Downloaded from orbit.dtu.dk on: Apr 27, 2024

DTU Library

# Demand response for a secure power system operation

**Bondy, Daniel Esteban Morales**

*Publication date:*
2017

*Document Version*
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

*Citation (APA):*
Bondy, D. E. M. (2017). *Demand response for a secure power system operation*. Technical University of Denmark.

**General rights**
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

DANIEL ESTEBAN MORALES BONDY

# DEMAND RESPONSE FOR A SECURE POWER SYSTEM OPERATION

## SERVICE SPECIFICATION, VALIDATION AND VERIFICATION IN VIEW OF DISTRIBUTED ENERGY SYSTEMS

PH.D. THESIS

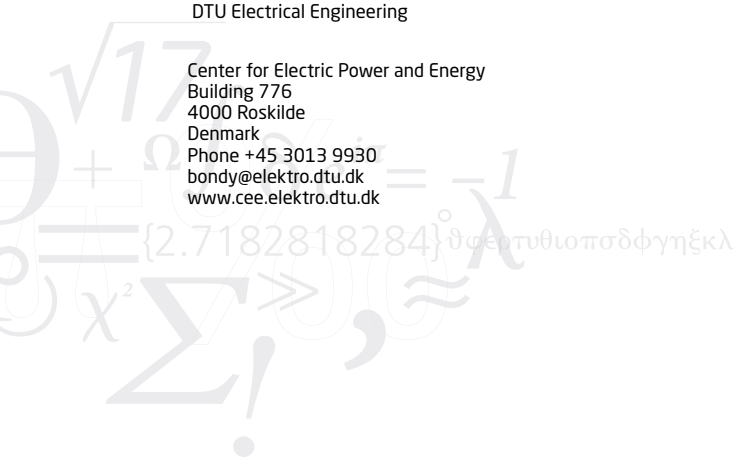
RISØ 2017

**DTU Center for Electric Power and Energy**
Department of Electrical Engineering

This document was typeset with XeLaTeX.
The book design is based on the Tufte-LaTeX document class and the DTU Compute PhD thesis template.

**DTU CEE**
**Center for Electric Power and Energy**
**Technical University of Denmark**
DTU Electrical Engineering

Center for Electric Power and Energy
Building 776
4000 Roskilde
Denmark
Phone +45 3013 9930
bondy@elektro.dtu.dk
www.cee.elektro.dtu.dk



# Preface

This thesis was prepared at the Energy Systems Operation and Management group under the Center for Electric Power and Energy, which is part of the department of Electrical Engineering at the Technical University of Denmark. The thesis is a requirement for acquiring the Ph.D. degree in engineering, and was funded by Innovation Fund Denmark through the Strategic Platform for Innovation and Research in Intelligent Power (iPower), the Programme for Energy Technology Development and Demonstration (EUDP) through PowerLabDK and the Technical University of Denmark (DTU).

The energy sector is moving away from fossil fueled electricity production to generation from intermittent renewable energy sources. In order to achieve a successful integration of these renewable sources in the power system, the use demand response is essential. Demand response is expected to deliver ancillary services to the power system and the demand response schemes must therefore be validated.

This thesis addresses the topic of aggregation of flexible units for verifiable demand response services in partial replacement of traditional ancillary service resources. The contribution to this topic is a revision of the validation procedure, the reformulation of service requirements, restructuring of ancillary service products and new metrics for verification of service delivery, all to account for the characteristics of demand response.

This thesis is multidisciplinary in its approach and draws upon concepts from fields such as: Power systems engineering, Control engineering, Software engineering, Systems engineering, Energy policy and regulation.

The project was supervised by Senior Scientist Henrik W. Bindner, and co-supervised by Associated Professor Hans Henrik Niemann, and Assistant Professor Kai Heussen, all three from DTU Electrical Engineering. Part of the research was conducted at the Lawrence Berkeley National Laboratory with Sıla Kılıççote as supervisor.

The thesis consists of a synthesis (along with adjustments and expansions) of the concepts presented in two journal papers and three conference papers written in the period 2012-2016.

Daniel Esteban Morales Bondy
March 2016

# Summary

Demand response will become important for the integration of renewable energy sources in the power system. By controlling large pools of small-sized consumption units, aggregators of demand response will provide ancillary services to Transmission System Operators and flexibility services to Distribution System Operators and Balance Responsible Parties. Since these services are essential for the secure operation of the grid, the aggregators must be validated. The process applied to traditional ancillary service resources can not be applied to aggregators since they are composed by geographically distributed heterogeneous resources.

Departing from the current methods employed by Transmission System Operators for validating ancillary service resources, this thesis presents a procedure for validating aggregators. The procedure consists of documentation of the aggregator capabilities and a conceptual framework for aggregator validation testing.

The documentation of aggregator capabilities is done through a Functional Reference Architecture for aggregators. The reference architecture identifies the basic functions that an aggregator must posses in order to do a successful service provision.

The conceptual validation framework defines the test setup for aggregator validation tests. The validation tests must capture the stochastic nature of the aggregator, therefore the validation procedure makes use of concepts from the field of statistical test design.

Benchmark scenarios and service requirements are the inputs to the validation tests. The service requirements are redefined in order to be inclusive of new technologies as ancillary service resources. This is done by redefining services in terms of performance, and removing requirements that assume service provision by large centralized generators.

Service performance evaluation and the service verification are the outputs of the validation tests. Also these concepts are redefined to suit the aggregator concept. Through general service models and service performance indices (inspired by the field of Control Performance Assessment), the service provision from aggregators can be evaluated.

# Resumé

Demand response (forbrugs fleksibilitet) vil være vigtig for integrationen af vedvarende energikilder i elnettet. Ved at styre store mængder af småforbrugsenheder vil aggregatorer af demand response levere systemydelser til transmissionsnetoperatører og fleksibilitetsydelser til distributionsnetoperatører og balanceansvarlige aktører. Da disse ydelser er essentielle for forsyningssikkerheden, skal aggregatorer valideres. Den valideringsproces, der bliver brugt på traditionelle enheder, kan ikke bruges på aggregatorer, da de består af geografisk distribuerede heterogene enheder.

Med udgangspunkt i den traditionelle metode brugt af transmissionsnetoperatører til validering af kraftværker præsenterer denne afhandling en metode for aggregatorvalidering. Metoden består af dokumentation af aggregatorens egenskaber og en konceptuel ramme for aggregatorvalideringstests.

Dokumentationen af aggregatoregenskaber udføres gennem en funktionel referencearkitektur til aggregatorer. Referencearkitekturen identificerer de basale funktioner, som en aggregator skal have for at kunne levere vellykkede ydelser.

Den konceptuelle ramme for validering definerer opstillingen for aggregatortests. Disse tests skal fange den stokastiske natur af aggregatorer, og valideringsmetoden må derfor gøre brug af koncepter fra statistisk testdesign.

Inputtet til valideringstestene sker i form af standard scenarier og ydelsesbetingelser. Ydelsesbetingelserne er omformuleret til at inkludere nye teknologier. Omformuleringen sker ved at definere systemydelser på basis af ydelsespræstation og ved at fjerne kravene, der antager, at ydelserne er leveret af centrale kraftværker.

Outputtet af valideringstestene er evalueringen af ydelsespræstation og ydelsesverificering. Også disse koncepter er gendefineret, så de passer til aggregator konceptet. Gennem generelle ydelsesmodeller og præstationsindikatorer (inspireret af koncepter fra reguleringsteknik) kan ydelseslevering fra aggregatorer evalueres.

# List of Publications

## Papers included in this thesis

The following papers can be found in the appendices in the order the concepts are presented in the thesis. Here they are listed chronologically.

### Papers presented at peer-reviewed conferences

D) D.E.M. Bondy, G.T. Costanzo, K. Heussen, and H.W. Bindner. "Performance assessment of aggregation control services for demand response". In: *Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2014 IEEE PES*. Istanbul, Turkey: IEEE, Oct. 2014, pp. 1–6. DOI: 10.1109/ISGTEurope.2014.7028779.

A) D.E.M. Bondy, K. Heussen, O. Gehrke, and A. Thavlov."A Functional Reference Architecture for Aggregators". eng. In: *Proceedings of the 20th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA 2015)*. Luxembourg, Luxembourg: IEEE, Sep. 2015. DOI: 10.1109/ETFA.2015.7301638

B) D.E.M. Bondy, O. Gehrke, A. Thavlov, K. Heussen, A.M. Kosek, and H.W. Bindner. "Procedure for Validation of Aggregators Providing Demand Response". *To be presented at Power Systems Computational Conference (PSCC 2016)*. Genoa, Italy. Jun. 2016.

### Paper submitted to peer-reviewed journal

E) D.E.M Bondy, A. Thavlov, and J.B.M. Tougaard and, Kai Heussen. "Performance Requirements Modeling and Assessment for Ancillary Services in view of Demand Response". *Submitted to IEEE Transactions on Smart Grids*.

### Draft journal paper

C) D.E.M. Bondy, J. MacDonald, E. Kara, O. Gehrke, K. Heussen, S. Kilic-cote and H.W. Bindner. "Redefining Requirements of Ancillary Services for Technology Agnostic Sources". *Draft*.

## Other publications not included in this thesis

The following publications have been prepared during the course of the PhD study. They are omitted from this thesis since they are not directly related to the primary objective of the PhD project, or the relevant contributions were subsequently published in the papers mentioned above.

- D.E.M. Bondy, G. Tarnowski, K. Heussen, and L.H. Hansen. "Operational Scenario: Manual Regulating Power". *Tecnical report for the iPower Consortium*. 2013.
- D.E.M. Bondy, and A. Thavlov. "FLECH PowerMax Service Requirement Specification". *Technical report for the iPower Consortium*. 2014.

### Papers as contributing author

- G.T. Costanzo, O. Gehrke, D.E.M. Bondy, F. Sossan, H.W. Bindner, J. Parvizi, and H. Madsen. "A Coordination Scheme for Distributed Model Predictive Control: Integration of Flexible DERs". In: *Proc. 4th IEEE PES Int. Conf. Innovative Smart Grid Technologies Europe*. Copenhagen, Denmark: IEEE, Oct. 2013. DOI: 10.1109/ISGTEurope.2013.6695474.
- K. Heussen, D.E.M. Bondy, J. Hu, O. Gehrke, L.H. Hansen. "A Clearinghouse Concept for Distribution-Level Flexibility Services". In: *Proc. 4th IEEE PES Int. Conf. Innovative Smart Grid Technologies Europe*. Copenhagen, Denmark: IEEE, Oct. 2013. DOI: 10.1109/ISGTEurope.2013.6695483.
- X. Han, A.M. Kosek, D.E.M. Bondy, H.W. Bindner, S. You, D.V. Tackie, J. Mehmedalic, and F. Thordarson. "Assessment of Distribution Grid Voltage Control Strategies in View of Deployment". In: *Intelligent Energy Systems (IWIES), 2014 IEEE International Workshop on*. IEEE. 2014, pp. 46–51. DOI: 10.1109/IWIES.2014.6957045.

# Acknowledgements

I started the work on my PhD project in autumn 2012, and through these 3+ years I have collaborated with many inspiring people, too many to mention all of them here. I am thankful to my colleagues at DTU for being a source of inspiration, especially my colleagues at ESOM, with whom I have shared ideas, discussed, worked, travelled and shared friendship with.

I want to thank Emre C. Kara, Jason MacDonald and Sıla Kılıççote from LBNL (and now SLAC) for their great collaboration during my stay in Berkeley and for making me feel welcome in the Grid Integration Group.

This thesis would not have been possible without the feedback and guidance of my supervisors, Henrik W. Bindner, Kai Heussen and Henrik Niemann. They have given me support in different, but complementary ways.

Last, but not least, I want to thank my friends and family for their support while working on this project: my girlfriend Cille has been patient and supportive in the last stages of the thesis writing, my father Isidro helped me understand my research in a geo-political context, and my mother Madeleine, apart from giving moral support, moved country and made it possible for me to do my external stay.

To my daughter Sofia, who shows me that playing is just as important as studying.

# Contents


Preface 1

Summary 3

Resumé 5

List of Publications 7

Acknowledgements 9

1 Introduction 23

1.1 On the justification of research in renewable generation 23

1.2 Changes in the power system 25

1.3 Problem statement, Delimitation and Contributions 29

2 The Aggregator 33

2.1 Background 33

2.2 Clarifying the Aggregator Concept 36

2.3 Advantages and Limitations of Aggregators 38

2.4 The Functional Aggregator Reference Architecture 40

2.5 Conclusions Regarding Aggregators 43

3 Validation of Aggregators 45

3.1 Background 45

3.2 The Validation Framework 47

3.3 Procedure for Validation of Aggregators 49
3.4 On Prequalification and Certification of Aggregators 51
3.5 Conclusions Regarding the Validation Framework 51

4 Service Modeling and Requirements 53
4.1 Background 53
4.2 Modeling of Service Requirements 57
4.3 Restructuring Ancillary Service Requirements 59
4.4 Conclusions on Service Requirements 64

5 Performance Assessment & Verification of Aggregator Services 65
5.1 Background 65
5.2 Quality of Service 67
5.3 The Aggregator Performance Indices 68
5.4 Application to Service Verification 75
5.5 Conclusions on Performance Assessment 76

6 Conclusion and Future Work 79
6.1 Future Work 80

A A Functional Reference Architecture for Aggregators 83
A.1 Introduction 83
A.2 Aggregation in Smart Grids 84
A.3 The Need for an Aggregator Reference Architecture 86
A.4 Functional decomposition 86
A.5 The Reference Architecture 88
A.6 Case studies 89
A.7 Conclusion and further work 90

B Procedure for Validation of Aggregators Providing Demand Response 93
B.1 Introduction 93
B.2 Conventional resource validation and aggregator differences 94
B.3 Requirements and proposed test procedure 96

B.4 Case Study of the Validation Procedure 99
B.5 Discussion 103
B.6 Conclusion 105

C Redefining Requirements of Ancillary Services for Technology Agnostic Sources 107
C.1 Introduction 107
C.2 Ancillary Services: Current Definitions and Requirements 108
C.3 Unconventional Resources 112
C.4 Restructuring the Ancillary Service Requirements 114
C.5 Case Study 118
C.6 Discussion 118
C.7 Conclusion 118

D Performance Assessment of Aggregation Control Services for Demand Response 119
D.1 Introduction 119
D.2 Background 120
D.3 DSM Performance Assessment 122
D.4 Case Study 125
D.5 Conclusion and Outlook 130

E Method for Ancillary Service Modeling and Performance Assessment 133
E.1 Introduction 134
E.2 Background 135
E.3 Method for service modeling 137
E.4 Performance Assessment of Service Delivery 139
E.5 Case Studies 143
E.6 Conclusion 146

# List of Figures

1.1 The Electric Power System as seen today. The power generated is first transmitted at high voltage levels to industrial consumers and transformer substations, from where it is distributed at medium/low voltage to medium size consumers and households. 26
1.2 The actors and relationships in the power market today. Note that the consumer buys electricity from a retailer, but has no further contact to the other market actors, i.e. the consumer has a passive role in the system. 26
1.3 The Electric Power System of tomorrow contains a large ICT infrastructure, which permits the flow of information and control between the system actors. Furthermore, the flow of electricity is not only from generator to consumer, but there is also intermittent electricity generation at distribution level. 28
1.4 The actors and relationships in the power market of tomorrow. Compared to the current market setup, the aggregator entity has been added, as well as the ability of DSOs to contract services from the aggregator. The aggregator delivers ancillary services to the TSO through a BRC. Also, the consumer becomes a player in the electricity markets through the aggregator. 28

2.1 The aggregator concept across domains. The aggregator is present in the physical domain with the ICT system, the legal domain through its service contracts, and the control domain through its decision making logic regarding the behavior of the flexibility assets. This figure is an adaptation of Figure A.1. 37
2.2 A hypothetical schematic of the aggregator as a service provider. This figure is a modified version of Figures 1.4 and E.1. 38
2.3 A visual representation of the proposed reference architecture. The symbols represent data types that are outputs to the functions, and the color represents the data-handling categorization of the functions. 42
2.4 Mapping the Open Energi aggregator to the functional reference architecture. In this case it is only blocks related to communication that appear both at the aggregator and asset side. 43

3.1 Schematic procedure for aggregator validation. The aggregator test must be done with aid of a simulation framework so that the variation in aggregator capabilities can be appropriately identified. 48

3.2 From all the possible state space the aggregator can operate in, it is only a subset that is considered nominal operation. Within this nominal operation, a test operation space is defined, where the stochastic variables that affect the aggregator performance are manipulated to test the aggregator reliability. These stochastic variables include, but are not limited to, weather conditions, communication failure and user behavior. 48
3.3 The validation framework, where the aggregator is the unit-to-test, is ideally composed of a software co-simulation platform with hardware-in-the-loop capabilities. The inputs are the validation test cases, and the output (i.e. the service) is verified and evaluated. The arrows represent information exchange. 49

4.1 This chapter focuses on the *service definition* block of the aggregator validation framework presented in Chapter 3. 53
4.2 The ancillary services are activated sequentially after a frequency contingency. In systems with high inertia the frequency nadir will occur at frequencies closer to the nominal frequency. 55
4.3 Error in reference tracking. 59
4.4 Error in band service. 59
4.5 Error in cap service. 59
4.6 In this case, the ramp of the ideal response is mainly determined by the system inertia and is to be sustained until FRR can be activated. 63

5.1 This chapter focuses on the *service verification and evaluation* block of the aggregator validation framework presented in Chapter 3. 65
5.2 Error and QoS for the three kinds of services, note that the acceptable band do not need to be symmetric. 68
5.3 Test of the QoS definition where five aggregators deliver the same service for the same time horizons and with different performance. 71
5.4 Test of the QoS definition where 2 aggregators deliver the same service for different time horizons and with the same magnitude of error. 72
5.5 Test of the QoS definition where five aggregators deliver the same service for different time horizons and with the same performance. 73
5.6 Test of the QoS definition where five aggregators deliver the same service over different time horizons and with different performance. 74

A.1 The aggregator concept across domains. 84
A.2 Each function outputs distinct kinds of data which are used by the other functions in different ways according to the aggregator implementation. 89
A.3 Distribution of functions for the Power Hub aggregator 90
A.4 Distribution of functions for the Open Energi aggregator 90

B.1 Schematic procedure for aggregator validation. The focus of this paper is on the relationship between the service requirements and the aggregator tests. 96
B.2 Schematic test setup where the test subject, the aggregator, is treated as a black box. 97
B.3 Flow diagram of the aggregator algorithm. 100
B.4 Simulation results of the 100% availability test for the whole portfolio. 103
B.5 Simulation results of the 70% availability test for the whole portfolio 104

C.1 The nadir of a system frequency excursion during contingency events is more pronounced if there is less inertia in the system. 109
C.2 In this case, the ramp of the ideal response is mainly determined by the system inertia and is to be sustained until secondary frequency control can be activated. 116

D.1 The setup of the power system with DSM. Note that the Aggregator can either be an independent entity or can be a role inside a BRP. 121
D.2 The PowerMax service requirements, where the red line represents the boundaries for the permissible error, and the shaded area represents the error in service delivery, which is within the limits established in the QoS. 124
D.3 The non-flexible load of the households under the transformer. The sample is statistically representative of Danish households. 126
D.4 Aggregated power flow at the point of common coupling for the reference case units without coordination: Demand Response based on day-ahead energy price. 126
D.5 The setup of the two Aggregation algorithms to be compared. 127
D.6 Simulation results for the D-MPC with $\alpha = 0.1$ 129
D.7 Simulation results for the C-MPC with $\alpha = 0.1$ 130
D.8 Power load at Point of Common Coupling for the controllable and non-controllable loads 130

E.1 The new player in the market for ancillary services is the aggregator, which sells consumption flexibility on the ancillary service markets through the Balance Responsible Consumer. Some markets allow the aggregators to participate directly in the ancillary service markets with the condition that they coordinate with their Balance Responsible Consumer. It also provides flexibility services to BRPs and DSOs. 136
E.2 Error and QoS for tracking services, note that the acceptable band do not need to be symmetric. 141
E.3 Error and QoS for band services. 141
E.4 QoS for a maximum cap service, a minimum cap service is defined similarly but with $x_{min}$ and $x_{acc,min}$ values. 141
E.5 Example of reference tracking error with performance $x_{meas}$, ideal performance $x_{ideal}$ and tolerance limits $x_{acc}$. 144
E.6 Simulation of a DK1 primary reserve power ramp response together with $x_{ideal}$ and $x_{acc}$ values. 145
E.7 $x_{ideal} = P^{Agg}_{max}$, $x_{acc} = P^{Agg}_{max,acc}$ for the considered *PowerMax* example. The activation Dead-band is the time period, where the aggregator is allowed to non-deliver. 147

# List of Tables

5.1 The values of $\eta$ for same service delivery horizons and different service performance. 71
5.2 The values of $\eta$ for different service delivery horizons and with the same magnitude of error. Since $\eta$ is normalized with time, the shorter service is evaluated worse than the service delivered over the long time horizon. 72
5.3 The values of $\eta$ for different service delivery horizons. A numerical difference appears in the third decimal due to rounding error. The precision of $\eta$ depends on the measurement equipment of the flexibility asset. 73
5.4 The values of $\eta$ over different service delivery horizons and different service performance. 74
5.5 The values of $\epsilon$ over different service delivery horizons with same delivery error, as shown in Figure 5.5. 75
5.6 The values of $\epsilon$ corresponding to Figure 5.6. 75

B.1 System services and their behavior 98
B.2 Performance of DTU-FlexServices 104

D.1 Results of three-day simulation 129
D.2 D-MPC Performance over different simulation lengths 130

# Abbreviations

*AGC* Automatic Generation Control. 54

*AMS* Asset Management Services. 56

*AS* ancillary services. 54

*BRC* Balance Responsible Consumer. 25

*BRP* Balance Responsible Party. 25

*CHP* combined heat and power generator. 27

*DER* distributed energy resource. 27

*DESS* distributed energy storage system. 27

*DG* distributed generation. 27

*DR* demand response. 27

*DSO* distribution system operator. 27

*EE* energy efficiency. 27

*ENTSO-E* European Network for Transmission System Operators for Electricity. 30

*EV* electric vehicle. 27

*FCR* frequency containment reserve. 54

*FLECH* Flexibilty Clearing House. 56

*FRR* frequency restoration reserve. 54

*HP* heat pump. 27

*ICT* information and communication technology. 27

*PV* photovoltaic cell. 27

*RR* replacement reserve. 54

*RTO* Regional Transmission Operator. 30

*SGAM* Smart Grids Architecture Model. 34

*SLA* Service Level Agreement. 57

*SLO* Service Level Objective. 57

*SOA* Service Oriented Architecture. 57

*SoC* State of Charge. 56

*TSO* transmission system operator. 26

*VPP* virtual power plant. 35

*WT* wind turbine. 27

*…while the individual man is an insoluble puzzle, in the aggregate he becomes a mathematical certainty. You can, for example, never foretell what any one man will do, but you can say with precision what an average number will be up to. Individuals vary, but percentages remain constant. So says the statistician.*

*Sherlock Holmes, in The Sign of Four*

# Chapter 1
# Introduction

Research has shown that climate change is a fact and that, with a 95% certainty, human activity is the main cause for global warming[1]. As a way to mitigate the increasing rate of climate change, the Danish government has set as an interim goal to reduce the national $CO_2$ emissions by 40% in 2020, in order to reach the target of 80% - 95% reduction by 2050[2]. This is to be done by covering 50% of the national electricity consumption with wind energy by 2020, fully covering the electricity and heating supply with renewable energy by 2035, and being completely fossil fuel independent by 2050. Investing in an intelligent and flexible power system is deemed to be important if we are to reach those goals[3].

Aggregators of flexible consumption resources are expected to be a key element in the secure operation of power systems with large penetration of intermittent renewable energy sources. They will facilitate ancillary services to the system operators through the control of distributed flexible resources. Thus, traditional views on service specification, requirements, validation procedures and verification must be adapted to the new power system paradigm.

The motivation for changing energy production to renewable sources is presented in Section 1.1. A general description of the changes to the power system is presented in Section 1.2. The technical problem formulation is presented in Section 1.3, where challenges are identified within this new power system framework, and this work's contributions to solve said challenges are summarized.

[1] [120] Stocker et al. *IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* 2013.

[2] [128] The Danish Government. *The Danish Climate Policy Plan - Towards a low carbon society.* 2013.

[3] [29] Danish Ministry of Climate, Energy and Building. *Smart Grid Strategy - The intelligent system of the future.* 2013.

## 1.1 On the justification of research in renewable generation

Energy is a pillar in the development of all countries. The access to energy is a necessity that traditionally is supplied by fossil fuels, but the use of fossil fuels carries consequences that have impacted nature and society in three different ways:

*Climate issues:* Reports like the Stern Review[4] have made it abundantly clear that global warming and climate change will have considerable negative impact on society. If the current trend continues, global temperatures will rise 2-3 °C within the next fifty years, but this number

[4] [119] Stern, Treasury, et al. *Stern Review: The economics of climate change.* 2006.

will increase by several degrees if emissions continue to grow. The consequences of global warming will impact society mainly through issues related to water:

- melting glaciers will increase flood risk and reduce water supplies
- declining crop yields
- rising sea levels will result in increased floods, as well as the disappearance of coastal regions and islands[5]
- changes in ecosystems may lead to the extinction of 15 - 40% of species, and the acidification of oceans may lead to decline in fish stocks

[5] According to one estimate, up to 200 million people may become permanently displaced due to these effects by mid-century[119].

*Health issues:* Air pollutants resulting from the use of fossil fuel for transportation and electricity generation has been shown to be responsible for large numbers of morbidity and mortality. For example, in a study focused on traffic-related air pollution on public health in Austria, France and Switzerland[6], it is found that in these three countries air pollution is directly attributable to:

- 6% of total mortality (40 000 cases)
- 25 000 new cases of chronic bronchitis in adults
- more than 290 000 episodes of bronchitis in children
- more that 0.5 million asthma attacks
- more than 16 million person-days of restricted activities.

[6] [77] Künzli et al. "Public-health impact of outdoor and traffic-related air pollution: a European assessment". 2000.

Other sources[7] estimate that air pollution leads to 3.3 - 3.7 million premature deaths per year, with the majority of the deaths occurring in Asia. Furthermore, climate change will have a direct impact on health through:

- increased frequency of and intensity of heat waves
- changes in distribution of vector-borne diseases
- increased floods and droughts[8]

[7] [83] Lelieveld et al. "The contribution of outdoor air pollution sources to premature mortality on a global scale". 2015; [136] World Health Organization. *Ambient (outdoor) air quality and health - Fact sheet N°313.*

[8] [51] Haines et al. "Climate change and human health: Impacts, vulnerability and public health". 2006.

*Geo-political issues:* The concept of energy security has changed from being a local issue to include new concepts with respect to the provision of energy services. While traditionally it was a simple question of supply, measured by the four As of energy security (availability, affordability, accessibility and acceptability), it now encompasses concepts as efficiency, environmentally benign, properly governed and socially acceptable energy services[9]. Also, governments around the world are taking steps to mitigate their vulnerability in energy supply, increasing the importance of sustainable energy generation.

[9] [100] Pasqualetti and Sovacool. "The importance of scale to energy security". 2012.

In the work by *Cherp & Jewel*[10], the authors make a compelling argument for a new method for addressing the concept of *energy security* by treating it as a case of general security. Thus, *energy security* must address the questions: 1) *Security for whom?*; 2) *Security for which values*; and, 3) *Security from what threats?*. Following the Danish government's climate policy plan and the goals of Energinet.dk[11], these three question can be answered in a Danish context as:

[10] [26] Cherp and Jewell. "The concept of energy security: Beyond the four As". 2014.

[11] Energinet.dk is the Danish Transmission System Operator, the entity responsible of maintaining a secure transmission power grid.

1. *Energy security* in Denmark means that the population and industry should have an *adequate* and *secure*[12] power system.
2. Given the previously cited climate policy plan, it is safe to assume that sustainability is the major value in what concerns *energy security*.
3. All the threats to *energy security* can not be outlined here, but they include concepts of resilience and vulnerabilities. An example of a new vulnerability is the dependence on intermittent energy sources.

In short, in order for the government to secure the future of the population against climate change, while ensuring that Denmark remains economically competitive through the research and export of green technologies, it is essential that a transition to a sustainable[13] energy system is achieved.

This thesis addresses one of the solutions proposed to deal with the vulnerability introduced by the increasing penetration of intermittent renewable energy sources in the grid, as well as the increased stress on the power system due to the electrification of the transport and heating sectors. The following section explains the changes that the grid experiences as a consequence of the transition to a sustainable power system.

[12] *System adequacy* refers to the power system's ability to supply electricity demand at all times and *system security* refers to the ability to withstand sudden disturbances [37].

[13] [17] Brundtland et al. "Our Common Future (*Brundtland report*)". 1987.

## 1.2 Changes in the power system

In order to understand the relevance of this research project, it is important to define how the power system is expected to change, and clarify the frame for the research. This section gives a general introduction to the changes expected in the power system. The main actors in the power system and their relationships (from a Danish perspective) are presented, which will help scoping the problem.

### The Traditional Power System: Produce as we Consume

This subsection is intended for readers who are not already familiar with the power system. It clarifies basic concepts such as production/consumption balance, system frequency, system operators, energy markets, etc.

The goal of the power system is to provide an adequate and secure electricity supply to the population. The electric power system today is composed of two layers (Figures 1.1-1.2):

*Physical grid* This is the level at which the electricity flows, going from generators to transmission system, to distribution system and finally to the end consumer.

*Market layer* This is where all the energy trade and business operations are made. This includes the sale of electricity from producers to Balance Responsible Consumers (BRCs). Retailers in turn buy electricity from the BRCs and sell it to the end consumer. Being a Balance Responsible Party (BRP), either as a consumer or as a producer, means that the actor is responsible for its own forecasts and must ensure the best possible that the actual production/consumption follows the planned schedules.

While market regulations can be adjusted or completely changed in order to cope with the large influx of renewable energy, the physical laws

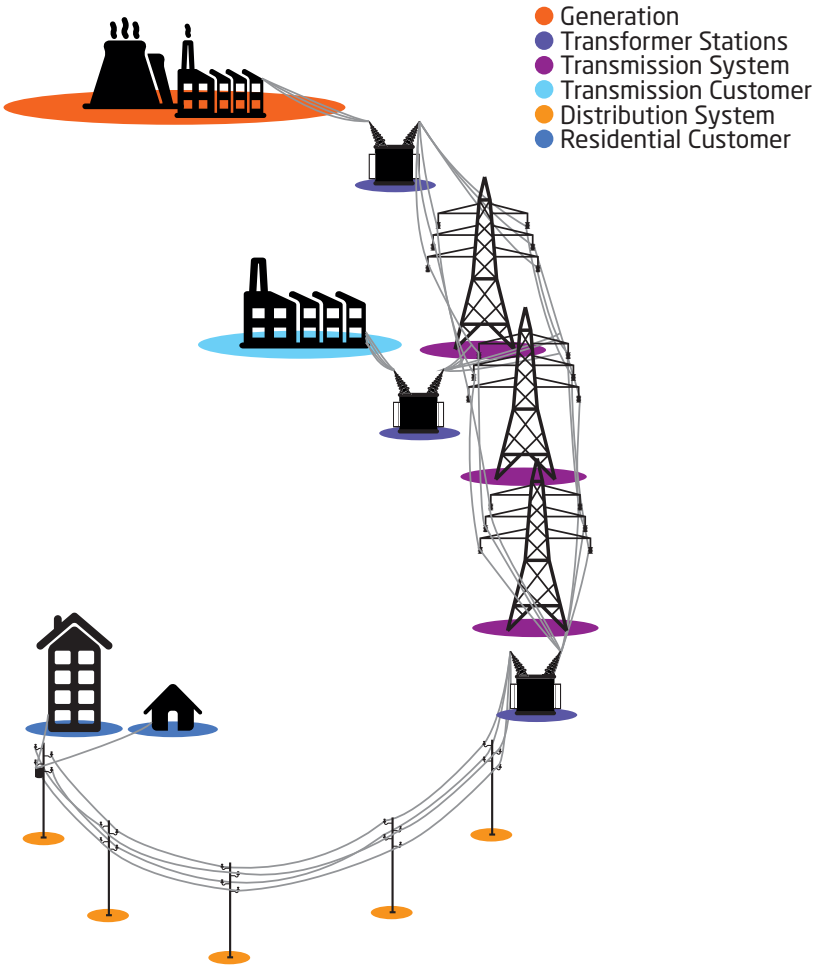



Figure 1.1: The Electric Power System as seen today. The power generated is first transmitted at high voltage levels to industrial consumers and transformer substations, from where it is distributed at medium/low voltage to medium size consumers and households.

cannot. When energy is produced it must also be consumed. With current technology it is unfeasible to store energy in large quantities, therefore energy companies must forecast how much energy consumers are going to need the next day and then buy energy accordingly. I.e., the production of energy must match the consumption of energy. If there is a surplus of energy in the system (production exceeds consumption), the system frequency increases[14], and might eventually damage electric components in the grid. Vice versa, a deficiency of energy in the system (consumption exceeds production) can lead to a blackout.

[14] The system frequency is a measure of the balance of the grid. Energy is traditionally produced with turbines which rotate synchronously in a given area. The system frequency, e.g. 50 Hz in Europe, is a measure of the balance of the system, with higher frequencies signaling a power surplus and lower frequencies signaling power deficit in the system.

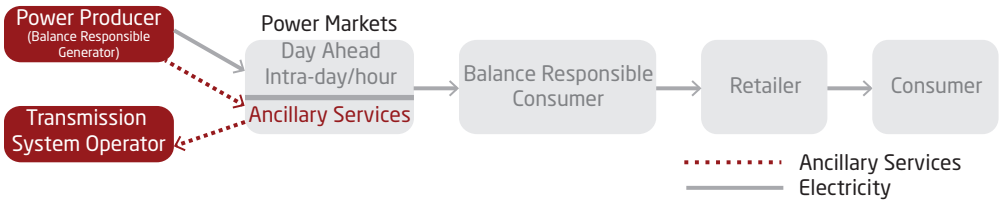



Figure 1.2: The actors and relationships in the power market today. Note that the consumer buys electricity from a retailer, but has no further contact to the other market actors, i.e. the consumer has a passive role in the system.

The consumption forecasts are imperfect, which leads to a constant imbalance between production and consumption of energy. The transmission system operator (TSO) is the entity responsible of resolving the imbalances of the system and maintaining the secure operation of the system. In order to do this, the TSO buys ancillary services from *certified generators*[15]

[15] The concept of certification of units to deliver ancillary services is central to this work and will be expanded upon in Chapter 3.

through the ancillary service markets. This market relationship is also reflected in Figure 1.2. There are different types of services, and thorough overviews and explanations of these can be found in the literature[16]. Here it suffices to say[17] that for most ancillary services, the TSO will pay generators to deviate from their planned production plans in order to bring the system back to balance. In the future, it is expected that the traditional sources of ancillary services, i.e. large central fossil-fuel powered generation plants, will be outphased in favor of smaller distributed and renewable generation. This means that new sources for ancillary services must be found.

[16] [39] ENTSO-E, UCTE. "Operational Handbook - P1-B". 2009; [110] Rebours. "A Comprehensive Assessement of Markets for Frequency and Voltage Control Ancillary Services". 2008.

[17] Further discussion on ancillary services will be presented in Chapter 4.

## The New Flexible Power System: Consume as we Produce

In the traditional power system, the uncertainty in consumption causes imbalances. With the increase of wind energy and solar generation, the uncertainty traditionally only associated with consumption spreads to the production side. Furthermore, traditional sources of ancillary services[18] are closing down, which means that TSOs must find new ways of balancing the system. Also, new problems will appear at the distribution system level, such as power congestion and voltage issues. These problems arise because of new consumption technologies appearing in the system, such as electric vehicles (EVs) and heat pumps (HPs), and because new generation units, e.g. wind turbines (WTs), small size combined heat and power generators (CHPs) and photovoltaic cells (PVs), are installed at distribution level. All these new units in the electric power system are commonly referred to as distributed energy resources (DERs)[19] or flexible resources. It is the responsibility of the distribution system operator (DSO) to resolve the problems arising due to the integration of the DERs, which can be the overloading of system components or voltage issues. These problems affect the quality of the power supply at residential level, but can also lead to issues at transmission level.

[18] Traditional sources of ancillary services are large centralized generators, typically based on fossil fuels. Some countries, e.g. Norway, can use hydroelectric turbines for ancillary services, but this technology is infeasible in countries that lack the appropriate geographical features, e.g. Denmark.

[19] In this work the concept of DER includes distributed generation (DG), energy efficiency (EE), demand response (DR) and distributed energy storage system (DESS), which is a combination of the the traditional definition of DER = DG + DESS (as seen in e.g. [134]) and the broader definition presented in [95].

The expected future smart grid can be seen in Figure 1.3, where not only the new DERs appear but an information and communication infrastructure coordinates the behaviour of the units for the benefit of the system. Smart metering is added at consumer level, and sensors are deployed at distribution level.

In order to address the new problems, both at transmission and distribution level, it is expected that consumers will become prosumers. In this context it means that the consumers will take an active role in the power markets by selling services to the system operators through an aggregator[20]. The aggregator will provide an asset management service to the end consumer, and by managing a pool of consumers, it will be able to control a large enough consumption volume to provide ancillary services to the system operators, or balancing services to the consumption BRP. The action of a consumer changing his or her consumption based upon an incentive, or direct control signal, is known as demand response (DR). The aggregator facilitates DR by providing the information and communication technology (ICT) and control infrastructure to DER owners, as well as statistical certainty of service delivery and legal responsibility towards the system op-

[20] The concept of the aggregator is further discussed in Chapter 2.

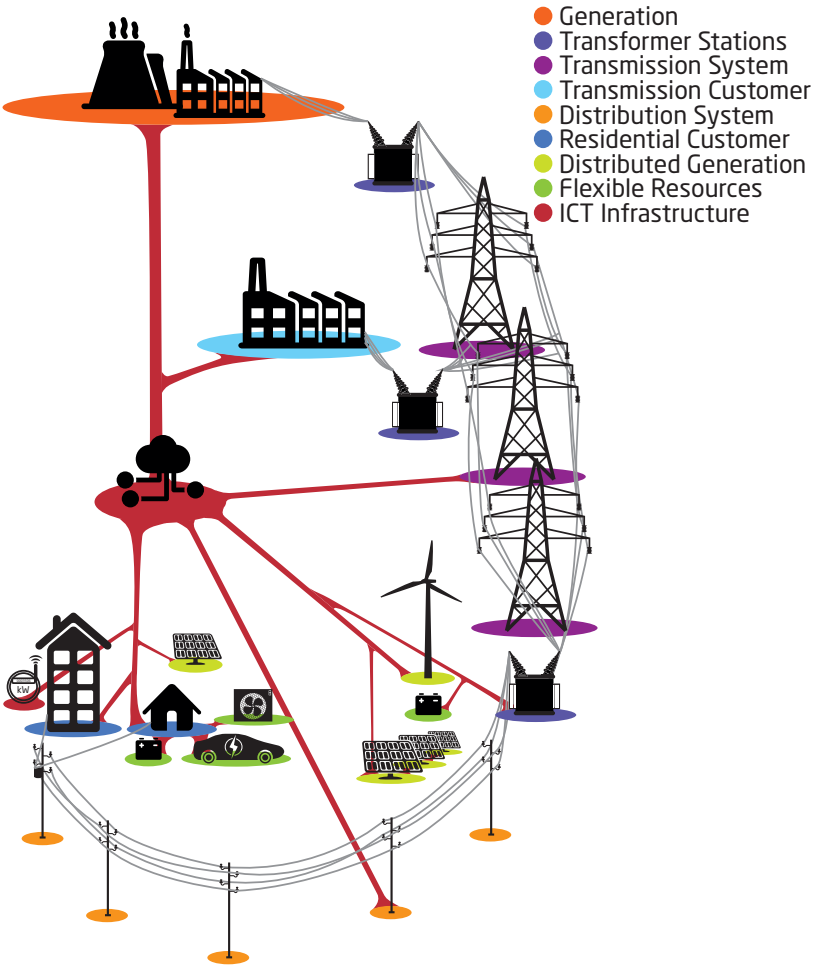



Figure 1.3: The Electric Power System of tomorrow contains a large ICT infrastructure, which permits the flow of information and control between the system actors. Furthermore, the flow of electricity is not only from generator to consumer, but there is also intermittent electricity generation at distribution level.

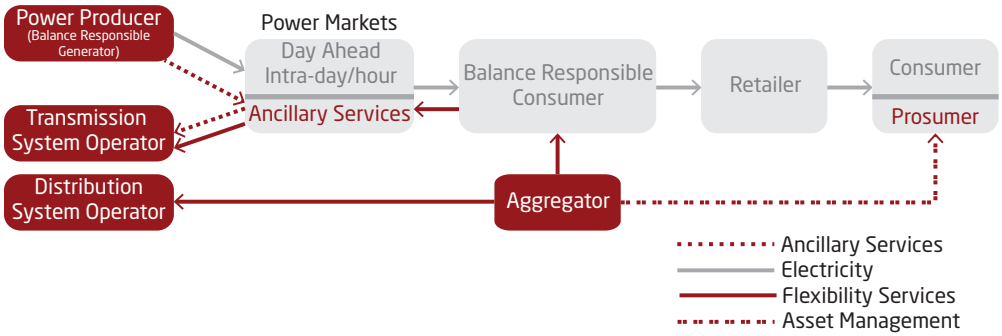



Figure 1.4: The actors and relationships in the power market of tomorrow. Compared to the current market setup, the aggregator entity has been added, as well as the ability of DSOs to contract services from the aggregator. The aggregator delivers ancillary services to the TSO through a BRC. Also, the consumer becomes a player in the electricity markets through the aggregator.

erators. The aggregator can be an independent commercial entity, or it can be a function inside one of the pre-existing market players. The new market setup can be seen in Figure 1.4, and it shows how the aggregator entity will interact with the existing market setup, and how the DSO will become a new player in the market, which will acquire services to resolve some of its problems.

In conclusion, we see the electric power system moving away from a *production-must-follow-consumption* pattern to *consumption-should-partly-follow-production* and hereby facilitate the integration of renewables and DERs. An integral part of achieving this change will be the use of control services to change the consumption behavior of units in the network. Given

that the units providing ancillary services to the grid are critical for the security of the system, system operators must be able to rely on that the units will behave as required. This is ensured by validating the new control algorithms and infrastructure through tests.

## 1.3 Problem statement, Delimitation and Contributions

Aggregators are expected to be new providers of ancillary services to the Transmission System Operators and flexibility services to the Distribution System Operators and Balance Responsible Parties by means of controlling flexibility. They must undergo the same prequalification/certification process that current providers of ancillary services must go through. Therefore, the control algorithms and architecture that constitute an aggregator must be validated. Several factors, such as the distributed nature and the modular composition of aggregators, make this problem non-trivial. The primary question this thesis seeks to answer is: *How can aggregators be validated, such that they can be relied on by the power market participants, and hereby actively help with the secure operation of the power system?*.

The Institute for Electrical and Electronics Engineers (IEEE) defines validation as: "The assurance that a product, service, or system meets the needs of the customer and other identified stakeholders."

Considering that *validation* is the key word in this question, the overarching question can be split into the following subproblems:

1. what are the needs the TSO wants fulfilled when acquiring services?
2. how can we measure how close the aggregator is to fulfilling those needs?
3. how can we establish a systematic procedure for assuring that the aggregator matches the required needs?

### Scoping and Methodology

The operation of the power system varies widely between countries and the range of ancillary services is wide. This thesis limits its focus to the following:

- In terms of the services considered, only those related to active power were analysed, i.e. Load Frequency Control at transmission level, congestion management at distribution level, and load balancing at BRP level. The topic of voltage related services was touched upon as a collaboration with *X. Han et al.*[21], but does not form part of the core of the presented research.
- The regulatory environment is assumed to be a liberalized market, such as the one in Denmark. Although some of the work is transferrable to other countries, e.g. the United States, the focus has been on solutions suited to the Nordic region.

[21] [57] Han et al. "Assessment of distribution grid voltage control strategies in view of deployment". 2014.

The methodology that has been followed has been pragmatic in nature. An understanding of the TSOs' empirical solutions to these problems has been achieved through the analysis of the regulations of Energinet.dk (the

Danish TSO), the European Network for Transmission System Operators for Electricity (ENTSO-E) and PJM (an American Regional Transmission Operator (RTO)), as well as through email correspondence and telephone interviews with representatives of Energinet.dk, PJM and CAISO (the Independent System Operator (ISO) of California). Departing from the status quo, for each research subquestion an analytical approach to identifying future requirements was taken, and a systems engineering approach in collecting these requirements. Solutions were designed with these requirements as goals or constraints.

## Contributions

The original contributions to the field are:

*Aggregator reference architecture:* The analysis of the aggregator in the power system through concepts from computer science and system engineering. Chapter 2 presents a candidate reference architecture for aggregators, which has the purpose of:

1. defining a standard lexicon around the aggregator, and
2. identifying the required functionality necessary for the effective working of an aggregator.

These two points are essential for understanding how the aggregator can fulfill the TSO needs.

*Aggregator validation procedure:* Validation methods for large central generation units are well developed, but must be adapted for aggregators of large quantities of small-scale distributed resources. The contribution is the expansion of the traditional method of validating generators to include statistical metrics for service requirements and performance, as well as statistical test design methods to the validation test procedure. Chapter 3 presents a framework for aggregator validation testing, as well as the outline of a procedure for validation tests. This contribution is important because it ensures capabilities of the aggregator are adequate for service provision.

*Definition and modeling of services:* Currently, the requirement definitions for ancillary services assume that the services will be provided by traditional units. In Chapter 4 a novel approach to ancillary service definitions is presented. Furthermore, a method for modeling service performance requirements was developed in order to facilitate service verification of aggregators. These contributions are important for the aggregator being able to deliver services in the power markets.

*Aggregator performance assessment:* In order to verify if an aggregator delivered a service according to its service contract, the performance of the aggregator must be assessed. In Chapter 5 concepts from Control Performance Assessment, i.e. from the field of process control, have been applied to the aggregator performance evaluation, leading to the

definition of a set of indices for performance assessment and service delivery. These indices are general measures for service delivery and are novel in the way that they are not defined for specific services, but can be applied interchangeably to the service models defined in Chapter 4.

## Thesis Structure

Each chapter contains the relevant state-of-the-art analysis for that topic and corresponding sub-conclusions. The thesis focuses on the theoretical concepts presented in the publications, but most of these concepts have been illustrated through cases studies in the articles. Most of these case studies are not presented within the body of the thesis, but the reader can refer to the relevant articles in the appendices for said case studies.

Chapter 2 discusses what an aggregator is and Chapter 3 presents the work on aggregator validation. Chapter 4 presents the work on services modeling and definition and Chapter 5 presents work on service verification and aggregator performance assessment. Overall conclusions and perspectives on future work are presented in Chapter 6, and the relevant articles forming the research content of the thesis are found as appendices.

# Chapter 2
# The Aggregator

The concept of aggregators has become widespread in the smart grid literature, yet the concept is still not clearly defined. This leads to a wide range of interpretations of the aggregator concept. This chapter contributes to the field of smart grid by analysing the aggregator through concepts from computer science and system engineering. Specifically, the concept of *functional reference architectures* is applied to the aggregator, making an encompassing definition of what an aggregator is, as well as defining an aggregator lexicon[1]. Furthermore, the candidate functional reference architecture establishes the essential functions aggregators must possess for effective service provision. This contribution is important for harmonizing the understanding of the aggregator, which enables the evaluation and comparison of aggregators. This will by extension ease the integration of aggregators in the power system. Most of the concepts presented here were originally presented as a work-in-progress conference paper[2] which can be found in Appendix A.

[1] In this context, an aggregator lexicon refers to the definition of a vocabulary related to the aggregator concept.

[2] [9] Bondy et al. "A Functional Reference Architecture for Aggregators". 2015.

## 2.1 Background

Aggregators have been designed and discussed widely in literature. This section presents the reference framework concept, a taxonomy[3] overview of current aggregator designs, and discusses the concept of flexibility.

[3] In this context, an aggregator taxonomy refers to a classification of aggregators based upon certain properties.

### 2.1.1 The Need for a Reference Architecture

A reference architecture "captures the essence of existing architectures, and the vision of the future needs and evolution to provide guidance to assist in developing new system architectures."[4]. It should provide:

- a common lexicon and taxonomy,
- modularization and the complementary context, and
- a common (architectural) vision.

The standard ISO/IEC/IEEE 42010 defines *architecture* as: fundamental concepts or properties of a system in its environment embodied in its elements, relationships, and in the principles of its design and evolution.

[4] [27] Cloutier et al. "The Concept of Reference Architectures". 2010.

This concept is not new in power systems, as can be seen from the draft technical report *IEC 62357-1: Power systems management and associated information exchange – Part 1: Reference architecture*, which focuses on the mapping of the interactions of all IEC standards related to the interactions of actors, components and systems within the power system. This is done with the aid of the Smart Grids Architecture Model (SGAM) which serves as a smart grid reference architecture[5]. These reference architectures arise due to the need of harmonizing the interactions within smart grids. Similarly, we propose a reference architecture to harmonize the understanding of the capabilities of aggregators. This is needed because:

[5] [23] CEN-CENELEC-ETSI. *Smart grid reference architecture*. 2012.

- Existing concepts and methods for benchmarking and generator validation cannot readily be translated from the generator based paradigm to the distributed paradigm of aggregators and flexibility services.

- Historically, ancillary services have been defined using a physical understanding of generator capabilities. There is a trend of changing the definitions towards technology-agnostic service models[6].

[6] As presented in Section 4.3.

- Service verification has been done through on-site measurements, which is infeasible with thousands of units participating in service provision.

The definition of a reference architecture for aggregators addresses these three issues, and enables benchmarking of aggregator architectures. Various types of aggregator implementation exist, realizing different design ideas for different sets of requirements. These requirements – and consequently the designs derived from them – are unlikely to converge towards a single solution because of the trade-offs involved, e.g. scalability and complexity. A common lexicon and taxonomy is a minimal precondition for aggregator comparison. If a reference architecture is to be used to describe many of these different designs, it must be highly modular. In practice, the essential functions of the aggregator must be distilled, in order for these functions to be usable as building blocks for the reconstruction of the *particular* functionality of a given implementation. The functions are arranged in a reference architecture such that metrics can be assigned to individual functions. In this way, the reference architecture can be used for the documentation of aggregator capabilities, which is part of the prequalification process[7].

[7] See Chapter 3.

### 2.1.2 Aggregator Taxonomy

In this work the concept of aggregation encompasses the creation and management of a portfolio of DERs[8] which seeks to provide the pooled flexibility in power consumption/production as a service or product to the power markets. This general definition covers most uses of the word in literature, but there is a large variation in the functionality that is expected from these aggregators. This can be seen by the wide variety of aggregator designs in literature[9]. The main reason for this has been that aggregators have been designed for specific kinds of units, for specific market rules and for specific services. An aggregator taxonomy is helpful for establishing a

[8] DERs can be referred to flexibility resources, or flexibility assets, when they form part of an aggregation portfolio

[9] See e.g. [74, 54, 116, 28].

common understanding of what an aggregator is (and is not) expected to do, and how it is anticipated to perform.

In some works[10] a distinction between aggregators is made in terms of which kind of task they perform. If they provide ancillary services they are categorized as Technical Virtual power plants (VPPs) and if they trade energy in the day-ahead energy market they are catalogued as Commercial VPPs. But recent work[11] proposes a *Dynamic VPP*, which is an aggregator that is designed to participate both in day-ahead markets and ancillary service markets. This type of advanced design could become commonplace in the future, making the *Commercial* vs. *Technical VPP* classification obsolete.

Other works classify[12] aggregators based upon their control paradigm into autonomous, direct, indirect and transactional control. While this classification is more robust towards future aggregator designs, it falls short on one main issue: *where is the intelligence located?*. In other words, the responsibility[13] and location of decision making is not taken into consideration in this classification. The responsibility and location of the decision making impacts the internal payment settlement of the aggregator, the scalability of the solution, the robustness towards communication faults and the response time, therefore it must be taken into account. In order to do this, a new taxonomy has been proposed[14], which identifies six classes of aggregator architectures ranging from fully centralized decision making, passing through diverse forms of distributed decision making, to fully autonomous[15]. Through this taxonomy, it is clear how the architecture of the aggregator will impact the performance in service provision.

[10] [46] FENIX. *Flexible Electricity Networks to Integrate the expected Energy Evolution*. 2009.

[11] [90] Nieße et al. "Conjoint dynamic aggregation and scheduling methods for dynamic virtual power plants". 2014.

[12] [76] Kosek et al. "An overview of demand side management control schemes for buildings in smart grids". 2013.

[13] The concept of responsibility is central to the operation of the power system, since it determines which market player is to pay for system imbalances.

[14] [56] Han et al. "A review on Distributed Control Strategies for Distributed Energy Resource Coordination".

[15] The control paradigm classification can be considered a further sub-classification within this taxonomy.

### 2.1.3 The Concept of Flexibility

Another concept that needs to be defined is the one of *flexibility*. The understanding of the term has evolved over time, but it has been used loosely as the amount of power consumption a unit is able to move in time, within the constraints set by the primary function of the unit, e.g. transportation in the case of an EV. The concept of flexibility within the operation of the power system has been discussed in the literature, e.g. *F. Sossan*[16] presents two types of flexibility:

*Type 1:* when customers change their consumption due to changes in prices, e.g. with time-of-use tariffs,

*Type 2:* the implicit flexibility in the process of a unit that allows it to change its consumption without affecting its primary function.

It is only *type 2* flexibility that is able to provide verifiable services, and therefore *type 1* flexibility will not be discussed for the rest of this work.

Within the frame of *type 2* flexibility, a taxonomy was proposed by *Petersen et al.*[17] where flexibility is catalogued as *buckets*, *batteries* and *bakeries*, and in *Hansen et al.*[18] it is shown how this taxonomy can be applied to a comprehensive set of DR schemes. While this taxonomy covers both the volume of power moved and time the consumption is moved, it is only on the time axis that it considers discrete changes. For example, a *bakery* unit must run for a fixed time, e.g. one hour, and its process can be moved

[16] [117] Sossan, Bindner, and Nørgård. "Indirect control of flexible demand for power system applications." 2014.

[17] [101] Petersen et al. "A taxonomy for modeling flexibility and a computationally efficient algorithm for dispatch in smart grids". 2013.

[18] [58] Hansen, Knudsen, and Annaswamy. "Demand response in smart grids: Participants, challenges, and a taxonomy". 2014.

in time but must always run for one hour from the moment it starts, while a *battery* type of unit can stretch or shrink its consumption on the time axis. This means that the taxonomy does not take the power granularity[19] of the flexibility into account. For example, some units can only be controlled through on/off switches and can therefore only provide fixed changes in consumption, while other units are able to provide a range of changes in consumption. Under the current assumption, i.e. that flexibility services will be provided by a large amount of small-size units, this limitation of the taxonomy seems irrelevant, since the power change limitation will be smoothed out because of the portfolio aggregations. However, if the volume requirements for services are reduced in the future, thus enabling aggregators with small portfolios, it might be relevant to further expand this flexibility taxonomy. This could, for example, be done by splitting the categories into *continuous* and *discrete* buckets, batteries and bakeries.

[19] Here granularity is meant as the discrete changes in power consumption a flexibility asset is able to realize.

Also, two more flexibility concepts must be considered: *deferred* and *curtailed* flexibility. When flexibility is provided as *curtailed* flexibility, it means that the consumption/generation is reduced/increased without the need to recuperate/shed the energy provided in the service[20]. Correspondingly, *deferred* flexibility is where the units providing the service need to return to a nominal state by recuperating or shedding energy. In demand response, the latter is the most common kind of flexibility used.

[20] The *buckets* in the taxonomy of *Petersen et al.*.

In essence, flexibility has two dimensions:

- A power (or energy) component, at a given volume and granularity, which the aggregator can offer, and
- A time component, which affects:
  - the time horizon over which the change in consumption/production can be sustained, and/or
  - the time granularity of the offered flexibility[21].

[21] The time and power granularities are important if the service is provided by units whose main process has time constraints, e.g. minimum on-time for compressors, and power constraint, e.g. only on-off capabilities of the full power rating.

The quantification of this flexibility is out of the scope of this work, but methods for this are being developed[22].

[22] [18] Bucher et al. "On Quantification of Flexibility in Power Systems". 2015; [117] Sossan, Bindner, and Nørgård. "Indirect control of flexible demand for power system applications." 2014.

A final observation with respect to flexibility is that system operators like Energinet.dk, the Danish TSO, are interested in acquiring flexibility from aggregators. The traditional approach to buying ancillary services is that the operator pays for an increase or decrease in power production, but flexibility has the extra time dimension. System operators expect flexibility services to fit in within the existing framework, but the two kinds of services can arguably be said to be essentially different. This might lead to undesired consequences such as the kickback effect[23].

[23] See the aggregator limitations presented in Section 2.3.

## 2.2 Clarifying the Aggregator Concept

Responsibility is a central concept within power systems. While it is technically possible to provide DR-ancillary services without an aggregator, it is impractical for each DER owner to enter into a contractual agreement for rendering services to the system operators[24]. In this

[24] The term *system operators* is used throughout this work to refer to both Transmission System Operators and Distribution System Operators.

sense, the aggregator becomes a legal entity that absorbs the legally binding responsibility of its customers and ensures that the aggregated portfolio follows an aggregated operation schedule. At the same time, the aggregator has an ICT infrastructure, which encompasses both the communication and decision making of the aggregator. It uses this infrastructure to coordinate the DERs/flexibility assets' behavior to match a service need of a higher volume than what an individual unit would be able to cover. Also, the aggregator entity will typically not own the flexibility assets it controls. This multi-domain approach to defining aggregators can be seen in Figure 2.1.

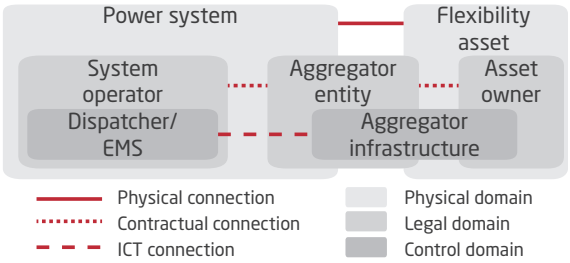



Figure 2.1: The aggregator concept across domains. The aggregator is present in the physical domain with the ICT system, the legal domain through its service contracts, and the control domain through its decision making logic regarding the behavior of the flexibility assets. This figure is an adaptation of Figure A.1.

These concepts lead to the following definitions:

*Aggregator role:* The role in the power system of performing aggregation with the purpose of selling the flexibility in consumption or production. The sale of flexibility can be a service to system operators or it can be traded in day ahead markets. The aggregator role can be assigned to a new player in the markets, or it can be assigned to an existing player, e.g. a Balance Responsible Party or a utility[25].

[25] How this role is integrated into the market is still an open question but proposals can be found in [133, 61].

*Aggregator entity:* The legal entity of the aggregator, which enters into contractual agreements with the other market players and flexibility asset owners. This entity is legally responsible for complying with the contractual agreements.

*Aggregator infrastructure:* The ICT and instrumentation infrastructure, both in terms of software and hardware, that the aggregator owns and operates in order to control the flexibility assets.

*Aggregator architecture:* How are the control elements and aggregator functions are related.

*Aggregator:* The term used to refer to a market player that has an aggregator role, entity and infrastructure.

Aggregators provide two kinds of services[26]:

[26] The terminology used in Appendices A and D varied slightly before settling on the terminology presented in this section.

*Flexibility services* which are provided to system operators and BRPs. These will take the form of ancillary services for the TSO, distribution system services for the DSO and portfolio balancing services for the BRP[27].

[27] The mentioned types of services will be expanded upon in Chapter 4.

*Asset management services* provided to the owners of the units, which consists of managing the flexibility asset for the owner, so that it can participate in the flexibility service provision, while still respecting the primary use/comfort settings of the asset owner.

This is shown in Figure 2.2, where the aggregator is selling services to the TSO through the Consumption BRP and directly to the DSO. This is a market setup which was concluded upon in the iPower project[28], although this project establishes the Flexibility Clearing House as a mediator between system operators and aggregators[29]. Other market setups allow for the aggregator to participate directly in the market, as long as they coordinate with their corresponding Consumption BRP, such that the aggregator avoids provoking imbalances at the level of the Consumption BRP.

[28] [98] Ølund and Hansen. *FLECH TSO Service - Fast Frequency Reserve*. 2014.

[29] [61] Heussen et al. "A Clearinghouse Concept for Distribution-Level Flexibility Services". 2013.

Figure 2.2: A hypothetical schematic of the aggregator as a service provider. This figure is a modified version of Figures 1.4 and E.1.

Finally, one of the central points of this work is that aggregators are essentially different from traditional generators. Aggregators differ from traditional generators in the sense that[30]:

[30] The differences pointed out here are further expanded upon in Section 3.1.1.

1. they are distributed systems where each unit has its own response properties, therefore the overall response behaves very differently than that of traditional generators;

2. they have no single point of measurement, which means the traditional measuring requirements can not be met;

3. reliability concepts must be adapted to their distributed nature, both in terms of communication reliability and service reliability;

4. aggregator architectures will vary widely, and each architecture will be sensitive to different operation scenarios;

Furthermore, traditional generators follow operational schedules, i.e. they have a baseline upon which the service is verified. Estimating the baseline for an aggregator is a difficult[31] and may lead to defining alternative methods for service verification.

[31] [7] Bode et al. "Incorporating residential AC load control into ancillary service markets: Measurement and settlement". 2013.

## 2.3 Advantages and Limitations of Aggregators

As stated in the previous section, the aggregator is a cross-domain entity. Most of the literature on aggregators and demand response focuses on the advances in the control domain which bring *operating*, *planning* and *economic* advantages to the power system[32]. In this work the focus is on the *operating* advantages and limitations introduced by aggregators.

[32] [96] O'Connell et al. "Benefits and challenges of electrical demand response: A critical review". 2014.

## Advantages

The operating advantages of aggregators can be divided into three categories: *scalability*, *reliability* and *responsiveness*.

In the control domain, the advantages of contracting an aggregator instead of a large amount of individual small sized units are similar to those of the legal domain. That is, the aggregator in its essence can be regarded as a solution for *scalability* of the smart control of flexible consumption or production. It would be possible for a system operator to directly engage all customers in order to buy services, but the coordination of such large quantities is impractical for the system operators. Thus, the system operators can request fewer services with large volume, and the aggregator will then supply this service with its portfolio of units.

The concepts described in this section are focused on aggregators as service providers, but the same concepts can be applied for aggregators trading in the day-ahead or intra-day markets.

Aggregators providing services through demand response can be more reliable that their traditional counterpart[33], i.e. large central fossil-fueled generators. A fault in one large generator will have a higher impact on the system than faults in several smaller-sized units. This means aggregators may improve the *reliability* of the power system. Also, DERs have limitations in their capabilities, e.g. cycling constraints for compressors, but by aggregating a sufficiently large pool of DERs, system operators will not be exposed to these limitations. A large pool of resources also means that the statistical certainty of the average behavior of the pool will be increased.

[33] [72] Kirby. "Load Response Fundamentally Matches Power System Reliability Requirements". 2007; [20] Callaway and Hiskens. "Achieving Controllability of Electric Loads". 2011.

Lastly, most DERs have very fast response times, which means that compared to traditional coal-fueled power plants, aggregators are able to provide very fast services. This implies that frequency excursions can be stopped faster and at a higher frequency nadir[34]. This leads to system operators requiring smaller reserves for maintaining the system security[35].

[34] [135] Vrettos, Ziras, and Andersson. "Integrating large shares of heterogeneous thermal loads in power system frequency control". 2015.

[35] [85] Makarov et al. *Assessing the value of regulation resources based on their time response characteristics*. 2008.

## Limitations

The main technical limitation of aggregators is that most DERs have as a main objective to satisfy the needs of it's owner, e.g. transportation in the case of EVs or heating in the case of HPs. Thus, the aggregator is constrained in its flexibility by the primary function of the DERs. Similarly, selling flexibility through aggregators is optional, so an aggregator must make a compelling business case, or other strong incentives, for the DER owner to participate in the service markets.

Another technical limitation is directly related to the kind of flexibility the aggregator provides. In most cases the aggregator will use *deferred* flexibility, where the units need to recuperate after the service delivery. If all units in the portfolio recuperate at the same time, the consumption spike that ensues may be a larger problem than the one the aggregator was contracted to solve. This is also known as the kick-back effect[36]. The problem of saturation can also be associated with this. DERs are usually only able to deliver services on short time horizons (compared to traditional generators) due to the limit size of the units. Once a minimum or maximum state has been reached, the flexibility of the unit disappears. This concept can be represented as a set of saturation curves[37], where asking for large

[36] [55] Han et al. *Identification of kick-back from normal situations due to activation of flexible demand*. 2014.

[37] [126] Thavlov. "Power System Integration of Flexible Demand in the Low Voltage Network". 2015.

volumes of power means the units can only deliver for short time periods and vice versa.

A non-technical limitation comes from market regulations. Market rules and ancillary service requirements are defined based on the capabilities of traditional generators. This means that aggregators are expected to behave as traditional generators, when they in essence are something completely different. This means that rules and requirements need to be changed if DER capabilities are to be fully exploited[38].

[38] This topic is addressed in depth in Chapter 4.

## 2.4 The Functional Aggregator Reference Architecture

Until now, the discussion on the aggregator has been focused on its role in the power system. In order to further the understanding of what an aggregator is, its functionality must be analyzed. One of the main contributions of the presented research is a *Functional Reference Architecture for Aggregators*. The objective of creating a reference architecture is to address the issue of benchmarking and validation/certification of aggregators[39]. The traditional approach to the certification of generators can not be applied to aggregators, and therefore new methods must be designed. Part of this method is to verify that an aggregator possesses the essential functionality for effective service provision. This essential functionality is defined in the proposed reference architecture.

[39] This topic will be discussed in depth in Chapter 3.

In order to formulate the aggregator reference architecture, a set of existing commercial and academic aggregators[40] were deconstructed into their basic functionality. The resulting functions of each aggregator were compared and clustered. From these clusters, a set of generic functions were formulated. The resulting functions are[41]:

[40] The analyzed aggregators were: Open Energi[99], PowerHub by DONG Energy[32], the Heterogenous Aggregator by Aalborg University[109] and the D-EMPC[28].

[41] For detailed explanations of each function see Section A.4.

A Service Interface: The function that translates information from the legal domain to the control domain[42]

[42] See Figure 2.1.

B Performance Monitoring: The function that evaluates and verifies the behavior of the client unit.

C Supervision and Resource Handling: The function that determines the availability and composition of the resource portfolio based upon the performance of the units.

D Operator Interface: The function that supports operator decision making.

E Control: The function that generates the appropriate control domain signals to manipulate the portfolio behavior.

F Flexibility Monitoring: The function that assesses the amount of flexible consumption/production available in the portfolio.

G Aggregator-internal Communication: The function that covers the internal communication within the distributed elements of the aggregator.

H Client Management: The function that determines the availability of the flexibility assets depending on their communication status.

I External Information Services: The function that pulls the necessary external data for the functioning of the aggregator, e.g. weather and price forecasts.

J Asset Interface: The function that translates between the control domain signals and the specific protocols used by the flexibility asset.

K Internal-information exchange: The function that enables the exchange of data and other information between the relevant functions of the aggregator.

These generic functions, with exception of *External Information Services*, are present in all aggregators. This exception is found in the cases where the aggregator delivers services based measurements of the grid, e.g. Frequency Containment Reserve[43]. The implementation of each function varies widely. Conceptually, each function has a specific purpose and task, but in the actual software implementation of the aggregator, one or more of these functions may be executed in the same module, or may be executed manually by an operator.

[43] See Chapter 4.

The functions can be classified in two different ways: a task based classification and a data-handling based classification. The task based classification divides the functions according to the kind of task the function executes:

*External interface:* The functions that provide information exchange with entities outside the aggregator infrastructure.

- Service Interface: Outputs a service model that sets the objective of the aggregator.
- Asset Interface[44]: Outputs the DER/flexibility asset data to the rest of the aggregator.

[44] The asset interface may also be considered part of the communication related functions, since it provides the communication translation between the flexibility asset and the rest of the aggregator infrastructure.

*Monitoring & Supervision:* The functions that parse information related to the unit portfolio.

- Performance Monitoring: Outputs the performance of the individual and aggregated flexibility assets. Can be used for internal purposes and/or service settlement purposes.
- Supervision and Resource Handling: Outputs the portfolio that is available for control, based upon the performance/compliance of the units.
- Operator Interface: Outputs manual decisions with respect to the portfolio and control.

*Control related functions:* The functions that involve the automated decision making with regards to the unit behavior manipulation.

- Control: Outputs control domain signals, e.g. activation or reference signals to the flexibility assets. This is highly dependent on the specific control architecture that the aggregator implements.
- Flexibility Monitoring: Outputs the state of the DERs/flexibility assets in terms of the flexibility available for control.

*Communication:* The functions that relate to the internal communication of the aggregator.

- Aggregator-internal Communication: Passes information
- Client Management: Outputs the portfolio of connected and responsive units.
- External Information Services[45]: Outputs the required external data.

*Knowlegde exchange:* This category only covers the Internal-information Exchange. It enables the information exchange between all parts of the aggregator. It can take any form, from a simple bus to highly developed data storage system, and has no specific output by itself.

This classification is represented in Figure 2.3 through the symbols marked on each function block.

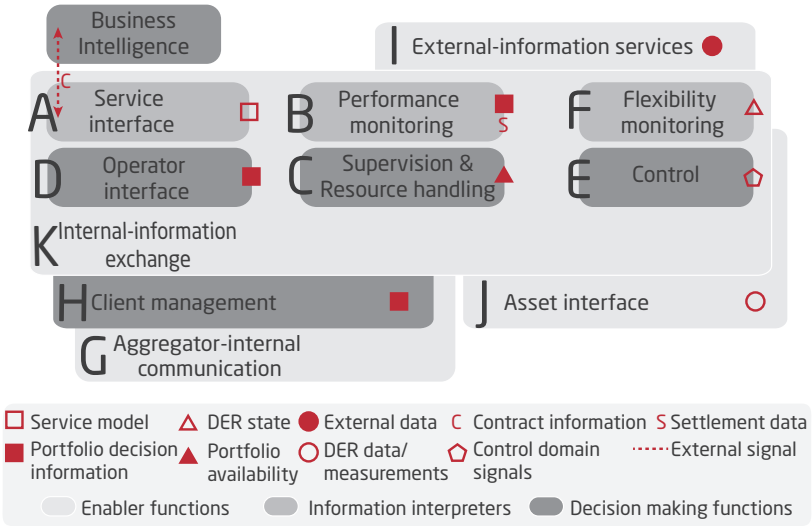
The data-handling based classification is done by grouping the functions based on how data/information is handled in the function. This classification is reflected in Figure 2.3 through the color code. The function classes are the following:

*Enabler functions:* Those functions that only pass the data/information on to other functions.

*Information interpreters:* Those functions that convert data into information.

*Decision making functions:* Those functions that use the information to make decisions.

The functions defined here abstract from any specific implementation of an aggregator[46], and provide the building blocks for the functional reference architecture shown in Figure 2.3[47]. This figure is a description of what the *aggregator infrastructure* block from Figure 2.1 encapsulates. Although the reference architecture abstracts from specific implementation, we present below an example of how an aggregator can be mapped to the reference architecture.




[45] Since the external information services only pull information from outside of the aggregator, and no information exchange is carried out, this function is not considered part of the external interface functions.

[46] Usually, a reference architecture must also define the relationship between its functions, but in this case, arrows representing data flow were avoided in the design, since they presuppose a specific aggregator architecture

[47] This diagram is a correction of the one presented in Appendix A.

Figure 2.3: A visual representation of the proposed reference architecture. The symbols represent data types that are outputs to the functions, and the color represents the data-handling categorization of the functions.

### 2.4.1 Application Example: Open Energi

As an example, the aggregator architecture for the British company Open Energi is described using the reference architecture. The information on how their aggregator functions was acquired through a publication[48] and unofficial email-interview with one of their developers.

Part of Open Energi's business model is to provide frequency response to National Grid, the TSO of the United Kingdom, by controlling the heating of bitumen tanks. This is achieved by installing a unit on each bitumen tank that can react upon changes in the grid frequency. How the control algorithm coordinates all units for an appropriate response is a company secret, but we were given enough information[49] to describe Open Energi with the reference architecture as presented in Figure 2.4.

[48] [25] Cheng et al. "Availability of load to provide frequency response in the great Britain power system". 2014.

[49] Note that Open Energi has not had the opportunity to review this example, and can therefore not be taken as an accurate example.

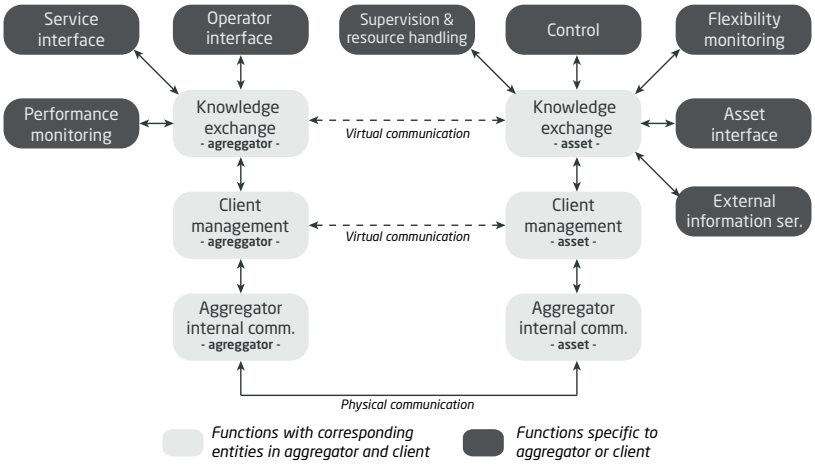



Figure 2.4: Mapping the Open Energi aggregator to the functional reference architecture. In this case it is only blocks related to communication that appear both at the aggregator and asset side.

In this case autonomous, fully-distributed architecture is represented by having most of the *decision making* and *information interpreter* functions on the asset side, and only having the operator interface, portfolio performance assessment and service interface on the aggregator side. When this mapping is complete, it can be used as documentation for the compliance of the aggregator with the reference architecture, which can serve as the documentation step of the prequalification process[50].

[50] See Chapter 3.

## 2.5 Conclusions Regarding Aggregators

The concept of aggregators is widespread in the smart grid literature, but the interpretations of what an aggregator is varies widely. One of the contributions presented in this chapter is to provide a common lexicon and reference architecture for aggregators so that discussion on the topic can be harmonized. This will hopefully lead to faster advances in the field. Also, the functional reference architecture is to be used in aggregator validation and certification[51]. A secondary use for the reference architecture is to serve as a guide for future designs of aggregators.

[51] The concept of aggregator validation and certification is discussed in depth in Section 3.4, where it is also discussed how this functional reference architecture can be useful.

When aggregators are discussed in the academic literature, the focus is usually on the control function. The functional reference architecture makes it clear that aggregators are more complex than single control block. There are several implicit functionalities, e.g. the *enabler functions* that are glossed over in academic studies, but are important for understanding the full capabilities the aggregator.

A current shortcoming of this work is that the presented functional reference architecture for aggregators is part of a work-in-progress paper and as such, needs to be refined and extended. Future work will include the design of key performance indices assigned to each function, such that the scores can be used to gain a better understanding of the capabilities of the aggregator and evaluate the maturity of the aggregator.

# Chapter 3
# Validation of Aggregators

Provision of ancillary services is essential for the security of the power system, and if aggregators are to provide these services, along with other flexibility services, they must undergo a prequalification process by the appropriate entity. This could be the TSO[1], a DSO or even an independent third party[2]. Traditionally, the prequalification process in Denmark has consisted of an initial submission of documentation describing the capabilities of the unit, and subsequently a test that validates the unit capabilities and communication. While this validation test is well established for large central generation units, how the test is to be applied to aggregators is still an open question. The solution to this question is of utmost importance if aggregators are to trusted for service delivery. In Chapter 2 the essential differences between aggregators and traditional generators are mentioned. In this chapter, these differences are expanded upon, and a framework for aggregator validation is presented. Furthermore, one of the main contributions of presented in this chapter is the expansion of the validation procedure to include statistical test methods and statistical measures for service requirements and performance. This procedure is originally presented in the conference paper[3] which can be found in Appendix B. The presented validation framework is original to this work. The validation process described here focuses on aggregator providing ancillary services, but can also be applied as a certification method for aggregators, such that they can participate with other products in the electricity market.

[1] In Denmark, Energinet.dk is the TSO and is in charge of the prequalification/approval process (described in [34]).

[2] For the rest of this chapter the responsible for carrying out the aggregator validation will be called the *testing entity*.

[3] [8] Bondy et al. "Procedure for Validation of Aggregators Providing Demand Response". 2016.

## 3.1 Background

In this section, conventional resource validation is briefly discussed and it is explained why the same method can not be applied to aggregators. Also a short section on the current work on aggregator testing is presented.

### 3.1.1 Conventional Resource Validation vs. Aggregator Validation

In Denmark, the prequalification process is divided into two steps:

The topic of conventional resource validation is discussed in more detail in Section B.2.1.

1. documentation for the unit is submitted to the TSO, and

2. a validation test where the unit's response to a signal from the TSO is evaluated.

The unit response tests serves two purposes: it validates that the response corresponds to the presented documentation, and it tests the communication system between the TSO control room and the unit. If the unit succeeds in the prequalification process, it is certified for participation in the ancillary service markets.

This process works on traditional generators because the dynamics of traditional generators are well understood. That is, generators can be described to a large degree of certainty through physical equations, and the unit response test serves to confirm the documented values of the equation variables[4]. This is not possible for aggregators because they behave fundamentally different from large generation units:

[4] The response test can also be seen as a system identification test.

1. The aggregator portfolio can either be of a heterogeneous or homogeneous nature. In both instances, the variance of the response of the portfolio units, along with the dynamic nature of the portfolio, means that it is difficult to describe the aggregator through physical equations and a single response test will give no insight into the overall response capabilities of the aggregator. This is aggravated by the fact that each DER will have its own set of requirements to satisfy its owner's needs.

A homogeneous aggregator is one which has a portfolio of same units, e.g. a fleet of EVs. A heterogeneous aggregator has a mix of units in its portfolio, e.g. EVs and thermostatically controlled loads.

2. Since the aggregator consists of geographically dispersed units, there is no single point of measurement. This means that the aggregated power profile does not represent a measurement at any single point in the power grid. This also means that traditional expensive measurement systems can not be used on aggregators.

3. The reliability concepts for distributed systems are different from those of single large units. Specifically, the failure modes are very different. The failure in a single unit in the aggregator has a much smaller impact on the overall aggregator performance compared to the failure of a subsystem in a generator. Also, communication reliability between the aggregator and the DER must be taken into account, as well as the added redundancy that stems from contracting a large pool of resources.

4. Aggregator architectures will vary widely depending on the control paradigm and the service requirements. An aggregator must be tested for a variety of operating conditions which are irrelevant for traditional generators.

It is both impractical and meaningless to validate every unit in an aggregator portfolio, since it is the statistical properties of the aggregated pool, not the individual unit, which makes the aggregator suitable for service delivery. The aggregator architecture must be tested as a whole, based upon statistical methods.

### 3.1.2 Aggregator Testing in Literature

There is currently no standardized procedure for prequalification of aggregators as there is with traditional generation units. Until now, the performance evaluation and testing of aggregators in academia has been ad-hoc to specific aggregator implementation[5], or the evaluation focus has been on computational or financial performance[6]. Similarly, a platform for simulation of aggregation strategy has been proposed[7], but the focus is on the simulation tool itself, which in turn focuses only on the demand side, and not on the process of validation. None have taken a systematic approach to generally evaluating the performance of the aggregators in terms of the contractual requirements of service delivery.

[5] See e.g. [135, 64, 82].

[6] [121] Su and Chow. "Performance evaluation of an EDA-based large-scale plug-in hybrid electric vehicle charging algorithm". 2012; [109] Rahnama et al. "Evaluation of Aggregators for Integration of Large-scale Consumers in Smart Grid". 2014.

[7] [31] Dittawit and Aagesen. "Demand side focused simulation platform for the evaluation of demand side management approaches". 2014.

### 3.1.3 Design of Experiments

The validation tests must be methodical and excite the aggregator such that the variance in its capabilities is well understood. Concepts from *Design of Experiments* are used for designing such tests, mainly[8]:

*Treatment:* A treatment is a specific combination of factor levels whose effect is to be compared with other treatments.

*Statistical Replication:* Replication can be defined as performing the same treatment combination more than once in an experiment. This is done in order to estimate the random error.

*Fractional Factorial Experiments:* Factors are the elements of a treatment, e.g. the baking treatment for a cake involves a given time at a given temperature[9]. In this case, time and temperature are factors that can be varied and will change the outcome of the treatment. Fractional factorial refers to taking a subset of the combinations of the factors.

The fractional factorial test presented in Appendix B follows the *off-line quality control* methods that were popularized by *G. Taguchi*[10]. Some aspects of these methods have been heavily criticized[11], but the methods presented in modern textbooks[12] have been adapted and changed according to these critiques. Thus, these methods seem appropriate to use for aggregator validation.

[8] [91]. *NIST/SEMATECH e-Handbook of Statistical Methods*. 2016.

[9] [97] Oehlert. *A first course in design and analysis of experiments*. 2010.

[10] [124] Taguchi and Wu. *Introduction to off-line quality control*. 1979.

[11] [16] Box, Bisgaard, and Fung. "An explanation and critique of Taguchi's contributions to quality engineering". 1988; [103] Pignatiello JR and Ramberg. "Top ten triumphs and tragedies of Genichi Taguchi". 1991.

[12] [97] Oehlert. *A first course in design and analysis of experiments*. 2010; [22] Cavazzuti. *Optimization Methods: From Theory to Design Scientific and Technological Aspects in Mechanics*. 2012.

## 3.2 The Validation Framework

The definition of a standardized validation procedure will become relevant as more aggregators, with a variety of architectures, appear in the power system and are willing to participate in the ancillary service markets. The process of validation for aggregators has three motivations:

- Allowing System Operators to contract aggregators that are able to provide adequate services (similar to the prequalification process that

current generators must undergo) by documenting the reliability of the aggregators.

- Ensuring balance responsible parties or other entities seeking to contract flexibility services that the aggregators are capable of reliably delivering electricity products.
- Allowing commercial entities interested in entering the aggregator market to test the design of their aggregator infrastructure and control algorithm before deployment.

The reliability of the aggregator depends on stochastic processes, e.g. consumer patterns and weather behavior. Therefore, it is natural that the validation procedure gives a statistical measure for the reliability. This means that the aggregator must undergo a series of validation test cases, as depicted in Figure 3.1. Formulating a set of test scenarios constrains the testing of the aggregator to a set of circumstances that the aggregator is expected to be able to handle, see Figure 3.2. These *treatments* must be reproducible and with sufficient sampling so that the validation can be backed up with statistical certainty. It is infeasible to carry out this procedure on the physical system. Therefore, this test process has to be carried out with aid of detailed simulations of the aggregator interaction with the electric power system and DERs, in combination with general models for communication.

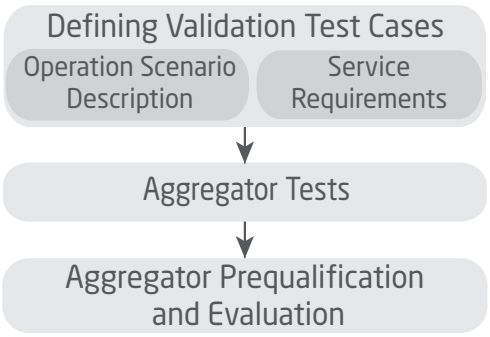



Figure 3.1: Schematic procedure for aggregator validation. The aggregator test must be done with aid of a simulation framework so that the variation in aggregator capabilities can be appropriately identified.

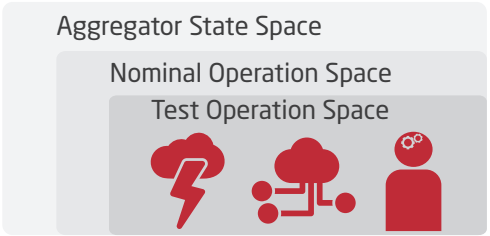



Figure 3.2: From all the possible state space the aggregator can operate in, it is only a subset that is considered nominal operation. Within this nominal operation, a test operation space is defined, where the stochastic variables that affect the aggregator performance are manipulated to test the aggregator reliability. These stochastic variables include, but are not limited to, weather conditions, communication failure and user behavior.

The proposed simulation tests should be carried out within a validation framework, as depicted in Figure 3.3. The service requirements[13] describe the goal the aggregator needs to achieve and the test scenarios define the normal operation disturbances that an aggregator should handle, see Figure 3.2. The aggregator will not be held responsible for service non-delivery when it is affected by major problems outside its responsibility domain, e.g. in case of severe grid faults.

The software framework needs to integrate the following models:

[13] The service requirements are discussed in depth in Chapter 4.

*Power System:* Depending on which kind of service the aggregator provides, either a transmission system or distribution system must be modeled, along with the relevant dynamics and appropriate time sampling.

*DERs:* For large scale aggregation a balance between model simplicity and accurate dynamics must be found.

*Communication Systems:* Time delays may have a large impact on the aggregator service performance, especially for those services that require fast response times.

The topic of integrating different simulation platforms for power system testing is being explored extensively[14] and the validation framework should be implemented by using and, if necessary, extending existing tools.

Finally, the *service verification and evaluation* block[15] must take measurements (either from simulation or field test) and evaluate the service provision compared to the established service requirements. This software module was implemented in SYSLAB for the iPower demonstration of the Flexibility Clearing House platform.

[14] [19] Buscher et al. "Towards Smart Grid system validation: Integrating the SmartEST and the SESA laboratories". 2015; [112] Schütte, Scherfke, and Sonnenschein. "mosaik-smart grid simulation API". 2012; [24] Chatzivasileiadis et al. "Cyber physical modeling of distributed resources for distribution system operations". 2015.

[15] The topic of service verification and evaluation is discussed further in Chapter 5.

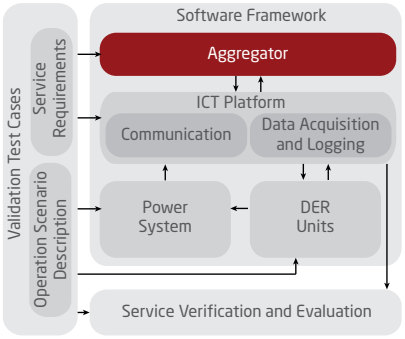



Figure 3.3: The validation framework, where the aggregator is the unit-to-test, is ideally composed of a software co-simulation platform with hardware-in-the-loop capabilities. The inputs are the validation test cases, and the output (i.e. the service) is verified and evaluated. The arrows represent information exchange.

A recording of the iPower demonstration can be found at [66].

## 3.3 Procedure for Validation of Aggregators

From the previous section it is clear that the service requirements form an essential part of the aggregator validation process. Service requirements are discussed in the depth in Chapter 4, but a set of test service requirement metrics have been formulated as part of the test method and are presented here.

A set of performance metrics must defined to measure how disturbances[16] affect service delivery. Based upon the current ancillary service definition, the chosen metrics are:

[16] See Figure B.3.2 for a visual representation of how disturbances affect the aggregator

*Time responsiveness:* how fast can the service be delivered from the moment the reference or measurement signal changes.

*Grid responsiveness:* how well can the aggregator follow changes in the grid state.

*Response accuracy:* how good is the aggregator at providing the full volume that is requested.

It is the TSO that defines the value of these metrics that signify a passed validation test. Since the tests are stochastic, the metric value should also have a stochastic component, this could for example be *time responsiveness* of service provision of 5 seconds with variance of $\pm$ 1 second. The metrics must be measured by an index and while literature has a wide array of indices for measuring performance, a specific index for aggregators is presented in Chapter 5.

The following steps are proposed for designing the test procedure:

1. The aggregator informs of the general composition of its portfolio, as well as the service it wants to be validated for.

2. The tester identifies the appropriate service requirements for the service to be tested for.

3. The tester identifies the expected normal operation of the aggregator.

4. The tester defines the test operation scenarios that the aggregator is expected to perform under. The scenarios must define the statistical properties, e.g. mean and variance, for the stochastic disturbances affecting the aggregator performance.

5. The tests are carried out on the aggregator:

   - simulation tests must be carried out, going through the appropriate factorial levels defined in the normal operation of the aggregator;
   - the simulation tests must be replicated with sufficient samples to capture how the error of the inputs propagates through the aggregator;

6. The aggregator performance is evaluated.

Depending on the excitation signals the aggregator is subject to, the tests are divided into two categories:

- step/ramp response, and
- continuous reference tracking.

The kind of test used for the validation will depend on the test scenario description.

In order to ensure that the simulations are correct, a limited selection of cases must be validated with field tests. This also ensures that the communication system between the aggregator and the system operator functions correctly.

To summarize, the validation procedure consists of a series of simulated tests, where the same excitation signal (be it a step/ramp response or a continuous signal) is replicated with enough samples, over a combination of factor levels, to identify the capabilities of the aggregator. A subsample of these tests must be validated through a field test.

An example of how the procedure is applied to an aggregator (without the final field test validation) can be found in Appendix B, Section B.4.

## 3.4 On Prequalification and Certification of Aggregators

It was previously mentioned that traditional generator prequalification consists of two steps, the documentation of the generator and the response test. The prequalification process must be adapted to aggregators. Parting from the concepts presented in this chapter, such a prequalification process could be the following:

1. *Documentation:* Description of the aggregator capabilities through the functional reference framework[17]. This can be used as check list for the basic required functionality.
2. *Validation test:* A set of response tests should be performed, along with the simulation aided validation procedure, in part to validate the aggregator reliability, but also to verify the communication between the TSO control center and the aggregator.
3. *Monitoring:* Furthermore, aggregator performance should be continually evaluated, and new validation tests should be carried out routinely. This is due to the dynamic nature of the aggregator portfolio, which may regularly change in size and composition, and due to the changes and updates that may come to the control algorithm.

[17] See Chapter 2.

The same process can be applied to a certification process, i.e. a process where a third part certifies the aggregator for participation in the different markets. In this case, the aggregator must be validated against other flexibility services, i.e. BRP portfolio balancing or distribution system services.

## 3.5 Conclusions Regarding the Validation Framework

The concept of validation of aggregators is important for the participation of aggregators in both ancillary services markets and other service markets. The original contribution of this work is the application of statistical method for validation test of aggregators. Also, the validation framework was presented, in which it is clear what are the elements that form part of aggregator validation.

In comparison with the traditional test method, the proposed validation procedure must capture the capabilities of a much more complex system, and therefore relies in part on simulations. A weakness in the proposed method is that the validation tests are highly dependable on the accuracy of the used models in the simulation. A way to mitigate this is to make the framework modular so that the tests can be run with hardware-in-the-loop (for model validation of individual units) or so that the framework can be connected to validated models, e.g. a Real-Time Digital Simulator (RTDS). The error between the used models and reality must be quantified[18] and taken into account for the final aggregator certification. Each block in the simulation must use validated models or software. This applies to the communication systems, the grid models and the DER models. The test architecture which validates the aggregators must also be validated.

[18] [118] Steinbrink and Lehnhoff. "Challenges and necessity of systematic uncertainty quantification in smart grid co-simulation". 2015.

Future work will consist of further refining the validation architecture, specifically defining the interfaces between modules, and implementing the software platform. Further exploration of the field of *Design of Experiments* may yield better methods than the fractional factorial method for quantifying the capabilities of the aggregator. Also, a set of realistic operation scenarios must be defined.

# Chapter 4
# Service Modeling and Requirements

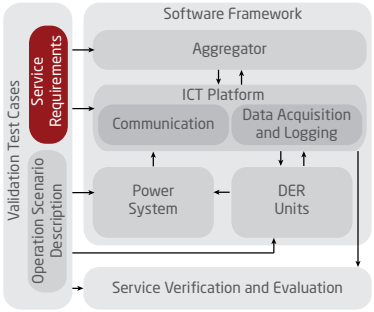



Figure 4.1: This chapter focuses on the *service definition* block of the aggregator validation framework presented in Chapter 3.

The requirements for ancillary services are in many countries defined, due to historical reasons, on the assumption that only generators provide ancillary services. It is clear that current service requirements are directly, or indirectly, blocking the integration of aggregators providing DR[1]. If aggregators are to be successfully integrated into the power system, the rules and requirements for participation must be changed. This chapter presents two contributions to integrating aggregators in the power system: a modeling method for service requirements, and a proposal for the restructuring of requirements for ancillary services. A method for modeling service requirements is important because the resulting models form the benchmark for the performance evaluation and verification of the aggregator (see Chapter 5), as well as being a direct input to the aggregator (see Chapter 2). The redefinition of ancillary service requirements is important since it will remove the barriers for aggregator participation in the AS markets, allowing system operators to utilize the properties of all available resources, both traditional and new, in an optimal way.

The concepts presented in this section are part of a submitted journal paper[2] which can be found in Appendix E and a draft journal paper[3] found in Appendix C, as well as work done as a collaborating author for a conference paper[4] and a technical report written for the iPower consortium[5]. Section 4.1 discusses different kinds of services that aggregators can provide and Section 4.2 presents how these can be modeled. These models are directly related to the *service requirements* block in the aggregator validation framework (see Figure 4.1). In Section 4.3 a proposal for how ancillary service requirements can be reformulated in order to be technology agnostic.

[1] [21] Cappers et al. "An assessment of market and policy barriers for demand response providing ancillary services in US electricity markets". 2013; [114] Smart Energy Demand Coalition (SEDC). "Mapping Demand Response in Europe Today". 2014.

[2] [13] Bondy et al. "Performance Requirements Modeling and Assessment for Ancillary Services in view of Demand Response". 2016.

[3] [10] Bondy et al. "Redefining Requirements of Ancillary Services for Technology Agnostic Sources". 2016.

[4] [61] Heussen et al. "A Clearinghouse Concept for Distribution-Level Flexibility Services". 2013.

[5] [12] Bondy and Thavlov. *FLECH PowerMax Service Requirement Specification*. 2014.

## 4.1 Background

The following section outlines concepts related to the definition and requirements of services at TSO and DSO level. While services

for the TSO (ancillary services) are well established, DSO and BRP services are a relatively new concepts which are being discussed, e.g. in the iPower project[6]. Also, the concept of Service Oriented Architectures is presented.

[6] [92] Nordentoft et al. *Development of a DSO-market on Flexibility Services*. 2013.

### 4.1.1 What are Ancillary Services?

Defining what ancillary services (AS) are, as well as which services the term includes, is difficult. This is due to both the differences in the way power systems are managed around the world and the differences in the terminology used to refer to such services. There is an overlap between the European and US definition[7] of AS in that both describe them as services used to ensure the reliability of the power system. In both European and US context reliability is addressed by considering *system adequacy* and *security* [8]. *System adequacy* is the power system's ability to supply the electricity demand at all times and *security* is the ability to withstand sudden disturbances.

Generally, maintaining an adequate and secure power system means maintaining the power system operating at nominal frequency and voltage. In cases where the power system deviates from nominal operation, either due to natural fluctuations in production/consumption or faults in the system, the system operators will activate ancillary services to restore normal operation.

Some countries, e.g. Denmark, consider voltage control, black start capabilities, short circuit control and reactive reserves as AS. This work focuses on those services that use active power to maintain the nominal frequency of the grid. In Europe[9] these services are frequency containment reserves (FCRs), frequency restoration reserves (FRRs) (either automatic[10] or manual), and replacement reserves (RRs)[11].

These reserves are activated as illustrated in Figure 4.2. The FCR is the fastest reserve and reacts automatically upon the grid measurements. Its role is to stop frequency excursions and its effectiveness can be measured by the *frequency nadir* [41]. The FRR is activated by tracking the Automatic Generation Control (AGC) signal, or through manual activation by the system operator, relieves the FCR (allowing the FCR to be available again) and restores the frequency to the nominal value. The RR relieve the FRR, usually through rescheduling of units or by bringing inactive units online.

[7] [130] Thermal Working Group. *Ancillary Services, Unbundling Electricity Products - an Emerging Market*. 2004; [132] US Federal Energy Regulatory Commission. *Order No. 888-A, Appendix B*. 1997.

[8] NERC also used the term system security, but in September 2001 security became synonymous with homeland protection in the US. Now it uses the term *operating reliability* [93]

[9] ENTSO-E changed in 2013 its nomenclature of AS, and the three presented here correspond roughly to the classical primary, secondary and tertiary reserves as presented in [110].

[10] In the United States, regulation is used for system balancing. This service corresponds to automatic FRR.

[11] [44] European network of transmission system operators for electricity (ENTSO-E). *Network Code on Load-Frequency Control and Reserves*. 2013.

### 4.1.2 Requirements for Ancillary Services

Because AS are essential for the secure operation of the system, the system operators have requirements and restrictions on the units providing AS. A super-set of requirements across different systems was presented by *Rebours*[12], and following his overview we classify requirements into three categories:

[12] [110] Rebours. "A Comprehensive Assessement of Markets for Frequency and Voltage Control Ancillary Services". 2008.

*temporal requirements* which relate to how fast and for how long a service must be delivered;

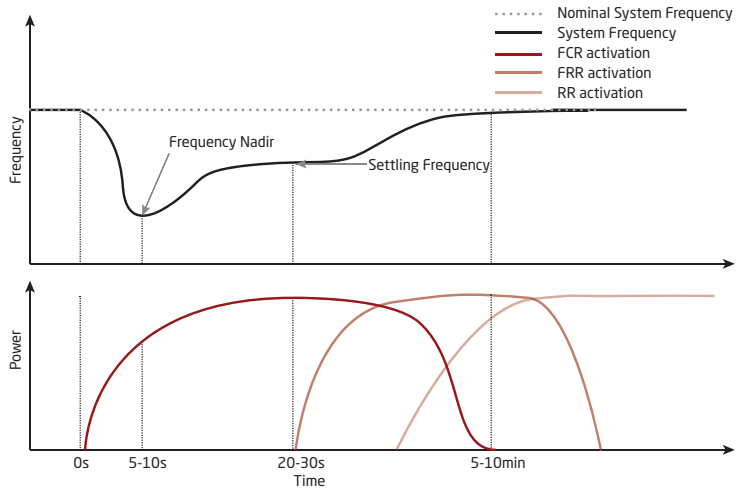



Figure 4.2: The ancillary services are activated sequentially after a frequency contingency. In systems with high inertia the frequency nadir will occur at frequencies closer to the nominal frequency.

*resource tuning requirements* which relate to specific values that tuning parameters in the resource must have;

*market requirements* which relate to bid sizes and similar parameters in systems where services are acquired through market mechanisms.

Of these three categories, only the temporal requirements relate to service performance. Furthermore, in most systems, the requirements are implicitly defined for traditional generation units. This means that most service requirements are oriented towards the least common denominator of service providers, e.g. a unit providing FCR should provide half of the service within 15 seconds and full response within 30 seconds[13]. A variety of generation and consumption units would be able to provide this service faster, but this quality is not rewarded. Another example is the requirement of having a PI-controller on units providing FRR in order to track the AGC signal. Such a controller is infeasible on distributed systems, but other modern controllers can provide offset-free control with similar properties. This means that the historical requirements for units participating in AS markets in many countries act implicitly, or explicitly, as barriers for new technologies to enter the market[14].

The concept of using demand side management to help the secure operation of the power grid has existed in different forms since the late 1970s[15]. But in recent years, the introduction of new consumption and generation technologies, i.e. DERs, along with the roll-out of a smart metering infrastructure and the advances in ICT, has lead to new opportunities in using smart control of small scale consumption/production as a service to the power grid. There is a large body of literature[16] concerning DR, and proposals to use it for AS[17].

### 4.1.3 Distribution System Services

As the amount of DERs installed at distribution level increases, the DSOs face new operational problems. Mainly, the increase in electric load will cause congestion and voltage issues. The traditional way of handling these

[13] [34] Energinet.dk. *Ancillary services to be delivered in Denmark - Tender conditions*. 2012.

[14] [21] Cappers et al. "An assessment of market and policy barriers for demand response providing ancillary services in US electricity markets". 2013; [114] Smart Energy Demand Coalition (SEDC). "Mapping Demand Response in Europe Today". 2014.

[15] [81] Lampropoulos et al. "History of demand side management and classification of demand response control schemes". 2013.

[16] [96] O'Connell et al. "Benefits and challenges of electrical demand response: A critical review". 2014.

[17] [135] Vrettos, Ziras, and Andersson. "Integrating large shares of heterogeneous thermal loads in power system frequency control". 2015; [87] Mathieu, Dyson, and Callaway. "Using Residential Electric Loads for Fast Demand Response: The Potential Resource and Revenues, the Costs, and Policy Recommendations". 2012; [137] Zarogiannis et al. "A Dynamic Behaviour Analysis on the Frequency Control Capability of Electric Vehicles". 2014.

are through reinforcement of the grid assets. Given the high cost of installing new cables, and the uncertainty in how the electricity consumption will change in the future, the use of flexibility services will be an attractive alternative.

One of the main outcomes of the iPower project was the definition of a set of flexibility services that demand aggregators can provide DSOs[18] for congestion management or voltage issues. The requirements for three of the congestion management services have been further detailed individually[19], and aggregator architectures have been designed to provide both congestion management[20] and voltage support[21]. At the same time, the concept of the Flexibilty Clearing House (FLECH) has been designed as a platform to enable the transparent contracting of flexibility[22].

An example of a flexibility service is the *PowerMax*[23] service, where the aggregator maintains the total consumption of its portfolio under a limit, within a specified period of time. This means that the aggregator is free to manipulate its portfolio as long as its peak load is below the limit specified by the DSO.

[18] [92] Nordentoft et al. *Development of a DSO-market on Flexibility Services*. 2013.

[19] The services requirements were detailed in the following technical reports[59, 4, 12].

[20] [64] Hu et al. "Coordinated Charging of Electric Vehicles for Congestion Prevention in the Distribution Grid". 2014.

[21] [57] Han et al. "Assessment of distribution grid voltage control strategies in view of deployment". 2014.

[22] [61] Heussen et al. "A Clearinghouse Concept for Distribution-Level Flexibility Services". 2013.

[23] [12] Bondy and Thavlov. *FLECH PowerMax Service Requirement Specification*. 2014.

### 4.1.4 BRP Portfolio Balance Service

The idea of using aggregators for internal balancing of BRPs was also explored in the iPower project[24]. The main objective of such a service is for the aggregator to provide services to BRPs so that they can avoid imbalance costs. This service is similar to FRR, except the contract is with a BRP, which means that contracted volumes will usually be smaller, and the acquisition horizon shorter. BRP portfolio balance services must be acquired 1 hour to 5 minutes before the activation period, and must be cheaper than the imbalance cost.

The concept of BRP services is also analysed by the Universal Smart Energy Framework (USEF) foundation. They propose four different BRP services[25]:

- Day-ahead portfolio optimization,
- Intraday portfolio optimization,
- Self-balancing,
- Generation optimization.

[24] [131] Tougaard, Hansen, and Sundström. *FLECH BRP Service - Portfolio Balancing*. 2015.

[25] [133] USEF foundation. *USEF: The Framework Explained*. 2015.

### 4.1.5 Asset Management Services

In Chapter 2 the concept of Asset Management Services (AMS) is introduced as the services that an aggregator provides to the owner of the DERs, or flexibility assets. An example of this is the case where the aggregator is an EV fleet operator that has the contractual responsibility of maintaining all EVs in the fleet within a certain State of Charge (SoC). The purpose of validating aggregators for these services is that flexibility asset owners can use the validation as a trust measure.

The main idea behind AMS is that the flexibility assets have a primary purpose, which is to satisfy the needs of their owner. The aggregator can use the flexibility of the units as long as the primary purpose is respected. Thus, from the perspective of customer comfort, an aggregator that is better at AMS is more desirable.

### 4.1.6 Service Oriented Architectures

The concept of Service Oriented Architecture (SOA) comes from the field of computer science and is an approach to creating software architectures based upon the concept of services. Parallels to this approach can be drawn with the way ancillary service provision works in unbundled energy markets.

Under the SOA paradigm, service is defined as: a logical representation of a repeatable business activity that has a specified outcome, is self-contained, may be composed of other services and is a *black box* to consumers of the service[26]. The following concepts are central to the SOAs:

*Standardized service contracts:* These can be interpreted as interfaces, in that they describe the service purpose and functionality[27].

*Service Level Agreement (SLA):* These are part of the standardized service contracts which define the service performance metrics with corresponding Service Level Objectives. SLAs can be interpreted as the requirements defined in an ancillary service contract.

*Service Level Objective (SLO):* These define the agreed means for measuring performance.

[26] [129] The Open Group. *Service-Oriented Architecture (SOA) - A white paper*. 2007.

[27] [40] Erl. *The SOA Magazine Service-Orientation and Object-Orientation Part I: A Comparison of Goals and Concepts*. 2008.

## 4.2 Modeling of Service Requirements

The validation framework presented in Chapter 3 uses the *service requirements* as a benchmark towards which the aggregator is evaluated. This is because the requirements contain the control objective of the aggregator. Currently, requirements for services are encoded within the contractual agreements between system operator and service provider. A standard method is needed for extracting this information and building a model that can be used for benchmarking, i.e. a *standardized service contracts* must be defined with their corresponding SLA and SLO.

By analysing the services presented in Section 4.1, the following requirements for the service requirements model are defined:

*M-R1* the model must clearly identify the SLOs of the service,

*M-R2* the model must incorporate both the ideal and acceptable service provision in a measurable/quantifiable way, i.e. performance metrics must be able to be applied to it,

*M-R3* the models must be technology agnostic,

*M-R4* since flexibility services imply a change of consumption pattern over a period of time, the models must consist of time series.

Based upon these requirements, a method for translating the contracts into a time series model has been developed. The method consists of the following six steps:

1. Identify physical parameters defining the service [M-R1].
   - e.g. Power production or consumption, measured grid frequency, time measurements, etc.
2. Identify the dynamic behaviors of the service related to system parameters (if any) [M-R1].
   - e.g. FCR expects a linear relation between a deviation from the nominal grid frequency and the generator set-point.
3. Identify the physical size of the service and the tolerated error [M-R2].
   - e.g. the volume of the bid for FCR.
4. Identify the ideal response time of the service and acceptable response [M-R2].
   - e.g. FCR in western Denmark must be 50 % of activated within 15 s and 100 % within 30 s.
5. Based on the dynamics, size and timing of the service, as well as the tolerated errors from points 1–4, develop a time series for ideal and acceptable service provision. The model will be a set of time series: $\mathbf{x}_{ideal}(t)$ for ideal response and $\mathbf{x}_{acc}(t)$ for acceptable response. Both time series can be a scalar or a vector, e.g. $\mathbf{x}_{acc}(t)$ can be formed by a set of upper and lower tolerance bounds or simply by an upper bound [M-R4].
6. Identify how the service error is to be measured [M-R1].

By only defining the SLA models in terms of performance, not in specific unit capabilities, the models implicitly comply with [M-R3].

Furthermore, the analysed services can be divided into three patterns:

*Reference Tracking:* Services where a reference signal must be followed, e.g. regulation in the United States.

*Band Service:* Services where the output is able vary between an upper and lower limit, e.g. smart charging of a fleet of EVs.

*Cap Service:* Services where the output must respect either a upper or lower bound, e.g. the *PowerMax* service.

Based upon the three kinds of service, the service error, i.e. step *6* in the modeling method, can be measured the following ways:

### Reference tracking

Reference tracking error can be calculated as:

$$e(t) = x_{meas}(t) - x_{ideal}(t), \tag{4.1}$$

where $x_{meas}(t)$ is the measured output, e.g. the total load of the aggregator portfolio, and $x_{ideal}(t)$ is the ideal response defined in the service model. This definition will lead $e < 0$ for measured values below the ideal and $e > 0$ for values above the ideal. In this case $\mathbf{x}_{acc}(t)$ will be a band around $x_{ideal}(t)$, and the values of $\mathbf{x}_{acc}(t)$ do not need to be symmetric.

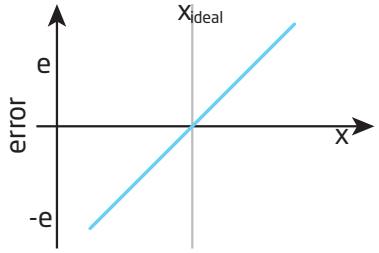


Figure 4.3: Error in reference tracking.

### Band service

The ideal response in a band service is defined as $\mathbf{x}_{ideal}(t) = [x_{min}(t), x_{max}(t)]$. The error in the band service can therefore be estimated by:

$$e(t) = \begin{cases} x_{meas}(t) - x_{min}(t), & x_{meas}(t) < x_{min}(t) \\ 0, & x_{min}(t) \leq x_{meas}(t) \leq x_{max}(t) \\ x_{meas}(t) - x_{max}(t), & x_{meas}(t) > x_{max}(t). \end{cases} \tag{4.2}$$

In this case, the $\mathbf{x}_{acc}(t)$ is a set of values that surrounds the band defined by $\mathbf{x}_{ideal}(t)$, as seen in Fig. E.5. The values of $\mathbf{x}_{acc}(t)$ do not need to be symmetric around the band.

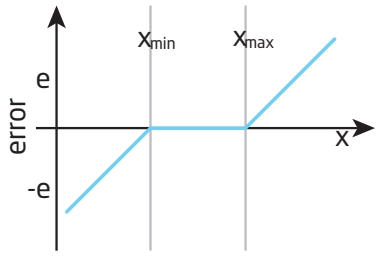


Figure 4.4: Error in band service.

### Cap service

In cap services, error is only tracked when $x_{meas}(t)$ is either above or below a given a limit value. Maximum cap error is calculated as shown in (4.3) and minimum cap can be similarly calculated. In (4.3), $x_{max}(t)$ is the ideal maximum limit according to the service contract:

$$e(t) = \begin{cases} x_{meas}(t) - x_{max}(t), & x_{meas}(t) > x_{max}(t) \\ 0, & x_{meas}(t) \leq x_{max}(t). \end{cases} \tag{4.3}$$

In the cap service, $x_{acc}(t)$ is a limit that either lies below $x_{min}(t)$ or above $x_{max}(t)$.

The applicability of these models is showcased in Section 5.2.

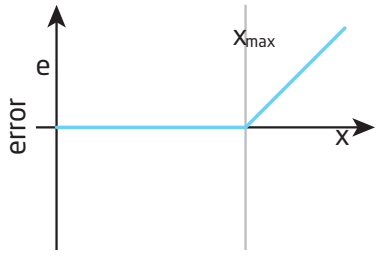


Figure 4.5: Error in cap service.

## 4.3 Restructuring Ancillary Service Requirements

This section is a summary of the draft paper in Appendix C. Some of the concepts missing in the draft paper have been introduced here.

Until now, system operators have been able to arrest frequency excursions fast enough because of the inherent system inertia. With the increasing penetration of wind power in the system, the electricity prices are lowered and operating fossil-fueled generator becomes economically unfeasible. This has the effect of reducing the system inertia, and reducing the availability of AS resources. Therefore new AS sources with faster response times are required. *Vrettos et al.*[28] show that if FCR is provided by DR (with a very

[28] [135] Vrettos, Ziras, and Andersson. "Integrating large shares of heterogeneous thermal loads in power system frequency control". 2015.

fast response), the frequency nadir occurs at higher frequencies. Also, *Makarov et al.*[29] argue that the value of regulation resources can be defined based upon the ramp capabilities of the service providing units. Faster reacting units are more valuable to system operators, since they help arrest the frequency excursion faster and at a higher nadir.

[29] [85] Makarov et al. *Assessing the value of regulation resources based on their time response characteristics*. 2008, p. 2cm.

AS requirements are specified by a system operator based on the desired control response for a particular power system, under the implicit assumption that the ideal unit response corresponds to a scalar fraction of the required system response. Today, these requirements — as reflected in the service definition — are not differentiated according to the capabilities of the unit providing the service. Therefore, service definitions are designed to accommodate the least capable unit in the portfolio. As a consequence, more capable units are not being fully utilized, leading to excess contracting of service providers.

In this section a new form of defining AS requirements is presented, which has as an objective to allow all units to participate in the AS markets on equal footing. The main assumption is that all units can be valuable for AS provision, even when they do not fully comply with current requirements, and that the system operators will be able to manage the system better if the capabilities of all available resources are utilized. Also, units should be remunerated based upon the value they represent to the operation of the system, and their performance compared to these expected values.

Regulative authorities have concluded that fast reacting units are valuable for the system operation, and started programs to benefit of these resources. An example of this is FERC order 755 (Pay for Performance) which has led to PJM splitting their regulation market product into RegA, for slow reacting units, and RegD for fast reacting units. The product differentiation approach has been a success for PJM, but splitting the market may still lead to suboptimal utilization of units and does not address the issue of new technologies being effectively excluded from certain markets. We propose instead to restructure the ancillary service definitions such that all types of service providers participate with the same market product defined by a set of optimal performance parameters, and not by minimum requirements. This means that the all entities providing a given ancillary service are optimally cleared under a single market.

### 4.3.1 The Overall Approach

The restructuring of AS requirements is formed by the following key concepts:

1. The system operator is able to formulate an overall *ideal AS response* that can be achieved through an optimal mix of resources. Any resource can make a bid for providing part of this ideal response.

2. The *parametrization of AS bid*, where the parameter values of each unit/bid reflect the service provider's capabilities to partially fulfill the ideal response. This avoids excluding units that may have useful capabilities in one parameter, e.g. very fast ramp rate, but low capabilities in

another parameters, e.g. only holding the response for a short time. Such a service definition allows compliance to be measured on a linear rather than a binary scale: In addition to compliance and noncompliance, different levels of partial compliance are possible.

3. Clearing all units under a *generalized single clearing-price auction* allows constructing an optimal portfolio, and enables competition between all resources, leading to lower prices.

4. *Performance-based remuneration* gives incentive to better AS provision and enables transparent performance-based clearing of the market.

These points are further described in the following subsections, although the focus of this work is on the *parametrization of AS bid* and the proposal for a market clearing mechanism. The concepts are explained in the rest of this section and Section 4.3.2 presents an example of the approach.

### Ideal Service Tender

*Makarov et al.*[30] define the ideal source for AS as one with "unlimited capabilities in terms of response time, energy output, ability to frequently reverse their output, ability to respond and follow the AGC setpoint changes, and size."[31] It is impossible for any one unit to possess these characteristics, but system operators aim at achieving this kind of system response by contracting several units.

In order to define an ideal tender the system operator must identify the needs of the system through metrics as those presented in Section 5.1.2, e.g. the *nadir-based frequency response* proposed by *Eto et al.*[32]. By understanding the specific needs the system, the system operator will be able to define which are the relevant parameters that will satisfy these needs.

[30] [85] Makarov et al. *Assessing the value of regulation resources based on their time response characteristics*. 2008.

[31] For this kind of response to be optimal, changes must be made to the AGC algorithm [102].

[32] [41] Eto, Joseph H et al. "Use of frequency response metrics to assess the planning and operating requirements for reliable integration of variable renewable generation". 2010.

### Service Parametrization

The overall ideal service requirements can be expressed as:

$$S^* = f_m(\mathbf{x}^*) \tag{4.4}$$

where $\mathbf{x}^*$ is a vector of ideal parameter values and $f_m(\cdot)$ a function that translates the parameters $\mathbf{x}$ into a model, e.g. into a time series as shown in Section 4.2. Furthermore, the system operator must inform how the parameters are valued with respect to the $S^*$, which is done through a capability value:

$$\kappa = g(\mathbf{x}) \tag{4.5}$$

$$\kappa \in [0, 1]. \tag{4.6}$$

The bids submitted on a parametrized tender must contain:

- the offered service parameters: $\mathbf{x}$
- the bid price: $P^{bid}$

- the estimated capability value[33]: $\kappa$

[33] This is proposed for cross-validation of the capability value

The bid parameters are service-specific and serve both for market-clearing and performance calculation.

### Market Mechanism

It is impossible to restructure the AS definition without addressing the market mechanism for determining the optimal set of resources. The market should be designed as a single clearing price auction, in which each resource bid is adjusted by *two* factors for bid quality: 1) the capability value $\kappa$ and 2) a historical performance[34] parameter $\eta^{hist}$. The historical performance parameter determines how close the unit has followed the properties defined in $\kappa$.

[34] The historical performance parameter can also be interpreted as an availability parameter, or certainty parameter. PJM also uses the historical performance score of unit for the settlement.

### Performance Based Remuneration

The estimation of the service provision performance can be done in different ways, depending on which parameters the system operator deems to be the most critical. The concept of performance assessment is discussed in Chapter 5, and performance index is introduced there. Here it suffices to say that the performance measurement is a function of the error in service delivery:

$$\eta = c(e(t)), \tag{4.7}$$

$$\eta \in [0, 1] \tag{4.8}$$

This value, along with the capability value $\kappa$ should form part of the service remuneration.

## 4.3.2 Approach Example

This section presents an implementation example of the AS requirements restructure. It must be noted that these are only examples and further research should be done in how best to implement the presented concepts.

### Ideal Tender

A system operator could determine that the ideal system response to a frequency excursion is the one that has a resulting frequency nadir at the settling frequency (thus minimizing the risk of tripping the under-frequency relays). Based upon the inertia in its system, the system operator determines the volume ($V_{tot}$) needed as well as the response characteristics needed to achieve this, see Figure 4.6.

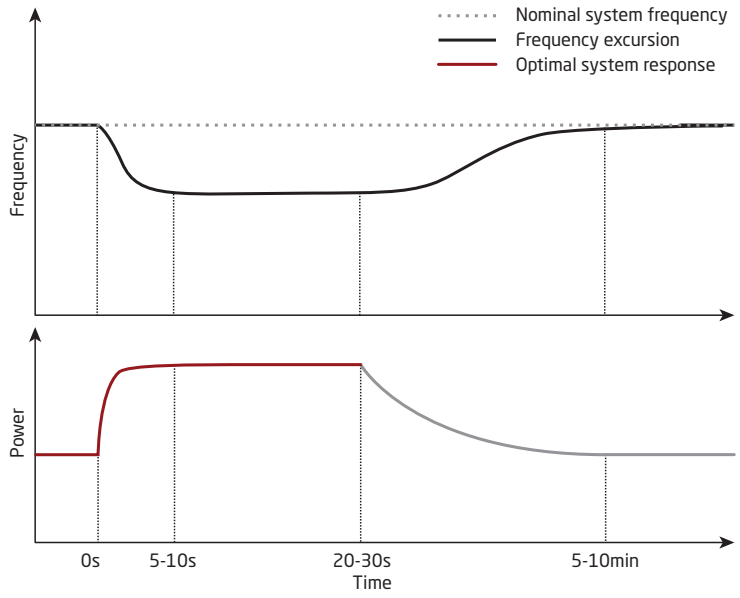



Figure 4.6: In this case, the ramp of the ideal response is mainly determined by the system inertia and is to be sustained until FRR can be activated.

## Service Parametrization

A system operator decides that the FCR in their market is defined by $\mathbf{x} = [\tau_r, \tau_d, V]$, where $\tau_r$ is the rise time of the service, $\tau_d$ is the duration of the service, and $V$ is the volume of the service. Due to the properties in its system, it decides that $\mathbf{x}^* = [30s, 20min, 90MW]$. The capability value of each bidder is calculated by:

$$\kappa_i = \alpha_1 \frac{\tau_{r,0}}{\max(\tau_{r,0}, \tau_{r,i})} + \alpha_2 \frac{\min(\tau_{d,0}, \tau_{d,i})}{\tau_{d,0}} + \alpha_3 \frac{V_i}{V_{tot}}, \quad \forall i \in \Omega \tag{4.9}$$

where $\tau_{r,0}$ and $\tau_{d,0}$ are a nominal value the system operator sets, $\tau_{r,i}$ and $\tau_{d,i}$ are the actual parameter values for each bidder, $\frac{V_i}{V_{tot}}$ is the bid contribution to the total required volume, and $\Omega$ is the pool of bids. Finally, $\sum_i \alpha_i = 1$, and in this case could be $\alpha_1, \alpha_2 = \frac{2}{5}, \alpha_3 = \frac{1}{5}$.

## Market Mechanism

The proposed clearing mechanism identifies a common clearing price based on the most expensive accepted bid[35]:

$$P^{\mathtt{clear}} = \max P_i^{\mathtt{bid}}, \quad i \in \Omega^{\mathtt{acc}} \tag{4.10}$$

[35] This is similar to the merit order lists used in e.g. the Nordic system for Manual Regulating Power[11].

where $\Omega^{acc} \subseteq \Omega$ is the subset of accepted bids of the set of received bids $\Omega$. The clearing mechanism selects the subset of bids which offer the cheapest overall clearing cost and meet the tender requirements with a given certainty of availability:

$$\Omega^{\mathtt{acc}} = \mathtt{argmin}_{\Omega^{\mathtt{hyp}} \in \mathcal{P}(\Omega)} \sum_{i \in \Omega^{\mathtt{hyp}}} \kappa_i P_{\Omega^{hyp}}^{\mathtt{clear}} \tag{4.11}$$

$$\text{subject to} \quad \sum_{i \in \Omega^{\mathtt{hyp}}} f_m(\mathbf{x_i}) \geq S^* \tag{4.12}$$

$$\eta_i^{hist} \geq \eta_{min}^{hist} \quad \forall i \in \Omega^{\mathtt{hyp}} \tag{4.13}$$

Where $\mathcal{P}(\Omega)$ denotes the Power Set of $\Omega$ and $\Omega^{hyp}$ is a (hypothesis) subset of the Power Set. $S^*$ is the ideal tender from Eq. (4.4) and $\eta_{min}^{hist}$ is the minimum historical performance requirement to participate in the market[36].

[36] This value represents how averse the system operator is to risk, and could also be considered part of the service parameters $\mathbf{x}$.

### Performance Based Remuneration

We propose that remuneration must be based on the performance evaluation of the service provision:

$$P_i^{\mathtt{rem}} = \eta_i \kappa_i P^{\mathtt{clear}} \qquad \forall\, i \in \Omega^{\mathtt{acc}}. \tag{4.14}$$

Thus, remuneration is based upon the value the resource has to the grid operator, how well it performs, and the most expensive activated resource.

## 4.4 Conclusions on Service Requirements

Aggregators have become possible sources for ancillary services and distribution system services. While system operators are aware of the potential in using flexibility for system balancing, the ancillary service requirements have not been changed in order to accommodate this new technology. This chapter presented a novel proposal for solving this issue, by restructuring the ancillary service requirements based upon a set of optimal parameters instead of the limiting minimum requirements found in many systems today.

Also, a method for modeling services was shown. The resulting models are relevant for the validation framework in that they provide the benchmark towards which aggregators must perform. In Chapter 5 it is shown how the service models can be used for performance assessment and verification of services.

The work presented in this chapter differs from the rest of the thesis, in that the concepts presented here are not focused on the aggregator itself. The service modeling method can be applied to any kind of service, not necessarily those provided to be provided by aggregators, and the objective of the ancillary service restructuring is to include any new technology, not only aggregators providing DR.

The AS restructuring is part of a draft paper and needs to be further refined. The implementation of the presented concepts needs further research, especially the mapping from service needs to parametrization, as well as a fair market mechanism.

# Chapter 5

# Performance Assessment & Verification of Aggregator Services

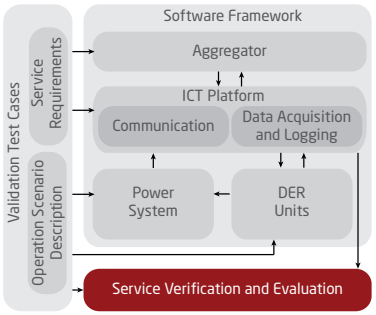



Figure 5.1: This chapter focuses on the *service verification and evaluation* block of the aggregator validation framework presented in Chapter 3.

The IEEE defines verification as:"confirmation, through the provision of objective evidence, that specified requirements have been fulfilled."

[1] [114] Smart Energy Demand Coalition (SEDC). "Mapping Demand Response in Europe Today". 2014.

Performance assessment is the process of quantifying and verifying the provision of a service according to the contractual specifications of the service. Performance assessment usually occurs at three stages[1]

- To qualify potential resources against service specifications as part of the validation/prequalification procedure.
- To verify service conformance to the service specifications during and after service delivery.
- To calculate the amount of service delivered by the resource as part of financial settlements.

The main contribution presented in this chapter is applying concepts from the field of Control Performance Assessment, which is a mature field within the process industry, to the power system. This is reflected in a set of performance indices developed for aggregator performance assessment which can be applied at the three stages outlined above. Coupled with the performance requirement models presented in Chapter cha:services these indices are a flexible method for performance assessment and verification of service delivery.

The initial work on aggregator performance assessment was presented in a conference paper[2] and further refined in a submitted journal paper[3]. These papers can be found in Appendix D and Appendix E.

[2] [14] Bondy et al. "Performance assessment of aggregation control services for demand response". 2014.

[3] [13] Bondy et al. "Performance Requirements Modeling and Assessment for Ancillary Services in view of Demand Response". 2016.

## 5.1 Background

Little attention has been given to the problem of performance assessment of aggregator controllers seen from a service-delivery perspective. As stated in Section 3.1.2, performance assessment of aggregators

has been mostly ad-hoc analysis specific to a problem the designers are trying to solve, but none have taken a systematic approach to the evaluation of aggregators in terms of established service requirements. In this section we present the concept of control performance assessment, performance indices in power systems and current service verification procedures, which are topics that serve as background for the rest of the chapter.

### 5.1.1 Control Performance Assessment

Control Performance Assessment (CPA) is already an established field within control engineering. Most of the applications within the field are found in the process industry[4], but since aggregators are a control system, and provide control services, it is natural to translate concepts of CPA to the power system. Usually, CPA methods fall within two types:

- benchmarking of controllers towards a theoretic optimum, taking stochasticity of the process into account; and
- benchmarking against deterministic properties required of the close-loop system.

Usually these indices are normed so that for an index $\eta$:

$$\eta \in [0, 1]. \tag{5.1}$$

[4] [67] Jelali. "An overview of control performance assessment technology and industrial applications". 2006.

### 5.1.2 Performance Indices in the Power System

Currently, the concepts of performance indices and evaluation criteria are used in the power system for the general assessment of how well the System Operator is managing the system. But these evaluation concepts are usually tailored to specific services, e.g. CPS1 and CPS2[5] used by NERC[6] for evaluating regulation, or the *nadir-based frequency response* metric[7] used for evaluating the quality of primary frequency control in an area. Other evaluation criteria have the power interruption to the end customer in focus, e.g. System Average Interruption Duration Index (SAIDI) and System Average Interruption Frequency Index (SAIFI)[8].

As part of the FERC order 755[9], PJM has introduced a performance score for the remuneration of services in the form of:

$$\text{Performance Score} = AS_A + BS_D + CS_P \tag{5.2}$$

where $S_A$ is an accuracy score, $S_D$ is a delay score, $S_P$ is a precision score, and $A + B + C = 1$ are scalar weights. These measure the delay and correlation between the regulation signal and the reaction of the unit, and the difference in energy requested vs. energy supplied [105]. While this is a detailed performance metric, it is tied to the way regulation is done in PJM (tracking of the regulation signal). Therefore, more general (and simple) models and performance metrics are needed to cover other frequency regulation services and the new flexibility services.

[5] An alternative to these two Control Performance Standards is formulated in [50].

[6] [89] NERC Resources Subcommittee. *Balancing and frequency control - A Technical Document*. 2011.

[7] [41] Eto, Joseph H et al. "Use of frequency response metrics to assess the planning and operating requirements for reliable integration of variable renewable generation". 2010.

[8] [79] LaCommare and Eto. "Cost of power interruptions to electricity consumers in the United States (US)". 2006.

[9] The order stipulates that all units providing regulation should get remunerated based upon their performance.

A measure for the performance of aggregators, that is not directed at a single service and that has service delivery in focus, is the topic of this chapter.

### 5.1.3 Service Verification Today

When contracted for service, units are subject to a set of requirements. First, units must pass a prequalification test. Second, certified metering instrumentation must be installed on the unit, and (expensive) telemetry equipment must be installed and connected to the system operator's Supervisory and Control Data Acquisition (SCADA) system.

For verifying reserve services, the system operator does random checks to see if the reserve is available at the unit[10]. With respect to regulation services, these are expected to be delivered within the required time requirements, and must be measured with acceptable accuracy. For example, for consumption units smaller than 1.5 MW acceptable accuracy is 2% of the load[11].

With the introduction of aggregators as providers of ancillary services, the AS specifications are being adapted to new resource types, but also prequalification and verification of service delivery need to be adapted to be suitable for the aggregated service delivery[12]. This is relevant both due to the change in ancillary service specifications and due to the introduction of new distribution system services[13].

[10] [34] Energinet.dk. *Ancillary services to be delivered in Denmark - Tender conditions*. 2012.

[11] [35] Energinet.dk. *Teknisk Forskrift 5.8.1 - Måledata til systemdriftsformål*. 201l.

[12] [7] Bode et al. "Incorporating residential AC load control into ancillary service markets: Measurement and settlement". 2013.

[13] [62] Heussen et al. "A clearinghouse concept for distribution-level flexibility services". 2013.

## 5.2 Quality of Service

The concept of Quality of Service (QoS) is closely related to the service models presented in Section 4.2. One of the elements of a service model is the definition of the service error. QoS is an instantaneous measure of how well the aggregator is delivering a service at any given time instant, and can be defined as the scaling of the error to the limits defined in the service model, i.e. :

This section relies heavily on the service modeling concepts presented in Section 4.2, and it is recommended that the reader familiarizes with that section before reading this section.

$$QoS(t) = e(t)C_n(t), \tag{5.3}$$

where $e(t)$ is the error in service delivery and $C_n(t)$ is a time varying normalization factor. This factor ensures that:

- $QoS \geq 0$,
- for $QoS \leq 1$ the service is considered delivered within the contractual constraints, and
- $QoS = 0$ is a perfect service delivery.

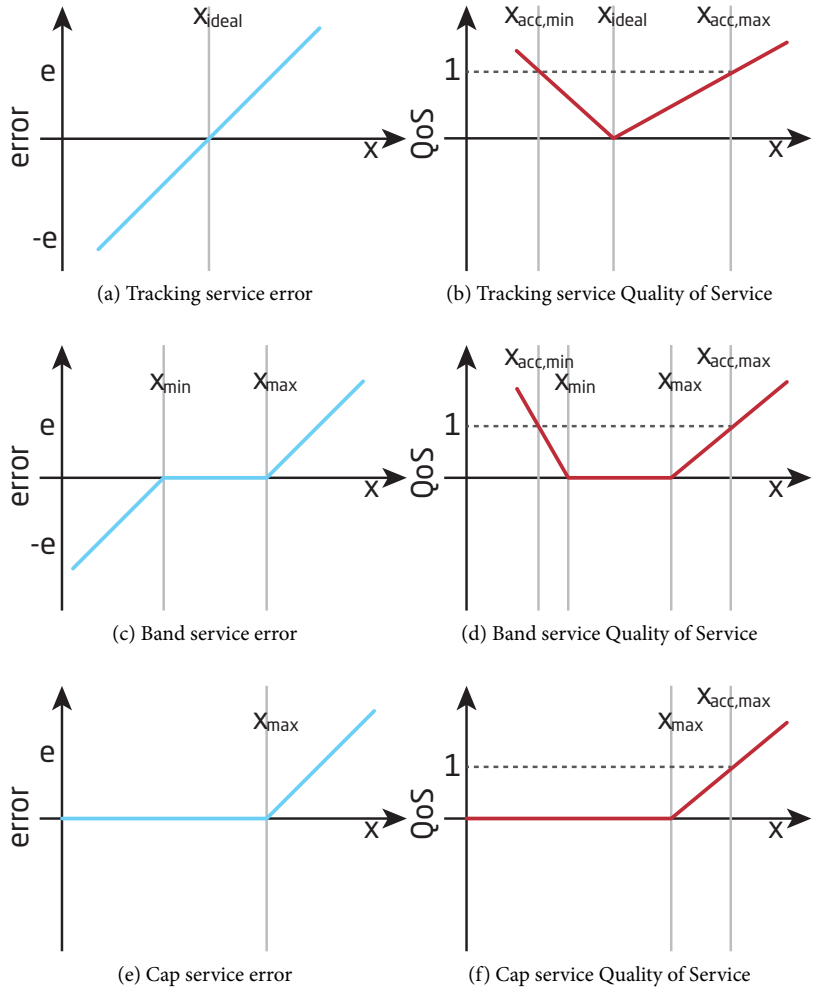
The original definition proposed in [14] assumed symmetric constraints around the acceptable provision, but in [13] this definition was expanded to

account for asymmetry, thus $C_n(t)$ is defined as:

$$C_n(t) = \begin{cases} \frac{1}{x_{acc,max}(t) - x_{max}(t)}, & e(t) \geq 0 \\ \frac{1}{x_{acc,min}(t) - x_{min}(t)}, & e(t) < 0. \end{cases} \tag{5.4}$$

where $x_{acc,max/min}$ and $x_{max/min}$ are part of the service model defined in Section 4.2. A visual representation of the error models and their corresponding *QoS* definition are shown in Figure 5.2. Note that in Equation (5.4), $C_n(t)$ is not defined for $x_{acc}(t) = x_{ideal}(t)$. This is a corner case, in which:

$$QoS(t) = e(t), \quad x_{acc}(t) = x_{ideal}(t) \tag{5.5}$$




(a) Tracking service error

(b) Tracking service Quality of Service

(c) Band service error

(d) Band service Quality of Service

(e) Cap service error

(f) Cap service Quality of Service

Figure 5.2: Error and QoS for the three kinds of services, note that the acceptable band do not need to be symmetric.

## 5.3 The Aggregator Performance Indices

Performance criteria used for evaluating controllers usually fall within three categories[14]: quality, reliability and energy efficiency. When assessing aggregators, service quality and service reliability define the performance of the aggregator. Three requirements are defined for the performance criteria of aggregators:

[14] [49] Green, Izadi-Zamanabadi, and Niemann. "On the choice of performance assessment criteria and their impact on the overall system performance - The refrigeration case study". 2010.

*P-R1* provide a *quality* measure normalized with respect to the contractual requirements (bounds) of a service and with respect to time:

$$\eta = f_P(x_{meas}, \mathbf{x}_{acc}, t), \quad \eta \in [0, 1], \tag{5.6}$$

*P-R2* provide a *reliability* measure in relation to service non-delivery, which is normalized with respect to time:

$$\epsilon = f_R(x_{meas}, \mathbf{x}_{acc}, t), \tag{5.7}$$

*P-R3* service quality and reliability evaluation must be applicable to entities providing multiple services:

$$\eta_M = \sum_{i \in M} f_M(\eta_i), \quad \eta_i \in [0, 1], \tag{5.8}$$

$$\epsilon_M = \sum_{i \in M} f_M(\epsilon_i), \tag{5.9}$$

where $\eta$ is a quality performance measure and $\epsilon$ is a reliability measure. $\eta_M$ and $\epsilon_M$ are the same measures applied to multiple services $M$. The measured output (or sum of outputs in the case of aggregation) is defined by $x_{meas}$, and the service bounds are defined by $\mathbf{x}_{acc}$, as defined in Section 4.2. $f_P(\cdot)$ is a function that evaluates service performance normalized to $\mathbf{x}_{acc}$ and time $t$. Similarly, $f_R(\cdot)$ is a function that evaluates service reliability based upon $\mathbf{x}_{acc}$ and normalized to time and $f_M(\cdot)$ is a function that gives an overall measure for multiple services

### 5.3.1 Service Performance Assessment index

The service performance assessment index consists of the weighted average of the normalized root mean square error (RMSE) of the service delivery[15]. The error is based upon the concept of QoS. Since reliability is measured separately, the performance assessment index measures the error as defined by:

$$QoS^{AS}(t) = \begin{cases} QoS^{AS}_{meas}(t), & \forall QoS^{AS}_{meas}(t) \leq 1, \forall t \\ 1, & \forall QoS^{AS}_{meas}(t) > 1, \forall t. \end{cases} \tag{5.10}$$

where $QoS^{AS}_{meas}(t)$ is the measured error in service delivery.

For evaluation of $K$ amount of ancillary services, over discrete time horizon of service delivery $N$, the index is defined:

$$\eta^{AS} = \sum_{i=1}^{K} W_i^{AS} \sqrt{\frac{\sum_{t=0}^{N_i} \left( {QoS^{AS}_{i,t}}^2 \right)}{N_i}} \tag{5.11}$$

$$\sum_{i=1}^{K} W_i^{AS} = 1 \tag{5.12}$$

where $QoS^{AS}_{i,t}$ is the truncated $QoS \in [0, 1]$ of the ancillary serviced delivery. This definition means that $\eta^{AS} \in [0, 1]$, where values close to 0 mean

[15] Originally, this index was defined in [14] as the integral square error (ISE) of the service delivery, which was then normalized to a maximum allowable error. This definition does note cope well when the service provision of several services are evaluated at the same time. Therefore, the index was reformulated as the RMSE.

a good service delivery, and values close to 1 mean a bad service delivery. It is expected that in most cases $K = 1$, but this definition allows for more services being evaluated at the same time. Also, $N_i$ denotes the individual time horizon of each service, which means that the weighted average is done on scalar values. This means that each service can be measured at their own time scale and over their own time horizon.

The index can be similarly defined for *M* amount of asset management services:

$$\eta^{AMS} = \sum_{i=1}^{M} W_i^{AMS} \sqrt{\frac{\sum_{t=0}^{N_i} \left( QoS_{i,t}^{AMS^2} \right)}{N_i}} \tag{5.13}$$

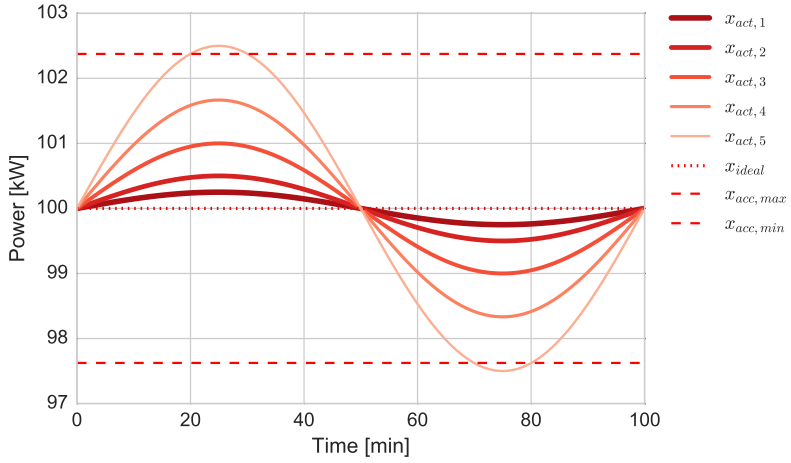
$$\sum_{i=1}^{M} W_i^{AMS} = 1 \tag{5.14}$$

It is likely that $M > 1$, e.g. if the aggregator is an EV fleet operator for a single large customer. Finally, if an aggregator desires to evaluate its own overall performance, e.g. as part of an internal reviewing process, it can combine both kinds of service provision in a weighted average:

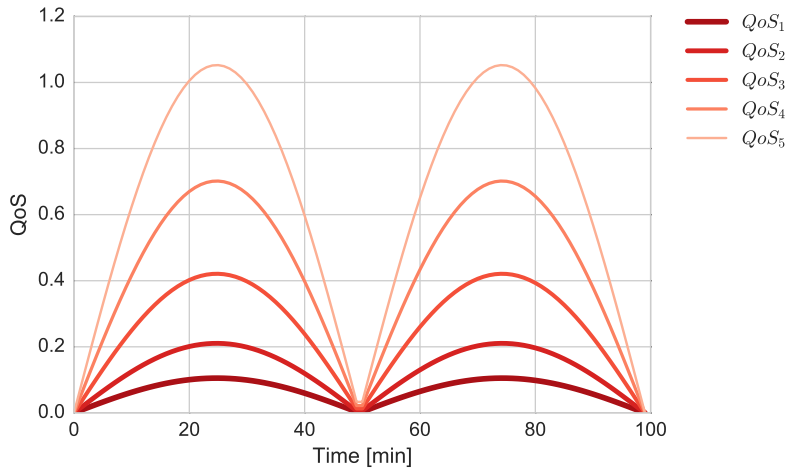
$$\eta_{tot} = \alpha \eta^{AS} + (1 - \alpha) \eta^{AMS}, \quad \alpha \in [0, 1] \tag{5.15}$$

where $\alpha$ is the weight ratio between the two kinds of service.

An example of how the service delivery could look for five different aggregators providing the same ancillary service, in this case a reference tracking service, with varying *QoS* is presented in Figure 5.3 and the corresponding performance evaluations are presented in Table 5.1. This scenario shows a wide spread of *QoS*, which is reflected in the $\eta$ values. The first aggregator has a relatively good performance and therefore has small $\eta$, while the worst performing aggregator has an $\eta$ ten times larger, i.e. worse performance.




(a) Simulated service delivery



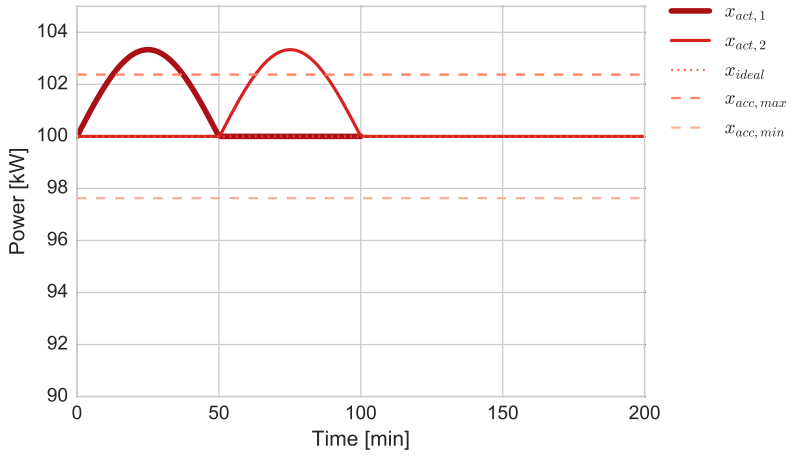

(b) QoS for the services

Figure 5.3: Test of the QoS definition where five aggregators deliver the same service for the same time horizons and with different performance.

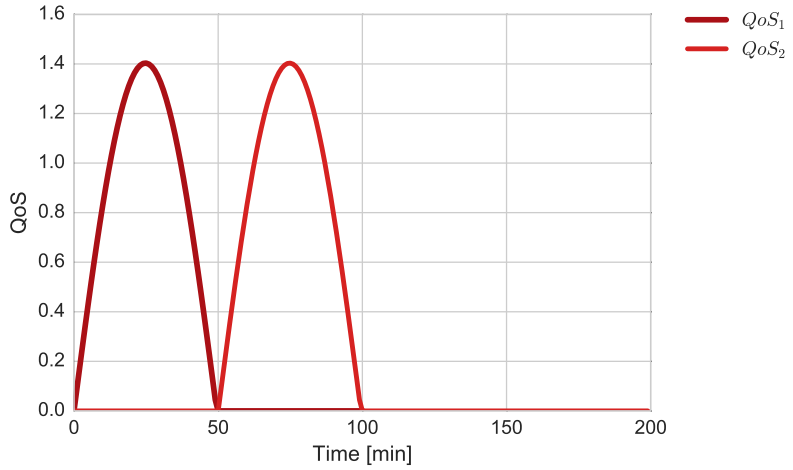
| Aggregator | $\eta$ |
|---|---|
| 1 | 0.0740 |
| 2 | 0.1481 |
| 3 | 0.2962 |
| 4 | 0.4937 |
| 5 | 0.7308 |

Table 5.1: The values of $\eta$ for same service delivery horizons and different service performance.

The definition of $\eta$ takes the time horizon of service provision into account. This is done, so that the performance assessment gives a result that is scaled to the time scale of service delivery. For example, if two aggregators perform with an error of equal magnitude, but one aggregator is contracted to deliver the service on a shorter time horizon, the assessment of this aggregator should be worse than the one which was contracted for a longer period. This is shown in Figure 5.4 and Table 5.2, where Aggregator 1 delivers the same service and the same error as Aggregator 2 but over half the delivery period.




(a) Simulated service delivery




(b) QoS for the services

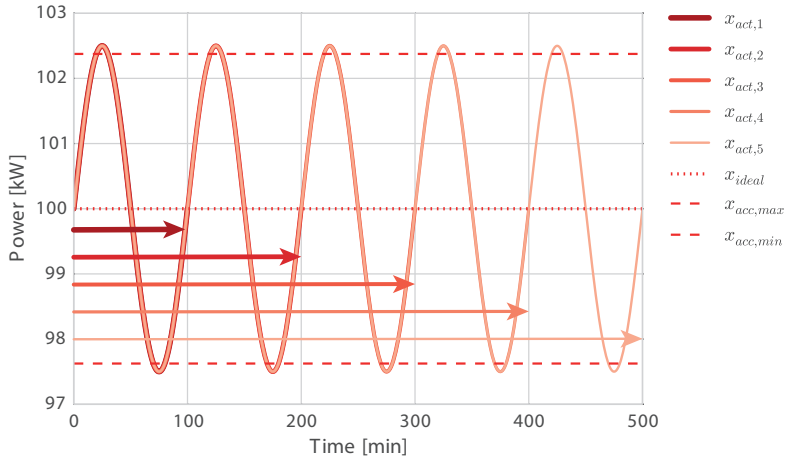
Figure 5.4: Test of the QoS definition where 2 aggregators deliver the same service for different time horizons and with the same magnitude of error.

| Aggregator | $\eta$ |
|---|---|
| 1 | 0.3491 |
| 2 | 0.2469 |

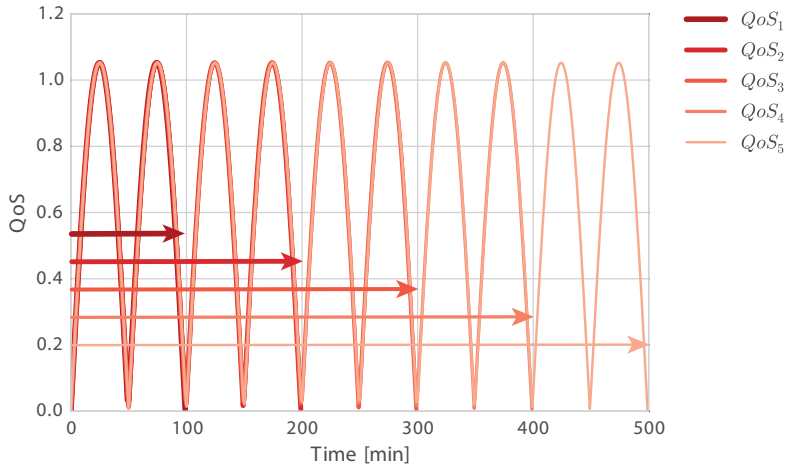
Table 5.2: The values of $\eta$ for different service delivery horizons and with the same magnitude of error. Since $\eta$ is normalized with time, the shorter service is evaluated worse than the service delivered over the long time horizon.

Conversely, if two aggregators deliver a service on different time horizons and both have an error in service delivery which is *proportionately* the same, the performance assessment will evaluate them to have equal performance. This is illustrated in Figure 5.5, where five aggregators deliver a reference tracking service over different time horizons. All aggregators have the same proportionate error[16] which leads to the same $\eta$, as can be seen in Table 5.3. There is a slight increase in the values of $\eta$ due to numerical round off, which means that the precision of $\eta$ depends on the precision of the measurements. The amount of significant digits should be determined by the system operator or whichever entity is in charge of the service performance assessment.

[16] This is ensured by using a sine as the "actual" performance of the aggregators, and each case is an extra period of the sine wave.




(a) Simulated service delivery



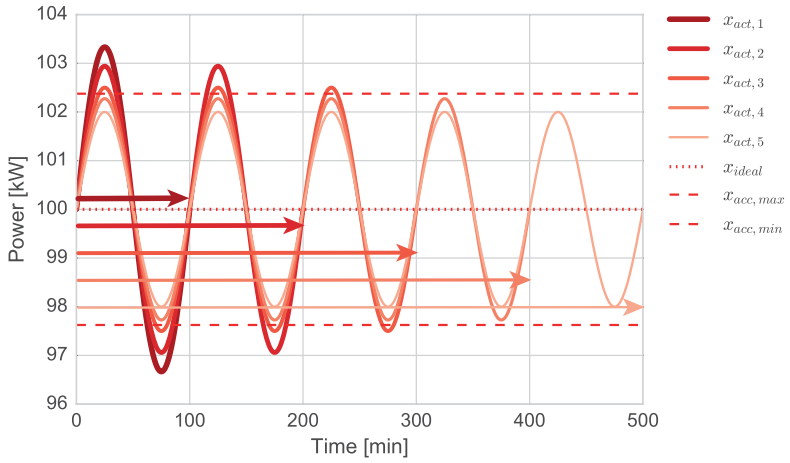

(b) QoS for the services

Figure 5.5: Test of the QoS definition where five aggregators deliver the same service for different time horizons and with the same performance.

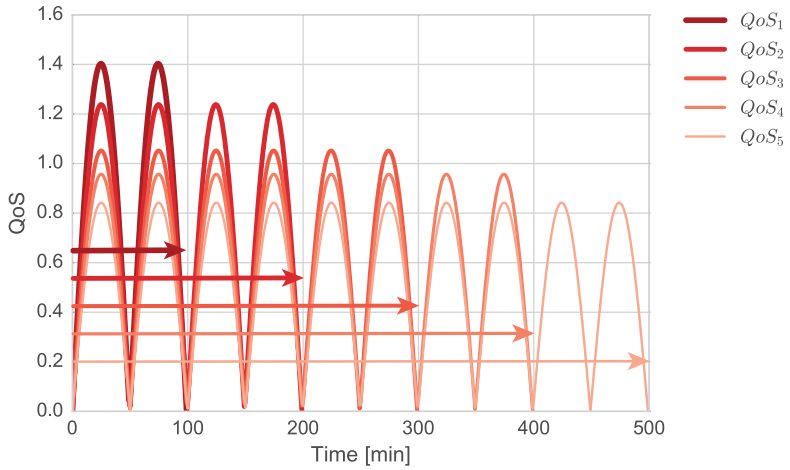
| Aggregator | $\eta$ |
|---|---|
| 1 | 0.7309 |
| 2 | 0.7327 |
| 3 | 0.7327 |
| 4 | 0.7333 |
| 5 | 0.7338 |

Table 5.3: The values of $\eta$ for different service delivery horizons. A numerical difference appears in the third decimal due to rounding error. The precision of $\eta$ depends on the measurement equipment of the flexibility asset.

Finally, Figure 5.6 and Table 5.4 show five aggregator delivering the same service for different time horizons and with different performance. It can be seen that truncating the *QoS* when calculating $\eta$ means that $\eta$ alone can not be used for assessing if a service was delivered. Thus, the service verification index, described in the next section, must be taken into account in order to give a complete idea of service performance and delivery.




(a) Simulated service delivery




(b) QoS for the services

Figure 5.6: Test of the QoS definition where five aggregators deliver the same service over different time horizons and with different performance.

| Aggregator | $\eta$ |
|---|---|
| 1 | 0.8199 |
| 2 | 0.7904 |
| 3 | 0.7333 |
| 4 | 0.6758 |
| 5 | 0.5949 |

Table 5.4: The values of $\eta$ over different service delivery horizons and different service performance.

### 5.3.2 Service Verification Index

Similar to the service performance assessment index, an index is defined for service verification[17] based upon *QoS* whenever $QoS_{meas} > 1$. A new non-delivery measure is introduced:

$$ND^{AS}(t) = \begin{cases} QoS^{AS}_{meas}(t) - 1, & \forall QoS^{AS}_{meas}(t) > 1, \forall t \\ 0, & \forall QoS^{AS}_{meas}(t) \leq 1, \forall t. \end{cases} \tag{5.16}$$

Using this new measure, the service verification index for $K$ amount of ancillary services, over a discrete time horizon of service delivery $N$, is defined as:

$$\epsilon^{AS} = \sum_{i=1}^{K} W_i^{AS} \sqrt{\frac{\sum_{t=0}^{N_i} \left( ND^{AS^2}_{i,t} \right)}{N_i}} \tag{5.17}$$

[17] This index can also be interpreted as an index measuring non-delivery.

where $\epsilon^{AS} \in [0, \infty]$ and $W_i^{AS}$ is the same as in Equation (5.12). Similar to $\eta$, $\epsilon$ is also normalized to time.

The service verification for the index can also be defined for asset management services:

$$\epsilon^{AMS} = \sum_{i=1}^{M} W_i^{AMS} \sqrt{\frac{\sum_{t=0}^{N_i} \left( ND_{i,t}^{AMS\,2} \right)}{N_i}} \tag{5.18}$$

where $M$ is the size of the unit portfolio and $W_i^{AMS}$ corresponds to Equation (5.14). Following the definition of $\eta$, the verification index scales with time. This can be seen through the verification index values corresponding to Figure 5.5 in Table 5.5. Again, due to numerical accuracy, only the second decimal number is significant.

| Aggregator | $\epsilon$ |
|---|---|
| 1 | 0.0838 |
| 2 | 0.0839 |
| 3 | 0.0840 |
| 4 | 0.0840 |
| 5 | 0.0841 |

Table 5.5: The values of $\epsilon$ over different service delivery horizons with same delivery error, as shown in Figure 5.5.

The two proposed indices have almost the same definition, the difference being that *ND* is not truncated when estimating $\epsilon$. Table 5.6 shows the $\epsilon$ values corresponding to the example in Figure 5.6, and shows clearly the definition of *ND*, i.e. values $QoS \leq 1$ are ignored, while $QoS > 1$ count as part of the non-delivery. Contrary to the service performance assessment index, the verification index does not have a clear limit for what is considered a verified serviced. For some services, e.g. ancillary services, it is critically important that $QoS(t) \ll 1$, which would mean a requirement of $\epsilon \approx 0$. In other cases, $\epsilon > 0$ is tolerable to certain extent. The tolerance limit for $\epsilon$ should be defined in the contract agreements between aggregator and the entity acquiring the services.

| Aggregator | $\epsilon$ |
|---|---|
| 1 | 0.3612 |
| 2 | 0.2512 |
| 3 | 0.0840 |
| 4 | 0.0 |
| 5 | 0.0 |

Table 5.6: The values of $\epsilon$ corresponding to Figure 5.6.

## 5.4 Application to Service Verification

The two indices defined in the previous section have different applications. As stated in the introduction to the chapter, verification occurs at the prequalification phase, at the operation phase and at the settlement phase. The work during this project has focused on the use of the indices for prequalification and settlement.

Currently, the verification of ancillary service delivery typically is based on a rigid performance assessment (pass/non-pass) of the units providing services[18]. PJM has implemented a pay-for-performance scheme by evaluating the performance of frequency regulation units, hereby changing the rigid verification procedures. Thus, PJM has established precedence in using performance metrics for verification of services, yet, as explained in Section 5.1.2, their *performance score* is tied to their regulation product. The general aggregator performance indices presented in this chapter, coupled with the service requirement models presented in Section 4.2, provide a flexible method to service verification.

[18] [34] Energinet.dk. *Ancillary services to be delivered in Denmark - Tender conditions*. 2012.

### 5.4.1 Prequalification Verification

In the case of prequalification, the indices form part of the assessment module of the validation framework presented in Figure 3.1. As such, the indices

can evaluate the simulated results and return the $\eta$ and $\epsilon$ results. Since the simulations will be based on statistical replication to gain a statistical certainty of the aggregator behavior, the verification of the simulated services will be a statistical value, e.g. with a mean and a distribution.

### 5.4.2 Settlement Verification

Verification of delivered services occurs as a post-delivery analysis. For an aggregator, the verification of a delivered service will be done by the entity who bought the service, or a third party metering company. It is still an open question how the specific DER consumption/production will be measured, since in most cases the resource will be behind a common metering point, e.g. the smart meter of a household. Also, current measurement requirements would force all resources to have expensive measurement equipment, of which the cost would far outweigh the profit of participating in the ancillary service markets. Assuming that a solution to this issue is found, e.g. through load disaggregation, virtual metering points[19] or certifying the integrated DER measurement instrumentation, the indices can be applied to the settlement verification.

[19] [133] USEF foundation. *USEF: The Framework Explained*. 2015.

Traditionally verification is done as a pass/non-pass assessment, where non delivery implies fines from the system operator. In this case, $\epsilon$ can be used as the hard constraint of service delivery, where exceeding a certain value of $\epsilon$ means a failure in deliver the service. Another option is for $\epsilon$ to be used for dimensioning the fine. Thus the final settlement from Equation (4.14) can be transformed into:

$$P_i^{\mathtt{rem}} = P^{\mathtt{clear}} \left(\eta_i \kappa_i - \epsilon_i\right) \qquad \forall\, i \in \Omega^{\mathtt{acc}}. \tag{5.19}$$

As part of an iPower demonstration event held at DTU Risø Campus in November 2014[20], a verification module was implemented using the first version of the performance assessment index[21] in the laboratory. The purpose was verification of a DSO service. In this case, the consumption of the participating DER came only from flexible heating, and the measured consumption could be used to verify the service. An integration of the verification script to the code running the demonstration also showed that the same code could be used as a simple online performance monitoring tool.

[20] [66]. *iPower FLECH Demonstration*. 2014.

[21] [14] Bondy et al. "Performance assessment of aggregation control services for demand response". 2014.

## 5.5 Conclusions on Performance Assessment

The topic of performance assessment is important for the prequalification of aggregators, the monitoring of aggregator service delivery and the settlement of services. The presented indices are a flexible tool that is useful for performance assessment in all three stages. Although the concept of performance indices are not new in the power system, the established indices evaluate the overall system performance and reliability. While these indices provide a useful tool for system operators to evaluate their performance in maintaining a secure grid, they are not suited for the evaluation of aggregators. Therefore, the definition of service performance assessment index

and the service verification index present a novel approach to the problem of evaluating the performance of aggregators. The indices are defined in terms of the weighted average of the RMSE of the quality of service. The RMSE returns a scalar value which takes into account the duration and time resolution of the service delivery. This makes it possible to compare service delivery across service definitions.

A weakness in the presented work is that for verification of services, both indices must be used. The value of $\eta$ will always fall within the acceptable range $\eta \in [0, 1]$ and it will be the value of $\epsilon$ which determines if the service is delivered. Still, $\eta$ gives an intuitive idea of how well the aggregator is performing within the limits stipulated within its service contract. $\epsilon$ does not have this same intuitive meaning, but in that sense it is not different from other fit measures used in statistics.

Future work will focus on a re-implementation of the verification module for the laboratory, incorporating the indices defined in [13]. Also, research must be done with respect to how to measure the individual DER consumption in an economically feasible way, with enough resolution to do settlement verification. Similarly, further research must be done with respect to when an aggregator is delivering different, specifically on how to distinguish which behavior corresponds to which service.

# Chapter 6
# Conclusion and Future Work

This PhD project focused on the question of how to validate consumption aggregators. Methods, concepts and procedures were developed around the creation of an aggregator validation framework. The classical procedure for generator validation consists of documentation of the generator capabilities and a limited set of validation tests. This procedure has been adapted to aggregators the following way:

*Documentation of aggregator capabilities:* A functional reference architecture for aggregators was formulated[1], which gives system operators an overview of the capabilities of the aggregator. The reference framework consists of 11 essential functions, which abstract from the specific implementation of the aggregator. The end objective is for each of these functions to have an associated key performance index, which helps the system operator asses the capabilities of the aggregator.

[1] [9] Bondy et al. "A Functional Reference Architecture for Aggregators". 2015.

*Validation tests:* A validation framework was defined in order to carry out the validation tests. Specifically, this thesis focused on three aspects of this framework:

- The definition and modeling of services[2]. These service models form the control objective of the aggregator and serve as benchmarks for the service performance evaluation and verification. This method is based upon identifying the relevant contractual parameters and defining a set of time series for ideal service delivery and acceptable service delivery. Furthermore, the concept of *Quality of Service* (QoS) was introduced as a measure of how well the aggregator provides a service. The QoS is defined as the error between the actual service delivery and the benchmark service model, scaled to the contractual limits. Thus, an acceptable service delivery has $QoS \in [0, 1]$.

[2] [13] Bondy et al. "Performance Requirements Modeling and Assessment for Ancillary Services in view of Demand Response". 2016.

- The definition of a Service Performance Index[3] ($\eta$) and a Service Verification Index[4] ($\epsilon$). These indices are metrics used for evaluating the aggregator, and are novel in that they are not made specifically for any single service, but can be used with service models in order to evaluate aggregators providing any service. The Service Performance Index is the root mean square of the QoS delivered by the aggregator. The Service Verification Index is also based upon the root mean

[3] [14] Bondy et al. "Performance assessment of aggregation control services for demand response". 2014.

[4] [13] Bondy et al. "Performance Requirements Modeling and Assessment for Ancillary Services in view of Demand Response". 2016.

square of the QoS, but it only measures how much the service delivery breaks the contractual limits, i.e. when $QoS > 1$.

- The procedure for carrying out the validation test was defined[5], based upon statistical concepts and expanding the service metrics from deterministic measures to statistical measures. Specifically, fractional factorial tests should be run with enough sampling and over adequate disturbance distributions in order to fully understand the capabilities of the aggregator that is being validated.

[5] [8] Bondy et al. "Procedure for Validation of Aggregators Providing Demand Response". 2016.

In an effort to ease the integration of aggregators into the ancillary service markets, a proposal for restructuring of ancillary service requirements was presented[6]. The new definitions consists of defining an ideal service tender, and parametrizing the bids so that each bid can fractionally fulfill the ideal service tender across one or more of the parameters. This means that units that have good capabilities in one parameter, but not in another, are still valuable for the system operator and can still participate in the market. The value of the resources is expressed by the capability value $\kappa$. An example of how this parametrization can be used in a market was presented, and it was shown how $\kappa$ and $\eta$ can be used for performance based remuneration.

[6] [10] Bondy et al. "Redefining Requirements of Ancillary Services for Technology Agnostic Sources". 2016.

The changes in the power system are leading to the decommissioning of traditional power plants, thus reducing the pool of available sources of ancillary services. At the same time, the increasing penetration of renewable intermittent generation will require units capable of providing faster balancing services. Aggregators seem to be able to solve part of this issue. Therefore, the validation of aggregators is important if they are to be used as sources for ancillary services. The research question of this thesis is in short: *how can aggregators be validated?*. The work presented here takes significant steps towards answering this question.

## 6.1 Future Work

There are mainly two subjects which have not been discussed here:

- The elaboration of operation scenario descriptions[7] that determine the situations that aggregators are expected to handle;
- Metering and measurement resolution of the DERs providing the services.

[7] See Figure 3.1.

These are important issues that must be addressed before aggregators can be fully integrated into the system.

The functional reference architecture was submitted as work-in-progress paper, and is missing one main feature. The key performance index for each function have not been assigned. This reference architecture must be completed in order for it to be useful for the system operators. As it is, it is useful as a guide for entrepreneurs seeking to open an aggregator business.

Similarly, the work on restructuring the ancillary service requirements is at a draft stage, and requires further work with respect to the study case. A project has already been set up for one of the collaborators from overseas to come to DTU and work on an experimental implementation of concept.

This project was mostly developed within the iPower project framework, which focused on flexibility services for DSOs and the Flexibility Clearing House. This has lead to an understanding of what DSOs expect from aggregators, which in turn lead to the use of DSO services as example cases in most of the papers I published. This project could have benefited of a closer collaboration with a TSO, so that the applicability of the method could be discussed with the people who might end up using it.

Finally, the concepts described here should be implemented in a software framework and an aggregator should be validated on it.

# Chapter A

# A Functional Reference Architecture for Aggregators

**Authors:**
Daniel Esteban Morales Bondy
Kai Heussen
Oliver Gehrke
Anders Thavlov
**Published at:**
Emerging Technologies and Factory Automation (ETFA), 2015 IEEE
Luxembourg, Luxembourg
**Abstract:**
Aggregators are considered to be a key enabling technology for harvesting power system services from distributed energy resources (DER). As a precondition for more widespread use of aggregators in power systems, methods for comparing and validating aggregator designs must be established. This paper proposes a functional reference architecture for aggregators to address this requirement.

## A.1 Introduction

The increase of electricity production from fluctuating renewable sources is creating a need for new ways of operating the power system. Demand response (DR), i.e. the exploitation of flexibility in electricity consumption, is considered a promising technology for mitigating this problem. However, a significant part of the DR potential exists in distributed, small and medium-sized loads. It is not practical for a power system operator to interact directly with all these flexibility assets. The role of aggregators is the creation and management of a portfolio of flexibility assets and representation of this combined flexibility to a system operator and/or market.

System operators today rely on generators for ancillary services to maintain reliable system operation. Generators undergo validation tests and continuous monitoring on the generator site. With ancillary services provided by aggregators, similar validation and performance requirements will have to be established. However, validation and monitoring requirements cannot effectively be translated from single site monitoring to distributed

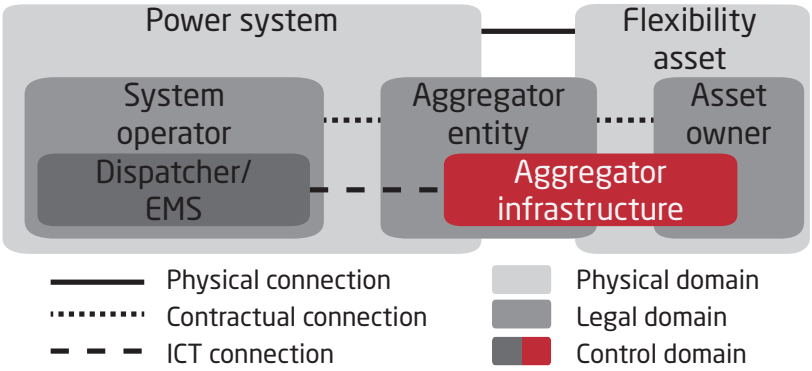



Figure A.1: The aggregator concept across domains.

aggregator control systems, and today's on-site monitoring cannot be scaled to distributed flexibility assets.

We propose a functional aggregator reference architecture that facilitates specification and validation of aggregator functional requirements and the generic modeling of contractual and verification performance requirements. Application of the proposed functional architecture to different aggregator designs suggests it as a meaningful benchmark for technology maturity.

## A.2 Aggregation in Smart Grids

We refer to the concept of aggregation as the creation and (commercial and technical) management of a portfolio of flexibility assets with the objective of offering the combined flexibility as a commercial service. The business role and technical function of performing aggregation is referred to as the Aggregator. In literature and business context use of these and related terms is not yet harmonized.

### A.2.1 Clarifying the Aggregator concept

The term *aggregation* has different relevant interpretations in business, information technology, control, as well as in the physical power system domain. Our concept of aggregators is illustrated in Fig. A.1, defining aggregators as a business role, aggregator entity, as well as a technical aggregator infrastructure.

The physical domain addresses the electrical interactions between flexibility assets (also referred to as DER) and power system. Whereas aggregation with respect to physical topology is a common concept (e.g. microgrids, cells), in our understanding, aggregators are not bound to aggregation with respect to physical network topology.

In the legal and business domain, an aggregator entity is an intermediary, maintaining contractual relations with flexibility asset owners and system operators (as receivers of flexibility services). The aggregator entity assumes legal responsibility for the delivery of a contracted service. The aggregator role may be filled by new independent market actors or be part of existing actors, such as utilities or balance responsible parties.

In the control domain, the aggregator infrastructure coordinates the behavior of flexibility assets. The control domain requirements are formulated as *flexibility services* to system operators and *asset management services* towards asset owners. Tracing these requirements for architectural validation and performance validation in the aggregator infrastructure is the focus of this paper.

The proposed aggregator concept is implementation agnostic and focused on formulation of functional requirements.

### A.2.2 The aggregator concept in technical literature

There is no unanimous definition in literature of what could be considered standard functionality of an aggregator. This is reflected by the wide variety of aggregator designs[74, 54, 116, 28], which differ in capabilities and purpose, and which use different (often implicit) criteria for classification.

Aggregators are commonly classified by control scheme into autonomous, indirect, transactional and direct control [76]. Another classification emphasizes the commercial or technical focus of aggregators, referring to *commercial* and *technical* virtual power plants (CVPP and TVPP) [46]; however, as both types require business and technical functionality, the CVPP/TVPP distinction expresses a difference in degree and is not categorical. An advanced aggregator realizing the full functionality spectrum as *Dynamic VPP* (DVPP) has been formulated in [90]. The proposed concept of aggregation encompasses all of the above but focuses on functional requirements for service provision, not business logic.

### A.2.3 Aggregator Business Harmonization and Standardization

Whereas aggregator functionality is becoming a shared concept, there are still many models describing a) which stakeholders may benefit from the flexibility service, b) the form of the flexibility service, c) which stakeholders (are allowed to) perform aggregation and who should receive compensation [78] and d) how to harmonize the interaction between aggregators and aggregated units.

With respect to a), market models are being revised and new service models introduced to assign a value to flexibility (either directly to system operators as ancillary service, or as enhancement of flexibility of existing portfolios). The form of the service, b), is often formulated as an abstract flexibility service, a trade-off between both grid needs and generalized resource characteristics. Regarding d), many aggregators use proprietary communication, loosely based on standards (e.g. IEC61850 or IEC 60870-5-104; increasingly also OPC-UA); harmonization efforts in Europe continue to be addressed in the Smart Grid Coordination Group (SGCG) under EU Mandate M/490. A successful interoperability effort in this domain is the OpenADR standard published also as IEC PAS 62746.10-1. Meanwhile the IEC TR 62357 *Reference Architecture to Smart Grid Information Exchange* is under revision. The reference architecture presented here focuses, within

the Smart Grid Architecture Model[115], on functional interoperability for aggregators (field to operation zones; DER and customer domains) supporting interactions with System Operators, market actors, and devices at process level.

## A.3 The Need for an Aggregator Reference Architecture

Existing concepts and methods for benchmarking and generator validation/certification cannot readily be translated from the (bulk) generator based paradigm to the distributed paradigm of aggregators and flexibility services. Historically, ancillary services have been defined using a physical understanding of generator capabilities. This definition is moving towards technology-agnostic service models. Service verification has been done through on-site measurements, which is infeasible with thousands of units participating in service provision.

The definition of a reference architecture for aggregators addresses these three issues, and enables benchmarking of aggregator architectures. A reference architecture "captures the essence of existing architectures, and the vision of the future needs and evolution to provide guidance to assist in developing new system architectures."[27]. It should provide:

- a common lexicon and taxonomy,
- modularization and the complementary context, and
- a common (architectural) vision.

Various types of aggregator implementation exist, realizing different design ideas for different sets of requirements. These requirements – and consequently the designs derived from them – are unlikely to converge towards a single solution because of the tradeoffs involved, e.g. scalability and complexity. A common lexicon and taxonomy is a minimal precondition for aggregator comparison. If a reference architecture is to be used to describe many of these different designs, it must be highly modular. In practice, the *general* functionality of an aggregator must be broken down into small enough functions in order for these functions to be usable as building blocks for the reconstruction of the *particular* functionality of a given implementation. The functions are arranged in a reference architecture such that metrics can be assigned to individual functions. In this way, the reference architecture can be used for validation of the aggregator. Our architectural vision accounts for the need for verifying distributed flexibility services.

## A.4 Functional decomposition

An aggregator is a complex system of interacting functions. In the following definitions, we abstract from implementation details, e.g. centralized vs. distributed systems, and focus purely on the purpose of the functions.

**A. Service Interface** The service interface translates the contractual agreements between the aggregator and its clients into a service model

containing quantifiable and measurable service requirements and a set of performance criteria. This service model is then used to map incoming service requests to control domain signals such as control variables, constraints or control parameters.

**B. Performance Monitoring** The performance monitoring function collects data from which the behaviour of individual clients can be derived. The data is analyzed to determine the performance of a client, and its compliance with the contracted flexibility service. This analysis may be internal to the performance monitoring function, or it may simply serve as a data gatherer for an external entity.

**C. Supervision and Resource Handling** The aggregator must maintain an overview of available client resources and their status. By comparing the communication status and monitored performance of individual clients to the control signals sent by the aggregator, the supervision function determines whether clients perform according to their contract. It may temporarily or permanently exclude non-compliant clients from the pool of available resources.

**D. Operator Interface** Although the power system is moving towards automated solutions, decision-making on critical issues is the responsibility of human operators. The aggregator architecture must support decision-making by presenting operators with the necessary information, and facilitating operator input and intervention.

**E. Control** The control function is in charge of generating the appropriate control domain signals for the portfolio. Depending on the control architecture, the control logic may be distributed between physical entities. The concept of a control domain signal covers several kinds of signals, including, but not limited to control inputs to DERs, coordination messages for distributed control and reference signals for hierarchical controllers.

**F. Flexibility Monitoring** In operation, the aggregator must assess the future flexibility of its portfolio in real time; this includes individual DER flexibility as well as the aggregated flexibility of the portfolio. The flexibility assessment can either be based on direct feedback from the DERs or entirely on estimation models (possibly stochastic) if direct feedback is not available.

**G. Aggregator-internal communication** Except for very few special cases, aggregation will almost always be implemented as a distributed computing system. In its basic form, such a system would consist of one aggregator and a number of clients. This may be extended by stacking multiple levels of aggregation etc. The internal communication function exchanges information between aggregator and clients.

**H. Client management** The client management function actively or passively tracks the availability of clients. It may also provide a mechanism for the dynamic addition and removal of clients, such as a discovery service, and maintain a protocol for temporarily disabling otherwise available resources. It contributes to resilience and graceful degradation of the portfolio.

**I. External Information Services** To be able to act optimally with respect to both control of its portfolio and trade of electricity in forward markets, aggregators will likely have to rely on different types of information services.

Such services include different types of forecasts and measurements in real time and may be provided by either internal processes or by a 3rd party.

**J. Asset interface** Most aggregators in a Smart Grid context will be used to harvest flexibility from existing energy resources. In most cases communication between aggregator and resource will use a fieldbus-style interface not designed for wide-area communication. The purpose of the asset interface is to maintain communication with a physical unit under aggregator control and provide abstraction from interface details.

**K. Information Exchange** Virtually any modular software framework contains a facility for information exchange between its components and storage of the overall system state: static data, dynamic data or both. A knowledge exchange may take many different forms, from a collection of object references towards a central or distributed database.

## A.5 The Reference Architecture

We have now established a set of functions to serve as building blocks for a reference architecture, but without concern for the relations between these blocks. Next, these relations will be examined; in other words: how could a practical aggregator infrastructure be composed from these function blocks?

### A.5.1 Function blocks and knowledge exchange diagram

The functions in section A.4 generally belong to one of the following categories:

- functions dedicated to communication between physically separate parts of the aggregator infrastructure or communication with 3rd party entities, i.e. enablers of the distributed nature of the system. functions which perform decisions with regards to flexibility asset behavior and portfolio composition.
- functions which interpret information and support the decision making functions.

These categories represent requirements for different architectural paradigms: Communication functions are layered or hierarchical, and, in the case of communication between aggregator and client, require an identically layered counterpart at the opposite end. The decision making and interpreter functions on the other hand require many hierarchical and non-hierarchical consumer-producer relations. Figure A.2 shows an overview of the relationships between functions according to the above concept.

### A.5.2 Principles of distribution of functions

While figure A.2 depicts the relationship between aggregator functions, it does not include information about the physical distribution of these

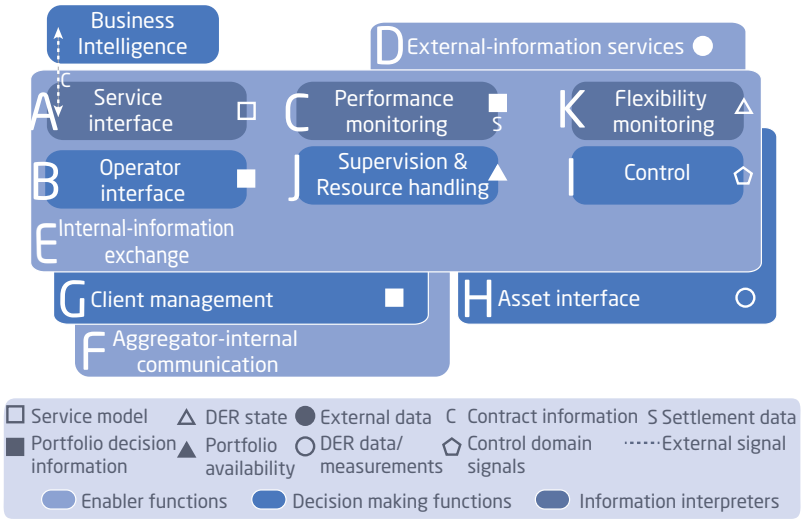



Figure A.2: Each function outputs distinct kinds of data which are used by the other functions in different ways according to the aggregator implementation.

functions between the asset side and the operator side of the aggregator infrastructure. This distribution is highly specific to the individual design and e.g. its degree of centralization (see section A.6).

In an actual implementation, several of these functions require corresponding instances on each side, effectively forming a communication stack.

The functions exhibiting these properties are:

- the internal aggregator communication function which provides the link between the two substacks. In many cases, this function will make use of a full OSI-layered stack in which the internal aggregator communication function provides the application layer,
- the client management function, implementing management protocols which would typically require a corresponding instance on the client side, and
- the knowledge exchange function which exchanges information with its client side counterpart independent of client management mechanisms.

All other functions, with the exception of the asset interface, may appear either on the operator side, on the asset side, or shared between both sides, depending on the implementation.

## A.6 Case studies

A number of existing aggregator designs – commercial as well as academic – have been mapped to the model in order to test its viability. Two cases with different design philosophies are presented here in order to illustrate the distribution of functionality between the operator and the asset side, and the information flow between the functions:

- *Power Hub* is an aggregator developed by Dong Energy in Denmark. It is used to control distributed generation and load in order to sell flexibility services to the ancillary power market (Figure A.3).

- *Open Energi* is a British company selling flexible consumption from industrial loads as an ancillary service. The aggregator functions are distributed between an operator node at a control center and asset nodes on custom hardware deployed at customer sites (Figure A.4).

The most significant difference between the two designs is the degree of autonomy of the asset node. The Power Hub concept is based on a centralized design which mainly uses the asset node as a communication gateway and places flexibility monitoring and control at the central operator site. This is also where external information such as market data is available through the information services function. The Open Energi controller acts on quantities measurable at the asset site and does not require external information; this allows control and flexibility monitoring to be placed at the asset node, leaving only supervisory functions at the operator site.

Both designs can be split into functions according to the subdivision proposed in section A.4.

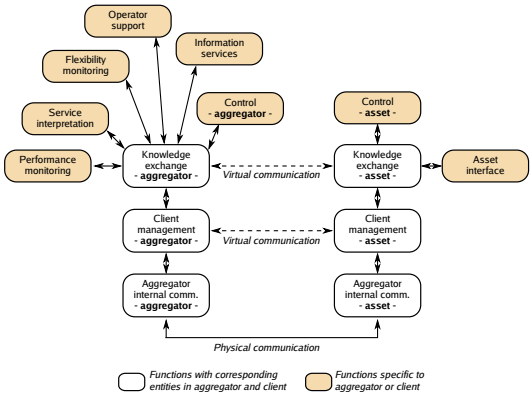



Figure A.3: Distribution of functions for the Power Hub aggregator

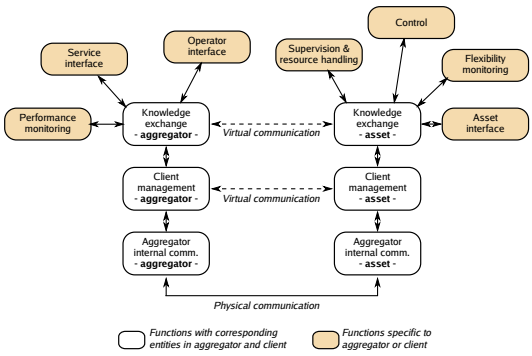



Figure A.4: Distribution of functions for the Open Energi aggregator

## A.7 Conclusion and further work

A reference architecture for the validation and comparison of aggregators has been presented. While the general framework has been established and successful mapping tests to a number of real-world aggregator designs have been performed, many details are still work in progress. The next steps towards completion will be the development of performance indicators for the individual functions and the establishment of a process for aggregator comparison and performance validation.

## Acknowledgment

Parts of this work are supported by the Programme for Energy Technology Development and Demonstration (EUDP) and Innovation Fund Denmark through the iPower project.

# Chapter B

# Procedure for Validation of Aggregators Providing Demand Response

**Authors:**
Daniel Esteban Morales Bondy
Oliver Gehrke
Anders Thavlov
Kai Heussen
Anna M. Kosek
Henrik W. Bindner
**To be presented at:**
Power Systems Computational Conference, 2016
Genoa, Italy
**Abstract:**
As aggregators become viable sources of ancillary services, they will be required to undergo a validation process similar to the prequalification process of traditional generators. Since aggregators are fundamentally different from traditional generators, a new test method must be designed for the aggregator validation. This work proposes a method for designing the tests necessary for the validation process. The method is exemplified with a study case and results are presented.

## B.1 Introduction

As renewable energy generation increasingly replaces conventional power plants, power system operators are looking for alternative sources for the ancillary services which were traditionally provided by these plants. It is expected that some ancillary services can be procured from distributed consumption and production units by making use of their unused operational flexibility. In order to provide coordinated services, such as demand response [84], from a large number of such distributed energy resources (DERs), a new actor is appearing in the power system: the aggregator [48].

Conventional sources for ancillary services must be certified before being able to offer control reserves in an ancillary services market [33]. It is

expected that aggregators will be required to undergo a similar prequalification process to ensure the appropriate performance of the provided service with respect to predefined requirements.

The achievable performance of aggregators, in terms of service provision, depends on its architecture [9], i.e. where the decision making is located [76] and its level of automation, the choice of hardware for implementation, the specification of communication protocols [71], how advanced the portfolio management is, etc. This means that some aggregator architectures will be better suited for a specific ancillary service than others [14].

It has been established that current requirements for participation in the ancillary services markets limit the participation of aggregators of demand response [21, 114]. Different research projects have looked into this problem, see e.g. [12].

Until now, the performance evaluation and testing of aggregators in academia has been ad-hoc to specific aggregator implementations [135, 109]. Aggregator test frameworks have been proposed [19], field tests have been carried out in order to validate DR schemes [70], and concepts regarding systematized testing have been exemplified [118].

This work addresses the gap between all these concepts, i.e. we present a procedure for the design of validation tests, which takes a systematical approach to aggregator testing, discussing the issue from input, i.e. service requirements, to output, i.e. performance metrics.

## B.2 Conventional resource validation and aggregator differences

Ancillary services are essential for the reliability of the power system. Because these services play such an important role in the safe operation of the system, it is essential that the units or entities providing a service perform according to the requirements set by the Transmission System Operator (TSO). These requirements and processes are typically specific to a particular TSO and influenced by national regulations, interconnection grid codes etc. Two examples are regulations established by Energinet.dk and PJM.

### B.2.1 Current requirements for prequalification

In Denmark, Energinet.dk, the danish TSO, ensures the appropriate service performance by requiring all units participating in the ancillary service markets to provide a documentation of their capabilities and go through an approval process [33]. This approval process consists of a test conducted at least three weeks prior to the service delivery date. The tests for Frequency Containment Reserve (FCR) generally involves the injection of a setpoint step into the plant's governor and the measurement of the response. The test for Automatic Frequency Restoration Reserve (A-FRR) involves the tracking of a reference signal from Energinet.dk. Currently, this procedures are not formally described. Demand resources are expected to provide a substantial amount of the ancillary services for the Danish grid in the future. Since

distribution system services are not widespread yet, the concept of unit certification is non-existent at the distribution level.

While Denmark is starting to open up to new sources of ancillary services and standardise its test procedures, PJM (a regional transmission operator in the United States) has a standardised prequalification procedure for regulating resources[1], which consists of three consecutive area regulation tests, where PJM Performance Compliance scores indicate how well the resource follows a simulated regulation signal. A single test lasts for 40 minutes and in order to pass it the unit must score at least 75% in three consecutive tests [104]. While this rule includes services provided by multiple generators at a single site, operators of demand resources are not required to be certified but must complete an initial training module on the requirements and business rules of the Regulation and Synchronized Reserve markets [123]. Currently demand resources are only allowed to form 25% of the total regulation [47] in PJM, and therefore their certification process is still not a large concern.

[1] Regulation in the US corresponds roughly to the FRR of ENTSO-E.

In both systems the validation tests have two goals: to ensure the communication with the units works correctly, and to validate the known performance model of the generators. Thus, a change of configuration in the setup requires a new certification of the generator. Also, a dedicated communication and measurement infrastructure between system operator and aggregator is required. The measurements must have high sampling frequency, e.g. better than 10 mHz, and high precision, e.g. sensitive to frequency deviations of $\pm$ 10 mHz. Measurement equipment that respect these requirements is expensive.

### B.2.2 Problems applying current validation methods to aggregators

The tests outlined above are specific to each system operator, but follow similar paradigms. The conventional test processes cannot be directly applied to portfolios of aggregated resources, mainly because a common assumption in the process is that the service delivery is performed by a single or small number of units. This allows inference of the unit's ramp capabilities through a response test, based on a known model. Also, a limited amount of precise and expensive measurement equipment needs to be installed.

An aggregator and the portfolio of units under its control behave fundamentally different from large generation units:

1. Individual generator units are well understood and models describing their static and dynamic properties are readily available. This is not the case for portfolios of aggregated units which are typically heterogeneous and can only be modelled through their statistical properties. This is aggravated by the fact that unit portfolios may be dynamically reconfigured during operation.

2. There is no direct equivalent to a single point of measurement: An aggregator's portfolio may consist of geographically dispersed units. Their aggregate power profile does not correspond to a measurement at any

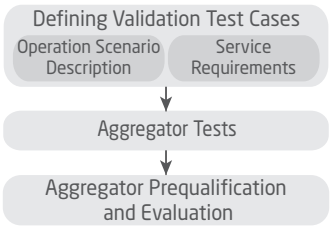



Figure B.1: Schematic procedure for aggregator validation. The focus of this paper is on the relationship between the service requirements and the aggregator tests.

single point of the grid. Coupled to this, it is economically infeasible to install the required expensive measurement equipment at each DER.

3. Aggregators, by definition, operate a distributed system (both in control and geographical terms) in which each unit has its own response properties and requirements. This leads to an aggregated response that behaves differently from that of conventional generators.

4. Reliability concepts for distributed systems are different; specifically, the failure modes are not the same. If a component of a monolithic generator unit fails, the whole unit may have to shut down. The failure of a single unit in an aggregator portfolio will often have a minor or negligible impact on the overall performance. In a large portfolio it will usually be possible to recruit an equivalent replacement unit providing the same services as the failed one.

For the above reasons, the same validation and service requirements cannot be applied to aggregators. This paper focuses on reinterpreting the validation tests to aggregators by adapting concepts from statistical testing to the power systems domain.

## B.3 Requirements and proposed test procedure

The objective of the current tests is to validate the parameters of a well understood model of generation units. For the reasons stated previously, the new tests need to identify an empirical behavior model of an uncertain and diverse entity: the aggregator control architecture and unit portfolio.

A test procedure is required that will allow the system operators to understand and predict the performance of a specific aggregator under a given set of operation conditions (Fig. B.1). In this section we present the underlying assumptions for such a procedure, the metrics used to measure the aggregator performance, and the proposed procedure.

### B.3.1 Test Procedure Assumptions

The critical assumption are:

1) A general test design must start from the assumption that the aggregator and its infrastructure are to be treated as a black box, in the sense that only the aggregator inputs and outputs are known but the details of the internal control architecture are unknown.

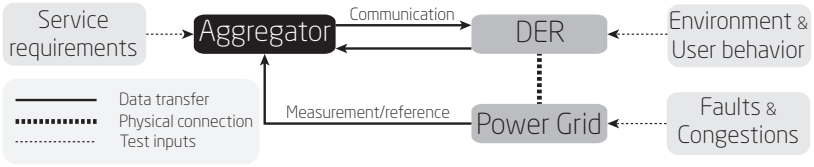



Figure B.2: Schematic test setup where the test subject, the aggregator, is treated as a black box.

2) In order to fully test aggregators under a number of relevant scenarios, and to capture the stochastic nature of their operation, aggregators must be tested with the aid of a simulation framework.

3) It is assumed that such a simulation framework is detailed enough in terms of power system models, DER models and information and communication technology (ICT) systems in order for the simulation results to reflect the real performance of a deployed aggregator with sufficient precision.

The assumptions that can be adjusted are:

4) The tests are defined by a set of operational scenarios and service requirements (Fig. B.1): a) the operational scenarios define the statistical distribution of the test disturbances; b) the service requirements define the expected behavior of the aggregator.

5) The mode of interaction between the test cases and the aggregator is defined in a test setup (Fig. B.2), where the disturbances (test inputs) defined in the operational scenarios affect the aggregator interaction with the DERs and the power grid.

6) The aggregator has two interfaces: *inputs* in the form of measurements and service reference signals received from the system operators; *outputs* in the form of control domain signals exchanged between the aggregator and the units in its portfolio.

7) The operational scenarios are not designed to cover aggregator operation under exceptional system conditions. This means that the aggregator will not be held accountable for non-delivery in cases where the cause is outside of the aggregator's influence, e.g. in the case of grid faults. If communication between the aggregator and the controlled units, or internally within the aggregator architecture, occurs over public telecommunication networks, the robustness to network outages must be tested, for example by simulating disturbances and delays.

8) The flexibility which the aggregator can offer is bounded by the contractual requirements between the aggregator and its clients.

9) Aggregator validation will be carried out by a third party test entity.

### B.3.2 Service Requirements - Test Metrics

In order to measure how the disturbances affect service delivery, a set of service performance metrics must be established. The main purpose of the current tests is to verify communication, responsiveness to frequency changes and tracking of a reference or AGC[2] signal. Coupling this with the performance requirements defined by the TSOs (e.g. [33, 92]), the expected

[2] Automatic Generation Control, see e.g.[39].

Table B.1: System services and their behavior

| System Operator | Service name | Service behavior |
|---|---|---|
| TSO | Frequency containment reserve (FCR) | autonomous response to frequency deviations ($\star$) |
| TSO | Frequency restoration reserve (FRR) | tracking of the AGC-signal |
| DSO | Congestion management | reference tracking |
| | | respecting a maximum feeder/transformer limit ($\star$) |
| | | demand response |
| | | grid state responsiveness ($\star$) |

behavior of the considered services was analyzed (Table B.1), and a set of requirement metrics were defined:

- Time responsiveness, i.e. how fast can the service be delivered from the moment the reference or measurement signal changes.
- Grid responsiveness, i.e. how well can the aggregator follow changes in the grid state (marked with $\star$ where relevant on Tab. B.1).
- Response accuracy, i.e. how good is the aggregator in providing the full volume that is requested.

These three metrics will constitute the measure with which an aggregator will be deemed to perform according to service requirements, and the tests must excite the aggregator such that it is possible to determine through the value of these metrics the performance of the aggregator. It must be pointed out that the grid responsiveness metric is only applicable to the evaluation of aggregators providing services that rely on direct measurement of the grid, e.g. FCR.

When system operators define the acceptable values of the service requirement metrics, the values should have a statistical component. An example could be that the time responsiveness of a service provision should in average of 5 seconds, with a variance of $\pm$ 1 second. The actual indices used for the proposed metrics are discussed in Sec.B.3.4.

### B.3.3 Aggregator Validation Procedure

The procedure for aggregator validation applies statistical principles for the evaluation of the aggregator performance. It consists of the following steps:

1. The general composition of the aggregator is established through documentation, and the service the aggregator is to be validated for is selected.
2. The testing entity identifies the appropriate service requirements for the selected service.

3. The testing entity identifies the expected normal operation of the aggregator based upon the service definition.
4. The testing entity defines the operation scenarios that the aggregator is expected to perform under.
5. The tests are carried out on the aggregator.
6. The aggregator performance is evaluated.

From the services analysed in this work, the tests are divided into two categories depending on their excitation signal:

- step response (like those for FCR),
- continuous reference tracking (like those for FRR).

The validation tests will use one of these excitation signals under a different set of circumstances defined in the operation scenarios. Sufficient sampling of the aggregator response to the excitation signal is important in order to ensure that the mean and variance of the performance metrics give a realistic impression of the aggregator performance under deployment.

### B.3.4 Evaluation of Test Results

The service requirement metrics (Sec. B.3.2) define the measure upon which the aggregator is evaluated. Different options exist that can be used to measure these metrics. One option is the aggregator performance index [14], which measures the error in service delivery for the services delivered to the system operators and the serviced delivered to the owners of the DERs. This metric captures both time responsiveness and response accuracy into a single value. A large set of performance indices exist within the field of control performance assessment, these can be utilised for the proposed validation method, see e.g. [67].

Given the stochastic nature of the tests, the indices will also be stochastic. The value of the performance indices is estimated at each iteration of the test, which means that the final value of the performance index reflects the stochasticity of the disturbances. For example, if the disturbances defined in the operation scenarios are Gaussian, the performance index will also have a mean and variance. These values need to be compared to those values defined in the service requirements. It will be the choice of the system operators what the service requirements should be, taking into consideration their risk adversity. Requiring a small variance on the performance indices minimizes the risk of not getting a full service delivery, but might also lead to more expensive services.

## B.4 Case Study of the Validation Procedure

In this section we apply the concepts outlined in Sec. B.3 on a simplified example of the validation procedure on an aggregator architecture similar to the one presented in [125]. While the design of these operating scenarios

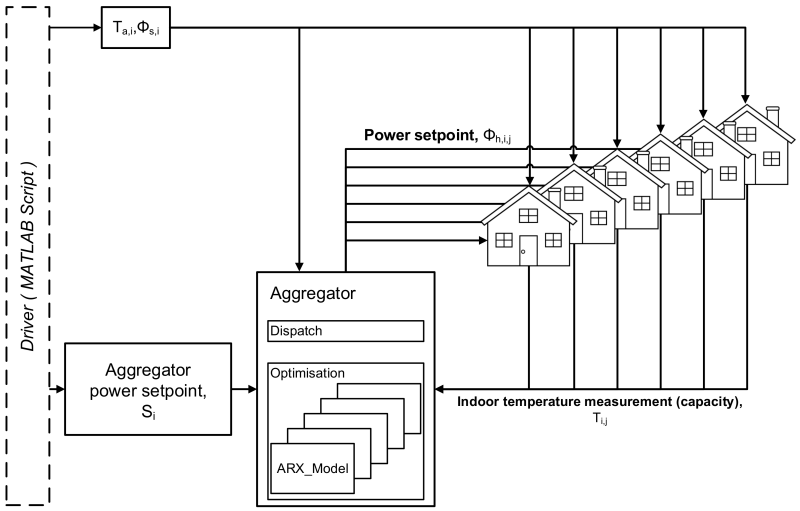



Figure B.3: Flow diagram of the aggregator algorithm.

is outside the scope of this paper, some overall assumptions have been made. The sample aggregator name is DTU-FlexServices, and it wants to sell ancillary services to the TSO called RisøGrid. The validation tests are carried out by the independent company AggTesters. The rest of this section presents the reference scenario, the example of the aggregator test and the evaluation process.

### B.4.1 Aggregator Framework & Portfolio

The objective of the DTU-FlexServices is to allocate a given amount of power, provided as a setpoint by the TSO, over a controllable portfolio of 100 resistive heating systems, each providing space heating to a detached household. The objective is subject to constraints on nominal power of the heating systems and indoor comfort, which is implemented as a tolerable band in which the temperature is allowed to vary given by the interval $[T_{min}, T_{max}]$. It is assumed that feedback on measured indoor temperature is available to the aggregator, such that the aggregator in real-time can assess the available capacity of the controlled heating system and ensure that indoor temperature constraints are not being violated during operation. Fig. B.3 presents the flow of data in the aggregator simulation framework. The aggregator uses a simple auto-regressive model with exogenous inputs (ARX) to assess the future available capacity of each individual households. The ARX model is given by,

$$T_{i+1} - a \cdot T_i = b \cdot T_{a,i} + c \cdot \Phi_{s,i} + \boldsymbol{d}^T \boldsymbol{\Phi}_{h,i:i-\tau_{lag}} \tag{B.1}$$

where $T_i$ is the measured indoor temperature of the household at time step $i$, $T_a$ is the outdoor temperature, $\Phi_s$ is the solar irradiance and $\boldsymbol{\Phi}_{h,i:i-\tau_{lag}}$ is a vector with the most recent observed power consumptions, i.e. $\left[\Phi_{h,i}, \Phi_{h,i-1} \cdots \Phi_{h,i-\tau_{lag}}\right]$. The lag parameter of the heat input, $\tau_{lag} \in \mathbb{N}_0$, is used to account for the potential time-lag that might exists between when heating is applied and

when it is observed in the indoor temperature. $a$, $b$, $c \in \mathbb{R}$, and $\boldsymbol{d} \in \mathbb{R}^{\tau_{lag}+1}$ are the unknown parameters of the ARX model, which are found using prior data for power consumption of the heating system. For simplicity $\tau \equiv 0$ is assumed in the following.

Each individual resistive heating system is assumed to be able to dispatch a continuous amount of power in the interval $[P_{min}, P_{max}]$, given by the nominal power of the heating system. Naturally, this is an approximation since resistive heating systems, in general, will only be able to dispatch power in discrete steps due the composition of resistive loads. However, considering a portfolio of many entities and following the law of large numbers, these discrete steps should level out and the assumption hold.

To allocate the amount of power over the portfolio of resistive heating system, following unit commitment problem is formulated,

$$\min \left| \sum_{j=1}^{N} \left(\Phi_{h,i,j}\right) - S_i \right| + \sum_{j=1}^{N} \Phi_{h,i,j} \mathrm{W}\left(T_{i+1,j}\right) \tag{B.2}$$

$$\text{s.t.} \quad P_{min,j} \leq \Phi_{h,i,j} \leq P_{max,j}$$

where the decision variable $\Phi_{h,i,j} \in \mathbb{R}$ is the amount of power being allocated to household $j$ at time step $i$, $N$ is the number of households in the portfolio, $S_i$ is the setpoint given to the aggregator and $\mathrm{W}\left(T_{i+1,j}\right)$ is a weight function of the predicted indoor temperature found from Equation (B.1). The weight function should be constructed such that $\mathrm{W}\left(\cdot\right) < -1$ for $T_{i+1,j} < T_{min}$, thus making the last term dominate the cost function and force the allocated power up for household $j$. Likewise, $\mathrm{W}\left(\cdot\right) > 1$ for $T_{i+1,j} > T_{max}$, thus forcing the power down. Following linear weight-function is proposed,

$$\mathrm{W}\left(T_{i+1,j}\right) = \frac{2\left(T_{i+1,j} - T_{min,j}\right)}{T_{max,j} - T_{min,j}} - 1 \tag{B.3}$$

The simulation model of the individual households is implemented as a stochastic linear state space model in discrete time, which is given by

$$\begin{aligned} \boldsymbol{T}_{i+1} &= \boldsymbol{A}\boldsymbol{T}_i + \boldsymbol{B}\boldsymbol{U} + \boldsymbol{\sigma}_i \\ T_i &= \boldsymbol{C}\boldsymbol{T} + e_i \end{aligned} \tag{B.4}$$

where $T_i \in \mathbb{R}$ is the locally measured indoor temperature which is assumed to be forwarded to the aggregator, $\boldsymbol{T}_i \in \mathbb{R}^n$ is the state vector and $\boldsymbol{U} \in \mathbb{R}^m$ is the input vector. $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times m}$ and $\boldsymbol{C} \in \mathbb{R}^{1\times m}$ are the system, input and output matrix, respectively. To account for unrecognized input and approximations, process noise, $\boldsymbol{\sigma}_i \in \mathbb{R}^n$, is added to the system equation, (B.4). In the following, $\boldsymbol{\sigma}$ is assumed to be a Gaussian white noise process. Furthermore, $n \equiv 1$ is assumed, i.e. only one temperature state is being simulated in the households; hence, since a Gaussian white noise process is fully characterized by its variance, the process noise is fully described by the variance $\sigma \in \mathbb{R}$.

The aggregator framework and simulation models, simulating the considered scenario, have been implemented in MATLAB and is presented in full detail in [125]. It is important to note that the aggregator is described in

this section for the purpose of the paper, but this description is contained within the conceptual black box described in Sec. B.3.1, and the testing entity only has access to the general composition of the aggregator portfolio.

### B.4.2 Service Requirements, Normal Operation and Operation Scenario

The DTU-FlexServices aggregator wants to participate in the ancillary service markets with a FRR up-regulation service with a volume of 250 kW. Since it is the first time DTU-FlexServices participates in the market for this service, RisøGrid requires DTU-FlexServices to go through the validation process. Following the steps outlined in Sec. B.3.3, the validation process consists of the following steps:

1. DTU-FlexServices presents the documentation for its portfolio.

2. RisøGrid sets the test service requirements as:

   - Response accuracy: $E[RMS] \leq 60\,kW$
   - The response durations: $\tau = 1\,h$

3. AggTesters identifies the normal operation scenario as:

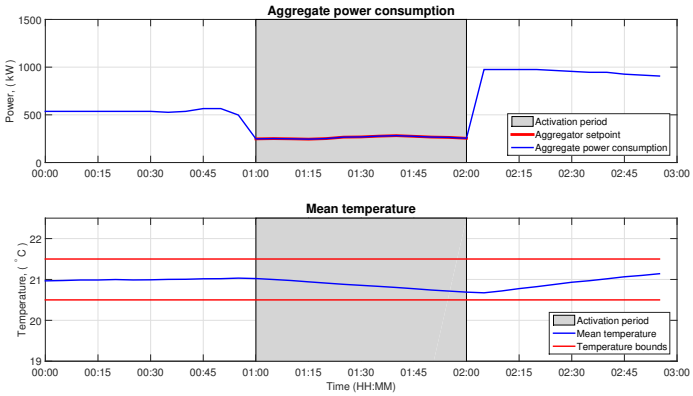
   - One source of uncertainty is the availability of the portfolio, which is a uniform distribution between 70% and 100%. This also accounts for minor changes in the portfolio size.
   - A second source of uncertainty is in the disturbances induced by unrecognized user behavior and inaccurate weather forecast in the house simulation model. This uncertainty is described by $\sigma$ in Eq. (B.4).

### B.4.3 Aggregator test

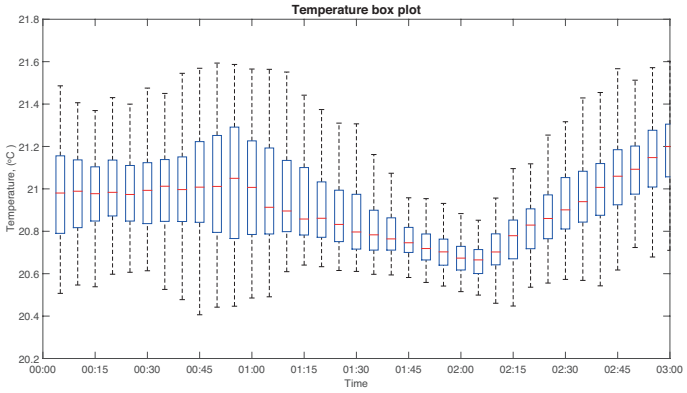
To test for different combinations of the two sources of uncertainties, a series of simulations are carried out with permutations of the two. Assuming the availability to be uniformly distributed, the tests are carried out in discrete steeps across the 70% – 100% spectrum of availability. Likewise, the variance of the noise process is tested in discrete steps in the 0.00 – 0.30 domain. Fig. B.4 and Fig. B.5 present the outcome of two different simulations for 100% and 70% availability, respectively, and $\sigma = 0.10$. Each permutation of the two noise sources is simulated 100 times.

The response accuracy of DTU-FlexServices and the average temperature of its portfolio can be seen in Fig. B.4a and Fig. B.5a. The distribution of the house temperatures can be seen in Fig. B.4b and Fig. B.5b, and it is clear that as the availability of the houses decreases, the flexibility for up-regulation is being saturated faster and the DTU-FlexServices is unable to track the FRR reference signal.

Having carried out the necessary test, RisøGrid proceeds to evaluate the results of the tests.




(a) Response accuracy




(b) House Temperatures

Figure B.4: Simulation results of the 100% availability test for the whole portfolio.

### B.4.4 Evaluation of test results

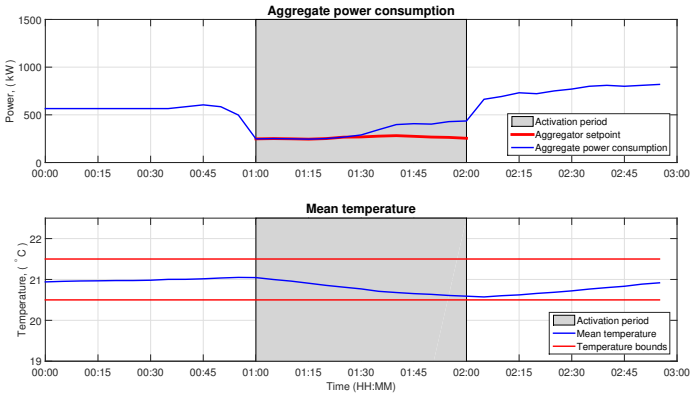
Since the case study looks at simplified setup, and the example does not take the time responsiveness metric into account, it does not make sense to use the aggregator performance metric mentioned in Sec. B.3.4. In Sec B.4.2, the root mean square (RMS) error is chosen to measure the response accuracy metric:

$$\eta_{RMS} = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\sum_{j=1}^{N}(\Phi_{h,i,j}) - S_i\right)^2} \tag{B.5}$$

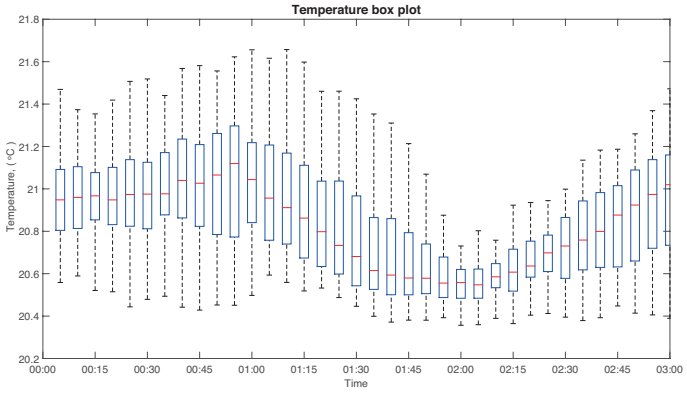
where $[1, M]$ are the iterations where the aggregator has been activated. The results of the test are presented in Table B.2, where it can be seen that $E[\eta_{RMS}] < 60\,kW$. Therefore the DTU-FlexServices is certified to provide FRR up-regulation service to RisøGrid.

## B.5 Discussion

Specific terminology has been introduced to describe the proposed method. This terminology can be mapped to that of the field of *Design of Experiments*, e.g. *definition of service requirements* maps to *definition of inner-noise factor* and *definition of test inputs* maps to *definition of outer-noise factors*.




(a) Response accuracy




(b) House Temperatures

Figure B.5: Simulation results of the 70% availability test for the whole portfolio

| | | Process noise, $\sigma$ | | | | | | | Avg. |
|---|---|---|---|---|---|---|---|---|---|
| | | 0.00 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | |
| Availability | 100% | 0.00 | 0.00 | 0.04 | 9.25 | 31.20 | 48.84 | 98.32 | 26.81 |
| | 95% | 0.03 | 0.00 | 1.51 | 19.19 | 37.97 | 66.56 | 102.05 | 32.47 |
| | 90% | 1.40 | 0.04 | 14.36 | 30.78 | 58.74 | 73.38 | 98.24 | 39.56 |
| | 85% | 1.10 | 35.23 | 4.06 | 45.28 | 83.06 | 83.40 | 115.11 | 52.46 |
| | 80% | 13.88 | 29.25 | 12.94 | 65.93 | 72.31 | 94.50 | 135.85 | 60.67 |
| | 75% | 54.28 | 40.74 | 39.91 | 75.22 | 86.14 | 114.13 | 135.76 | 78.03 |
| | 70% | 45.63 | 90.90 | 85.41 | 99.02 | 93.68 | 123.82 | 142.64 | 97.30 |
| Avg. | | 16.62 | 28.02 | 22.60 | 49.24 | 66.16 | 86.38 | 118.28 | 55.33 |

Table B.2: Performance of DTU-FlexServices

Specifically, the method resembles *fractional factorial methods for off-line quality control*, see e.g.[97]. In the case study presented in Sec. B.4, the inner factor, or controllable variable, is kept at a single level, i.e. the same activation signal is sent to the aggregator for each run of the experiment. The two outer factors, or noise variables, were varied over a distribution dictated by the operational scenario, i.e. the availability of the portfolio was varied on seven levels and, likewise, the process noise in the house simulation models was varied on seven levels. An important contribution of this work is applying this kind of formal test procedures to the problem of aggregator validation. The field of Design of Experiments is broad, and a further revision on the topic may yield a better method proposals than the one proposed here.

In this paper we focus only on the two uncertainty sources mentioned above, therefore the test for time responsiveness, i.e. delay in the communications systems between the aggregator and a DER is not considered. This means that the test design presented in the case study is a simplified version of what an actual aggregator validation test would require. Future research must identify the relevant variables that need to be tested under the relevant operation scenarios.

In comparison with the traditional test method, this validation procedure must capture the capabilities of a much more complex system, and therefore relies in part on simulations. As presented in [118], the error between the used models and reality must be quantified and taken into account for the final aggregator certification. Each block in the simulation must use validated models or software. This applies to the communication systems, the grid models and the DER models. The test architecture, e.g. the one presented in [19], which validates the aggregators must also be validated.

There are still several open issues that need to be investigated with regards to aggregator validation. For example, the definition of the operation scenarios was only briefly discussed, and heuristics must be developed in order to define scenarios that are effective when testing aggregators.

Aggregator validation must be an ongoing process, that should be carried out periodically or whenever the aggregator portfolio or architecture changes significantly. Furthermore, aggregators are expected to participate in different electricity markets. Due to these reasons, along with the complexity of designing appropriate simulations, we believe that the task of validating aggregators should not carried out by the system operators, but by an independent third party.

## B.6 Conclusion

This work presents an initial approach to establishing a method for designing aggregator validation tests. This method differs from the traditional generator certification tests in that it relies on a statistical approach. Specifically, it reinterprets the generator certification tests to aggregators by adapting concepts from statistical testing to the problem. The validation test must be carried out with the aid of simulations, so that the stochasticity of the real world disturbances affecting the aggregator can be taken into account.

While several of the concepts that form the proposed validation procedure, e.g. software framework for aggregator tests and aggregator performance assessment, have been addressed before, this work describes how these concepts can be unified in order to do a systematic testing of aggregators.

The validation procedure was shown through a simplified case study on an existing aggregator design. While the example shows a fictive setup, it appropriately represents the procedure.

An important step for the development of the validation method is the implementation of a complete test architecture with validated component models. With such a simulation framework, with realistic communication

and DER models, communication delays can be implemented in order to test aggregators for time responsiveness.

We consider the work presented here an important element of enabling aggregators in the smart grid, thus enabling consumption to actively participate in the secure operation of the power system. This will help the integration of renewable energy sources into the power system.

# Chapter C

# Redefining Requirements of Ancillary Services for Technology Agnostic Sources

**Authors:**
Daniel Esteban Morales Bondy
Jason S. MacDonald
Emre C. Kara
Oliver Gehrke
Kai Heussen
Sıla Kılıççote
Henrik W. Bindner
**Published at:**
Draft
**Abstract:**
Main points of the paper:

- DR and other technologies have unused potential that TSOs are not utilizing
- This sub-optimality is due to current requirements oriented towards the lowest common denominator
- Two options for optimal utilization of new resources: split market or unify under new conditions
- We choose to unify by parametrizing the service definitions/requirements

## C.1 Introduction

The requirements for ancillary services (AS) in many countries are defined, due to historical reasons, on the assumption that only generators provide ancillary services. With the increase in adoption of distributed energy resources (DERs) and controllable smart loads, as well as the emergence of schemes for utilizing consumption flexibility, such as demand response (DR), new sources for ancillary services from the demand-side are available.

These sources posses qualities that in many cases match the performance needs of the system better than traditional generators, yet their participation in the ancillary service markets is restricted due to requirements barriers. Since there are both economic and technical benefits in exploiting these qualities, a method must be designed so that system operators can readily utilize the positive qualities of both traditional and new ancillary services sources. In this paper we propose new frequency ancillary service requirements, focused on service performance, which are source/technology independent.

By changing the AS requirements to focus on performance rather than unit capabilities and utilizing new technologies as ancillary service providers, system operators will be able to maintain better system reliability[94], and increase participation in the ancillary service markets. Furthermore, service verification and settlement will benefit those players that are able to provide better quality services.

The rest of the paper is organized as follows: section C.2 presents the current ancillary service definitions and requirements; section C.4 presents the new ancillary service requirements; section C.3 presents the performance properties of different new technologies that make them suitable for frequency ancillary service provision. Section C.5 presents a case study of the impact of the new requirements, and section C.7 presents conclusions and thoughts for future research.

## C.2 Ancillary Services: Current Definitions and Requirements

### C.2.1 Power system operation and requirements

In the electricity system, supply and demand must be kept balanced at all times and two metrics are commonly used to evaluate the current imbalance in a power system: system frequency and the area control error (ACE). System frequency is a measure of the speed at which all interconnected, synchronous generators are rotating. ACE is a measure of the deviation in scheduled power exchanges between interconnected electricity systems.

System imbalances have two causes:

1. Expected imbalances due to deviations between the planned generation and the actual electricity demand.

2. Unexpected imbalances due to system contingency.

System operators procure ancillary services[1] in order to deal with these imbalances in their daily operation. The structure of the ancillary services varies between systems but can generally be divided into primary, secondary and tertiary control [2][110]. This work will focus specifically on the primary and secondary frequency control.

An illustration of the power system frequency during a contingency is shown in Fig. C.1. When a system contingency occurs, such as the loss of a large generator or a transmission line, there is a sudden loss in generation that is made up for by the small amount of inherent storage in the rotating inertia of the remaining synchronous generators. This sudden loss

[1] Ancillary services are also used to solve other operating issues in the power system, such as voltage problems, but this work will focus specifically on frequency services.

[2] Recently, ENTSO-E has changed its terminology for AS providing Load-Frequency Control to Frequency Containment Reserves, Frequency Restoration Reserves (either automatic or manual), and Replacement Reserves[44]. This classification matches roughly into the framework presented in [110].

reduces the rotational speed of said generators which results in a frequency excursion whose slope is determined by the total inertia of the system. The inertia of the system is determined by the amount of kinetic energy in the synchronous generators in the system, and as these generators are decommissioned, the inertia in the system will decrease, thus increasing the volatility of the system.

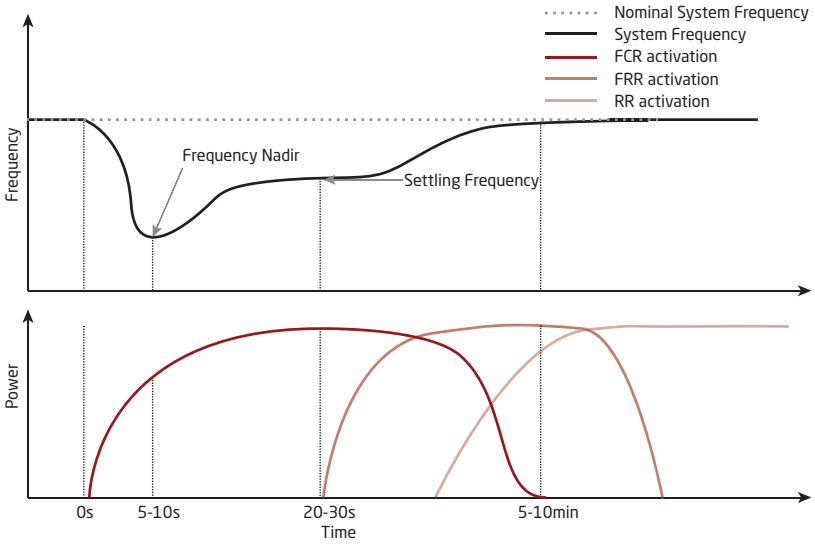



Figure C.1: The nadir of a system frequency excursion during contingency events is more pronounced if there is less inertia in the system.

## C.2.2 Ancillary services

The objective of ancillary services can generally be defined as: maintaining an adequate and secure power system. This means maintaining the power system operating at nominal frequency and voltage. In cases where the power system deviates from nominal operation, either due to natural fluctuations in consumption or faults in the system, the system operators will activate ancillary services to restore normal operation.

Power produced by renewable energy sources (RESs) has a low marginal price, which pushes the overall electricity prices down in markets with high RESs penetration. This means that the operation of traditional generators is becoming economically challenging, and will lead to the decommissioning of fossil-fueled generators. In other instances the regulatory framework prioritizes renewable production, which also leads to the decline of fossil-fueled generators[42]. Both cases lead to a system with less inertia and leaves the system operators with fewer sources for ancillary services. New sources of ancillary services will become important, and it is in the interest of the system operators that the remaining sources for ancillary services, both from production and consumption sides, are exploited optimally.

### Primary Frequency Control

Primary frequency control is the fastest response (in the seconds range) and is used to arrest and begin reversing frequency excursions occurring due to sudden imbalances between supply and demand, often caused by contingency events. Primary frequency control is traditionally performed by generators under "droop" control, in which a change in power output is made proportional to locally measured frequency. The reason for this is that the response to the frequency excursion must occur as fast as possible and be proportional to the size of the excursion so that the system frequency stabilizes within an acceptable time frame.

### Secondary Frequency Control

Secondary frequency control is a slower response (in the seconds to minute range) that takes over for the primary frequency control and returns the system to nominal frequency by controlling the output of participating resources. Much more urgently, it restores the full bipolar range of the primary reserve, so the system gets back to nominal $n-1$ redundancy. Usually an entity estimates the control reference signals based upon the ACE and system frequency to simultaneously resolve the imbalance at the interconnection and maintain stable operation (see, e.g. [89, 38], for more details). The control algorithm that directly controls the output of resources providing this service is often called Automatic Generation Control (AGC)[3]. The generators will have either a proportional controller or a proportional-integral (PI) controller to track the AGC signal. The secondary response also needs to occur as fast as possible, yet due to its centralised control approach, it is not able to provide as fast a response as primary control. The underlying need for the secondary service is to supply a fast reference tracking response without overshoot.

[3] This balancing control logic can have a centralized, pluralistic or hierarchical architecture [38] to determine the individual reference signals for the generators.

### Service Requirements

Because AS are essential for the secure operation of the system, the system operators also have requirements and restrictions on the units providing AS. A super-set of requirements across different systems is defined in [110]. These requirements can roughly be classified into three categories: *temporal requirements*, which relate to how fast and for how long a service must be delivered; *resource tuning requirements*, which relate to specific values that tuning parameters in the resource must have; and *market requirements*, which relate to bid sizes and similar parameters in systems where services are acquired through market mechanisms. Of these three categories, only the temporal requirements relate to service performance. Furthermore, in most systems are the requirements implicitly defined for traditional generation units. This means that most service requirements are oriented towards the least common denominator of service providers, e.g. a unit providing primary frequency control should provide half of the service

within 15 seconds and full response within 30 seconds[34]. A variety of generation and consumption units would be able to provide this service faster, but this quality is not rewarded. Another example is the requirement of having a PI-controller on units providing secondary frequency control, in order to track the AGC signal. Such a controller is infeasible on distributed systems, but other modern controllers can provide offset-free control with similar properties.

### C.2.3 Problem statement

Until now, system operators have been able to arrest frequency excursions fast enough because of the inherent system inertia, but as the inertia decreases, faster response times are required of the primary frequency control.

The system operator must have enough primary reserves to arrest the frequency as fast as possible, before the system enters a state where a blackout is inevitable. A metric for how effective the procurement of reserve is the *frequency nadir* [41], and it is desirable that the value is as close as possible to the nominal frequency of the system.

Similarly, the system operator should ensure that the secondary reserves act as fast as possible to relieve the primary reserves and also bring the frequency from the settling frequency back to the nominal frequency.

In [135] it is shown that if primary frequency response is provided by demand response (with a very fast response), the frequency nadir occurs at higher frequencies. Also, in [85], the authors argue that the value of regulation resources can be defined based upon the ramp capabilities of the service providing units. Faster reacting units are more valuable to system operators, since they help arrest the frequency excursion faster and at a higher nadir. It does require changes to the AGC in order to utilize the fast response, but this would also lead to the need for fewer reserves.

In short, the historical definitions for service requirements results in the suboptimal use of today's AS resources. Due to the legacy definitions there is an implicit bias for traditional resources, and alternative technologies, such as demand response, are are restricted in their contribution to AS provision, and their favorable properties are not utilized or undervalued.

While system operators have been able to maintain a secure system using traditional resources, the changes in the power system, i.e. the decrease of system inertia and increased fluctuation due to RES, require units that react faster than the current minimum requirements. Also, an increased overall volume of balancing resources will be required due to the larger deviations caused by the RES. Units that provide a faster response but are not able to provide the full response duration should be enabled to contribute to AS provision and be valued accordingly.

If these technologies, both the underutilized and the ones restricted from providing services, are used optimally for ancillary service delivery, it follows from the conclusions presented in [85, 135] that frequency excursions could be arrested at higher frequency nadir, thus lessening the required amount of reserves, which leads to a lower-cost operation of the system. Regulative authorities have concluded that fast reacting units are valuable

for the system operation, and started programs to benefit of these resources. An example of this is FERC order 755 (Pay for Performance) which has led to PJM splitting their regulation market product into RegA, for slow reacting units, and RegD for fast reacting units. The product differentiation approach has been a success for PJM, but splitting the market into different products does not address two points: 1) the overall pool of resources will not be optimally utilized, and 2) as other new technologies appear in the system, the market might fragment further, also leading to non-optimal utilization of resources. We propose instead to restructure the ancillary service definitions such that all types of service providers participate with the same market product defined by a set of optimal performance parameters, and not by minimum requirements. This means that the all entities providing a given ancillary service, e.g. primary frequency control, are optimally cleared under a single market. The service restructuring is detailed in the following section.

## C.3 Unconventional Resources

### C.3.1 Related Work

There is growing evidence that demand-side resources (DSRs) can participate in ancillary services, thus substituting the need for traditional ancillary service resources. However, the DSRs that can provide ancillary services vary greatly in composition, and have distinct properties. Most of the research on DSRs to provide transmission-level services has focused on specific services using a particular set of loads connected to the grid. This is partly due to varying characteristics of DSRs and partly to the suitable control architecture for the proposed services.

A common set of resources studied in connection to DR are thermostatically controlled loads [88, 68, 127, 87], such as electric space heating [87, 127], residential and industrial refrigeration [80], and space heating using heat pumps [52]. Thermostatically controlled loads are valued for their ability to provide ancillary services because of the inherent thermal inertia present in the systems. The thermal inertia acts as energy storage, permitting the curtailment or deferral of power consumption. The application of TCLs as a DR mechanism can also be seen in industrial settings such as large refrigeration systems [108], the heating of bitumen tanks [25], and indoor climate control using HVAC [6].

Batteries can also provide ancillary services through demand response. Electric vehicles (EVs) can be considered as mobile batteries with additional time varying constraints. By changing their charge patterns while guaranteeing the mobility needs of the owner, EVs can offer the demand-side flexibility needed to provide ancillary services to the grid [137, 69].

The potential of using the dimming of lighting in office buildings for DR is presented in [111]. A pilot project in Denmark also used the lighting system in an industrial green house for DR, showing the potential of using DR to manage congestion in the distribution system.

Water pumps–used in wastewater treatment systems and agriculture–are also considered a promising resource for ancillary services. Specifically, in [53], the authors suggest that water can be temporarily stored in pipes and tanks, hence delaying the transportation for treatment in waste water treatment systems. The load flexibility of agricultural water pumps stems from the inherent flexibility in the time of irrigation.

In this paper, we examine the use of DR when system reliability is jeopardized. A great deal of research has focused on DSR-specific controller design and limitations due to load characteristics and comfort needs. Specifically, many aggregation frameworks exist in the literature that overcome cycling constraints and response frequency limitations. Hence, instead of focusing on the design of such DSR-specific controllers, we assume that AS provided by DSRs will be sold to system operators by an aggregator, and that the aggregator is responsible for control accuracy. Our objective in this paper is to discuss and formulate ideal performance requirements for ancillary services in a number of relevant features, and to provide a market clearing mechanism that selects a portfolio of resources in a resource-agnostic and performance-oriented way. By doing so, we propose a strategy in which (i) we remove the barriers preventing unconventional resources from participating in AS markets due to the static nature of AS market definitions and requirements, and (ii) we provide a fair and performance-based market clearing structure in which the unused potential of DSRs can be easily incorporated.

### C.3.2 Properties

The identified DSR parameters are given as follows:

*Response time* This is the time it,takes for a unit to receive a DR signal and react upon it.

*Response duration* How long is a pool of these units able to sustain service provision: short, medium or long.

*Response magnitude* This the *amount* of load used by the DR resources that can be increased or decreased. The increase capability is defined as the *take* magnitude and the decrease capability is defined as the *shed* magnitude.

### C.3.3 Unused Potential of DSRs and Barriers to DSR Participation

Although DSRs provide additional freedom to help shape response compared to traditional AS providers, the existing ancillary service market rules and requirements are a strong barrier to DSR market participation. A recent study identifies such barriers in the US[21]. The rules and requirements that limit resource participation in different markets are not consistent among different RTOs and ISOs; however, the authors identify three major groups of these rules: rules on the size of the resource, rules on the measurement

and telemetry of the resource, and rules on market bidding time. Out of six different ISOs and RTOs in the US, only one allows load aggregations to provide regulation services, and only two allow aggregation participation as a spinning reserve provider. Furthermore, only two ISOs and RTOs allow aggregate telemetry. Providing telemetry at an individual resource level increases the overall cost of metering, making it challenging for DSRs to provide cost-competitive AS. Finally, most of the AS markets procure in day-ahead markets, and day-ahead DSR participation is harder due to increasing uncertainty in DSR flexibility forecasts.

In order to accommodate the slow-ramping resources as well as DSRs in the AS markets, there is an increasing need to either split AS into different service classes or parametrize the service definition so that the resources are selected only by their ability to satisfy the system needs. The ideal resource to satisfy the system need is one with "unlimited capabilities in terms of response time, energy output, ability to frequently reverse their output, ability to respond and follow the AGC setpoint changes, and size."[85][4] To include and incentivize the participation of technologies that in some parameters are closer to the ideal than those defined by the current service and market requirements, two methods can be utilized: product differentiation and product restructuring.

[4] For this kind of response to be optimal, changes must be made to the AGC algorithm [102].

Some transmission system operators, like PJM, have already suggested that better service performance is more valuable than simply adhering to traditional rules and requirements, and split their regulation market into a slow service product and a fast service product. This work explores the alternative: restructuring the market so that all technologies can participate in the same market, and the system operator can optimize the use of the resources based upon their capabilities. This entails reformulating the temporal and market requirements, and removing the requirements that implicitly assume that the services are provided by traditional generators, thus making the requirements technology-agnostic.

## C.4 Restructuring the Ancillary Service Requirements

### C.4.1 Overall approach

The proposed restructuring assumes that system operators acquire AS reserves through a market, and that potential AS providers bid their reserve capacity in that market. The restructuring is based on the following four key concepts (which are expanded upon throughout this section):

- The formulation of an *ideal ancillary service response* that the system operator desires for the system. This formulation will be strongly dependent on the needs of the system operator, e.g. very fast response in case of low system inertia, and will be submitted as a tender to the market.
- The *parametrization of the AS bids*, where the parameters reflect the service providers' capabilities to partially fulfill the ideal service response. This removes the minimum-requirements-barriers on new technologies, thus enabling any useful unit to participate in the AS provision, which facilitates market liquidity.

- Clearing all units under a *generalized single clearing-price auction*, provides incentives to bid actual marginal cost. In this auction, the capability value of each service provider and their historical performance is taken into account.
- *Performance-based remuneration* gives incentive to better AS provision and enables transparent performance-based clearing of the market.

Based on an assessment of the complete decision process, we merge the four key concepts outlined before into a novel approach to an ancillary service definition that accounts both for performance of resources and the actual spectrum of system needs. The holistic assessment includes:

- *Planning*: Assessment of system need, parametrization of resource performance and specification of tender conditions.
- *Scheduling*: Quantification of AS tender volume, AS bid submission, and market clearing.
- *Operation*: Reserves dispatch/activation and monitoring.
- *Settlement*: Verification of service delivery and remuneration.

As outlined above, for effective inclusion of DR (or any other unconventional resource) in AS markets, a revision of each phase is required. Our proposal focuses on a new *parametrization of services* (Sec. C.4.3), which affects in particular *market clearing* (Sec. C.4.4) and *remuneration* (Sec. C.4.5).

In Section C.5 we illustrate the impact of this reformulation in comparison with present market mechanisms, and in C.6, the alignment with present mechanisms and its applicability to novel ancillary service models (REF WARRINGTON/policy based) is reflected.

### C.4.2 Ideal service tender

The ideal source for AS is one with "unlimited capabilities in terms of response time, energy output, ability to frequently reverse their output, ability to respond and follow the AGC setpoint changes, and size ."[85][5] It is impossible for any one unit to possess these characteristics, but system operators aim at achieving this kind of system response by contracting several units.

[5] For this kind of response to be optimal, changes must be made to the AGC algorithm [102].

In existing AS, there is an implicit assumption that ideal unit response corresponds to a scalar fraction of the required system response. In contrast, in presence of a diverse resource portfolio, the commonly expected fast response is secondary to an overall cheaper mixed portfolio which delivers a better system response, e.g. by combination of a fast duration-limited and slower unlimited response time resources.

For example, a system operator could determine that the ideal system response to a frequency excursion is the one that has a resulting frequency nadir at the settling frequency (thus minimizing the risk of tripping the under-frequency relays). Based upon the inertia in its system, the system

operator determines the volume ($V_{tot}$) needed as well as the response characteristics needed to achieve this, see Figure C.2.

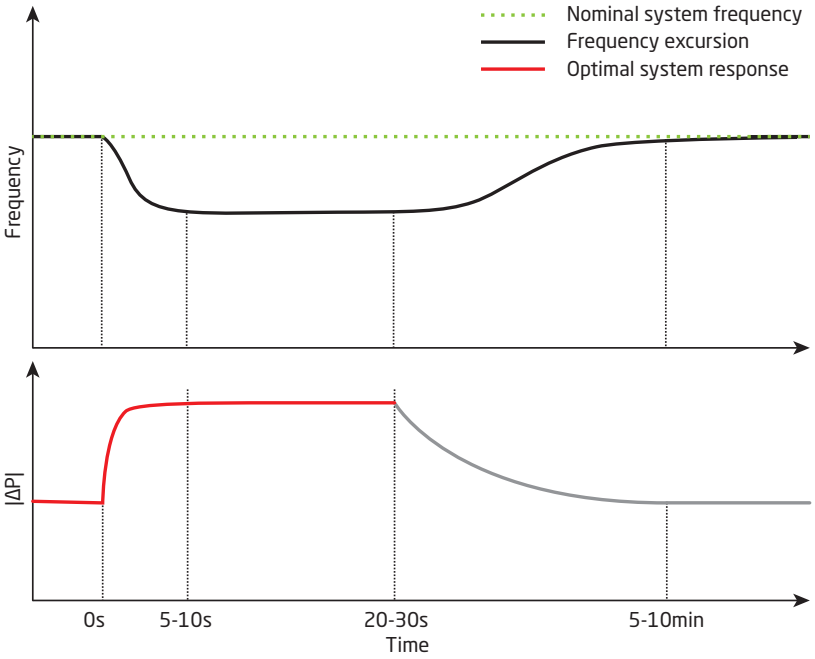



Figure C.2: In this case, the ramp of the ideal response is mainly determined by the system inertia and is to be sustained until secondary frequency control can be activated.

## C.4.3 Parametrization of service performance

Ancillary service requirements are specified by a system operator based on the desired control response for a particular power system. Today, these requirements — as reflected in the service definition — are not differentiated according to the capabilities of the unit providing the service. Therefore, service definitions are designed to accommodate the least capable unit in the portfolio. As a consequence, more capable units are not being fully utilized, leading to excess contracting of service providers. This suboptimal allocation of resources could be addressed by introducing a performance dependent definition of ancillary services, i.e. a service definition which allows compliance to be measured on a linear rather than a binary scale: In addition to compliance and noncompliance, different levels of partial compliance are possible. In this context, services will be defined such that the best possible performance of the most capable unit corresponds to full compliance.

One of the challenges with such an approach is to achieve a useful definition of partial compliance. Depending on the complexity of the service, many parameters of DR resources may have to be included in a performance comparison to determine their relative value. For example, resources with identical response magnitudes, ramp rates and endurances may represent a significantly different value to the buyer of a cyclic service if one resource requires a high recovery time between cycles. A performance model is therefore needed to provide a mapping between the multidimen-

sional parameter space of a DR resource and its capability to comply with a given service, expressed on a linear scale.

We introduce the following definition of a capability value:

$$\kappa = g(\mathbf{x}) \tag{C.1}$$

$$\kappa \in [0, 1] \tag{C.2}$$

where $\mathbf{x}$ is a vector of the resource parameters relevant to a particular service, and $g(\cdot)$ is a function mapping the parameter space to a scalar value according to resource utility. This mapping function is highly specific to a particular service and must therefore be developed by the service requester, e.g. a TSO. The function is then communicated to the resources as part of the service definition included in a tender. $\kappa$ for a particular resource can then be calculated by its operator prior to bidding.

### C.4.4 Market mechanism

In order to leverage the proposed AS restructuring, the market clearing mechanism needs to be changed. The clearing should take the *capability value* of the service providers into account, and ideally also the probability of availability (certainty in service). There are many different ways of formulating such a market clearing mechanism and here we present an example of a market that utilizes the service parametrization to form an ideal service response.

The market is designed as a single clearing price auction, in which each resource bid is adjusted by *two* factors for bid quality: 1) a shape-matching parameter $\kappa_i$ and 2) a historic performance parameter $\eta_i^{hist}$. The clearing mechanism identifies a common clearing price based on the most expensive accepted bid.

$$P^{\mathtt{clear}} = \max P_i^{\mathtt{bid}}, \quad i \in \Omega^{\mathtt{acc}} \tag{C.3}$$

where $\Omega^{acc} \subseteq \Omega$ is the subset of accepted bids of the set of received bids $\Omega$. The clearing mechanism selects the subset of bids which offer the cheapest overall clearing cost and meet the tender requirements:

$$\Omega^{\mathtt{acc}} = \quad \mathtt{argmin}_{\Omega^{\mathtt{hyp}} \in \mathcal{P}(\Omega)} \sum_{i \in \Omega^{\mathtt{hyp}}} \kappa_i P_{\Omega^{\mathtt{hyp}}}^{\mathtt{clear}} \tag{C.4}$$

$$\mathtt{s.t.} \quad \sum_{i \in \Omega^{\mathtt{hyp}}} V_i \geq V_{tot} \tag{C.5}$$

$$\eta_i^{hist} \geq \eta_{min}^{hist} \quad \forall i \in \Omega^{\mathtt{hyp}} \tag{C.6}$$

$$\sum_{i \in \Omega^{\mathtt{hyp}}} \eta_i V_i / V_{tot} \geq \eta^{AS} \tag{C.7}$$

Where $\mathcal{P}(\Omega)$ denotes the Power Set of $\Omega$.

The specification of tender and bid parametrization needs to be aligned with the mechanisms applied during real-time operation the resource dispatch and activation. Resource performance is monitored with respect to the behaviour expected from bid parametrization, and is further expanded upon in the next subsection.

### C.4.5 Performance-based remuneration

Performance-based remuneration has already been introduced in United States through the FERC order 755. Similarly, in this work we propose that service providers are paid according to how close they follow the capability parameters they bid to the market. The estimation of the service provision performance can be done in different ways, depending on which parameters the system operator deems to be the most critical. A service performance index is proposed in [13], where service performance is defined as the root mean square error of the actual service delivery compared to the ideal model:

$$\eta^{post} = \sqrt{\frac{\sum_{t=0}^{N} \left(QoS_t{}^2\right)}{N}}, \tag{C.8}$$

$$\eta^{post} \in [0, 1], \tag{C.9}$$

where $N$ is the time horizon over which the service is delivered and $QoS \in [0, 1]$ is the *Quality of Service* of the ancillary service, which is the error in service delivery scaled to the tolerance limits defined by the system operator. This leads to the final settlement price of service provision being defined as:

$$P_i^{\mathtt{rem}} = \eta_i^{post} \kappa_i P^{\mathtt{clear}} \qquad \forall i \in \Omega^{\mathtt{acc}}. \tag{C.10}$$

## C.5 Case Study

To be finished

## C.6 Discussion

To be finished

## C.7 Conclusion

To be finished

## Acknowledgments

We thank the following people for interview/feedback:

- Scott Baker (PJM)
- Joe Eto (LBNL)
- Preben Nyeng, Peter Bruhn (Energinet.dk)

Parts of this work are supported by the Programme for Energy Technology Development and Demonstration (EUDP) and Innovation Fund Denmark through the iPower project.

# Chapter D

# Performance Assessment of Aggregation Control Services for Demand Response

**Authors:**
Daniel Esteban Morales Bondy
Giuseppe Tommaso Costanzo
Kai Heussen
Henrik W. Bindner
**Published at:**
Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2014 IEEE PES
Istanbul, Turkey
**Abstract:**
Aggregation algorithms that provide services to the grid via demand side management are moving from research ideas to the market. With the diversity of the technology delivering such services, it becomes essential to establish transparent performance standards from a service delivery perspective. This paper formulates performance measures and an index to evaluate in hindsight the quality of service delivery by an aggregator, both with respect to ancillary service and asset management service.

The index is based on requirements formulated in service contracts and provides an overall assessment of the quality of service provided by an aggregation control algorithm. By a detailed case study we present and an application of the index, comparing the performance of two different control architectures for demand side management delivering a distribution grid service.

## D.1 Introduction

The future increase in energy production from Renewable Energy Sources (RES) may lead to a power system where production is distributed, and where the Transmission System Operators (TSOs) require a larger amount of balancing services. At the same time, the increase in Distributed Energy Resources (DERs) brings new challenges to the Distribution System Op-

erators (DSOs), which may need new kinds of ancillary services[92]. It is anticipated that DER owners will be able to provide services to the system operators via Demand Side Management (DSM).

An Aggregator is a market player, or market role, whose business case is to manage DER units in its portfolio and use their inherent consumption flexibility to participate in the ancillary service markets, i.e. it controls units in order to perform DSM. A general classification of different aggregation methods is presented in [76], an example of direct control can be found in [5], and an analysis and evaluation of indirect control architectures can be found in [63].

Since the Aggregator has contractual obligations with customers and system operators, it is important that the control algorithm the Aggregator uses proves suitable for the task. From a service perspective, an aggregation algorithm is considered suitable if the performance, i.e. the quality of service (QoS), it delivers is within the contractual limits. The Aggregator must therefore control its DER portfolio in such a way that it fulfills the needs of both the DER owners and the System Operator.

Little attention has been given to the problem of performance assessment of aggregator controllers seen from a service-delivery perspective. This paper approaches the problem by presenting two main ideas:

- both ancillary services and DSM have minimum QoS requirements that need to be respected. In this work we propose a way of modeling the service requirements so that the quality of service delivery can be measured;
- a performance index suitable for evaluating the quality of aggregation control algorithms from point of view of the Aggregator.

The paper is organized as follows: Section D.2 gives a general description of concepts relevant to the definition of the index, while the index itself is defined in Section D.3. A case study is presented in Section D.4 and further research is discussed in Section D.5.

## D.2 Background

### D.2.1 Ancillary Services

Ancillary services are acquired by TSOs in order to ensure the stability of the system and they can generally be divided into primary, secondary and tertiary ancillary services [110]. Each class of ancillary services has a different purpose in grid operation and works on different time scales.

In the Danish system, producers, represented by a Balance Responsible Party (BRP), are allowed to bid into the ancillary-services market once they have been approved by the TSO. In order to be approved, the producers must prove that they are able to deliver the relevant services within the requirements defined in [34, 11]. Here, the TSO defines the bounds of error in service delivery, e.g. how much deviation with respect to a reference power schedule can be accepted before the service is considered non-delivered. In this work, the QoS measures the deviations from the contracted behavior.

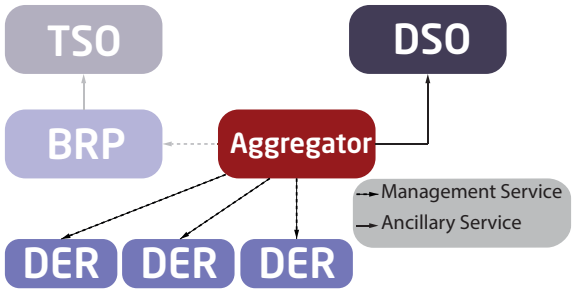



Figure D.1: The setup of the power system with DSM. Note that the Aggregator can either be an independent entity or can be a role inside a BRP.

Furthermore, it is expected that new ancillary services will appear in the near future [92]. The two main problems that the DSO seeks to solve are congestion issues, i.e. overloading of cables or transformers, and voltage issues. Throughout this paper, the recurring example of an ancillary service is the *PowerMax*, one of the new DSO services. This service is discussed further in Sec.D.3.3.

### D.2.2 Asset Management Service

Since the flexibility of individual DERs is too small to provide services to the system operators, an Aggregator pools the flexibility of the units, and presents their flexibility in the market as a single entity, see Fig.D.1. Thus, the Aggregator is responsible for managing the DER units according to certain requirements defined by the owners, hereby providing an asset management service. This service must respect the primary function of the DER.

By changing the consumption behavior of DER units, the Aggregator performs Demand Side Management (DSM), providing ancillary services to the DSO or, through a BRP, the TSO. The Aggregator and the BRP could be the same entity, but if they are not, the Aggregator should not work against the balancing responsibilities of the BRP.

### D.2.3 Control Performance Assessment

There is a field of theory on evaluation of controllers: Control Performance Assessment (CPA). Applications of this theory are found mostly in the process industry; for a thorough overview of its applications we refer to [67, 49].

Typically, CPA methods fall within two approaches. One approach, first introduced in [60], is to benchmark controller performance against a theoretical optimum, while taking the stochasticity of the system into account. The second approach is to benchmark against deterministic properties the closed-loop system must have, e.g. settling time and steady-state error [1]. In both cases, the index is usually scaled such that:

$$\zeta = \frac{J_{opt}}{J_{act}}, \tag{D.1}$$

where $J_{opt}$ is the theoretical optimal (minimum) value of the performance criterion $J$ (which is usually impossible to achieve in reality), and $J_{act}$ is the actual measured value of the criterion. Since $J_{opt} < J_{act}$, then $\zeta \in [0, 1]$.

According to [49], performance criteria used to evaluate a controller usually fall within three categories: Quality, Reliability, and Energy. Quality and reliability are concepts that can be directly related to ancillary service provision. The interpretation of energy-related criteria may be suitable for asset-management purposes but is considered out of scope in this work.

## D.3 DSM Performance Assessment

We identify four requirements for performance assessment of DSM:

R1 Provide a *quality* measure normalized to the contractual requirements (bounds) of a service. By normalizing the quality measure to the bounds, the *QoS* value for both ancillary services and asset-management services will have comparable dimensions.

R2 The measure should be normalized with respect to time.

R3 Provide a *reliability* measure in relation to service non-delivery.

R4 Each service must have a separate, individually verifiable, measure. For example, to evaluate service delivery w.r.t. ancillary-service delivery, the asset-management quality is irrelevant.

To satisfy these requirements, we propose a performance index quantifying the quality of ancillary services and asset-management services, and a non-delivery counter (NDC) which increases every time the QoS is out of bounds. Normalization is based on a scaling factor modeled after the contractual limits of the respective service. The limits are defined via a contract with the entity requesting the service. Thus, the performance index is specifically designed to evaluate how well the service provision conforms to the contractual boundaries.

### D.3.1 Definition of the performance index

In previous sections we have defined the concept of QoS as a deviation, $e(t)$, from a contracted behavior. Since there is a contractual limit on the allowed deviation, the error is normed to be a percentage of this limit such that:

$$QoS_s(t) = |e(t)|C_s(t); \quad QoS_s(t) \in [0, 1] \tag{D.2}$$

where *s* is either *AS* for ancillary service or *AMS* for asset-management service, and $C_s(t)$ is the corresponding normalization factor derived from the service model. When $QoS_{AS}(t) \geq 1$, the measure for reliability NDC is increased.

Using the square root of the Integral Square Error index (i.e. the 2-norm, as defined in e.g. [113]), the following performance criterion is defined for

service delivery seen from the Aggregator perspective:

$$J(N) = \sqrt{\int_0^N \left( \sum_{k=1}^{M} QoS_{AMS,k}(t)^2 + QoS_{AS}(t)^2 \right) dt} \tag{D.3}$$

where $QoS_{AM,k}(t)$ and $QoS_{AS}(t)$ are the time-dependent measures of service quality for the asset-management service and the ancillary service, respectively. The units controlled by the Aggregator are denoted by the index *k*, the unit portfolio is of size *M*, and *N* is the time horizon over which the services are provided. While the index (D.1) benchmarks the actual performance criterion against a theoretical minimum, we benchmark it against the worst case scenario $J_{max}$, such that the performance index is given by:

$$\eta = \frac{J_{act}(N)}{J_{max}(N)} \tag{D.4}$$

where $\eta \in [0, 1)$ for a valid service delivery and for which values close to zero represent good performance of service delivery. If $\eta \geq 1$ the Aggregator does not perform according to its service contract.

Normalization with respect to time is achieved when benchmarking against $J_{max}(N)$, since $J_{max}(N)$ is estimated by integrating over the service delivery period. Contrary to index (D.1), which gives an intuition of how close performance is to the optimum, index (D.4) gives an intuition of how far performance is from the worst case scenario. The index is designed this way because the theoretical optimum of service delivery is $J_{opt} = 0$, i.e. no error in service delivery.

### D.3.2 Calculating the index

Having defined what the performance index measures, we will proceed with establishing how to obtain the required values to estimate the index. Calculating the performance index requires the following steps:

1. Identify and model the service requirements and errors in service provision, giving the scaling factor $C(t)_s$.
2. Estimate $J_{act}(N)$.
3. Calculate *J(N)* for operation on the requirement boundaries ($J_{max}(N)$).
4. Calculate $\eta$ by benchmarking $J_{act}(N)$ with $J_{max}(N)$.

For the first step, the service requirements must be defined and translated into measurable errors. For some services, the error can be stated as a tracking error, e.g. $e = y_{ref} - y_{meas}$. In other cases, service requirements are defined by operation within bands, which may lead to an error defined as:

$$e(x) = \begin{cases} x_{min} - x & \text{if } x \leq x_{min} \\ 0 & \text{if } x_{min} \leq x \leq x_{max} \\ x - x_{max} & \text{if } x \geq x_{max} \end{cases} \tag{D.5}$$

This step is a service-specific problem and is non trivial.

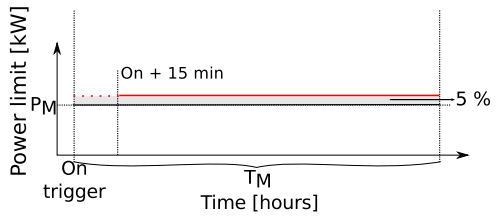



Figure D.2: The PowerMax service requirements, where the red line represents the boundaries for the permissible error, and the shaded area represents the error in service delivery, which is within the limits established in the QoS.

The second step requires computing $J_{act}(N)$ using measurement data from the unit portfolio. This can be a challenge for evaluation in field deployment. In this paper it is assumed that the measurement data is available, either through a DSO or a third-party metering company.

The third step requires the calculation of *J(N)* along the contractual boundaries for service delivery, in this way, the maximum allowed error is found for the service. The boundaries are based on the service models presented in the first step. By adding the maximum permissible error for all services, $J_{max}(N)$ is obtained. The following subsection present an example of how to determine $J_{max}(N)$.

### D.3.3 An example: DSO Service PowerMax

For demonstration purposes, in this section $J_{max}(N)$ for the PowerMax service is calculated. Typically, the service will be contracted several months ahead of the actual delivery. The activation schedule (On and Off triggers), the maximum power cap ($P_M$), the maximum duration of the service per activation ($T_M$), and the quality of service (*QoS*) are defined when contracting the service. The contract is valid for a period of several months, where the Aggregator is obliged to follow the established schedule.

The limits specified for the QoS[92] of the PowerMax service are presented here:

- Deviation from On trigger: $\pm$ 15 min. per day
- Deviation in size of service (dependent on $P_M$): Max. $\pm 5\% P_M$
- Acceptable no. of unsatisfactory activations(non-delivery): NDC $= 4$

A graphical representation of these service requirements is depicted in Fig. D.2. It is clear that the maximum acceptable error in service delivery is the shaded area. Note that the limit for non-delivery of service during the first 15 minutes of activation is dotted due to the fact that non-delivery is not counted during this period. The specifications for counting unsatisfactory activations are not clarified in [92], so it is assumed that breaking the QoS limits on one sampling period counts as one non-delivery. In the case where the service is not respected in three consecutive (or non-consecutive) sampling periods, NDC $= 3$.

For example, in the case where $P_M = 5\,kW$, $T_M = 4\,h$ and the power is measured once an hour, $J_{max}(N) = 2$, as it represents the square root of the square of the maximum (when $J_{act}(N) = 1$) permissible error over 4 hours.

## D.4 Case Study

This case study presents the aggregation of multiple flexible DERs via coordinated operation: 75 DERs installed in a suburban residential area, which are all connected to the same feeder leading to a 10/0.4 kV transformer. The transformer is rated to a maximum power flow of 200 kVA, which is sufficient under the current load circumstances, but will be a constraint in the future.

This case study addresses a scenario with high electric-vehicle (EV) penetration, low photo-voltaic (PV) penetration and electric space heating in all households. Furthermore all DERs connected to the same LV feeder offer their flexibility to the same Aggregator. Then, the proposed performance index for service provision is evaluated for two different aggregation control algorithms: Centralized soft Model Predictive Control (C-MPC) and Distributed soft Model Predictive Control (D-MPC).

### D.4.1 The reference case: without units coordination

In this section we make a scenario hypothesis for year 2050 regarding PV and EV penetration in a distribution feeder in a rural area and present simulation results. The following units are connected to the LV transformer:

- 40 buildings with electric climate control: resistive space heating with maximum load of 10 kW and air conditioning with a maximum load of 5 kW.
- 20 large EVs, with a battery size of 25 kWh, 11 kW.
- 10 small EVs, with a battery size of 14 kWh, 3.3 kW.
- 5 PV (polycrystalline) installations of 6 kW rated power each.

The PV installations provide forecasts of the production for one day ahead. To simulate uncertainty in the forecasts, Gaussian noise has been added to real data of PV production according to:

$$P_{PV-F,t} = P_{PV-T,t} + v_t, \quad v_t \sim N\left(0, \alpha\sqrt{P_{PV-T,t}}\right) \tag{D.6}$$

where $P_{PV-F,t}$ is the forecasted PV power production at time $t$, and $P_{PV-T,t}$ is the actual power production at time $t$ (from historical data). The term $\alpha$ is an uncertainty factor, which defines the variance of the noise as a percentage of the actual PV production, e.g. $\alpha = 0.1$ corresponds to a 10% forecast error. Uncertainty in solar radiation and ambient temperature are modeled in the same way. The actual power production time-series used in this case covers the same days as [28].

The load related to households is divided into climate control (flexible load) and everything else (non-flexible load). The building climate control is operated on MPC basis for minimum deviation from the temperature set point. Regarding the non flexible household loads, a five-day (one-hour-sampled) profile of the non-flexible load of 40 households is depicted in Fig. D.3.

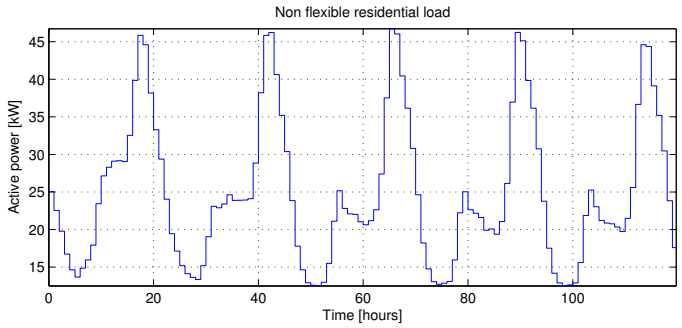



Figure D.3: The non-flexible load of the households under the transformer. The sample is statistically representative of Danish households.

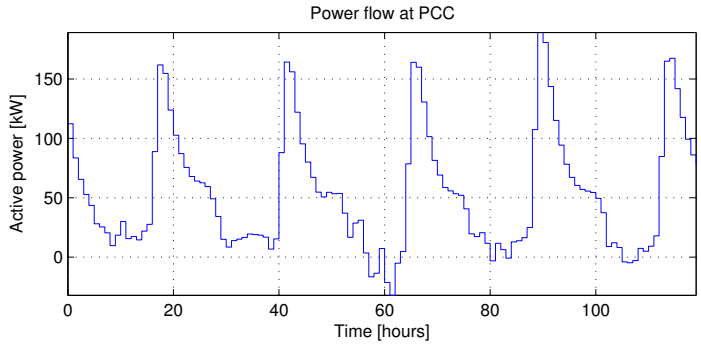
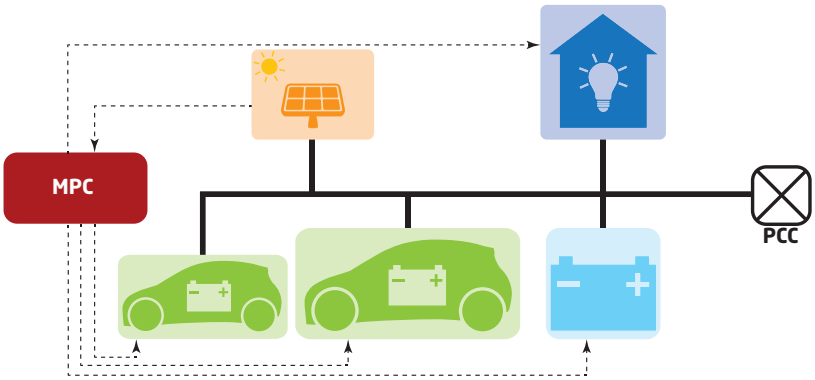



Figure D.4: Aggregated power flow at the point of common coupling for the reference case units without coordination: Demand Response based on day-ahead energy price.

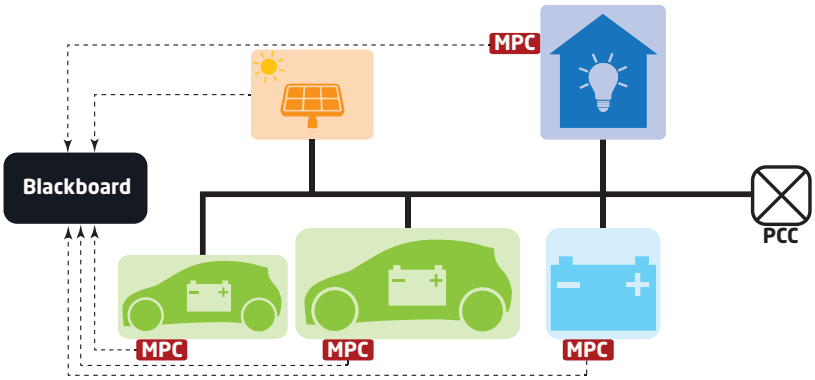
The EVs leave the charging station at a uniform randomly distributed time between 6:00 and 8:00, and are plugged again at a uniform, randomly distributed time between 16:00 and 18:00. The EVs operate on dumb charging, i.e. they try to fully charge as soon as they are connected to the grid. By running a simulation of the described scenario without units coordination, the results shown in Fig. D.4 are obtained.

EVs operating on dumb charging can cause peak consumption up to 190 kW at the point of common coupling (PCC). Given that the transformer capacity is 200 kW and it is customary to reserve 30% of the transformer capacity for emergency operations [36], the DSO aims at keeping the load below 140 kW and limiting the inverse power flow at the substation. Thus, the DSO can sign a contract for PowerMax service (see Sec. D.3.3) with an Aggregator which, at any time, operates Demand Response via Direct Load Control (DLC) [76] in order to limit the power flow at the transformer. The maximum capacity available at the transformer is therefore 140 kW for direct power flow and -10 kW for inverse power flow.

The rest of this section presents the C-MPC and D-MPC formulations. For the formulation of the mathematical models we refer to [86] for the battery model and to [2] for the building space heating model (modified, as proposed in [28]). For the modeling of the services, we apply the method described in Sec.D.3.2. A discussion on the simulation results concludes this section.




(a) The setup of the Centralized MPC scheme.




(b) The setup of the DMPC scheme as seen in [28].

Figure D.5: The setup of the two Aggregation algorithms to be compared.

### D.4.2 The Centralized Model Predictive Control scheme

In this scheme the Aggregator contains the control algorithm to centrally manage all the units in its portfolio (Fig. D.5(a)). Since the Aggregator optimizes its portfolio's consumption through MPC, it has detailed knowledge of the state and dynamics underneath. The units portfolio is the same as of the reference case. The C-MPC control problem is formulated as quadratic optimization with soft constraints (as seen in e.g. [106]):

$$\min_{u_t,\vartheta_t} \; J = \sum_{t=1}^{N} \left[ \|y_t - r_t\|_Q^2 + \rho\vartheta_t + \psi\gamma_t \right] \tag{D.7a}$$

$$subject\ to:$$

$$x_{t+1} = Ax_t + Bu_t + Ed_t \tag{D.7b}$$

$$y_t = Cx_t + Du_t \tag{D.7c}$$

$$u_{\min,t} \leq u_t \leq u_{\max,t} \tag{D.7d}$$

$$y_{\min,t} - \gamma_t \leq y_t \leq y_{\max,t} + \gamma_t \tag{D.7e}$$

$$PCC_{\min,t} - \vartheta_t \leq u_t \leq PCC_{\max,t} + \vartheta_t \tag{D.7f}$$

$$\vartheta_t, \gamma_t \geq 0 \tag{D.7g}$$

where $r_t$ and $y_t$ are the output reference and system outputs (internal house temperature and battery state of charge) respectively over the prediction horizon $t = 1..N$, $\psi$ is the weight for output soft constraints, with $\gamma$ being the corresponding slack variable, and $\rho$ penalizes the *power over max* defined in Eq. D.7f. Since this MPC controller is centralized, the state space system matrices in Eq. (D.7b) and Eq. (D.7c) are formed by block

diagonal-adding each of the systems' respective matrices. With the set of units $\mathcal{S} = \{1..N\}$, it follows:

$$
\begin{aligned}
x &= \begin{bmatrix} x_1 \\ x_j \end{bmatrix}, u = \begin{bmatrix} u_1 \\ u_j \end{bmatrix}, d = \begin{bmatrix} d_1 \\ d_j \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_j \end{bmatrix} \\
A &= \begin{bmatrix} A_1 & 0 \\ 0 & A_j \end{bmatrix}, \quad B = \begin{bmatrix} B_1 & 0 \\ 0 & B_j \end{bmatrix} \\
C &= \begin{bmatrix} C_1 & 0 \\ 0 & C_j \end{bmatrix}, \quad D = \begin{bmatrix} D_1 & 0 \\ 0 & D_j \end{bmatrix} \\
E &= \begin{bmatrix} E_1 & 0 \\ 0 & E_j \end{bmatrix}, \quad \vartheta = \begin{bmatrix} \vartheta_1 \\ \vartheta_j \end{bmatrix}, \gamma = \begin{bmatrix} \gamma_1 \\ \gamma_j \end{bmatrix}
\end{aligned}
\tag{D.8}
$$

where the index $j \in \mathcal{S}$ and the system in Eq. (D.7b) and Eq. (D.7c) is extended with all the units belonging to the set $\mathcal{S}$.

### D.4.3 The Distributed Model Predictive Control scheme

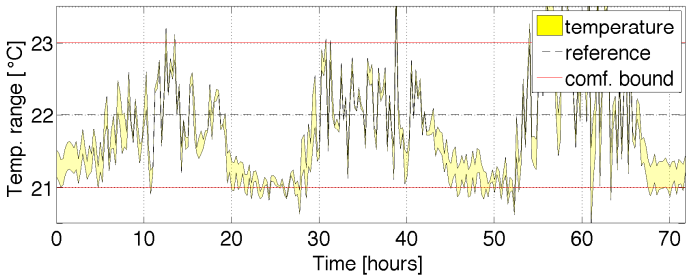
In the D-MPC formulation, units within the same cluster retrieve the power plans of the other units, compute their own plan (over a prediction horizon) accordingly and publish it on a blackboard. Note that in this case study, in contrast to what has been proposed in [28], the unit controllers have soft constraints on the outputs (temperature for buildings and State of Charge (SOC) for batteries and EVs). In this algorithm, as soon as the units publish their consumption plan, the available power at the PCC decreases in such a way that the subsequent units communicating with the blackboard tend to adjust their plan accordingly. After a negotiation period the units are entitled to operate according to the power plan that has been published in the blackboard for the next time frame. Figure D.5(b) shows the configuration for the D-MPC. This is an example of transactional control [76], where the unit power consumption is negotiated.

### D.4.4 Comparison and discussion of results

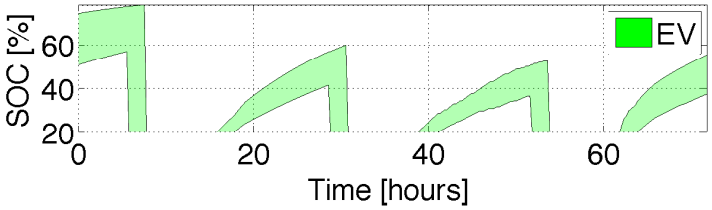
Certain assumptions have been made with regards to controllers:

The EVs are preferably kept operating in the range $SOC = [0.2, 0.9]$ due to battery life concerns[86], although it is possible to operate in $SOC = [0.0, 1.0]$. The comfort band for the households lies in the band $T_{ref} = 22°C \pm 1°C$. The concept of non-delivery is not used in the asset-management services, but the absolute boundaries for user-comfort bands lie on $T_{ref} = 22°C \pm 1.5°C$.

The required PowerMax service is of $P_M = 90kW$ each day in the periods of 16:30 to 20:30. The time sampling of the simulation is of 15 minutes and the power plans are computed for a horizon of 23 hours (i.e. the MPC prediction horizon). The EVs are not capable of providing Vehicle-to-Grid (V2G) services, i.e. EVs only charge.




(a) Households temperatures




(b) EV State of Charge

Figure D.6: Simulation results for the D-MPC with $\alpha = 0.1$

These assumptions lead to the results presented Figs. D.6-D.8 and Tables D.1-D.2. The following conclusions can be made:

1) from Figs. D.6 and D.7 it can be seen that both controllers are quite good at staying within the QoS limits of the DSO and EV owners, which can be seen in the fact that none of the controllers have non-delivery and $\eta$ is small. It is clear that the value of $\eta$ comes from the behavior of the household heating, where the C-MPC delivers a better quality service to end users than the D-MPC, although it might not be obvious from the figures.

2) controller performance is sensitive to prediction uncertainties, as can be seen in the varying values of $\eta$ depending on the uncertainty $\alpha$ (see Eq. (D.6)), which is shown in Table D.1.

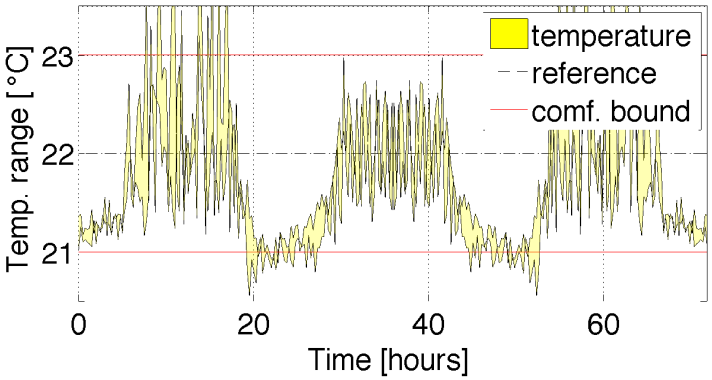
3) in terms of service provision, the C-MPC outperforms the D-MPC. This arises from the fact that the C-MPC has absolute control of all units and determines a global optimum.

4) due to the behavior difference between the local EV controllers in the D-MPC scheme, and the behavior of the C-MPC, the power consumption of the EV is very different (compare Fig. D.6(b) and Fig. D.7(b)). This also leads to a vast difference in the power flow at PCC (see Fig. D.8).

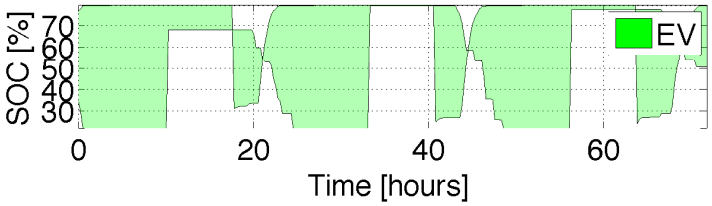
5) from the values in Table D.2, it can be seen that the values of $\eta$ are in the same order of magnitude when simulations are done for varying numbers of days. This is caused by the normalization of $J_{act}(N)$ over time (reflected in $J_{max}(N)$). This means $\eta$ evaluates the aggregation algorithm taking service provision time into account, and gives an overall assessment of the algorithm, dependent on the length of time the Aggregator must sustain the service provision.

| | D-MPC | | C-MPC | |
|---|---|---|---|---|
| $\alpha$ | 0.1 | 0.2 | 0.1 | 0.2 |
| NDC | 0 | 0 | 0 | 0 |
| $\eta$ | 0.0075 | 0.0160 | 0.0054 | 0.0153 |

Table D.1: Results of three-day simulation

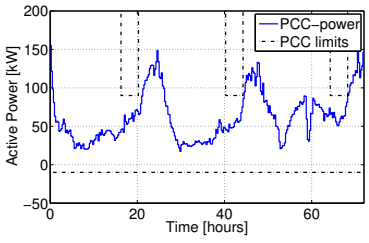



(a) Households temperatures

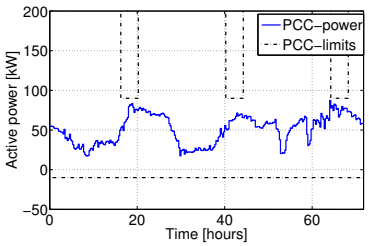



(b) EV State of Charge

Figure D.7: Simulation results for the C-MPC with $\alpha = 0.1$




(a) Total power load for C-MPC




(b) Total power load for D-MPC

Figure D.8: Power load at Point of Common Coupling for the controllable and non-controllable loads

## D.5 Conclusion and Outlook

Drawing inspiration from the field of Control Performance Assessment, this study proposes a performance index for the evaluation of control services for DER aggregation. The index is useful for the systematic evaluation of the adequacy of different control architectures providing ancillary services. It was shown how the index is computed, and a case study was presented in which two different control algorithms were evaluated. The results were presented and discussed, showing that the C-MPC in this case is capable of providing a better QoS. In order to do a successful evaluation of an aggregation algorithm, it is important that the QoS specifications of the future ancillary services are well defined. This is a challenge in itself since many of the ancillary services assume a production baseline, which is easy to establish in traditional generators, but proves to be difficult for small households (see e.g. [15]). Research effort should be put into redefining ancillary-

| Days simulated | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $\eta$ | 0.0013 | 0.0018 | 0.0082 | 0.0052 | 0.0062 |

Table D.2: D-MPC Performance over different simulation lengths

service requirements to suit DSM, taking into account the probabilistic nature of managing a large number of units.

The evaluation of aggregation control algorithms is an important part of a general validation framework for Aggregators. Future work will include further development of this Aggregator validation framework, where controllers can be tested under different grid and communication network topologies, as well as a diverse set of fault scenarios.

## Acknowledgment

The authors acknowledge the financial support of iPower (www.ipower-net.dk). The authors thank Shi You from DTU Elektro for providing data for the non-controllable loads.

# Chapter E

# Method for Ancillary Service Modeling and Performance Assessment

**Authors:**
Daniel Esteban Morales Bondy
Anders Thavlov
Janus Bundsgaard Mosbæk Tougaard
Kai Heussen
**Submitted to:**
IEEE Transactions on Smart Grid
**Abstract:**
Aggregation of large quantities of consumption units is expected to be a new source of power system ancillary services. For large and conventional generation units the dynamic response is well understood and detailed individual measurement is feasible, which factors in to the straightforward performance requirements applied today. Aggregation-based ancillary service delivery can be very responsive and fast, but the dynamic response can also be uncertain, subject to both variations in aggregator infrastructure and algorithms as well as diversity of flexibility resources. For secure power system operation, a reliable service delivery is required, yet it may not be appropriate to apply conventional performance requirements. The service performance requirements and assessment method therefore need to be adapted to service provision from aggregators. This paper develops a modeling method for ancillary services performance requirements applicable to aggregation portfolios, including performance and verification indices. The use of the modeling method and the indices is exemplified in two case studies.
**Keywords:**
Ancillary Service Modeling; Performance Assessment; Aggregator; Demand Side Management

## E.1 Introduction

The power industry is experiencing a significant shift away from being based on fossil fuels towards more generation from Renewable Energy Sources (RES). The tendency is a substantial increase in the amount of RES, often as distributed energy resources (DER), and a growing electrification of the heating and transportation sectors [65, 43]. The non-dispatchable and stochastic nature of RES and the increasing electrification of consumption call for new sources of ancillary services, as conventional generation is pushed out of the market. This alters the traditional distribution of flexibility resources in the sector, where relatively few large power plants provides electric power and ancillary services (AS). A new AS resources will be demand response (DR) from small-scale entities, such as commercial buildings or private households, whose flexibility in consumption will be harnessed by aggregators [107, 73, 3, 135, 87, 122]. With the introduction of aggregators as providers of ancillary services, the AS specifications are being adapted to new resource types, but also prequalification and verification of service delivery need to be adapted to be suitable for the aggregated service delivery [7]. This is relevant both due to the change in ancillary service specifications and due to the introduction of new distribution system services [62].

Currently, the verification of ancillary service delivery typically is based on a rigid performance assessment (pass/non-pass) of the units providing services [34]. Based upon the FERC order 755 [45], PJM, an American regional transmission operator, has implemented a pay-for-performance scheme by evaluating the performance of frequency regulation units, thus changing the rigid verification procedures. While performance criteria have been formulated for specific services, e.g. load frequency control [50] or primary frequency control [41], these focus on the overall performance of the reserves seen from a grid perspective, i.e. the criteria do not evaluate the individual performance of each service-providing unit. PJM has introduced a performance score which evaluates the unit performance, but its definition is tied to the regulation (reference tracking) service [104].

It is clear from the literature that the interpretation of performance assesment of aggregators varies widely. This is examplified by the ad-hoc evaluations of specific aggregator implementations, e.g. [135], or the when the performance evaluation focuses on non-service related metrics like computational or financial performance, e.g. [121, 109]. This inconsistency in performance evaluation is a consequence of the lack of clear service requirements for aggregators. If aggregator are to deliver ancillary services, it must be clear on what grounds they are being verified. This issue must be addressed through the formulation of standard performance requirement models.

This paper presents a method for modeling a generic set of active power ancillary services requirements, as well as requirements for distribution system congestion management services. Furthermore, it is shown how these models can be used for performance assessment of aggregators. The article refines the concept of a service performance assessment index and Quality of Service, both treated in previous work, and further introduces

a new metric for assessing the non-delivery of a service. The non-delivery assessment is proposed for verification of the aggregator service delivery.

The rest of the article is organized as follows: background on AS verification and changing characteristics are presented in Sec. E.2, the modeling method is presented in Sec. E.3 and the service performance assessment and verification indices are presented in Sec. E.4. The use of the service modeling and indices are shown through two case studies in Sec. E.5 and concluding remarks are presented in Sec. E.6.

## E.2 Background

In this section, the terminology and concepts utilized throughout the paper are defined.

### E.2.1 Ancillary services provision from unconventional resources

Ancillary services are utilized by transmission system operators (TSOs) to ensure a adequate and secure operation of the power system. Currently, frequency containment reserve (primary reserve) and frequency restoration reserve (secondary reserves) [39] are widely utilized by TSOs to ensure frequency stability. In the future, such services are expected to be delivered by aggregators [107, 135]. Furthermore, with the introduction of aggregators, new possibilities arise for solving problems at distribution system level, e.g. congestion issues, leading to new flexibility services being defined for distribution system operators (DSOs), as presented in [30].

With an increased electrification of the energy system due to the introduction of electric vehicles (EVs), heat pumps and local generation, DERs are expected to deliver an increasing amount of ancillary services in the future power grid. The DERs that can be utilized for ancillary services are those which can provide flexibility in consumption or generation without significantly impacting their primary energy service, e.g. battery state of charge and indoor temperature comfort [28, 52].

### E.2.2 Roles in the market for ancillary services

Ancillary services are acquired in a single-buyer auction by a transmission system operator (TSO)[1] via an open market, where approved participants can bid their reserves. Balance responsible parties (BRPs) are responsible for the balance of power production or consumption within their portfolio, with respect to the schedule of traded energy and the respective. The actors and their relationships can be seen in Fig. E.1. Apart from the required approval, also a minimum bid size limits market entry, as most DERs have a smaller capacity [75].

Aggregators, who pool large numbers of DERs, can represent these as aggregate resource to a system operator. The aggregator can deliver TSO and DSO services as well as flexibility services for balance responsible

[1] a European TSO corresponds largely to an ISO (independent system operator, with the limitation that a TSO does not host the energy markets.

parties (BRPs) e.g. [131, 133]. Finally, in order to avoid the aggregator creating imbalances for the Balance Responsible Consumer, all aggregator-service sales must occur through the Balance Responsible Consumer.

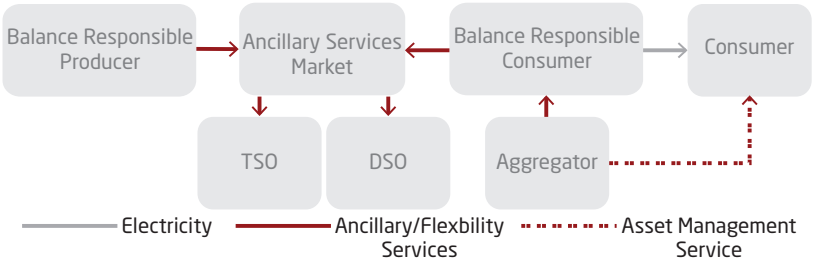



Figure E.1: The new player in the market for ancillary services is the aggregator, which sells consumption flexibility on the ancillary service markets through the Balance Responsible Consumer. Some markets allow the aggregators to participate directly in the ancillary service markets with the condition that they coordinate with their Balance Responsible Consumer. It also provides flexibility services to BRPs and DSOs.

### E.2.3 Service verification today

When contracted for service, units are subject to a set of requirements. First, units must pass a prequalification test.Second, certified metering instrumentation must be installed on the unit, and telemetry equipment must be installed and connected to the system operator's Supervisory and Control Data Acquisition (SCADA) system.

For verifying reserve services, the system operator does random checks to see if the reserve is available at the unit [34]. With respect to regulation services, these are expected to be delivered within the required time requirements, and must be measured with acceptable accuracy. For example, for consumption units smaller than 1.5 MW acceptable accuracy is 2% of the load [35].

### E.2.4 The need for service requirements modeling

The concept of verification of ancillary services is moving towards a more flexible definition. Furthermore, new types of flexibility services are appearing, e.g. [131, 62]. Part of the lessons learned from a demonstration of one of these new DSO services [66, 12] is that the services, along with their requirements, need to be clearly defined if aggregators are to deliver them. The main problem being that the verification of services delivered by aggregators is complex. Also, research points at the need for change of the traditional service requirements if aggregators are to be enabled [84, 75]. Thus, models that translate service contract requirements into benchmarks for performance assessment are needed.

From the market setup described in the previous section, parallels can be drawn to the concept of *Service-Oriented Architectures* (SOAs) from the field of computer science. Under that paradigm, service is defined as: a logical representation of a repeatable business activity that has a specified outcome, is self-contained, may be composed of other services and is a *black box* to consumers of the service [129]. An element in SOAs are *Standardized service contracts*, which contain service level agreements (SLAs). The

SLAs can be interpreted as the requirements defined in an ancillary service contract. SLAs must define service performance metrics with corresponding service level objectives (SLOs), which are the agreed means for measuring performance.

PJM has established precedence in using performance metrics for verification of services. Their *performance score* consists of a weighted average of three measures: delay, accuracy and precision. These measure the delay and correlation between the regulation signal and the reaction of the unit, and the difference in energy requested vs. energy supplied [105]. These measures are tailored to the way frequency regulation is done at PJM (tracking of the regulation signal). Therefore, more general (and simple) models and performance metrics are needed to cover other frequency regulation services and the new flexibility services.

## E.3 Method for service modeling

This section presents the requirements for the service models, as well as a method for deriving these models.

### E.3.1 Requirements of service models

Through analysis of the TSO services defined in [34], the potential DSO services defined in [30] and asset management services, a set of requirements for the models have been established:

M-R1 the model must clearly identify the SLOs of the service,

M-R2 the model must incorporate both the ideal and acceptable service provision in a measurable/quantifiable way, i.e. performance metrics must be able to be applied to it,

M-R3 the models must be technology agnostic,

M-R4 since flexibility services imply a change of consumption pattern over a period of time, the models must consist of time series

Furthermore, based upon previous work of the authors [14], the following requirements are defined for the performance metrics:

P-R1 provide a *quality* measure normalized with respect to the contractual requirements (bounds) of a service and with respect to time,

P-R2 provide a *reliability* measure in relation to service non-delivery, which is normalized with respect to time,

P-R3 service quality and reliability evaluation must be applicable to entities providing multiple services.

### E.3.2 Method for formulation of requirements model

Based upon the requirements [M-R1..M-R4], the method for SLA modeling is defined by the following six steps:

1. Identify physical parameters defining the service [M-R1],

2. Identify the dynamic behaviors of the service related to system parameters (if any) [M-R1],

3. Identify the physical size of the service and the tolerated error [M-R2],

4. Identify the ideal response time of the service and acceptable response [M-R2]

5. Based on the dynamics, size and timing of the service, as well as the tolerated errors from points 1–4, develop a time series for ideal and acceptable service provision. The model will be a set of time series: $\mathbf{x}_{ideal}(t)$ for ideal response and $\mathbf{x}_{acc}(t)$ for acceptable response. Both time series can be a scalar or a vector, e.g. $\mathbf{x}_{acc}(t)$ can be formed by a set of upper and lower tolerance bounds or simply by an upper bound [M-R4],

6. Identify how the service error is to be measured [M-R1].

By only defining the SLA models in terms of performance, not in specific unit capabilities, the models implicitly comply with [M-R3].

### E.3.3 Generic model components

We identify three service model patterns: reference tracking, band service or a maximum/minimum cap. The error measure, $e(t) \in \mathbb{R}$, for each of the service types is defined in the following subsections. This approach was initially introduced in [14], and is further refined in this work.

#### Reference tracking

Reference tracking error can be calculated as:

$$e(t) = x_{meas}(t) - x_{ideal}(t), \tag{E.1}$$

where $x_{meas}(t)$ is the measured load/generation and $x_{ideal}(t)$ is the signal to be tracked. This definition will lead $e < 0$ for measured values below the ideal and $e > 0$ for values above the ideal. In this case $\mathbf{x}_{acc}(t)$ will be a band around $x_{ideal}(t)$, and the values of $\mathbf{x}_{acc}(t)$ do not need to be symmetric.

### Band service

The ideal response in a band service is defined as $\mathbf{x}_{ideal}(t) = [x_{min}(t), x_{max}(t)]$. The error in the band service can therefore be estimated by:

$$e(t) = \begin{cases} x_{meas}(t) - x_{min}(t), & x_{meas}(t) < x_{min}(t) \\ 0, & x_{min}(t) \leq x_{meas}(t) \leq x_{max}(t) \\ x_{meas}(t) - x_{max}(t), & x_{meas}(t) > x_{max}(t). \end{cases} \quad \text{(E.2)}$$

In this case, the $\mathbf{x}_{acc}(t) = [\mathbf{x}_{acc,min}(t), \mathbf{x}_{acc,max}(t)]$ is a set of values that surround the band defined by $\mathbf{x}_{ideal}(t)$, as seen in Fig. E.5. The values of $\mathbf{x}_{acc}(t)$ do not need to be symmetric around the band.

### Cap service

In cap services, error is only tracked when $x_{meas}(t)$ is either above or below a given a limit value. Maximum cap error is calculated as shown in (E.3) and minimum cap can be similarly calculated. In (E.3), $x_{max}(t)$ is the ideal maximum limit according to the service contract:

$$e(t) = \begin{cases} x_{meas}(t) - x_{max}(t), & x_{meas}(t) > x_{max}(t) \\ 0, & x_{meas}(t) \leq x_{max}(t). \end{cases} \quad \text{(E.3)}$$

In the cap service, $x_{acc}(t)$ is a limit that either lies below $x_{min}(t)$ or above $x_{max}(t)$.

## E.4 Performance Assessment of Service Delivery

In Sec. E.2.4 three requirements for performance metrics are presented. These can be expressed formally as:

$$\text{[P-R1]:} \quad \eta = f_P(x_{meas}, \mathbf{x}_{acc}, t), \quad \eta \in [0, 1], \quad \text{(E.4)}$$

$$\text{[P-R2]:} \quad \epsilon = f_R(x_{meas}, \mathbf{x}_{acc}, t), \quad \text{(E.5)}$$

$$\text{[P-R3a]:} \quad \eta_K = \sum_{i \in M} f_M(\eta_i), \quad \eta_i \in [0, 1], \quad \text{(E.6)}$$

$$\text{[P-R3b]:} \quad \epsilon_M = \sum_{i \in M} f_M(\epsilon_i), \quad \text{(E.7)}$$

$$\quad \text{(E.8)}$$

where $\eta$ is a quality performance measure, $\epsilon$ is a reliability measure, $\eta_M$ and $\epsilon_M$ are the same measures applied to multiple services $M$. The measured output (or sum of outputs in the case of aggregation) is defined by $x_{meas}$, and the service bounds are defined by $\mathbf{x}_{acc}$, as defined in Sec. E.3.2. $f_P(\cdot)$ is a function that evaluates service performance normalized to $\mathbf{x}_{acc}$ and time $t$. Similarly, $f_R(\cdot)$ is a function that evaluates service reliability based upon $\mathbf{x}_{acc}$ and normalized to time and $f_M(\cdot)$ is a function that gives an overall measure for multiple services.

These concepts were originally presented in [14], but are revised and expanded upon following concepts from [126]. In order to asses service performance three concepts are introduced in this section:

- Quality of Service, which is an instantaneous measure of how well the aggregator is delivering one service within the contract constraints;
- service performance assessment index, which describes the overall performance of the aggregator over the delivery period for the services, or subset of services, it is providing; and a
- service verification index, which describes how much an aggregator is breaking the contractual agreements (non-delivering) of the services, or a subset of services, it provides.

Differently from the previous work, the service delivery index is split into measures of the ancillary services (AS) delivered to system operators and the AMS delivered to unit owners (see Sec. E.2.1). In this way, a system operator (or a third party certification company) can use the index for certification of aggregators, for which the AMS evaluation is irrelevant. Furthermore, the service verification index is introduced, and a new way of defining the quality of service is presented.

### E.4.1 Quality of Service

Quality of service (QoS) is a measure defined in [14], where it is used to assess the quality of a power system service at any given time. QoS at any given time is given by:

$$QoS(t) = e(t)C_n(t) \tag{E.9}$$

where *e(t)* is the error in service delivery introduced in Sec. E.3, and $C_n(t)$ is a normalization factor that can be time varying. Following [P-R1], we define:

- $QoS \geq 0$;
- for $QoS \leq 1$ the service is considered delivered within the contractual constraints;
- and $QoS = 0$ is a perfect service delivery.

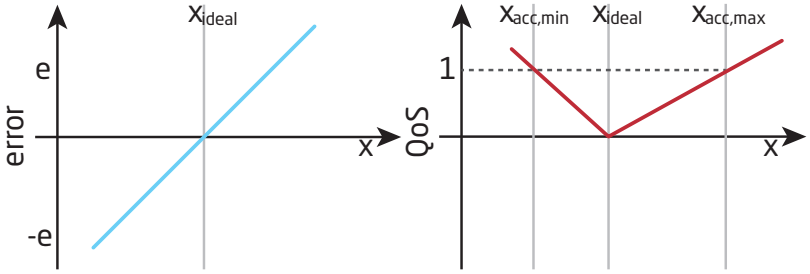
In order to achieve these definitions, the normalization factor $C_n(t)$ must be calculated from $\mathbf{x}_{acc}(t)$ thus:

$$C_n(t) = \begin{cases} \frac{1}{x_{acc,max}(t) - x_{max}(t)}, & e(t) \geq 0 \\ \frac{1}{x_{acc,min}(t) - x_{min}(t)}, & e(t) < 0. \end{cases} \tag{E.10}$$

where $x_{acc,max/min}$ and $x_{max/min}$ are part of the service model defined in Sec. E.3. By defining $C_n(t)$ in this way, we take into account the possibility of asymmetry in the values of $x_{acc}$, and ensure that QoS is a positive value. A visual representation of this scaling can be seen in Fig. E.2–Fig. E.4, where the QoS for the three kinds of services are presented. In general, the rate with which $QoS(t)$ increases depends on the difference between $x_{acc}(t)$ and $x_{ideal}(t)$.

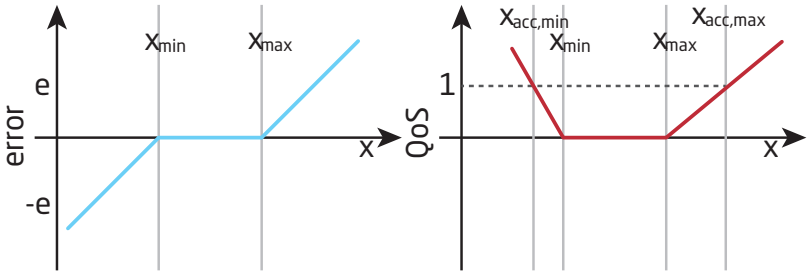
Note that in (E.10), $C_n(t)$ is not defined for $x_{acc}(t) = x_{ideal}(t)$. This is a corner case, in which:

$$QoS(t) = e(t), \quad x_{acc}(t) = x_{ideal}(t) \tag{E.11}$$



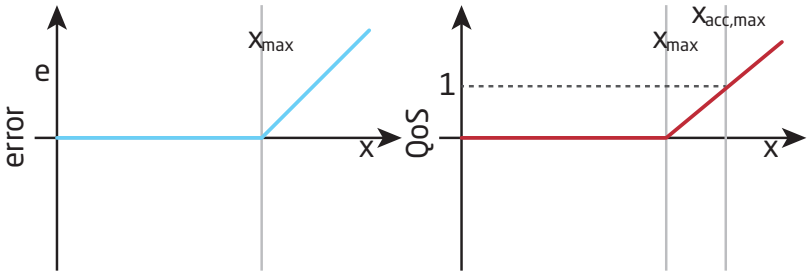

(a) Error (b) Quality of Service

Figure E.2: Error and QoS for tracking services, note that the acceptable band do not need to be symmetric.




(a) Error (b) Quality of Service

Figure E.3: Error and QoS for band services.




(a) Error (b) Quality of Service

Figure E.4: QoS for a maximum cap service, a minimum cap service is defined similarly but with $x_{min}$ and $x_{acc,min}$ values.

### E.4.2 Assessing service delivery

Based on the above instantaneous measure for the quality of individual services, we can evaluate the aggregator as a whole based upon the quality of all the services it delivers.

The overall service delivery index of AS is defined by $\eta^{AS}$ in Eq. (E.13), but before calculating the index, the non-delivery incidents (which are measured apart) must be sorted out. This is done by restricting $QoS^{AS}_{K,meas}(t)$ (the measured quality of service for the $K$ ancillary services the aggregator is providing) such that it does not account for $QoS > 1$:

$$QoS^{AS}(t) = \begin{cases} QoS^{AS}_{meas}(t), & \forall QoS^{AS}_{meas}(t) \leq 1, \forall t \\ 1, & \forall QoS^{AS}_{meas}(t) > 1, \forall t. \end{cases} \tag{E.12}$$

where $K$ is the total number of AS the aggregator provide. This restriction is not done in [14] since that work did not use a separate reliability index. This means that $\eta_{AS}$ is only a measure of the service provision performance within the contractual limits.

While the previous definitions have been established in continuous time, the actual measurement and calculations are done in discrete time. This leads to $\eta_{AS}$ being estimated for $K$ amount of AS over each corresponding discrete time horizon $N_K$:

$$\eta^{AS} = \sum_{i=1}^{K} W_i \sqrt{\frac{\sum_{t=0}^{N_i} \left( {QoS^{AS}_{i,t}}^2 \right)}{N_i}} \tag{E.13}$$

$$\sum_{i=1}^{K} W_i^{AS} = 1 \tag{E.14}$$

where $W^{AS}_K$ is the assigned weight to each AS, leading to $\eta^{AS} \in [0, 1]$, and $\eta$ close to zero representing good performance while $\eta$ close to 1 representing a barely acceptable performance. This means that the service performance assessment index for all the AS the aggregator provides is a weighted average of the root mean square (RMS) of the error in all service deliveries, thus satisfying [P-R3a]. With this index it is possible to evaluate aggregators that deliver more than one AS at a time, e.g. a frequency containment reserve and a replacement reserve, and assign a hierarchy of importance with respect to the services. However, how to do distinguish measurements to verify the services, and how to evaluate which service is more important, is out of scope of this work, but the definition of Eq. (E.13) takes the possibilities into account. In the case where only a single service delivery is considered, Eq. (E.13) is simply the RMS of the error in service delivery:

$$\eta^{AS} = \sqrt{\frac{\sum_{t=0}^{N} \left( {QoS^{AS}_{t}}^2 \right)}{N}}, \tag{E.15}$$

which satisfies [P-R1].

Eq. (E.13) gives an idea of the performance of the aggregator where the duration of time delivery is taken into account. This means that two service provisions are evaluated equally when their error in service delivery compared to the duration of the service delivery are the same.

### E.4.3 Verifying service delivery

Requirement *P-R2* defines a reliability measure. To address this requirement, an index $\epsilon^{AS}$, similar to the service performance assessment index, is defined for verifying the delivery of AS[2]. Also, a non-delivery measure for the AS provision, $ND^{AS}$, is defined according to the expression:

[2] This can also be interpreted as evaluating non-delivery of service.

$$ND^{AS}(t) = \begin{cases} QoS^{AS}_{meas}(t) - 1, & \forall QoS^{AS}_{meas}(t) > 1, \forall t \\ 0, & \forall QoS^{AS}(t) \leq 1, \forall t. \end{cases} \tag{E.16}$$

Eq. (E.16) shows that whenever the QoS of a service exceeds 1, i.e. the limit of what is an acceptable service provision, the amount with which it breaks the acceptable constraint is measured by *ND*. $\epsilon^{AS}$ is calculated in the same way as $\eta^{AS}$ using $ND^{AS}_K(t)$ instead of $QoS^{AS}_K(t)$:

$$\epsilon^{AS} = \sum_{i=1}^{K} W_i \sqrt{\frac{\sum_{t=0}^{N_i} \left( ND^{AS^2}_{i,t} \right)}{N_i}} \tag{E.17}$$

where $\epsilon^{AS} \in [0, \infty]$. This expression satisfies [P-R3b], and in the case where *K=1* it also satisfies [P-R2].

Thus, $\epsilon$ is used to asses the severity of non-delivery events. For some systems it is critically important that $QoS(t) \leq 1$ at any time, in which case $\epsilon$ should be close to zero for the contract to be considered respected. Other systems can tolerate $QoS(t) > 1$ for some period, which leads to a higher acceptable $\epsilon$. A service delivery is verified if $\epsilon \leq \epsilon_{max}$, and this contractual limit, i.e. the value of $\epsilon_{max}$, must be assessed individually depending on the nature of the system.

In [14], non-delivery is assesed using a non-delivery counter (NDC). $\epsilon$ differs from the NDC in that it both captures the time span of non-delivery and the magnitude of the violation, whereas NDC only captures the amount of time samples where non-delivery is detected. $\epsilon$ might prove advantageous over the NDC as a service verification index for some systems. A disadvantage of $\epsilon$ is that it might be a less intuitive measure to communicate to the service providers compared to the NDC.

Fig. E.5 shows an example of reference tracking error service performance assessment. Deviations between $x_{meas}$ and $x_{ideal}$ inside the band defined by $x_{acc}$ will lead to $QoS < 1$, while deviations outside the $x_{acc}$ band will lead to $QoS > 1$. For this particular example $\eta^{AS} = 0.7501$ and the service verification index is $\epsilon^{AS} = 0.2324$, which indicates that generally the service provision is bad at following the reference, and also has a moderate amount of non-delivery. The service acquirer will have to decide whether this verification index value is acceptable or if it should lead to economical penalization or contract termination.

## E.5 Case Studies

Using two different ancillary services and an asset management service as cases, we will illustrate the utility of the generic service modeling method, the service performance index and the service verification index. The first

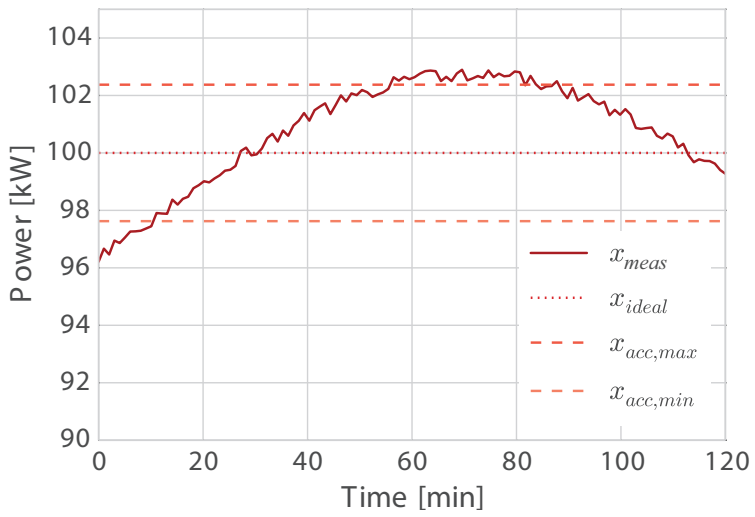



Figure E.5: Example of reference tracking error with performance $x_{meas}$, ideal performance $x_{ideal}$ and tolerance limits $x_{acc}$.

case study focuses on frequency containment reserve in western Denmark, the second focuses on the theoretical PowerMax DSO service, and the third focuses on the temperature management of a residential house.

### E.5.1 Frequency Containment Reserve in Western Denmark

Frequency Containment Reserve (FCR) is utilized to contain frequency excursions deviating from the nominal 50 Hz in *ENTSO-E RG Continental Europe's synchronous area* of which western Denmark (DK1) is part of. The Danish TSO, Energinet.dk, is obliged to provide a proportional share of $\pm$ 23 MW [34] out of the total synchronous area need of $\pm$ 3000 MW. Energinet.dk buys these reserves at daily auctions. The service specifications are defined in [34].

The six steps outlined in Sec. E.3 are used to model the ideal and tolerated service response. 1) The physical parameters are grid frequency (accuracy of $\pm$ 10 mHz or better), generator reserve power output, and timing of service delivery (accuracy of 1 s or better). 2) The reserve must be supplied linearly at deviations of $\pm$ 200 mHz relative to 50 Hz, with a $\pm$ 20 mHz dead-band around 50 Hz. 3) The physical size of the service depends on the reserve bid size. This work will look at a generic reserve bid. According to the discussion from Eq. (E.9), $x_{ideal}$ cannot be equal to $\mathbf{x}_{acc}$. Therefore, a $\pm$ 1% tolerance band of $x_{ideal}$ is assumed. 4) The first 50% of the service must be supplied within 15 s and 100% must be supplied within 30 s. The ideal response can be defined as a response with an instant 100% power ramp [85]. 5) The ideal and tolerated response of this service provision is plotted as $x_{ideal}$, $x_{acc,min}$ and $x_{acc,max}$ in Fig. E.6, which assumes that a reserve power set-point has already been established based on the values from step 2.

Fig. E.6 shows a simulation of primary regulation active power ramp $x_{act}$ for the time interval $[-5, 35]$ s. The service delivery performance index

and non-delivery verification index are $\eta^{AS} = 0.4257$ and $\epsilon^{AS} = 0.1392$, calculated using Eq. (E.13) and Eq. (E.17). The TSO must determine a threshold $\epsilon_{max}$, such that the service provider is penalized or the contract is terminated if $\epsilon^{AS} > \epsilon^{AS}_{max}$. It is not the scope of this work to asses a suitable value of $\epsilon^{AS}_{max}$.

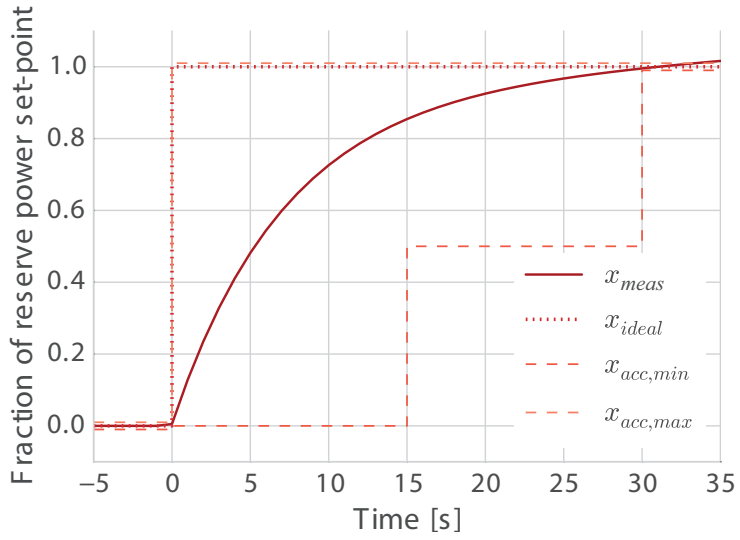



Figure E.6: Simulation of a DK1 primary reserve power ramp response together with $x_{ideal}$ and $x_{acc}$ values.

## E.5.2 PowerMax in a distribution system

The *PowerMax* service was first described in [30] and further specified in [12]. It is a DSO service, where the DSO can make a tender for a load reduction $\Delta P^{DSO}$ to a max level $P^{DSO}_{max}$ in parts of the distribution system that are forecasted to experience congestions during some periods (e.g. hours 17-20 during winter months). The motivation for *PowerMax* is that the service could be an economically beneficial alternative to grid reinforcements in some situations. This is both due to saved interest and depreciation on investments plus the avoided risk of over-sizing equipment in case of future energy savings or if the disappearance of a large consumer makes the reinforcement unnecessary.

In order to identify its service needs, it is assumed that the DSO is able to separate the total consumption forecast $\hat{P}_{tot}$ in the congested part of the distribution grid into a controllable load forecast $\hat{P}_{CL}$ and a base load forecast $\hat{P}_{BL}$:

$$\hat{P}_{tot} = \hat{P}_{CL} + \hat{P}_{BL} \tag{E.18}$$

$$\hat{P}_{CL} = \sum_{Agg} \hat{P}_{CL,Agg}, \quad Agg \in \mathbf{A} \tag{E.19}$$

where $\mathbf{A}$ is the set of all aggregators in the considered part of the grid. Only the aggregators *Agg* that bid for the service tender make up $\hat{P}_{CL}$, while the rest of $\mathbf{A}$ is part of $\hat{P}_{BL}$. The aggregators must be contracted to deliver a total power reduction $\Delta P$, such that the system operational limit $\bar{P}_{sys}$ is

not violated by the peak base load forecast and the peak controllable load forecast:

$$\hat{\bar{P}}_{BL} + \hat{\bar{P}}_{CL} - \Delta P \leq \bar{P}_{sys}. \tag{E.20}$$

This inequality can be fulfilled by setting a peak limit $\bar{P}_{CL}$:

$$\bar{P}_{CL} = \hat{\bar{P}}_{CL} - \Delta P \tag{E.21}$$

where $\Delta P$ and $\bar{P}_{CL}$ are the variables for the DSO service tender. In order to formulate a service tender, the magnitude of these variables must be estimated taking into account the uncertainty of the forecasts, giving the following expressions:

$$\Delta P^{DSO} = \sum_{Agg} \Delta \hat{P}_{CL,Agg} + \text{Risk}\{\hat{P}_{CL} + \hat{P}_{BL}\} \tag{E.22}$$

$$P^{DSO}_{max} = \hat{\bar{P}}_{CL} - \Delta P_{DSO} \tag{E.23}$$

where $\Delta \hat{P}_{CL,Agg}$ is the estimated power reduction for the individual aggregator bid, Risk$\{\hat{P}_{CL} + \hat{P}_{BL}\}$ is the risk associated to the load forecast uncertainty. $Agg \in \mathbf{A_C}$ and $\mathbf{A_C} \subseteq \mathbf{A}$, i.e. $\mathbf{A_C}$ is the subset of aggregators that bid on the tender. After the DSO has identified a suitable $P^{DSO}_{max}$ and $\Delta P^{DSO}$ to solve the congestion issue, the DSO formulates a service tender for which aggregators can bid their corresponding $\Delta P^{Agg}$ and $P^{Agg}_{max}$.

The method from Sec. E.3 is used to model *PowerMax* ideal and acceptable response. 1) The physical parameters are $P^{Agg}_{max}$, $\Delta P^{Agg}$ and months/-days/hours the service shall be delivered. 2) The system does not posses a dynamic behaviour related to system parameters. 3) As an example, the service tender defines $P^{Agg}_{max} = 200$ kW and $+1\%$ allowed deviation $P^{Agg}_{max,acc}$. 4) In this example we use 120 min service provision time with allowed non-delivery in the first 15 min, and the last 5 min, of the service delivery (following the service definition in [30]) and the ideal service delivery is the one that respects $P^{DSO}_{max}$. 5) Figure E.7 plots $x_{ideal}$ and $x_{acc}$. The *Activation Dead-band* indicates the regions where the aggregator is not obliged to deliver the service because of the tolerances defined under step 4. 6) The service is a maximum cap service and the error is measured as in Eq. (E.3).

An example of a load curve $P_{Agg} = x_{act}$ is presented in Fig. E.7. The service delivery and verification are evaluated using Eq. (E.13) and Eq. (E.17), yielding $\eta^{AS} = 0.5074$ and $\epsilon^{AS} = 0.2701$ respectively. As with the performance assessment of the FCR in DK1, it is not within the scope of this paper to asses the value of $\epsilon^{AS}_{max}$, yet a qualified assessment can be made. To asses $\epsilon^{AS}_{max}$, the DSO must analyze the dynamics of the problem the service is helping relieve. For *PowerMax*, the dynamics are governed by the heating of the overloaded equipment (e.g. transformer or cable), which deteriorates over time due to overheating. A feeder might be tolerant to short term overloads and therefore the DSO might set $\epsilon^{AS}_{max}$ higher than in the FCR case.

## E.6 Conclusion

This paper presents a method for modeling ancillary service requirements. These models can be used for evaluating the performance of a service pro-

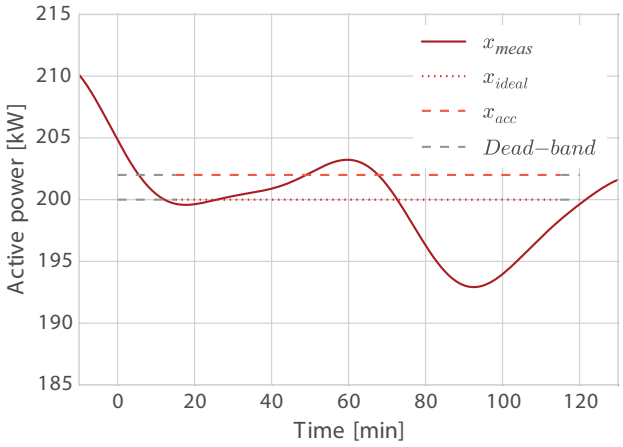



Figure E.7: $x_{ideal} = P_{max}^{Agg}, x_{acc} = P_{max,acc}^{Agg}$ for the considered *PowerMax* example. The activation Dead-band is the time period, where the aggregator is allowed to non-deliver.

vision. The performance is assessed by means of a service performance assessment index and a service verification index. The use of the modeling method and the indices are illustrated with two case studies covering a traditional ancillary service and a new distribution system service. The main purpose behind the development of the modeling method and the indices is to expand the current service verification methods to be suitable for aggregated demand response. The performance assessment of aggregators in terms of the services they are to provide is an important element in integrating new sources of ancillary services in the power system. These new sources are expected to play an important part in the security of the future power system.

## Acknowledgment

Parts of this work are supported by the Programme for Energy Technology Development and Demonstration (EUDP) through PowerLabDK and Innovation Fund Denmark through the iPower project. The authors thank Antonio Zecchino and Henrik W. Bindner for reviewing the draft of the paper.

# Bibliography


[1] KJ Åstrom. "Assessment of achievable performance of simple feedback loops". English. In: *International journal of adaptive control and signal processing* 5.1 (1991), pp. 3–19. ISSN: 08906327, 10991115.

[2] Peder Bacher and Henrik Madsen. "Identifying suitable models for the heat dynamics of buildings". In: *Energy and Buildings* 43.7 (2011), pp. 1511–1522. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2011.02.005.

[3] B. Biegel, L. H. Hansen, P. Andersen, and J. Stoustrup. "Primary Control by ON/OFF Demand-Side Devices". In: *IEEE Transactions on Smart Grid* 4.4 (Dec. 2013), pp. 2061–2071. ISSN: 1949-3053. DOI: 10.1109/TSG.2013.2257892.

[4] Benjamin Biegel, Lars Henrik Hansen, Peder Cajar, and Anders Birke. *FLECH PowerCap Service Requirement Specification*. Tech. rep. 2014.

[5] Benjamin Biegel, Jakob Stoustrup, Jan Bendtsen, and Palle Andersen. "Model Predictive Control for Power Flows in Networks with Limited Capacity". In: *2012 American Control Conference*. 2012, pp. 2959–2964.

[6] David H. Blum. "Ancillary Service Demand Response Strategies for Variable Air Volume HVAC Systems". MA thesis. Massachusetts Institute of Technology, 2013.

[7] Josh L Bode, Michael J Sullivan, Dries Berghman, and Joseph H Eto. "Incorporating residential AC load control into ancillary service markets: Measurement and settlement". In: *Energy Policy* 56 (2013), pp. 175–185.

[8] Daniel Esteban Morales Bondy, Oliver Gehrke, Anders Thavlov, Kai Heussen, Anna M. Kosek, and Henrik W. Bindner. "Procedure for Validation of Aggregators Providing Demand Response". In: *Proceedings of the Nineteenth Power Systems Computation Conference (PSCC)*. 2016.

[9] Daniel Esteban Morales Bondy, Kai Heussen, Oliver Gehrke, and Anders Thavlov. "A Functional Reference Architecture for Aggregators". In: *Proceedings of the 20th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA 2015)* (2015).

[10] Daniel Esteban Morales Bondy, Jason MacDonald, Emre Kara, Oliver Gehrke, Kai Heussen, Sila Kiliccote, and Henrik W. Bindner. "Redefining Requirements of Ancillary Services for Technology Agnostic Sources". In: *Draft* (2016).

[11] Daniel Esteban Morales Bondy, Germán Tarnowski, Kai Heussen, and Hansen Lars Henrik Hansen. *Operational Scenario: Manual Regulating Power*. Tech. rep. from the Strategic Platform for Innovation and Research in Intelligent Power (iPower) project. DTU Elektro, 2013.

[12] Daniel Esteban Morales Bondy and Anders Thavlov. *FLECH PowerMax Service Requirement Specification*. Tech. rep. DTU CEE, 2014.

[13] Daniel Esteban Morales Bondy, Anders Thavlov, Janus B. M. Tougaard, and Kai Heussen. "Performance Requirements Modeling and Assessment for Ancillary Services in view of Demand Response". In: (2016). submitted to: IEEE Transactions on Smart Grids.

[14] D.E.M. Bondy, G.T. Costanzo, K. Heussen, and H.W. Bindner. "Performance assessment of aggregation control services for demand response". In: *Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2014 IEEE PES*. Oct. 2014, pp. 1–6. DOI: 10.1109/ISGTEurope.2014.7028779.

[15] Severin Borenstein. *Money for Nothing*. www.energypolicyblog.com/2014/05/14/money-for-nothing/. Blog. 2014.

[16] George Box, Søren Bisgaard, and Conrad Fung. "An explanation and critique of Taguchi's contributions to quality engineering". In: *Quality and reliability engineering international* 4.2 (1988), pp. 123–131.

[17] Gru Brundtland, Mansour Khalid, Susanna Agnelli, Sali Al-Athel, Bernard Chidzero, Lamina Fadika, Volker Hauff, Istvan Lang, Ma Shijun, Margarita Morino de Botero, et al. "Our Common Future (*Brundtland report*)". In: (1987).

[18] Matthias A. Bucher, Stefanos Delikaraoglou, Kai Heussen, Pierre Pinson, and Göran Andersson. "On Quantification of Flexibility in Power Systems". In: *Proceedings of PowerTech Eindhoven 2015*. IEEE, 2015. ISBN: 9781479976935. DOI: 10.1109/PTC.2015.7232514.

[19] M. Buscher, K. Piech, S. Lehnhoff, S. Rohjans, C. Steinbrink, J. Velasquez, F. Andren, and T. Strasser. "Towards Smart Grid system validation: Integrating the SmartEST and the SESA laboratories". In: *Industrial Electronics (ISIE), 2015 IEEE 24th International Symposium on*. June 2015, pp. 1144–1150. DOI: 10.1109/ISIE.2015.7281633.

[20] Duncan S. Callaway and Ian A. Hiskens. "Achieving Controllability of Electric Loads". und. In: *Proceedings of the IEEE* 99.1 (2011), pp. 184–199. ISSN: 15582256, 00189219. DOI: 10.1109/jproc.2010.2081652.

[21] Peter Cappers, Jason MacDonald, Charles Goldman, and Ookie Ma. "An assessment of market and policy barriers for demand response providing ancillary services in US electricity markets". In: *Energy Policy* 62 (2013), pp. 1031–1039.

[22] Marco Cavazzuti. *Optimization Methods: From Theory to Design Scientific and Technological Aspects in Mechanics*. Springer Science & Business Media, 2012.

[23] Smart Grid Coordination Group CEN-CENELEC-ETSI. *Smart grid reference architecture*. Tech. rep. 2012.

[24] Spyros Chatzivasileiadis, Marco Bonvini, Javier Matanza, Rongxin Yin, Zhenhua Liu, Thierry Nouidui, Emre C Kara, Rajiv Parmar, David Lorenzetti, Michael Wetter, et al. "Cyber physical modeling of distributed resources for distribution system operations". In: *arXiv preprint arXiv:1505.00078* (2015).

[25] Meng Cheng, Jianzhong Wu, S. Galsworthy, N. Jenkins, and W. Hung. "Availability of load to provide frequency response in the great Britain power system". In: *Power Systems Computation Conference (PSCC), 2014*. Aug. 2014, pp. 1–7. DOI: 10.1109/PSCC.2014.7038294.

[26] Aleh Cherp and Jessica Jewell. "The concept of energy security: Beyond the four As". In: *Energy Policy* 75 (2014), pp. 415–421.

[27] Robert Cloutier, Gerrit Muller, Dinesh Verma, Roshanak Nilchiani, Eirik Hole, and Mary Bone. "The Concept of Reference Architectures". In: *Systems Engineering* 13.1 (2010), pp. 14–27. ISSN: 1520-6858. DOI: 10.1002/sys.20129. URL: http://dx.doi.org/10.1002/sys.20129.

[28] Giuseppe Tommaso Costanzo, Oliver Gehrke, Daniel Esteban Morales Bondy, Fabrizio Sossan, Henrik W. Bindner, Jacopo Parvizi, and Henrik Madsen. "A Coordination Scheme for Distributed Model Predictive Control: Integration of Flexible DERs". In: *Proc. 4th IEEE PES Int. Conf. Innovative Smart Grid Technologies Europe*. Copenhagen, Denmark: IEEE, Oct. 2013.

[29] Danish Ministry of Climate, Energy and Building. *Smart Grid Strategy - The intelligent system of the future*. http://www.ens.dk/sites/ens.dk/files/undergrund-forsyning/el-naturgas-varmeforsyning/Smart-grid/smartgridstrategy_uk_web.pdf. May 2013.

[30] Yi Ding, Lars Henrik Hansen, Peder Dybdal Cajar, Poul Brath, Henrik W Bindner, Chunyu Zhang, and Niels Christian Nordentoft. *Development of a DSO-market on flexibility services*. Tech. rep. 2013.

[31] Kornschnok Dittawit and Finn Arve Aagesen. "Demand side focused simulation platform for the evaluation of demand side management approaches". In: *Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2014 IEEE PES*. IEEE. 2014, pp. 1–6.

[32] *DONG Energy's PowerHub*. URL: https://www.powerhub.dk.

[33] Energinet.dk. *Ancillary services to be delivered in Denmark - Tender Conditions*. Tech. rep. 2012.

[34] Energinet.dk. *Ancillary services to be delivered in Denmark - Tender conditions*. On-line. Oct. 2012. (Visited on 11/01/2013).

[35] Energinet.dk. *Teknisk Forskrift 5.8.1 - Måledata til systemdriftsformål*. On-line. June 201l.

[36] Trine Madsen Engel. *Optimal utilization of the electricity distribution network*. Tech. rep. in danish. RisøDTU, 2011.

[37] ENTSO-E. *Continental Europe Operation Handbook - G Glossary*. Tech. rep. European network of transmission system operators for electricity, 2014.

[38] ENTSO-E. *Continental Europe Operation Handbook - P1 Load-Frequency and Control Performance*. Tech. rep. European network of transmission system operators for electricity, 2014.

[39] ENTSO-E, UCTE. "Operational Handbook - P1-B". In: *Policy* (2009).

[40] Thomas Erl. *The SOA Magazine Service-Orientation and Object-Orientation Part I: A Comparison of Goals and Concepts*. 2008.

[41] Eto, Joseph H, Undrill, John, Mackin, Peter, Daschmans, Ron, Williams, Ben, Haney, Brian, Hunt, Randall, Ellis, Jeff, Illian, Howard, Martinez, Carlos, O'Malley, Mark, Coughlin, Katie, and LaCommare, Kristina H. "Use of frequency response metrics to assess the planning and operating requirements for reliable integration of variable renewable generation". In: *Lawrence Berkeley National Laboratory* (2010).

[42] EURELECTRIC. *A sector in transformation: electricity industry trends and figures*. Tech. rep. Jan. 2015.

[43] Eurelectric. *Electrification of Heating and Cooling*. 2011.

[44] European network of transmission system operators for electricity (ENTSO-E). *Network Code on Load-Frequency Control and Reserves*. Tech. rep. 2013.

[45] Federal Energy Regulatory Commission. *Order 755: Frequency regulation compensation in organized wholesale power markets*. Oct. 2011.

[46] FENIX. *Flexible Electricity Networks to Integrate the expected Energy Evolution*. Tech. rep. FENIX, 2009.

[47] Forward Market Operations. *PJM Manual 11: Energy & Ancillary Services Market Operations*. Tech. rep. PJM, 2015.

[48] Lazaros Gkatzikis, Iordanis Koutsopoulos, and Theodoros Salonidis. "The Role of Aggregators in Smart Grid Demand Response Markets". English. In: *IEEE Journal on Selected Areas in Communications* 31.7 (2013), pp. 1247–1257. ISSN: 07338716, 15580008. DOI: 10.1109/JSAC.2013.130708.

[49] Torben Green, Roozbeh Izadi-Zamanabadi, and Hans Henrik Niemann. "On the choice of performance assessment criteria and their impact on the overall system performance - The refrigeration case study". In: *Proc. IEEE Conf. Control and Fault Tolerant Systems (SysTol'10)*. Niece, France, 2010, pp. 624–629.

[50] George Gross and Jeong Woo Lee. "Analysis of Load Frequency Control Performance Assessment Criteria". In: *Power Systems, IEEE Transactions on* 16.3 (2001), pp. 520–525.

[51] A. Haines, R.S. Kovats, D. Campbell-Lendrum, and C. Corvalan. "Climate change and human health: Impacts, vulnerability and public health". In: *Public Health* 120.7 (2006), pp. 585–596. ISSN: 0033-3506. DOI: http://dx.doi.org/10.1016/j.puhe.2006.01.002. URL: http://www.sciencedirect.com/science/article/pii/S0033350606000059.

[52] Rasmus Halvgaard, Niels Kjølstad Poulsen, Henrik Madsen, and John Bagterp Jørgensen. "Economic Model Predictive Control for Building Climate Control in a Smart Grid". In: *Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES*. IEEE. 2012, pp. 1–6.

[53] Rasmus Halvgaard, Peter Tychsen, Thomas Munk-Nielsen, Morten Grum, and Henrik Madsen. "Waste Water Treatment Plants and the Smart Grid". In: *Abstract Book-Dtu Sustain Conference 2014*. Technical University of Denmark. 2014.

[54] Sekyung Han, Soohee Han, and K. Sezaki. "Development of an Optimal Vehicle-to-Grid Aggregator for Frequency Regulation". In: *Smart Grid, IEEE Transactions on* 1.1 (June 2010), pp. 65–72. ISSN: 1949-3053. DOI: 10.1109/TSG.2010.2045163.

[55] Xue Han, Henrik W Bindner, Henrik Hansen, Kåre Seest Rasmussen, Peder Dybdahl Cajar, Yi Ding, and David Victor Tackie. *Identification of kick-back from normal situations due to activation of flexible demand*. Tech. rep. iPower Consortium, 2014.

[56] Xue Han, Kai Heussen, Oliver Gehrke, Henrik W. Bindner, and Benjamin Kroposki. "A review on Distributed Control Strategies for Distributed Energy Resource Coordination". In: (). Submitted to IEEE Trans. on Smart Grids.

[57] Xue Han, Anna Magdalena Kosek, Daniel Esteban Morales Bondy, Henrik W Bindner, Shi You, David Victor Tackie, Jasmin Mehmedalic, and Fannar Thordarson. "Assessment of distribution grid voltage control strategies in view of deployment". In: *Intelligent Energy Systems (IWIES), 2014 IEEE International Workshop on*. IEEE. 2014, pp. 46–51.

[58] Jacob Hansen, Jesper Knudsen, and Anuradha M Annaswamy. "Demand response in smart grids: Participants, challenges, and a taxonomy". In: *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*. IEEE. 2014, pp. 4045–4052.

[59] Lars Henrik Hansen, Olle Sundström, Silas Harbo, and Rasmus Villefrance. *FLECH – Technical requirement specification*. Tech. rep. 2013.

[60] TJ Harris. "Assessment of control loop performance". English. In: *Canadian Journal Of Chemical Engineering* 67.5 (1989), pp. 856–861. ISSN: 00084034, 1939019x.

[61] Kai Heussen, Daniel Esteban Morales Bondy, Junjie Hu, Oliver Gehrke, and Lars Henrik Hansen. "A Clearinghouse Concept for Distribution-Level Flexibility Services". eng. In: *Proceedings of IEEE Innovative Grid Technologies Europe 2013* (2013), p. 6695483. DOI: 10.1109/ISGTEurope.2013.6695483.

[62] Kai Heussen, Daniel Esteban Morales Bondy, Junjie Hu, Oliver Gehrke, and Lars Henrik Hansen. "A clearinghouse concept for distribution-level flexibility services". In: *Innovative Smart Grid Technologies Europe (ISGT EUROPE), 2013 4th IEEE/PES*. IEEE. 2013, pp. 1–5.

[63] Kai Heussen, Shi You, Benjamin Biegel, Lars Henrik Hansen, and Katrine Bech Andersen. "Indirect control for demand side management - A conceptual introduction". In: *Proc. 3rd IEEE PES Int. Conf. Innovative Smart Grid Technologies Europe*. Berlin, Germany, Oct. 2012.

[64] Junjie Hu, Shi You, Morten Lind, and Jacob Østergaard. "Coordinated Charging of Electric Vehicles for Congestion Prevention in the Distribution Grid". In: *IEEE Transactions on Smart Grid* (2014).

[65] International Energy Agency. *Energy Technology Perspectives 2012*. 2012. DOI: 10.1787/energy_tech-2012-en.

[66] *iPower FLECH Demonstration*. Dec. 2014. URL: https://www.youtube.com/watch?v=skgmbUqjBiw.

[67] Mohieddine Jelali. "An overview of control performance assessment technology and industrial applications". In: *Control Engineering Practice* 14.5 (2006). Intelligent Control Systems and Signal Processing ICONS 2003, IFAC International Conference on Intelligent Control Systems and Signal Processing, pp. 441–466. ISSN: 0967-0661. DOI: http://dx.doi.org/10.1016/j.conengprac.2005.11.005.

[68] E.C. Kara, M. Berges, B. Krogh, and S. Kar. "Using smart devices for system-level management and control in the smart grid: A reinforcement learning framework". In: *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*. Nov. 2012, pp. 85–90. DOI: 10.1109/SmartGridComm.2012.6485964.

[69] Emre C Kara, Jason S Macdonald, Douglas Black, Mario Bérges, Gabriela Hug, and Sila Kiliccote. "Estimating the benefits of electric vehicle smart charging at non-residential locations: A data-driven approach". In: *Applied Energy* 155 (2015), pp. 515–525.

[70] Sila Kiliccote. "Field testing of automated demand response for integration of renewable resources in California's ancillary services market for regulation products". In: (2013).

[71] Sila Kiliccote. "Open automated demand response communications in demand response for wholesale ancillary services". In: *Lawrence Berkeley National Laboratory* (2010).

[72] B.J. Kirby. "Load Response Fundamentally Matches Power System Reliability Requirements". In: *Power Engineering Society General Meeting, 2007. IEEE*. June 2007, pp. 1–6. DOI: 10.1109/PES.2007.386227.

[73] Stephan Koch, Johanna L Mathieu, and Duncan S Callaway. "Modeling and control of aggregated heterogeneous thermostatically controlled loads for ancillary services". In: *Proc. PSCC*. 2011, pp. 1–7.

[74] J Koen Kok, Cor J Warmer, and IG Kamphuis. "PowerMatcher: multiagent control in the electricity infrastructure". In: *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*. ACM. 2005, pp. 75–82.

[75] Elta Koliou, Cherrelle Eid, José Pablo Chaves-Ávila, and Rudi A. Hakvoort. "Demand response in liberalized electricity markets: Analysis of aggregated load participation in the German balancing mechanism". In: *Energy* 71 (2014), pp. 245–254. ISSN: 0360-5442. DOI: http://dx.doi.org/10.1016/j.energy.2014.04.067. URL: http://www.sciencedirect.com/science/article/pii/S0360544214004800.

[76] Anna Kosek, Giuseppe Tommaso Costanzo, Henrik Bindner, and Oliver Gehrke. "An overview of demand side management control schemes for buildings in smart grids". In: *Proc. IEEE Int. Conf. Smart Energy Grid Eng. (SEGE'13)*. Oshawa, Canada, Aug. 2013.

[77] N. Künzli, R. Kaiser, S. Medina, M. Studnicka, O. Chanel, P. Filliger, M. Herry, F. Horak Jr., V. Puybonnieux-Texier, P. Quénel, J. Schneider, R. Seethaler, J-C. Vergnaud, and H. Sommer. "Public-health impact of outdoor and traffic-related air pollution: a European assessment". In: *The Lancet* 356.9232 (2000), pp. 795–801. ISSN: 0140-6736. DOI: http://dx.doi.org/10.1016/S0140-6736(00)02653-2. URL: http://www.sciencedirect.com/science/article/pii/S0140673600026532.

[78] Marion Labatut, Pavla Mandatova, and Charlotte Renaud. *Designing fair and equitable market rules for demand response aggregation*. Tech. rep. EURELECTRIC, 2015. URL: http://www.eurelectric.org/media/169872/0310_missing_links_paper_final_ml-2015-030-0155-01-e.pdf.

[79] Kristina Hamachi LaCommare and Joseph H. Eto. "Cost of power interruptions to electricity consumers in the United States (US)". In: *Energy* 31.12 (2006), pp. 1845–1855. ISSN: 0360-5442. DOI: http://dx.doi.org/10.1016/j.energy.2006.02.008. URL: http://www.sciencedirect.com/science/article/pii/S036054420600051X.

[80] Venkatachalam Lakshmanan, Kristian Gudmand-Høyer, Mattia Marinelli, Anna Magdalena Kosek, and Per Bromand Nørgård. "Energy shift estimation of demand response activation on domestic refrigerators – A field test study". In: *Proceedings of the 49th International Universities Power Engineering Conference (UPEC 2014)*. IEEE, 2014. ISBN: 978-1-4799-6557-1. DOI: 10.1109/UPEC.2014.6934681.

[81] I. Lampropoulos, W. L. Kling, P. F. Ribeiro, and J. van den Berg. "History of demand side management and classification of demand response control schemes". In: *Power and Energy Society General Meeting (PES), 2013 IEEE*. July 2013, pp. 1–5. DOI: 10.1109/PESMG.2013.6672715.

[82] N. Leemput, F. Geth, B. Claessens, J. Van Roy, R. Ponnette, and J. Driesen. "A case study of coordinated electric vehicle charging for peak shaving on a low voltage grid". In: *Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on*. Oct. 2012, pp. 1–7. DOI: 10.1109/ISGTEurope.2012.6465656.

[83] J. Lelieveld, J. S. Evans, M. Fnais, D. Giannadaki, and A. Pozzer. "The contribution of outdoor air pollution sources to premature mortality on a global scale". English. In: *Nature (London)* 525.7569 (2015), pp. 367–371. ISSN: 00280836, 14764687. DOI: 10.1038/nature15371.

[84] Jason MacDonald, Peter Cappers, Duncan S Callaway, and Sila Kiliccote. "Demand Response Providing Ancillary Services: A Comparison of Opportunities and Challenges in the US Wholesale Markets". In: *Grid Interop* (2012).

[85] Yuri V Makarov, Shuai Lu, Jian Ma, Tony B Nguyen, California Energy Commission, et al. *Assessing the value of regulation resources based on their time response characteristics*. Tech. rep. 2008.

[86] F. Marra, Guang Ya Yang, C. Traholt, E. Larsen, C.N. Rasmussen, and Shi You. "Demand profile study of battery electric vehicle under different charging options". In: *Power and Energy Society General Meeting, 2012 IEEE*. 2012, pp. 1–7. DOI: 10.1109/PESGM.2012.6345063.

[87] Johanna L. Mathieu, Mark Dyson, and Duncan S. Callaway. "Using Residential Electric Loads for Fast Demand Response: The Potential Resource and Revenues, the Costs, and Policy Recommendations". In: *ACEEE Summer Study on Energy Efficiency in Buildings* (2012).

[88] A. Molina-Garcia, F. Bouffard, and D.S. Kirschen. "Decentralized Demand-Side Contribution to Primary Frequency Control". In: *Power Systems, IEEE Transactions on* 26.1 (Feb. 2011), pp. 411–419. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2010.2048223.

[89] NERC Resources Subcommittee. *Balancing and frequency control - A Technical Document*. Tech. rep. North American Reliability Corporation, 2011.

[90] Astrid Nieße, Sebastian Beer, Jorg Bremer, Christian Hinrichs, Ontje Lunsdorf, and Michael Sonnenschein. "Conjoint dynamic aggregation and scheduling methods for dynamic virtual power plants". In: *Computer Science and Information Systems (FedCSIS), 2014 Federated Conference on*. IEEE. 2014, pp. 1505–1514.

[91] *NIST/SEMATECH e-Handbook of Statistical Methods*. 2016. URL: http://www.itl.nist.gov/div898/handbook/.

[92] Niels Christian Nordentoft, Chunyu Zhang, Yi Ding, Lars Henrik Hansen, Peder Cjar, Henrik Bindner, and Poul Brath. *Development of a DSO-market on Flexibility Services*. Tech. rep. iPower, Mar. 2013.

[93] North American Electric Reliability Corporation (NERC). *Definition of "Adequate Level of Reliability"*. Tech. rep. 2007. URL: http://www.nerc.com/docs/pc/Definition-of-ALR-approved-at-Dec-07-OC-PC-mtgs.pdf.

[94] Mark Norton, Hans Vanderbroucke, Emil Larsen, Craig Dyke, Susana Benares, Chloe Latour, and Thong Vu Van. *Demand Side Response Policy Paper*. Tech. rep. European network of transmission system operators for electricity (ENTSO-E), 2014.

[95] NYS Department of Public Service. *Reforming the Energy Vision*. Tech. rep. 2014.

[96] Niamh O'Connell, Pierre Pinson, Henrik Madsen, and Mark O□Malley. "Benefits and challenges of electrical demand response: A critical review". In: *Renewable and Sustainable Energy Reviews* 39 (2014), pp. 686–699. ISSN: 1364-0321. DOI: 10.1016/j.rser.2014.07.098.

[97] Gary W Oehlert. *A first course in design and analysis of experiments*. 2010. ISBN: 0-7167-3510-5.

[98] Michael Ølund and Lars Henrik Hansen. *FLECH TSO Service - Fast Frequency Reserve*. Tech. rep. iPower Consortium, 2014.

[99] *Open Energi*. URL: http://www.openenergi.com.

[100] Martin J Pasqualetti and Benjamin K Sovacool. "The importance of scale to energy security". In: *Journal of Integrative Environmental Sciences* 9.3 (2012), pp. 167–180.

[101] M. K. Petersen, Kristian Edlund, Lars Henrik Hansen, Jan Bendtsen, and Jakob Stoustrup. "A taxonomy for modeling flexibility and a computationally efficient algorithm for dispatch in smart grids". In: *American Control Conference (ACC), 2013*. IEEE. 2013, pp. 1150–1156.

[102] Mansoureh Peydayesh and Ross Baldick. "The effects of very fasty response to frequency fluctuation". In: *International Association for Energy Economics* (2012).

[103] Joseph J Pignatiello JR and John S Ramberg. "Top ten triumphs and tragedies of Genichi Taguchi". In: *Quality Engineering* 4.2 (1991), pp. 211–225.

[104] Chris Pilong. *PJM Manual 12: Balancing Operations*. Tech. rep. PJM, 2015.

[105] PJM. *Regulation Performance Evaluation Process*. URL: https://www.pjm.com/~/media/committees-groups/task-forces/rpstf/20110810/20110810-item-03-regulation-performance-evaluation-process.ashx.

[106] Guru Prasath and John Bagterp Jørgensen. "Soft Constraints for Robust MPC of Uncertain Systems". English. In: *International Symposium on Advanced Control of Chemical Processes* (2009), p. 54.

[107] D. Pudjianto, C. Ramsay, and G. Strbac. "Virtual Power Plant and System Integration of Distributed Energy Resources". In: *Renewable Power Generation, IET* 1.1 (Mar. 2007), pp. 10–16. ISSN: 1752-1416. DOI: 10.1049/iet-rpg:20060023.

[108] S. Rahnama, J. Stoustrup, and H. Rasmussen. "Integration of heterogeneous industrial consumers to provide regulating power to the smart grid". In: *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*. Dec. 2013, pp. 6268–6273. DOI: 10.1109/CDC.2013.6760880.

[109] Samira Rahnama, S Ehsan Shafiei, Jakob Stoustrup, Henrik Rasmussen, and Jan Bendtsen. "Evaluation of Aggregators for Integration of Large-scale Consumers in Smart Grid". In: *Proc. 19th IFAC Word Congress, Cape Town, South Africa*. 2014.

[110] Yann Rebours. "A Comprehensive Assessement of Markets for Frequency and Voltage Control Ancillary Services". PhD thesis. University of Manchester, School of Electrical and Electronic Engineering, 2008.

[111] Francis M. Rubinstein and Sila Kiliccote. "Demand Responsive Lighting: A Scoping Study". In: (2007).

[112] Steffen Schütte, Stefan Scherfke, and Michael Sonnenschein. "mosaik-smart grid simulation API". In: *Proceedings of SMARTGREENS* (2012), pp. 14–24.

[113] Sigurd Skogestad and Ian Postlethwaite. *Multivariable Feedback Control- Analysis and Design*. 2nd ed. Chichester, England: Wiley, 2005.

[114] Smart Energy Demand Coalition (SEDC). "Mapping Demand Response in Europe Today". In: *Tracking Compliance with Article* 15 (2014).

[115] Smart Grid Coordination Group. *SGAM User Manual - Applying, testing & refining the Smart Grid Architecture Model (SGAM)*. CEN-CENELEC-ETSI. 2014.

[116] E. Sortomme and M.A. El-Sharkawi. "Optimal Charging Strategies for Unidirectional Vehicle-to-Grid". In: *Smart Grid, IEEE Transactions on* 2.1 (Mar. 2011), pp. 131–138. ISSN: 1949-3053. DOI: 10.1109/TSG.2010.2090910.

[117] Fabrizio Sossan, Henrik W. Bindner, and Per Bromand Nørgård. "Indirect control of flexible demand for power system applications." PhD thesis. 2014.

[118] C. Steinbrink and S. Lehnhoff. "Challenges and necessity of systematic uncertainty quantification in smart grid co-simulation". In: *EUROCON 2015 - International Conference on Computer as a Tool (EUROCON), IEEE*. Sept. 2015, pp. 1–6. DOI: 10.1109/EUROCON.2015.7313679.

[119] Nicholas Herbert Stern, Her Majesty's Treasury, et al. *Stern Review: The economics of climate change*. Vol. 30. HM treasury London, 2006.

[120] *IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA, 2013, 1535 pp.

[121] Wencong Su and Mo-Yuen Chow. "Performance evaluation of an EDA-based large-scale plug-in hybrid electric vehicle charging algorithm". In: *Smart Grid, IEEE Transactions on* 3.1 (2012), pp. 308–315.

[122] Michael Sullivan, Josh Bode, Bashar Kellow, Sarah Woehleke, and Joe Eto. "Using residential AC load control in grid operations: PG&E's ancillary service pilot". In: *Smart Grid, IEEE Transactions on* 4.2 (2013), pp. 1162–1170.

[123] System Operations Division Markets Division. *PJM Manual 40: Training and Certification Requirements*. Tech. rep. PJM, 2015.

[124] Genʊichi Taguchi and Yuin Wu. *Introduction to off-line quality control*. Central Japan Quality Control Assoc., 1979.

[125] A. Thavlov and H.W. Bindner. "An aggregation model for households connected in the low-voltage grid using a VPP interface". In: *Innovative Smart Grid Technologies Europe (ISGT EUROPE), 2013 4th IEEE/PES*. Oct. 2013, pp. 1–5. DOI: 10.1109/ISGTEurope.2013.6695360.

[126] Anders Thavlov. "Power System Integration of Flexible Demand in the Low Voltage Network". PhD thesis. 2015.

[127] Anders Thavlov and Henrik W. Bindner. "Utilization of Flexible Demand in a Virtual Power Plant Set-Up". In: *IEEE Transactions on Smart Grid* (2014). ISSN: 1949-3053. DOI: 10.1109/TSG.2014.2363498.

[128] The Danish Government. *The Danish Climate Policy Plan - Towards a low carbon society*. http://www.ens.dk/sites/ens.dk/files/policy/danish-climate-energy-policy/danishclimatepolicyplan_uk.pdf. Aug. 2013.

[129] The Open Group. *Service-Oriented Architecture (SOA) - A white paper*. 2007. URL: https://www2.opengroup.org/ogsys/catalog/w074.

[130] Thermal Working Group. *Ancillary Services, Unbundling Electricity Products - an Emerging Market*. Tech. rep. Union of the Electricity Industry - EURELECTRIC, 2004.

[131] Janus B. M. Tougaard, Lars Henrik Hansen, and Olle Sundström. *FLECH BRP Service - Portfolio Balancing*. Tech. rep. iPower Consortium, 2015.

[132] US Federal Energy Regulatory Commission. *Order No. 888-A, Appendix B*. Online:http://www.ferc.gov/legal/maj-ord-reg/land-docs/rm95-8p4-000.txt. Mar. 1997.

[133] USEF foundation. *USEF: The Framework Explained*. Tech. rep. 2015.

[134] *Using Distributed Energy Resources: A How-To Guide for Federal Facility Managers*. Tech. rep. National Renewable Energy Laboratory, 2002. URL: http://www.nrel.gov/docs/fy02osti/31570.pdf.

[135] Evangelos Vrettos, Charalampos Ziras, and Goran Andersson. "Integrating large shares of heterogeneous thermal loads in power system frequency control". In: *PowerTech, 2015 IEEE Eindhoven*. IEEE. 2015, pp. 1–6.

[136] World Health Organization. *Ambient (outdoor) air quality and health - Fact sheet N°313*. URL: http://www.who.int/mediacentre/factsheets/fs313/en/.

[137] Athanasios Zarogiannis, Mattia Marinelli, Chresten Træholt, Katarina Knezovic, and Peter Bach Andersen. "A Dynamic Behaviour Analysis on the Frequency Control Capability of Electric Vehicles". In: *Proceedings of International Universities' Power Engineering Conference (UPEC) 2014*. IEEE, 2014. ISBN: 9781479965571. DOI: 10.1109/UPEC.2014.6934763.