approximated with piecewise linear functions. Another example is the cost function of generators, which is modelled as a piecewise linear cost curve (or a piecewise constant marginal cost curve) rather than a (relatively) smooth one. These are workable approximations in most but not all circumstances.

Since the early 1980s a new class of "interior point" methods has emerged as a strong competitor to the Simplex algorithm. One advantage of these methods is that they can be relatively easily adapted to non-linear problems. Today, very large scale problems both linear and non-linear, can be reliably and accurately solved. With this powerful new technology, new approaches can be envisaged which would have been considered impractical when NEMDE was first developed in the late 1990s.

6.2 Power System Modelling Approaches

We can classify power system models (which lie at the core of any electricity market dispatch and pricing engine) into three main types.

- Regional system model;

  NEMDE implements this approach. The system is divided into regions, within which network security constraints seldom apply. Only interconnectors between regions are modelled, which in turn are also compressed into single "virtual" lines. Simple energy conservation rules are applied. Fixed marginal loss factors are applied within regions, which must be determined off-line. Security constraints must be expressed as a combination of injections and offtakes in the model and must also be generated off-line, although some dynamic elements have been introduced over time. Although the network is simplified, a dispatch and pricing model based on a regional simplification retains significant levels of size and complexity in the modelling of individual dispatchable units, their offer structures and network security constraints.

- Full network model using a DC power flow approximation

  The engineering equations governing AC power flow in a network can be simplified in many applications to an acceptable degree of accuracy. It turns out that if voltages are maintained within reasonable bounds by whatever means, the system can be simplified into a set of relationships that resemble much simpler DC power flow relationships (even though the flow is AC). The simplest version of this is a lossless model but we can also model losses with piecewise linear functions or with continuous non-linear functions if we use a non-linear solver.

- Full network model using AC power flow modelling

  If we wish to consider voltages and reactive power as market elements, or at least as elements that might be valued by market outcomes, we need to move to a full AC power flow model. We can then model voltages and reactive power explicitly, but a challenge is then how to value voltage deviations and the real and reactive power relationships within plant. The model is also “more non-linear” than a DC power flow model, and sometimes challenging to solve, although this is less of an issue today.

Why use an explicit network model? In the DC power flow case at least, there is a direct mapping between network line flows and injections and offtakes (exploited even in the regional model) so the full network is not computationally necessary. However, including the network does make the model a lot more transparent and can reduce the need for off-line calculations and associated approximations. It should also be noted that a non-linear solver also allows non-network elements of the dispatch engine to be modelled non-linearly with advantage, as will be noted in the following sub-sections.

6.3 Overview of Options

Table 3 below summarises how the options described in Section 3 could be implemented with different optimiser technology.

Table 3: Modelling Approaches Suitable for Proposed Improvements

Network Model	Regional	DC Power Flow	DC Power Flow	AC Power flow
Optimiser Technology	Linear	Linear	Non-linear	Non-linear
Generate Dynamic Marginal Loss Factors		X	X	X
Generate Dynamic Security Constraints	X	X	X	X
Optimise Contingency FCAS Cost Allocation	X	X	X	X
Smooth Out Offer Profiles			X	X
Value Reactive Power and Voltage Control				X

The possible variations are simplified as there are further subdivisions of solver technology and of dispatch model approaches, but we attempt to capture the main ones. The crosses in each cell indicate that implementation of the option is practical with that solver technology. Details and discussion follow.

6.4 Determination of Marginal Loss Factors

Calculation of dynamic marginal loss factors requires a network model of some sort to be included in the dispatch and pricing model. Line losses could be modelled in one of three ways:

- as piecewise linear line loss functions in a DC power flow model, using a linear optimiser;
- as continuous (usually quadratic) line loss functions in a DC power flow model, using a non-linear optimiser; or
- as continuous line loss functions involving voltage and angle differences in an AC power flow model, using a non-linear optimiser.

These approaches are presented in order of likely accuracy, but all would be significantly more accurate than a constant loss factor that is updated periodically using projected loads and generation. As noted in Section 3.1, the approach has the added advantage of eliminating periodic updates by AEMO based on projected loads and generation.

In a dispatch model solution in the absence of network constraints, the marginal loss factors (MLFs) would be the ratio of the marginal value of energy at each connection point, divided by the marginal value at the associated regional reference node. In fact, in this special case the outcome would be equivalent to nodal pricing. Where intra-regional constraints are binding, the effect of these constraints would be removed in post optimal processing, giving a regional rather than a nodal outcome.

Where MLFs are small, the effect of dynamic MLFs would also be small and relatively predictable, and so readily accounted for in contracting. Where flows are more variable and even tidal, MLFs would vary more, as they ideally should. Participants might then need to consider modified contractual arrangements.

An additional advantage of modelling losses with continuous, non-linear functions using a non-linear optimiser is that it resolves the so-called non-physical loss problem when using piecewise linear loss model approximations in a linear solver. Non-physical losses occur when prices become negative. A piecewise linear loss model then tries to use the highest loss segments first, which is not physically realistic.

If the losses are modelled as a continuous, non-linear function, this cannot occur. However, the system then becomes non-convex, a mathematical term which means there are potentially many local minima. This property is inherent to the situation. The natural approach in this case is to seek to move to the nearest local solution. This is not only likely to be nearly optimal (and very difficult to prove that it is not), but also practical in that negative prices seldom last very long and it is pointless to disturb the system more than necessary to deal with this temporary issue.

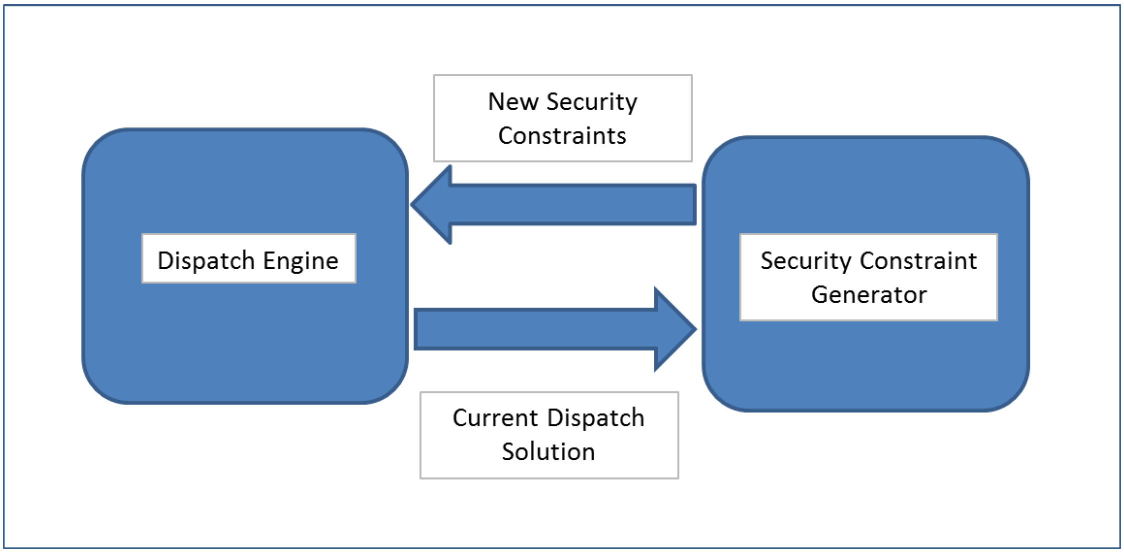
6.5 Treatment of Security Constraints

While constraint generation has been improved in NEMDE over the years, much work is still done off line as described in Section 3.2. Possible contingencies have to be anticipated and sets of constraints developed off line to deal with them. More recent designs of dispatch systems often have a dispatch model which typically a DC power flow model, arranged in a tight loop with a dynamic constraint generator, which will typically contain an AC power flow model as illustrated in Figure 17. The process typically works as follows:

1. The solution from the previous dispatch interval (say 5 minutes ago) and its constraint set are taken as starting points. The current constraint coefficients are refreshed as necessary to account for the new state.
2. The dispatch engine is solved with the new offers, other current data and the current constraint set.
3. The current solution is passed to the constraint generator which seeks out any security violations and generates constraints to deal with them.
4. If there were no security violations found, the solution is considered valid and the process completes; otherwise the new constraints are passed back to the dispatch engine and the process returns to Step 2.

With hot starts, a range of analytical smarts and diligent housekeeping, this process can be made very efficient.

Figure 17: Dynamic Generation of Security Constraints



This system does not actually require a network model in the dispatch engine; the security constraint generator has its own network model and can develop constraints in the generic constraint format if required. Constraints generates can take account of dynamic real time data or estimates of the state of equipment (e.g. Its temperature) either directly measured or estimated using a technique such as Kalman filtering.

6.6 Contingency FCAS Cost Allocation

While we envisage that most if not all of the financial flows in regulation FCAS could move to a real time market, contingency FCAS is more in the nature of insurance and some "premium" or enablement payment may still be required to ensure that the capability is present when needed. However, the real time market should lessen the additional incentive required to be available for service.

Given the contingency enablement services may still be required but at potentially lower costs than at present, a more efficient cost allocation mechanism can be envisaged. In fact, the following proposal is not new, although it has not been implemented thus far. It could be implemented in the current dispatch engine and any superset of it, with only minor modifications.

AEMO calculates a requirement for each contingency FCAS service that takes account of the largest current contingency and any automatic load relief if the contingency occurs. In broad terms, NEMDE includes relationships along the lines of:

FCAS_Requirement >= Size_of_contingency – Load_relief

The requirement must be large enough to cover every credible contingency. The contingency can be a generator, load or interconnector that would cause network separation. Where separation is possible, the load relief has to be based on the islanded condition. Setting aside the details, in principle there should be one such constraint for every contingency and impacted element in the system. In practice some would never bind and are omitted, but there could still be a significant set that could be active.

For a specific service, it is likely that the optimisation process will constrain the size of one or more contingencies to be of equal size. These constraints will have non-zero shadow prices and their sum will be the marginal cost of the service i.e. the marginal cost of supply will equal the marginal costs of all the critical contingencies and the market will balance. The shadow price of each binding constraint is the marginal contribution of each contingency to the FCAS cost. They need not be identical between contingencies, but the solution constructs them in such a way that the energy and FCAS costs together would deliver a dispatch consistent with participant offers, as is required unless an intra-regional network constraint is binding.

In asserting that the market will balance, we note that the expected load relief enters the equation and should be part of the settlement process. Where a network element is a critical contingency it should pay also, through a charge on the settlement residue. This line of thinking can be extended to services not yet part of the FCAS suite, such as fast frequency response.

Perhaps surprisingly, the approach could also be quickly and usefully implemented for the provision of inertia, with minimal new development. Take South Australia as a practical and very current example. In that region, a critical contingency that involves inertia in South Australia is a possible failure of the Heywood interconnector. AEMO now seeks to have sufficient inertia in South Australia to restrict the Rate of Change of Frequency (RoCoF) in

South Australia to not more than 3 Hz/second immediately after that contingency. How could this be enforced in the dispatch engine?

Examination of the relationship between acceleration, contingency size and inertia leads to a remarkably simple constraint that could be applied:

Size_of_Heywood_Contingency <= Constant x Inertia_in_SA

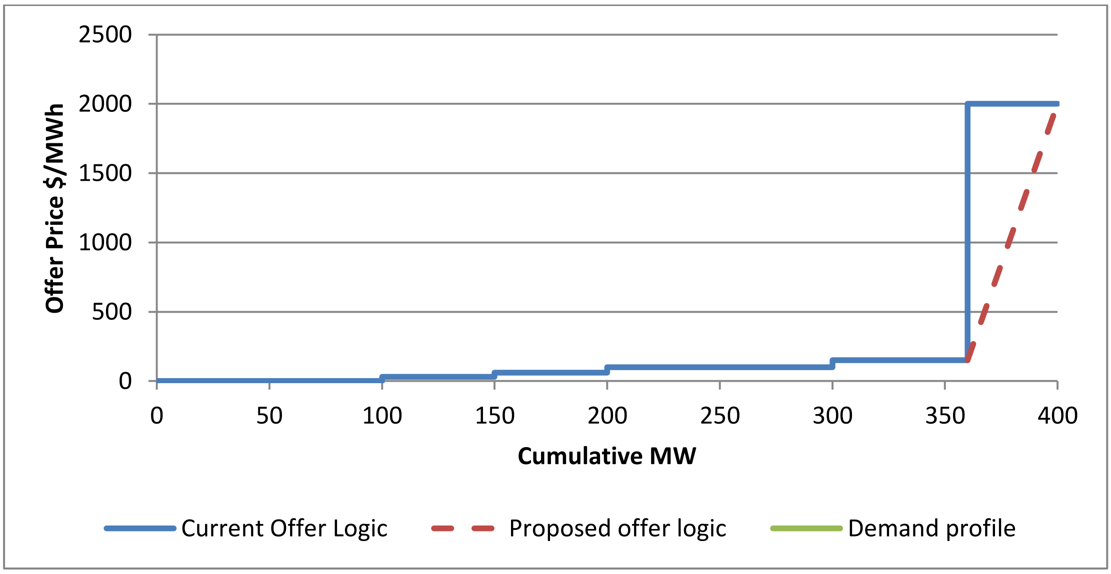
If this constraint would be violated, the dispatch engine would constrain the flow until the constraint is satisfied, as long as there is sufficient available generation to make up the difference. Alternatively, some small-scale generation such as GTs might find it possible to enter the dispatch without risk of spoiling the price in the process. The marginal value of the constraint could be paid to the providers of inertia and the cost charged to the interconnector net settlement surplus (the interconnector outage being the contingency to be managed).

Of course, inertia is not really dispatchable, so some form of forward-looking process may be required for a workable system (the NEM uses Predispatch, STPASA and MTPASA, so these could be amended). Nevertheless, payment and cost allocation based in the shadow price of this constraint, or a projected shadow price, could be a useful supplement to a longer term procurement process for inertia in SA. However, it may well be that the proposed real time market, topped up with the income from this constraint, may be sufficient to encourage the needed inertia to be on line without any over-riding long term contracting.

6.7 Modelling of Offer Profiles and Regional Demand

Offer profiles are modelled as piecewise constant marginal cost blocks, which correspond to a piecewise linear offer cost curve. Only 10 blocks are supported by NEMDE. The blue line in Figure 18 illustrates a simplified offer curve.

Figure 18: Modifying an Offer Band to be Continuous

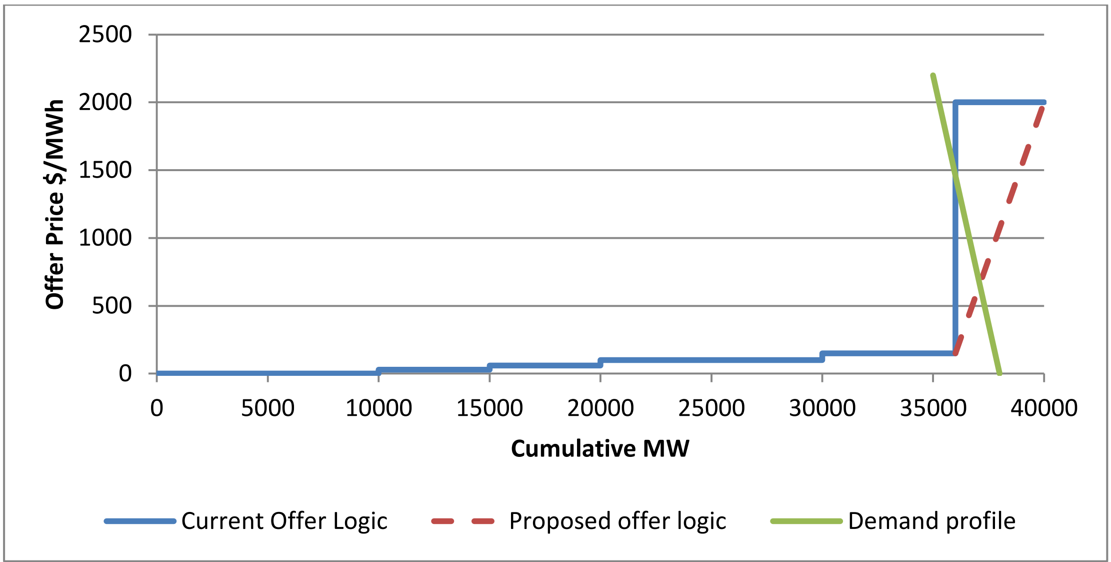


This arrangement can generally support any desired curve to reasonable accuracy. However, at the "top end", the steps can be relatively large, as shown. This in turn makes the acceptance or rejection of an offer band more of a gamble than it could be, not only for the participant but also for the market as whole. One way to deal with this could be to implement a software switch that optionally converts the offer into a continuous curve as shown as the red dotted line in the figure, which has been applied to the last band only. In this way, a participant offer could be accepted at any MW from 360 to 400, instead of either one or the other (except if the offer happens to be marginal). Some participants could see an advantage in having this flexibility.

This facility could be made available individually for some bands band or all bands. It is easily implemented with a non-linear (in fact, quadratic) solver by adjusting a linear term and adding a quadratic term in the objective of the solver – simple and transparent. As now, the only requirement is that the offer function be non-decreasing.

Such an approach could also improve the modelling of demand. Non-scheduled load can be price sensitive as retailers and loads have access to the dispatch price 5 minutes ahead, as well as price forecasts over longer periods. As discussed in Section 3.4, AEMO could model this price sensitivity as a function of the current variables used in forecasting supplemented by regional reserve, which can be calculated each dispatch interval.

Figure 19: Offer Stacks with Elastic Demand



To illustrate this approach, assume the simplified offer profile of Figure 18 approximates an offer profile for the whole (unconstrained) market, as modelled in NEMDE. Figure 19 shows this offer profile against an elastic demand profile shown in green.

This figure illustrates that, when demand is elastic and offer prices are all piecewise constant blocks as shown in blue, there will typically be NO offer price in the stack that equilibrates with demand; the system *cannot reach equilibrium*. Normally, this is no issue as any small discrepancy is taken up with regulation; price errors would be small and not even noticeable. However, at the top end price gaps between bands can be larger. This is one reason[14] why AEMO has difficulty forecasting demand and price when reserve margins are low.

This problem is partially resolved if AEMO could actually model its regional demand with an elastic curve as shown. This would allow the system to equilibrate at a price and demand where the where blue and green lines cross. An even better equilibrium would be a situation where the green and red dotted lines cross, when both supply and demand costs and benefits are both modelled as continuous functions.

Elastic demand could also be modelled by adding a quadratic term to the objective function when elasticity is assessed to be high through statistical studies; the two suggestions in this sub-section are mirror images of each other.

This sub-package would act to stabilise 5-minute dispatch prices especially when reserve margins are low, and should improve the efficiency and stability of dispatch at those times.

Under the proposed package of improvements, the auction would, in time, solicit demand elasticity, reducing the likelihood of non-price rationing at times of potential supply shortfall. This non-regulated, non-scheduled demand elasticity can be measured and reliably included in the forecast.

6.8 Valuing Reactive Power, Voltage and System Strength

In a dispatch and pricing system using an AC power flow network model, voltage becomes a variable rather than being fixed. Reactive power flows also become a variable, to be used largely to manage voltages. There is a weak but still important link between reactive power and real power flow in the network.

When voltage becomes variable, acceptable limits and how these are to be modelled – as hard or soft constraints[15] - must be decided. Further, various reactive devices such as synchronous condensers as well as the operating envelope for reactive and real power in generating units need to be accounted for. The proliferation of these small but important decisions to define the model is one reason why AC power flow models are not yet commonly used in dispatch and pricing engines. AC power flow modelling can and is used for constraint generation described in Section 3.2. While accounting for reactive power, such a process does not optimise reactive power flows with real power flows.

A prototype AC power flow dispatch and pricing engine would nevertheless be a useful tool to evaluate strategies involving reactive power. An example would be to appreciate and perhaps relax technical requirements on reactive power provision where there is a beneficial trade-off with real power and vice versa. Another would be to help evaluate system strength issues that are becoming critical. In the context of an AC power flow model, a fault can be regarded as a contingent event.

---

[14] The other is that AEMO does not take into account the current state of the system (specifically, regional reserve margin) for its single point forecast.

[15] A hard constraint is a bound on the variable; a soft constraint we can define as a smooth function (e.g. a quadratic) that tends to limit the variable within the region of a central nominal value.

6.9 Implementation Issues _Toc478724212

Table 4: Modelling Approaches Suitable for Proposed Improvements

Network Model	Regional	DC Power Flow	DC Power Flow	AC Power flow
Optimiser Technology	Linear	Linear	Non-linear	Non-linear
Generate Dynamic Marginal Loss Factors		X	X	X
Generate Dynamic Security Constraints	X	X	X	X
Optimise Contingency FCAS Cost Allocation	X	X	X	X
Smooth Out Offer Profiles			X	X
Value Reactive Power and Voltage Control				X

Table 4 repeats Table 3 in the earlier Section 6.2. We can see that some proposals can be addressed in the context of the current NEMDE infrastructure while others require a significant change to implement a network model and/or to use a non-linear solver. Further, all proposals warrant further study and refinement in a prototyping environment[16]. It is not immediately clear what the priorities should be and whether some options might be overtaken by events.

We note that a non-linear interior point solver can also solve purely linear models with good efficiency. Therefore, we lose little or nothing by building our prototype on a non-linear solver platform. Further, we do not need to start from scratch; we can port the current NEMDE to the nonlinear environment to provide a solid starting point. This suggests the following strategy to make rapid progress in evaluating these options in a prototype environment.

- Port the current system to a non-linear solver. This might require some adjustment to interfaces but should otherwise be straightforward as the core model and code is unchanged.
- If not done already, modularise the code so that each component (e.g. the network) is clearly delineated from other components (e.g. generator models or security constraints). The aim is to be able to upgrade specific modules by implementing a software switch that substitutes one code block for another. The code involved would:
    - ✓ generate the model in standard format for the solver; and
    - ✓ process the results of the solver into usable form.
- Chose a specific improvement option and write the code for that option as a module which can be switched in and out.

[16] We commend a prototyping strategy for market development options, not only for internal AEMO and AEMC evaluation but also to promote participant understanding and constructive comment.

- Test and refine the code as required, in parallel with the real system to be able to compare results.
- Implement when the module has stabilised.
- Implement other modules in the same way in accordance with priorities set.

This approach should work even for an apparently radical change such as implementing an AC power flow model. In this case:

- First implement a DC power flow model to replace the regional model, with voltages modelled but held constant at reference values.
- To implement AC power flow, make all voltages variable and write the additional relationships for reactive power (similar to real power) as well as the relationships that link real power to reactive power. While this is a major addition, a great deal of the existing code base need not be touched.
- Implement the software switch to the new code, as suggested.

There are great advantages in maintaining a single code base for prototype development.

7 Allocation of Fees and Charges

7.1 Charging Principles

Consistent with NER Clause 3.2.4(a) for the dispatch process which is reproduced in Appendix A, the principles for the efficient allocation of charges are simply summarised as:

1. Implement a two sided market in the product or service wherever possible.
2. Where a service must be centrally procured or managed, allocate costs so that they can be avoided by participants thereby reducing the need for and cost of the service.
3. Where approach 2 is not possible or practical, allocate costs where they do the least damage to efficiency, normally where no response to that cost is likely.

7.2 Suggested Adjustments

Previous sections have outlined ways in which some NEM functions could be more efficient and their costs allocated better if centrally procured. We list them here for completeness.

- With the proposed real time market implemented as proposed in Section 5, the burden on the current FCAS enablement markets would be lowered or, in the case of regulation at least, even eliminated for practical purposes.
- To the extent that contingency services may still need support with an enablement process, that burden would be less than before because of the incentives offered by the real time market. This would allow FCAS enablement costs to be allocated more efficiently through the dispatch and pricing process as proposed in Section 6.6.

In view of the current supply concerns in the NEM but also as a general principle, the allocation of fees and charges should also be reviewed with a view to allocating them where response is likely to be least, perhaps on a NMI basis. For example, market fees and SRAS costs cannot in practice be avoided by participants and therefore serve no marginal cost signal, and should be allocated by NMI instead.

One important case is the component of market fees that ought to be allocated to market development. In its recent Directions Paper on the maintaining system security options, the AEMC claims that it may take 3 years to develop market approaches to proving some services such as inertia and fast frequency response. In the meantime a heavy handed approach under the control of NSPs is proposed[17]. One possible reason for this leisurely timetable is the lack of development that either AEMO or AEMC has undertaken on possible market solutions to date. This can in turn can be explained by the Rule change, summarised in Appendix B, that removed from AEMO any obligation to undertake investigations into improving the dispatch process. Market development would likely be an early casualty when participants apply pressure on AEMO to reduce its costs and associated fees.

In view of the urgency developing around the need to improve the NEM, research, development and demonstration of options would best be funded directly by governments.

[17] http://www.aemc.gov.au/Markets-Reviews-Advice/System-Security-Market-Frameworks-Review

8 The Market Auction Upgrade Package

8.1 Package Overview

Our focus is to implement marginal pricing further into the NEM than applies at present, to increase the scope for freedom of action by participants and to reduce the pressure for heavy-handed regulation to deal with emerging challenges in the NEM. Our proposed upgrade package has the following principal elements:

Fixing Pricing within the Half-Hour

1. Quickly implement a system to implement robust marginal pricing within the half hour, focused initially on scheduled units that are monitored with SCADA, and based on the control logic now used for regulation FCAS. The package would:
   a. leave untouched existing core system used for dispatch and settlement;
   b. implement 5-minute pricing by introducing a Ramping Ancillary Service (RAS);
   c. remove artificial price steps which could cause market and system instability;
   d. implement a SCADA-based real time market consistent with FCAS regulation control logic and integrated with the RAS;
   e. amend the obligation to comply with dispatch instructions; and
   f. remove some generators from scheduling systems such as AGC.

Further Enhancing Pricing within the Half-Hour

2. After gaining operating experience and undertaking further R&D, extend the initial implementation of this system to:
   a. extend application to faster-acting services such as inertia, fast frequency response and the faster contingency FCAS services by extending the pricing engine to recognise these services and by using suitably programmed revenue metering instead of SCADA to value performance; and
   b. extend participation to non-scheduled customers, including very small ones.

Upgrading NEMDE capability to Improve Pricing and Cost Allocation

3. By upgrading solver technology used in the NEMDE to one with non-linear capability, progressively prototype, run in parallel and eventually implement a range of NEMDE improvements which could include:
   a. the ability to offer smooth and more finely tuned offer and bid profiles;
   b. full co-optimisation of contingency FCAS enablement services;
   c. the generating dynamic security constraints to improve network transfer capability;
   d. support a move to dynamic Marginal Loss Factors (MLFs) instead of constant static ones that are updated by AEMO; and
   e. support an option to make the treatment of voltage control and reactive provision more transparent.

Improving Cost and Fee Allocation

4. Improve the allocation of costs and fees by:
   a. Noting that the real time market should drive FCAS enablement cost lower, allocate FCAS contingency enablement costs according to the shadow prices of the FCAS requirement constraint in NEMDE, when those requirement constraints are implemented;
   b. In clause 3.8.1 of the NER (See Appendix B), introduce a *requirement* to replace the current *option* for AEMO to investigate and develop the dispatch engine and associated processes, the cost of this function to be borne by governments rather than participants; and
   c. Reconsider other cost allocations in the light of:
      i. the principle set out in Clause 3.1.4(8) of the NER (see Appendix A) or, more simply, the NEO, applied more broadly to all fees;
         (1) (rather than the complex set of principles AEMO has regard to under the present Rules);
      ii. the expenses participants will incur during implementation; and
      iii. in the current stressed environment of the NEM, the desirability of keeping existing plant in operation.

8.2 Implementation Strategy

The proposed upgrade package can be implemented incrementally and key parts relatively quickly and easily once a decision to proceed is made. The following strategy focuses on minimal disturbance to existing systems and incremental additions, ultimately leading to much improved marginal pricing and greater opportunities for a wider range of participants and technologies in the NEM.

Benefits and risks will be considered in later subsections.

Fixing Pricing within the Half-Hour

1. While the objective of the current Sun Metals rule change proposal on 5-minute settlement has merit, the currently proposed form of implementation has risks associated with it (see following sub-section) and would, if implemented, preclude a more complete solution. Therefore, we should seek to:
   a. persuade the AEMC a solution along the lines of proposal outlined in this report is "more preferable" to the Sun Metals solution or minor variations on it; or, if that is considered to be too much of a departure from the Sun Metals proposal,
   b. argue against adoption of the Sun Metals rule change with a view to submitting an alternative rule change proposal along the lines outlined in this report.
2. To support a better understanding of the proposal in this report before a decision is made, consider:
   a. publishing a set of calculations/charts demonstrating how the RAS/real time market would have been settled based on historical dispatch outcomes;

   b. Implementing a simple web-based demonstration and training prototype[18] and run associated training courses. This could as an option run off real dispatch and the Regulation FCAS SCADA data used by causer pays in a demonstration mode.

3. When a decision is made to proceed with the production system:
   a. the prototype system should be upgraded or re-written to a production standard - complexity is of the same order as the current Causer Pays system, which is relatively modest;
   b. because the RAS real time system is an addition rather than a modification to the existing settlement logic and because it is essentially a balanced arrangement, the settlement of the RAS/RFUM system could be phased in by, say, 20% increments (0%, 20%, 40%, 60%, 80%, 100%) over a period, to allow participants to become familiar with operations under the upgraded arrangements without significant financial consequences in the early stages; and
   c. the RAS real time market non-financial outcomes should be published as part of AEMO's usual information dissemination activity.

Further Enhancing Pricing within the Half-Hour

4. In parallel with the initial RAS real time market implementation, undertake an R&D project to support fast acting responses (based on local measurements) and possible non-scheduled operation of some facilities. Specifically:
   a. Investigate, develop and test a revenue meter/frequency meter package which can be programmed to implement and calculations suitable for RAS real time market settlement;
   b. Investigate, develop and test a meter/frequency meter package which can be programmed to provide a user interface for participants operating under an enhanced RAS real time market system.
   c. Using small samples of participants of different types, investigate the response patterns and pricing parameters and settlement arrangements that induce stable, market driven responses.
5. Review the rule changes that may be necessary to support implementation of an enhanced RAS real time market system [19], including:
   a. MDP management of real time market metering data, bearing in mind the need for AEMO to maintain good visibility over the patterns of response in order to tune its pricing algorithm;
   b. the boundaries between scheduled and non-scheduled operation;
   c. The relationship between a small customer who may want to participate in the real time market and its retail energy supplier.
6. When a decision is made to proceed with the production system:

18 A demonstration mode is where there are no settlement consequences even though a settlement calculation is done. The prototype need not be expensive to implement as it can be run independently of current AEMO systems and need not have many of the features required of a production system.
19 Some current AEMC rule change proposals are relevant.

a. the prototype systems should be upgraded or re-written to a production standard. The complexity is of the same order as the current Causer Pays system;

given that scheduled units will already be on the 4 second SCADA-based initial RAS/RFUM system, the new settlement system based on specially programmed revenue metering could be phased in by, say, 20% increments (0%, 20%, 40%, 60%, 80%, 100%) over a period, with settlement based on the SCADA based system making up the remainder[20]. This phasing would leave settlement of the slower acting elements unchanged, but gradually phase in the faster acting ones that can be captured by the specially programmed revenue metering;

b. the phase-in process can also be used to fine tune progressively the pricing model that drives the real time market:

   i. This pricing model would be a multi-dimensional version of the FCAS regulation implied pricing model (used in Causer Pays) that takes account not only of slow acting responses, but also of responses at other time scales (60 seconds, 6 seconds, and sub-second (Including load damping, fast frequency response and inertia).

c. When these arrangements are bedded in:

   i. carry out trials and adjust regulatory arrangements to allow participation in the real time market by non-scheduled participants, specifically small scale ones;

   ii. review the scope for loosening the conformance requirements for scheduled units, given that the real time market provides better incentives to maintain frequency stability.

d. The enhanced RAS/ real time market non-financial outcomes should be published as part of AEMO's usual information dissemination activity.

Upgrading NEMDE capability to Improve Pricing and Cost Allocation

7. Establish a new solver platform to prototype enhancements to NEMDE.

   a. a sparse nonlinear solver would allow porting of the current NEMDE linear model directly into a non-linear environment with only interface changes;

8. As a first step, enhance the current regional dispatch model to exploit the new nonlinear capability, including:

   a. prototyping interconnector losses as smooth quadratic functions rather than piecewise linear ones[21];

   b. prototyping offer bands modelled with continuous rather than stepwise prices to help stabilise ex ante spot prices, especially when prices are high;

   c. prototyping provision for some demand forecast price elasticity when reserve margins are very low;

20 This phasing would leave settlement of the slower acting elements unchanged, but gradually phase in the faster acting ones that can be captured by the specially programmed revenue metering;

21 This may in fact reduce solve times and, by using warms starts, effectively deal with the non-physical loss problem.

d. prototyping logic that sets contingency FCAS requirements to the set of largest contingencies, and assign costs to participants on the basis of the resulting shadow prices for the requirement constraint[22]; and

e. depending on the outcome of the prototyping and any required rule changes, port the prospective enhancements to the production NEMDE.

9. As a second step, prototype an upgrade to the regional model within NEMDE to a DC power flow network model[23], if necessary interfaced with an AC power flow model operating in a tight loop, to support addition NEMDE enhancements, including:

   a. an option to implement dynamic loss factors rather than constant, static ones that are updated periodically by AEMO;

   b. opportunities to generate security constraints more dynamically to take better account of current system conditions, thereby expanding the secure operating envelope of the network; and

   c. depending on the outcome of the prototyping and any required rule changes, port the prospective enhancements to the production NEMDE.

10. As a third step, prototype modelling voltages and reactive power in addition to the real power and angles of the DC power flow approximation, in order to research improvements to the management of voltages and reactive power in the network[24].

11. Recognise that all these capabilities could be built into a single prototype dispatch model, with the ability to switch specific functions in and out as desired.

Improving Cost and Fee Allocation

12. Investigations, prototyping and production implementations of the proposed and ongoing pricing and dispatch process improvements should be funded directly by governments to ensure performance in an effective and timely manner; specifically:

    a. the prime role should be taken by AEMO as the market operator, with inputs from AEMC, current and prospective participants and external sources of expertise.

    b. the requirement of Clause 3.8.1(f) of the NER should be changed from "may" to "must"

13. Opportunities to improve cost allocation to parties who can manage the issue, for example, as proposed in 8.d above for contingency FCAS enablement, should be taken once the systems to support them have been demonstrated.

14. The AEMC should review other cost allocations applying in the NEM to ensure better adherence to efficiency principles.

[22] The practicality of this approach may be influenced by the extent to which the Real Time FCAS Usage Market acts to reduce the burden on (and prices in) the FCAS enablement markets.

[23] A full network model within NEMDE does not imply full nodal pricing, which is a conceptually separate matter. Regional prices are set as the prices at the regional nodes in the network. We do not recommend any change to regional pricing in this report.

[24] Although competitive markets in reactive power may be difficult to make work in practice, the ability value reactive and dispatch it flexibly could offer market benefits not currently realised.

8.3 Benefits of the Proposed Package of Improvements

The main benefit of the proposed package that offers a coherent, forward looking and market-oriented approach (either fully or in part) to a wide range of issues currently under consideration in the NEM (by the AEMC, AEMO, the Finkel Review and other activity), The main drivers are system security, reliability and cost concerns that have emerged only recently in the public eye. A fair comment is that the institutions that run and govern the NEM have been caught without options for workable and tested market solutions to these challenges.

We give three brief examples, taken from the AEMC's current list of rule changes and reviews, where the proposed package could provide a full or substantial solution.

Five Minute Settlement –AEMC Reference ERC0201

As described earlier in this report, this proposed rule change would be partial solution to a longstanding NEM issue. However, if implemented as current discussion papers are suggesting, it would over incentivise load and fast generator response, to the point of delivering unstable physical and price outcomes. Why? Because the proposed solution does not deliver marginal pricing designed to keep the system stable. If today's problem is poor marginal price signals across a half hour, under this proposal tomorrow's problem will be poor marginal price signals at a dispatch boundary.

Needless to say, submissions to the discussion paper are polarised depending on whether the participant is exposed to fast technologies (batteries) or slower technologies (OCGTs). The argument is now about batteries versus gas plant, which is simplistic and unhelpful.

Further the design would appear to require a significant revamp of AEMO and participant systems, as well as to upset the existing derivative contract market, upsetting more participants than necessary.

The proposed intra-half hour component of the improvement package specifically addresses these issues by:

- integrating energy and FCAS services within the half hour into a coherent, marginal pricing package designed to ensure system stability as well as relative freedom of participant operations; and
- separating these pricing arrangements out as a separate service (RAS), so as not to affect existing systems directly.

With this more comprehensive yet modular approach, the benefits of marginal pricing within the half hour can be realised more easily and more quickly.

Demand side obligations to bid into central dispatch: AEMC Reference ERC0189

This rule change proposal has been under consideration for some time. The proposal is to require at least the larger price responsive loads to make offers of load reduction and to be scheduled, with the aim of improving AEMO's load and price forecasting capability. This of course comes at the expense of a heavy compliance burden on loads, likely to dampen their enthusiasm for participation given that load responsiveness is not their core business.

Consideration of this matter also seems oblivious to the potential behavioural changes that a successful resolution of the 5-minute pricing issue would bring – perhaps the hope or expectation among some participants is that the 5 minute settlement rule change will eventually lapse.

If marginal price signals could be made clearer, stronger and more accurate within the half hour as we propose, we would expect more predictable unscheduled load responses. We would expect our proposals to encourage response when the system needs it. Rather than attempt to shackle these responses, a more creative approach is to improve marginal price signals and forecasting and, to the extent necessary, include an element of price elasticity in the demand forecast.

Load forecasting could be improved by including a measure of regional reserve into the forecast. Analysis would show that very low reserve in a region implies higher prices and a reasonably predictable pattern of price responsiveness in regional demand. With a slight modification within a non-linear NEMDE framework as proposed, this elasticity could be recognised in the dispatch and pricing system, leading to more stable outcomes.

All this could be done objectively under a documented procedure, to remove any hint that AEMO is participating in the market. One cannot hope to remove the uncertainty in load forecasts, but one should aim to remove any consistent biases in those forecasts.

System Security Market Frameworks Review: AEMC Reference EPR0053

The AEMC and AEMO are conducting as System Security Review and have recently released a discussion paper outlining a proposed approach[25].

The Review has identified two key issues becoming important as non-synchronous generation becomes more dominant and synchronous units retire, especially in cases where smaller regions can become isolated, but ultimately throughout the whole system.

- System inertia is declining, which means that the system can lose frequency at a faster rate than current arrangements can deal with, especially in regions that can become isolated.
- “System strength” is also declining in some locations; voltages cannot be sufficiently sustained during a fault, which may then trip equipment and threaten system collapse.

The AEMC’s discussion paper proposes an immediate approach which would give networks the responsibility to either invest themselves or contract with participants to deal with these problems; longer term “market” solutions are seen as being practical only three or more years into the future.

The package proposed in this paper provides a much faster path to a market solution of the frequency control problem. A focussed effort could see a Stage 1 solution in place well

[25] The Discussion Paper can be downloaded from http://www.aemc.gov.au/Major-Pages/System-Security-Review

before next summer and a real prospect of at least trial sites of a solution that would effectively price system inertia and fast frequency response from these sites. The danger of going down a regulated path such as network obligations and generator technical standards, initially is that it may be very hard to reverse. It is also far more likely to lead to an excessive level of investment than a market path.

It is a greater challenge to envisage market arrangements for boosting system strength. However, development of a prototype AC power flow dispatch model as proposed in this report could provide a platform for evaluating solutions in the context of the broader market.

8.4 Risk Considerations

There are risks from following a course of action and risks from not following it. In the case of the proposed NEM upgrade package, we aim to minimise the risks associated with implementation and to highlight the risks of not doing so.

The proposed NEM dispatch, pricing and implementation package has the following risk management features:

- As a general proposition, no existing systems would be removed or directly replaced, at least in the short term. For example, we implement an effective 5-minute settlement arrangement (and more) but retain the existing half hour spot and contract arrangements. The aim is to add a more efficient overlay to the current arrangements, which have served the NEM well historically. Over time, some existing systems may become effectively redundant, although they might still be retained for security (e.g. regulation enablement).
- We propose a prototyping phase followed by implementation when issues practical issues have been ironed out. Prototyping can be used to demonstrate the approach not only for internal development purposes, but also to get feedback from a wide range of parties before implementation. Prototyping is also a fast and cheap way to make real progress on a concept. In the absence of a prototype, issues can be raised with no real hope of resolution.
- Even at the implementation phase, we can phase in the new systems. For example, the proposed real time market operating within the half hour can be phased in at modest and gradually increasing levels to allow a learning phase and to deal with operational issues (such as ensuring that pricing delivers stable outcomes, as it is designed to do) before they have significant operational or financial consequences. We apply the same concept to the dispatch engine improvements. New features can be switched on one at a time, once proved up in the prototype phase.
- The package is proactive rather than reactive. For example, with a clear focus the first phase of an intra-half hour package could be implemented relatively quickly, avoiding or at least greatly reducing the need for embedding a heavy-handed regulatory approach to deal with system security issues such as maintaining sufficient inertia and fast frequency response.

- The proposal for government direct funding of these developments circumvents the inability of current governance arrangements to take a proactive approach to the market design challenges facing the NEM. Removing inefficient fees from participants should allow capital to be reinvested in changes to their own systems and to improve their behaviour in response to the improved marginal price signals.

9 Conclusions

In this report we have proposed a package of improvements to the NEM auction processes. The advantages of the proposed package are that it:

- is focused on improving marginal pricing signals, and is so consistent with the NEM objective of maximising freedom of operation by participants to deliver secure and efficient outcomes;
- deals explicitly with many of the technical challenges currently facing the NEM and, where applicable, is superior to the short term regulatory interventions now being proposed; and
- can be prototyped and phased in without disrupting existing systems and trading.

We commend the package for consideration and early implementation.

Appendix A: Market Design Principles in the NER

The following extract is taken for the current Version 89 of the National Electricity Rules (NER), Chapter 3 on Market Rules.

3.1.4 Market design principles

(a) This Chapter is intended to give effect to the following market design principles:

(1) minimisation of *AEMO* decision-making to allow *Market Participants* the greatest amount of commercial freedom to decide how they will operate in the *market*;

(2) maximum level of *market* transparency in the interests of achieving a very high degree of *market* efficiency, including by providing accurate, reliable and timely forecast information to *Market Participants*, in order to allow for responses that reflect underlying conditions of supply and demand;

(3) avoidance of any special treatment in respect of different technologies used by *Market Participants*;

(4) consistency between *central dispatch* and pricing;

(5) equal access to the market for existing and prospective *Market Participants*;

(6) *market ancillary services* should, to the extent that it is efficient, be acquired through competitive market arrangements and as far as practicable determined on a dynamic basis. Where dynamic determination is not practicable, competitive commercial contracts between AEMO and service providers should be used in preference to bilaterally negotiated arrangements;

(7) the relevant action under section 116 of the *National Electricity Law* or direction under clause 4.8.9 must not be affected by competitive market arrangements;

(8) where arrangements require participants to pay a proportion of *AEMO* costs for *ancillary services*, charges should where possible be allocated to provide incentives to lower overall costs of the *NEM*. Costs unable to be reasonably allocated this way should be apportioned as broadly as possible whilst minimising distortions to production, consumption and investment decisions; and

(9) where arrangements provide for *AEMO* to acquire an *ancillary service*, *AEMO* should be responsible for settlement of the service.

Appendix B: Requirement to Develop the Dispatch Algorithm

Extract From National Electricity Rules, Version 89

3.8.1 Central Dispatch

……..

(f) AEMO ***may*** investigate from time to time:

(1) the scope for further development of the dispatch algorithm beyond the minimum requirements specified in clause 3.8.1(b); and

(2) the sufficiency of the dispatch algorithm in meeting the minimum requirements specified in clause 3.8.1(b),

and following compliance with the Rules consultation procedures, publish a report setting out its recommendations.

Extract From National Electricity Code, Version 01

Note: NEMMCO is the Market Operator, predecessor to AEMO in the electricity sector.
NECA is the National Electricity Code Administrator, predecessor to AEMC in the electricity sector

3.8.1 Central Dispatch

……..

(f) *NEMMCO* ***must*** investigate from time to time:

(1) the scope for further development of the *dispatch algorithm* beyond the minimum requirements specified in clause 3.8.1(b); and

(2) the sufficiency of the *dispatch algorithm* in meeting the minimum requirements specified in clause 3.8.1(b),

and following compliance with the *Code consultation procedures,* submit its recommendations in a report to *NECA* no later than 2 years after *market commencement.*