



U.S. DEPARTMENT OF  
**ENERGY**

PNNL-20120

Prepared for the U.S. Department of Energy  
under Contract DE-AC05-76RL01830

# Low Probability Tail Event Analysis and Mitigation in the BPA Control Area

S Lu            CA McKinstry  
AJ Brothers   S Jin

October 2010



**Pacific Northwest**  
NATIONAL LABORATORY

*Proudly Operated by **Battelle** Since 1965*

## DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY

*operated by*

BATTELLE

*for the*

UNITED STATES DEPARTMENT OF ENERGY

*under Contract DE-AC05-76RL01830*

Printed in the United States of America

Available to DOE and DOE contractors from the

Office of Scientific and Technical Information,

P.O. Box 62, Oak Ridge, TN 37831-0062;

ph: (865) 576-8401

fax: (865) 576-5728

email: [reports@adonis.osti.gov](mailto:reports@adonis.osti.gov)

Available to the public from the National Technical Information Service,  
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161

ph: (800) 553-6847

fax: (703) 605-6900

email: [orders@ntis.fedworld.gov](mailto:orders@ntis.fedworld.gov)

online ordering: <http://www.ntis.gov/ordering.htm>



This document was printed on recycled paper.

(9/2003)

## Low Probability Tail Event Analysis and Mitigation in the BPA Control Area

S Lu, Ph.D.: Senior Research Engineer, Project Manager  
AJ Brothers, Ph.D.: Senior Research Scientist  
CA McKinstry: Senior Research Scientist  
S Jin, Ph.D.: Research Engineer

Advisor: YV Makarov, Ph.D.: Chief Scientist - Power Systems

October 2010

Prepared for  
U.S. Department of Energy  
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory  
Richland, Washington 99352

## SUMMARY

This report investigates the uncertainties influencing the operations of a power system and their contributions to so-called tail events (the rare occurrences of large system imbalance usually caused by a combination of uncertainties), especially under high penetration of wind energy in the system. A Bayesian network model has been developed to quantify the impact of these uncertainties on system imbalance. A framework is presented for a decision support tool, which can help system operators to better estimate the need in balancing reserves and to prepare for tail events.

There are many uncertainties in a power system affecting generation and load balance. Load variations, transmission outages, and generator startup failures are a few examples. With more intermittent resources integrated into the system, balancing the system becomes more challenging. This report analyzes what was involved in events resulting in large system imbalance (or tail events), using the Electric Reliability Council of Texas (ERCOT) reliability event on February 26, 2008 as the study case. As a widely reported event showing the impact of wind generation, analysis of this event can provide helpful insight to the sources of problems and actions that can improve system responses to tail events.

In seeking approaches helping BPA operators to deal with uncertainties, this report uses yearly and hourly distributions of balancing requirement and tail events, respectively, to establish an overview of potential system imbalances, and to identify those hours when problems are most likely to occur. For real-time prediction, a Bayesian network model is constructed to reflect the statistical relationships between system imbalance and forecast errors, generation control errors, and other influential factors. Reserve requirements and the probability of tail events can be derived from the model to determine the sufficiency of system balancing reserves.

Finally, the report presents a framework for a decision support tool to predict balancing resources needed in each operation hour, and suggest dispatch actions, such as acquiring more reserve, curtailing scheduled delivery, and reducing generation. Transmission impacts from suggested actions and the cost of actions or no actions can also be assessed. The tool can serve as a platform to integrate all balancing-related information (load forecast, wind power forecast, forecast errors, ramp prediction, etc) to provide real-time decision support.

## ACKNOWLEDGEMENT

The authors gratefully acknowledge the contributions of John Pease, BPA Project Manager, Technology Innovation Group, for his advice and help to the team, as well as the following BPA colleagues for sharing their experience and providing data for the study:

Fran Halpin, Manager, Duty (hydro) Scheduling, PGSD  
Bart McManus, Team Lead, Transmission Operation, TOT  
Cynthia Bowline, Mechanical Engineer (hydro planning) PGPO  
Michael Van Tress, Meteorologist, PGPW  
Susan Baker-Holden, Manager, Transmission Scheduling, TSS

The authors would like to thank Evan Jones and Ross Guttromson, Pacific Northwest National Laboratory, for providing the project team with valuable suggestions and their support to this work.

The authors would also like to express their thanks to Susan Arey, Pacific Northwest National Laboratory, for her highly effective administrative support on the project and editorial review to the report.

# CONTENTS

SUMMARY .....	iii
ACKNOWLEDGEMENT .....	iv
I. Introduction .....	7
II. Lessons Learned from ERCOT Feb 26, 2008 Event.....	9
2.1 Contributing Factors .....	9
2.2 Services and Actions Involved.....	9
2.3 What Has Helped Prominently?.....	10
2.4 What Should Be Improved?.....	10
2.4.1 Load Forecast.....	11
2.4.2 Wind Forecast .....	11
2.4.3 Prediction of Adequacy of Balancing Reserve .....	12
2.5 Summary of Lessons Learned.....	13
III. Analysis of Tail Events Distributions .....	14
3.1 Yearly Distribution of Tail Events.....	14
3.2 Hourly Distribution of Tail Events .....	15
IV. Real-time Prediction of Balancing Reserve Requirements and Tail Events .....	17
4.1 Bayesian Network Models .....	17
4.2 Building a Bayesian Network Model for the BPA Power System .....	17
4.3 Preprocessing the Data.....	19
4.4 Learning Probabilities from the Data.....	19
4.5 Generation of Forecast Data .....	21
4.6 Bayesian Network Model Output .....	22
4.7 Validation Studies.....	23
4.7.1 Comparison with Naïve Persistence Forecasts .....	23
4.7.2 Prediction of Tail Event Probability .....	25
4.8 Potential Improvements on the Bayesian Network Model .....	26
4.8.1 Improve Prediction Accuracy .....	26
4.8.2 Adapt the Prediction Model to Changed System Composition .....	27
V. Framework of a Decision Support Tool for Within-hour Balance of the System .....	28
VI. Conclusion .....	30
VII. References .....	31

## FIGURES

Figure 1. LaaR deployment during February 26, 2008 event [1] .....	10
Figure 2. 02/26/08 16:00 – 22:00 load and day-ahead load forecast (from ERCOT Operations Report [1]) .....	11
Figure 3. 02/26/08 15:00 – 21:00 total wind output, wind forecast in the 16:00 day-ahead resource plan [1] .....	12
Figure 4. 2007 BPA system regulation capacity requirement and tail event distribution	14
Figure 5. 2010 BPA system regulation capacity requirement and tail event distribution	15
Figure 6. Average regulation capacity shortage of all tail events: distribution corresponding to each hour during a day .....	16
Figure 7. Bayes net model for predicting system imbalance. ....	18
Figure 8. Screen shot of discretized temperature using GeNIe software.....	19
Figure 9. Screen shot of DBN for predicting system imbalance as implemented in GeNIe. .....	20
Figure 10. First order dynamic Bayes net used to predict system imbalance.....	21
Figure 11. Output generated by GeNIe for predicting system imbalance. ....	22
Figure 12. Prediction of system imbalance for the next hour. ....	23
Figure 13. Prediction of system imbalance for the second hour. ....	23
Figure 14. Example demonstrating method for converting BN model forecasts from posterior probability distribution on system imbalance to point-estimates. ....	24
Figure 15. MAE comparison between BN and NP models forecasting at one, two and three hours: positive system imbalance cases .....	25
Figure 16. MAE comparison between BN and NP models forecasting at one, two and three hours: negative system imbalance cases .....	25
Figure 17. Observed probability of tail events vs. predicted probability by the BN model: positive system imbalance cases. ....	26
Figure 18. Observed probability of tail events vs. predicted probability by the BN model: negative system imbalance cases. ....	26
Figure 19. Decision support tool for balancing operations based on Bayesian network model.....	29
Figure 20. Block diagram of the intelligent system model in the decision support tool. .	29

## I. Introduction

*Tail event* refers to the situation in a power system when the imbalance between pre-scheduled generation and actual load becomes very significant. Tail events can be caused by unfavorable combination of load and wind forecast errors, and unpredicted fast load and wind ramps. This type of events occurs infrequently and appears on the “tails” of the probability distribution of system imbalance; therefore, they are referred to as tail events.

*Balancing reserve* or *balancing capacity* refers to the flexible generation capacity used to balance the difference between hourly generation schedules, and actual load and control area interchange variations within the hour. Balancing capacity can be further separated into load following and regulation capacity, according to the time scale it is operated. Load following is the process to dispatch balancing resources every 5 to 15 minutes, while regulation process does it minute by minute. Tail events occur when balancing capacity is insufficient to match the difference between scheduled hourly generation and real-time needs.

With increasing penetration of variable renewable resources, such as wind and solar generation, tail events have been occurring more frequently, and magnitudes of imbalance in these events are getting larger. Maintaining sufficient balancing reserves, both upward and downward, becomes increasingly challenging.

One objective of this project is analyzing previously reported events, in which significant amount of wind generation was involved, to identify sources of problems and solutions. Section II analyzes such an event in the Electric Reliability Council of Texas (ERCOT) system. Analysis specifically on BPA system events was not conducted because of the lack of record.

To understand the potential severity of tail events in the BPA system, this report includes an analysis of yearly and hourly distributions of regulation requirements and the size of tail events. Yearly distributions show the occurring frequency versus MW level of system imbalance; hourly distributions show the regulation requirement and average MW level of tail events corresponding to each of the 24 hours of a day.

In order to procure sufficient balancing reserves and enforce procedures to cope with the tail events, it is important to develop approaches to predicting these events. A Bayesian network (BN) model is built to predict the distribution of balancing capacity requirement in real-time operation and the probability of tail events. Statistical relationships between system imbalance and forecast errors, generation schedule control errors, and other influential factors such as weather, temperature, wind speed, etc., are established using historical data. The proposed tool acts similarly to an experienced system operator estimating the difference between hourly generation schedule and actual generation needs, based on his/her experience.

The framework for a new decision support tool, based on the BN model, is developed in this project. The framework describes major components of the tool that predicts balancing reserve requirements and suggests dispatch actions. The tool can serve as a platform to integrate all



balancing-related information (load forecasts, wind power forecasts, forecast errors, ramp predictions, etc) to aid real-time decision-making.

## II. Lessons Learned from ERCOT Feb 26, 2008 Event

On February 26, 2008, the ERCOT system implemented step two of its emergency electric curtailment plan (EECP) during an evening load ramp due to the increasing imbalance between generation and load. The event was widely reported because wind generation forecast error contributed partially in the event. The event time line is listed in detail in ERCOT operations report [1]. Description of the EECP steps can be found in [2]. Main findings on this event that are related to the purpose of this project are summarized in the following.

### 2.1 Contributing Factors

There are three contributors to this event: *load forecast error* during a larger than expected evening load ramp-up from 17:00 to 19:00, *wind forecast error* as large as 1000 MW during a ramp-down of wind generation from 15:00 to 21:00, and unexpected *loss and unavailability of conventional generation*. The first two are major causes, with the third one adding complexity to the problem. Collectively these factors led to ERCOT calling on reserve capacity, including curtailments of “loads acting as a resource” (LaaR) – large industrial and commercial electricity users who have agreed to allow ERCOT to curtail their electricity supply in exchange for economic compensation.

### 2.2 Services and Actions Involved

Main system services and actions, involved in the ERCOT event, include the following (listed in the order of actual implementation):

- (1) Up balancing energy service
- (2) Regulation up service
- (3) Non-spinning reserve service
- (4) Dispatch of out of merit energy and capacity
- (5) Deployment of responsive reserve
- (6) Reducing distribution voltage to reduce load
- (7) Requesting emergency energy from other systems
- (8) Deployment of LaaR.

In ERCOT, each market participant is obligated to provide ancillary services (AS). Market participants may provide the AS themselves, or rely on the ERCOT to acquire the AS through a centralized auction. ERCOT operates the following day-ahead AS markets for the balancing between generation and load: balancing energy service, regulation down service (RgDn), regulation up services (RgUp), responsive reserves services (RRS), non-spinning reserve services (NSRS), replacement reserve services (RPRS) and out of merit energy and capacity (OOME and OOMC) [3].

Regulation down and regulation up services are used to control the power output of resources to maintain the frequency within predetermined limits. RRS is intended to help restore the frequency of the interconnected transmission system within the first few minutes of an event that causes a significant deviation from nominal frequency. RRS is similar to the “spinning reserve

service” in other electricity markets. NSRS is a service that is provided through utilization of the portion of off-line generation capacity capable of being synchronized and ramped to a specified output level within thirty minutes (or load that can be interrupted within thirty minutes and that is capable of running (or being interrupted) at a specified output level for at least one hour). NSRS may also be provided from unloaded on-line capacity that meets the above response requirements and that is not participating in any other activity, including ERCOT market, self-generation and other energy transaction. RPRS is a service that is procured from generation resources units planned to be off-line, but which can be brought on-line, and loads acting as a resource that are available for interruption. OOMC and OOME are used by ERCOT to provide for the availability of sufficient capacity so that balancing energy bids are available to solve capacity insufficiency, congestion, or other reliability needs.

### 2.3 What Has Helped Prominently?

Demand response implemented in the ERCOT system effectively stopped the fast ramp-up during the event. Before the deployment of LaaR, load was ramping up very quickly at an unpredicted rate of ~60MW/min (refer to Figure 2). After the deployment of LaaR, load almost flattened. With the deployment of LaaR, the event lasted less than two hours. No customers lost power involuntarily. Also, according to [1], the response of LaaRs to deployment signal was generally good. Only two loads failed to deploy within 10 minutes. It appears to be the deployment of LaaR which halted frequency decline and restored ERCOT to stable operation.

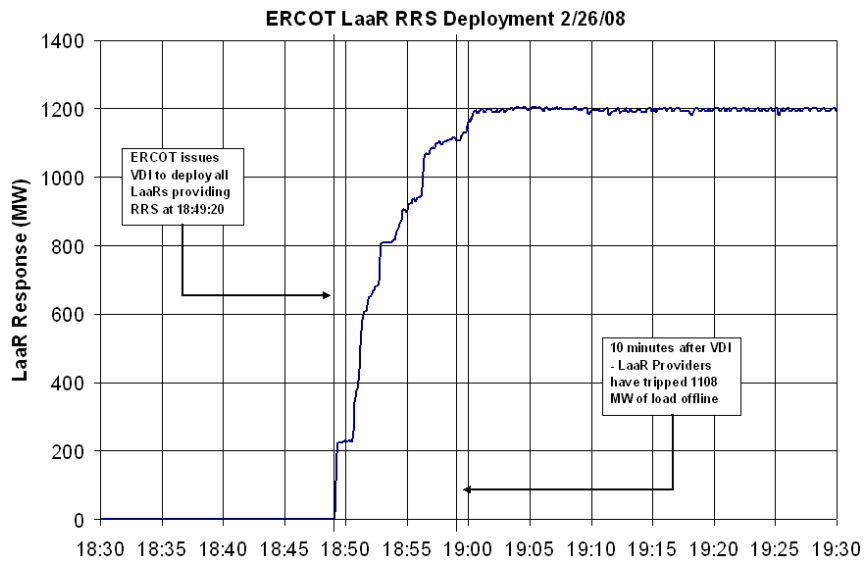


Figure 1. LaaR deployment during February 26, 2008 event [1]

### 2.4 What Should Be Improved?

ERCOT relies on resource plans provided by qualified scheduling entities (QSE’s) for its assessment of available capacity. The *unpredicted large load ramp* and *wind forecast error* were two major contributors to the event. The combination of these two factors made the procured balancing reserve appear insufficient.

### 2.4.1 Load Forecast

As shown in Figure 2, the day-ahead load forecasts in hour ending at 1800 and hour ending at 1900 were both higher than the actual load. However, the ramp rate predicted in the day-ahead forecast was ~1750 MW/hour (35650 MW at 1900 minus 33900 MW at 1800), while the actual load between 18:00 and 18:41 grew from ~33000 MW to ~35550 MW or an increase of ~2550 MW in 40 minutes. In real-time operations, the generation schedule might have been adjusted downward to meet the actual load during hours before 1800. Then the combination of unpredicted high load ramp rate during the hour ending at 1900 and the shortage of 900 MW from wind caused difficulties in balancing the system. Associated with the underestimation of combined load and wind ramp rate, ERCOT seemed not to have obtained sufficient balancing reserve for that period. During this period ERCOT exhausted its available regulation up service (~1000 MW) and up balancing energy service.

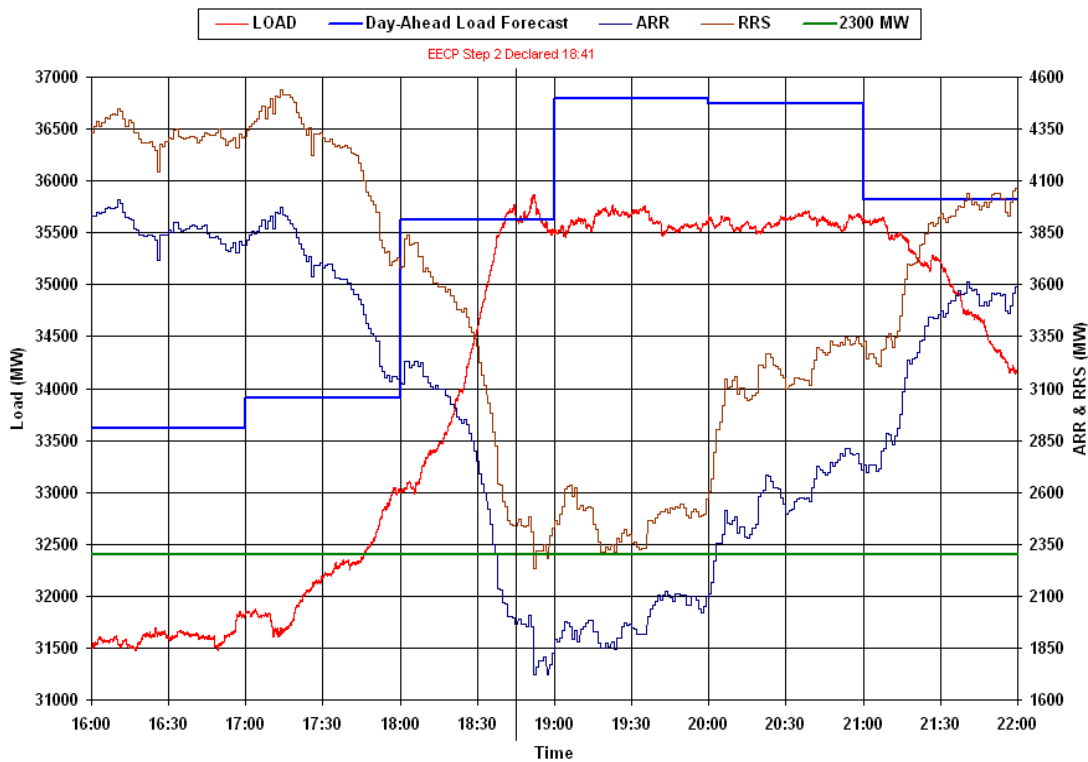


Figure 2. 02/26/08 16:00 – 22:00 load and day-ahead load forecast (from ERCOT Operations Report [1])

### 2.4.2 Wind Forecast

The ERCOT day-ahead resource plan was not able to forecast the magnitude of the drop in wind. The wind power was about 1300 MW in forecast, while it was 400MW during the operating hour. This resulted in a 900 MW shortage of generation from wind resources.

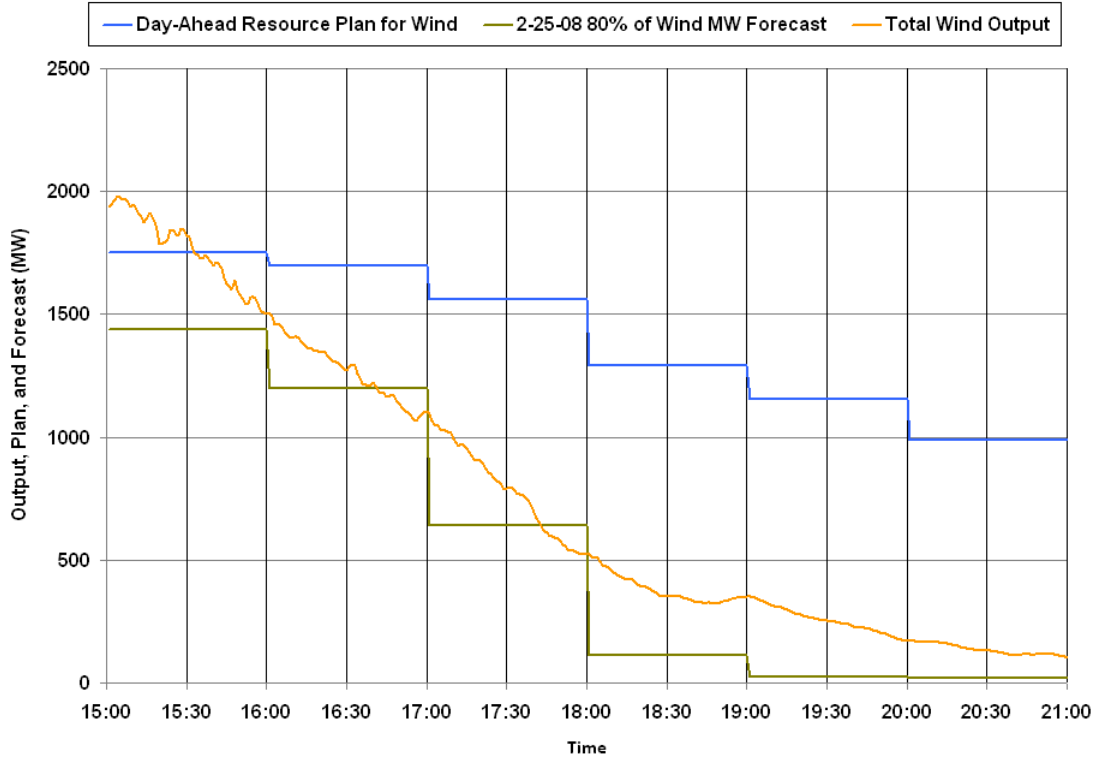


Figure 3. 02/26/08 15:00 – 21:00 total wind output, wind forecast in the 16:00 day-ahead resource plan [1]

However, the 80% wind forecast shown in green did predict the wind output with good fidelity [1]. This forecast looking ahead to evaluate system resource adequacy could mitigate the severity of consequences resulting from wind forecast error.

### 2.4.3 Prediction of Adequacy of Balancing Reserve

ERCOT system purchases balancing capacity and energy to cover the variability of the system, caused by load forecast errors, wind power forecast errors, uninstructed generation deviations, etc., and to maintain system frequency.

Regulation service is a part of the overall balancing capacity addressing minute-to-minute variability. To evaluate the regulation service requirements, ERCOT collects monthly historic deployed regulation service data. This data is used to calculate average historically deployed regulation service for one-minute periods. By calculating the 98.8 percentile of the deployed regulation up and regulation down by hour, ERCOT estimates the expected needs for similar months [4].

It is, therefore, by design that for a certain amount of time the system will be lacking of sufficient balancing capacity, or experience load-generation imbalance, because of the exhaustion of available resource. The imbalance can be small and negligible, but it can also get very large and cause intolerable frequency deviation in the system. ERCOT Feb 26, 2008 event is an occasion of the later. In these occasions, it would be of great help if the operators can have a prediction of system imbalance based on forecast and system resource information updated in real time. When the operators are aware of the possible forthcoming event, and amount of shortage in resources,

they can prepare ahead of the time to avoid large frequency deviations. A methodology to provide such real-time prediction is described in Section IV of this report as well as in [5].

## ***2.5 Summary of Lessons Learned***

This section analyzes the ERCOT reliability event on February 26, 2008, which was widely reported because of the involvement of wind generation. The following lessons can be derived from the analysis:

Large mismatches between generation and load can be caused by *load forecast errors, wind forecast errors, and control errors on traditional generators*, or a combination of all of the above.

Usually forecast errors in capacity or MW could cause balancing capacity insufficiency in real-time operation. Unpredicted ramp rates can also cause problems, even if the forecasted generation capacity requirement is higher than the actual need. Therefore, *the system balancing resources should be evaluated and procured based on their capacity (MW) and ramping capability (MW/min)*. This flexibility requirement should be met by the day-ahead and hour-ahead unit commitment and scheduling processes.

*Unexpected ramps caused by load and wind can cause serious problems with the system balance.* Power system operators are usually concerned by the fast ramps caused by wind power plants. The ERCOT event on February 26, 2008 shows that unexpected fast load ramps can also cause problems.

*A predictive tool evaluating system balancing requirements and comparing them with available system characteristics will be very helpful.* This tool could take inputs, including load forecast, wind power forecast, predicted availability of system balancing resources, and then evaluate the capability of system resources to meet system capacity and ramping requirements for the next few hours.

*Demand response can effectively reduce load-generation mismatches in an emergency situation.* Demand response in the ERCOT event only involved load reduction (LaaR). The response from LaaR effectively flattened the upward ramp of evening load in the ERCOT system during the event, alleviating the generation insufficiency problem and terminating frequency drop.

### III. Analysis of Tail Events Distributions

Analysis of probability distributions of tail events provides an overview of system within-hour balancing requirements and the expected frequency of tail events at various MW levels. The analysis contains two steps:

1. Yearly distributions showing the occurrence frequency versus MW level of system imbalances;
2. Hourly distributions showing the balancing requirement and average MW level of tail events corresponding to each of the 24 hours of a day;

The analysis uses the approach developed in our previous studies for BPA [6, 7]. This approach assumes that the within-hour balancing capacity is required to compensate for system imbalances in 99.5% of the cases. By definition, a tail event occurs when the balancing capacity needed is larger than the amount available. The study is focused on capacity requirements of the reserves; however, other types of requirements, such as ramp rate and ramp duration [7], can be analyzed in a similar fashion. All of the results shown in this paper were generated based on a set of 21-month data provided by BPA.

#### 3.1 Yearly Distribution of Tail Events

Yearly distributions are based on the analysis of system balancing requirements, which is calculated from system load and wind data using the methodology developed in [7]. Required balancing capacity is defined as the MW level that can cover 99.5% of the cases that have been simulated. By subtracting available balancing reserve from the balancing requirement, we get the MW shortage of balancing capacity in the system. Figure 4 and Figure 5 show the distribution of the regulation capacity requirement and distribution of MW shortage in regulation in the BPA system in 2007 and 2010, respectively. As the other component of balancing reserve, sufficiency of load following capacity can be analyzed similarly but was not considered in Figure 4 and Figure 5. The 2010 results were obtained based on forecasted load and wind data in the BPA data set.

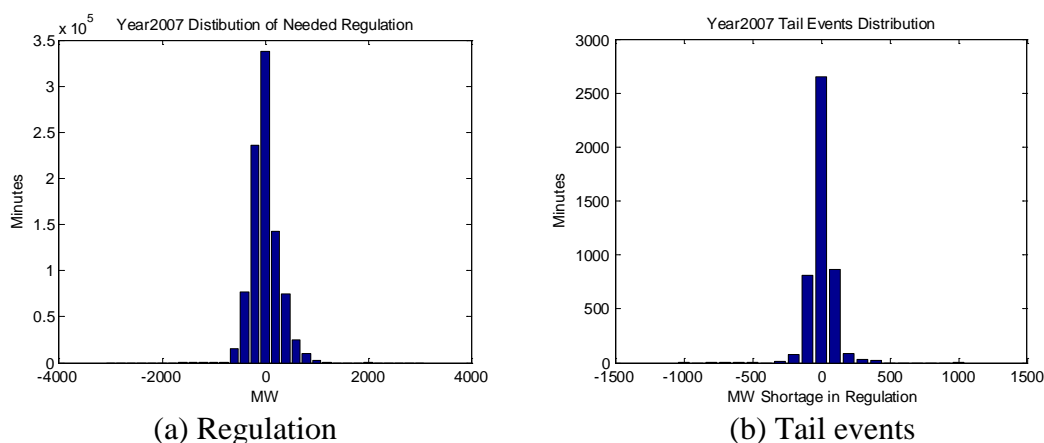
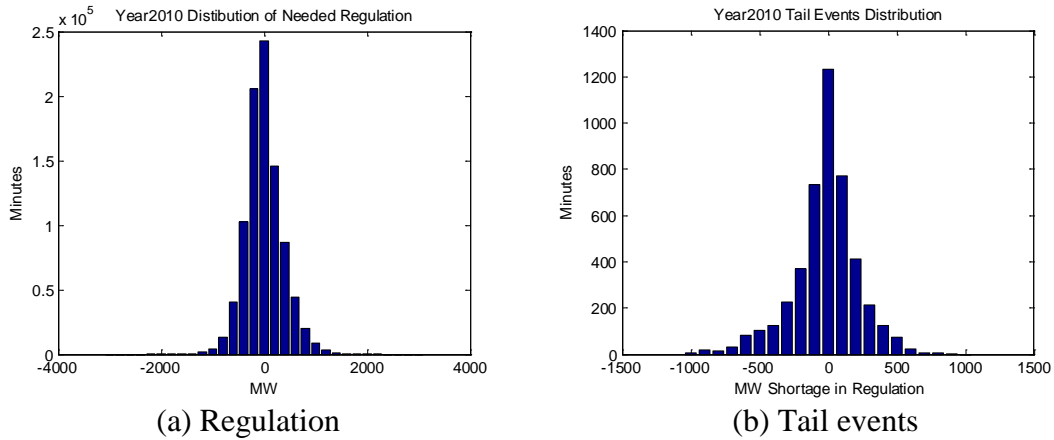


Figure 4. 2007 BPA system regulation capacity requirement and tail event distribution



**Figure 5. 2010 BPA system regulation capacity requirement and tail event distribution**

From the distribution of tail events, the number of minutes during the year when the system will be in shortage of regulation-up or regulation-down capacity at any specific MW level can be determined. For example, from plot (b) of Figure 4 and Figure 5 it can be seen that in 2007, under the assumptions in the analysis, there were only several minutes when the system experienced a shortage of 500 MW regulation-down capacity, while in 2010, the same situation will be occurring for around 100 minutes in total.

Yearly distribution plots provide an overview of the balancing requirement and frequency of tail events in the system under study, enabling a quick estimate of the degree of risk associated with any given level of reserve capacity.

### ***3.2 Hourly Distribution of Tail Events***

Hourly distributions are also generated based on system balancing requirement and available balancing reserve at 99.5% level. The process contains the following steps:

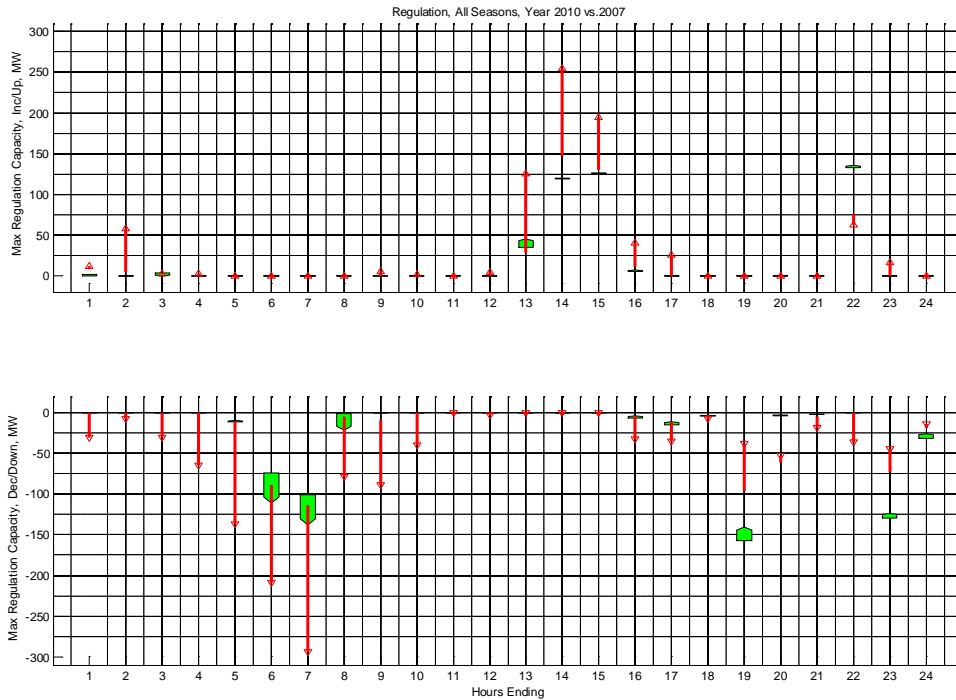
1. Data series of balancing capacity requirement are generated using the methodology developed in [7].
2. A time series representing the MW balancing capacity shortage, both upward and downward, is derived by subtracting the available balancing reserve from the balancing requirement data series.
3. The derived data points are then grouped into 24 hours of a day based on when the shortage occurred.
4. Data points allocated into the same hour are averaged to represent the average MW level of balancing capacity shortage.

The hourly distribution provides general information on which hours operators should watch carefully for the sufficiency of system balancing reserve.

In Figure 6, hourly distribution of regulation requirements and average MW shortage during each hour is plotted (again, sufficiency of load following capacity can be analyzed similarly but was not considered here). To also show the effect of wind power on the magnitudes of system tail



events, scenarios with wind and without wind in year 2007 (historical year) and 2010 (future year) in the BPA system are plotted in the same figure.



**Figure 6. Average regulation capacity shortage of all tail events: distribution corresponding to each hour during a day**

In Figure 6, the average MW shortage in tail events is shown hour by hour for 24 hours of a day. The green bars are results for 2007, and red lines for year 2010. The tip points of the green bars and the red arrows correspond to “with wind” conditions, while the flat end of the green bars or the red lines correspond to “without wind” conditions. The length of the bars and the lines indicate the contribution of wind generation to the MW shortage of regulating reserve in the BPA system. By comparing the tip points of the bars and the lines, one can see the expected increment of tail event MW level for the corresponding hours.

## IV. Real-time Prediction of Balancing Reserve Requirements and Tail Events

In this section, statistical relationships between system imbalance and forecast errors, generation schedule control errors and other influential factors such as temperature, wind speed, etc., are modeled using an approach called Bayesian network (BN). This model preserves the statistical characteristics obtained from system historical data and uses them as the basis for the prediction of the future. It is similar to the process when an experienced system operator estimates the difference between generation hourly schedules and the actual generation requirements in the system, based on the operating experience he/she has accumulated. Naturally, the prediction given by the BN model is more quantitative than human estimates. The model is expected to be able to help system operators to determine the sufficiency of balancing capacity in the system in real time.

### 4.1 Bayesian Network Models

Bayesian network (also called Bayes net) models are used to represent relationships among uncertain variables. The graphical representation of a BN consists of *nodes* and *directed links* or *arrows*. The nodes represent variables, and the arrows show the inter-dependencies between these variables. Arrows point from *parent* nodes to their *child* nodes; they show the direction of conditional dependence. Child nodes are conditionally dependent on their parents and are conditionally independent of their non-descendants given their parents.

Dynamic systems that change continuously through time, such as power systems, can be modeled using dynamic Bayesian network (DBN) models, where the state space of the system is modeled for successive time intervals. There are two simplifying assumptions that are typically used in constructing DBNs. One is stationarity—the probabilities of transitions from one state to another within each time slice are the same. The other is the Markov condition—the transition probabilities between time slices depend only on a finite number of states in previous periods. For a first order DBN they would depend only on one previous period states. Additional information on BN models can be found in [8-10].

### 4.2 Building a Bayesian Network Model for the BPA Power System

The proposed DBN model is a prediction tool for real-time operation of the BPA system. This model, shown in Figure 7, forecasts the value or state of system imbalance (SI) for future time intervals depending on system parameters observed in the current and future intervals. The nodes representing variables with uncertainty are depicted in Figure 7 as ovals. The model may also include decisions that might be made depending on the forecasted system imbalance. These decisions can be curtailment operations, for example, shown as rectangles. The model could also identify potential outcomes resulting from the predicted system imbalance, e.g., transmission congestion problems and control performance standard (CPS) violations, which are shown as hexagons in Figure 7.

As can be seen in Figure 7, system imbalance in the proposed model has three primary causes: load forecast error (LFE), wind forecast error (WFE), and generator scheduling control error (SCE). All three are stochastic and have strong serial correlation. The LFE at time (t+1) is forecasted from the LFE at time (t) as well as from the load and temperature and the wind and storm information at that time step. LFE is affected by diurnal and seasonal cycles, as well as meteorological events such as the passage of fronts (storms) with associated rapid changes in temperature and wind velocity. WFE at time (t+1) is analogously predicted from WFE, wind power, and storm information at time (t). SCE is also predicted from relevant variables in a previous time step as shown in Figure 7.

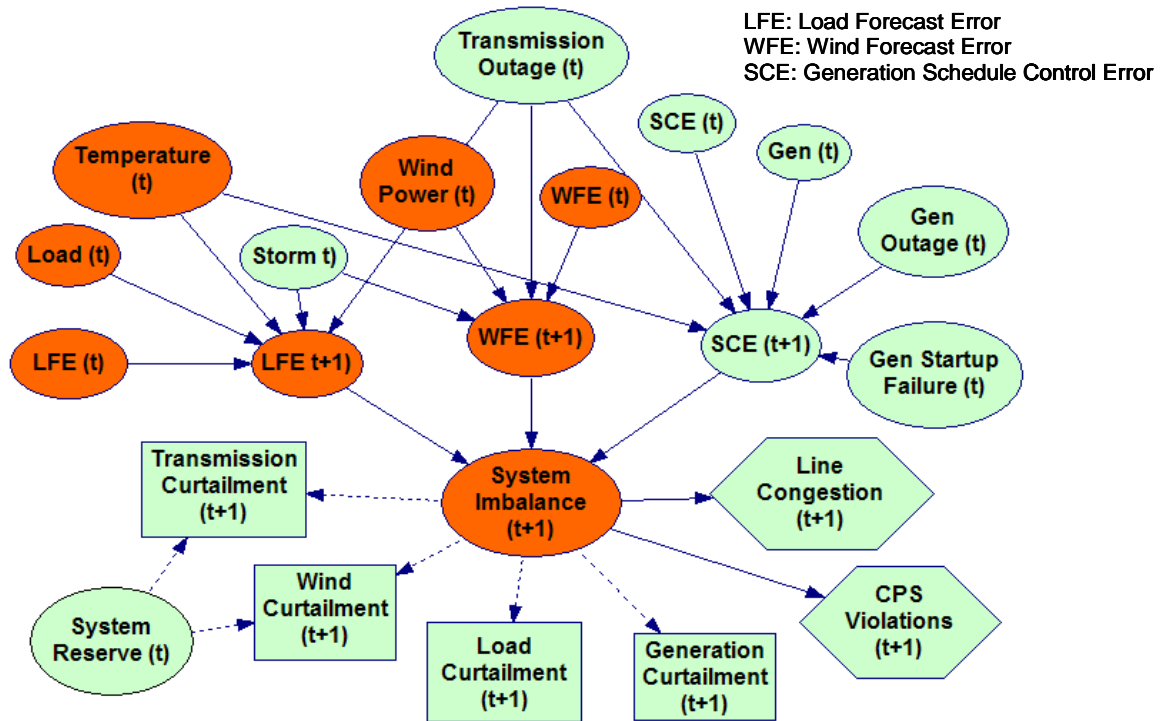


Figure 7. Bayes net model for predicting system imbalance.

The system can be perturbed by events such as transmission and generation outages. The occurrence of these events is unpredictable, but their effects on the system imbalance are deterministic in nature and are incorporated into the model to show their potential impact on LFE, WFE, SCE and subsequently on the system imbalance.

The hierarchy in Figure 7 shows the dependency relationship between nodes. Nodes for load, load forecast error, temperature, wind power, and wind forecast error observed at the current time step (t), are used in forecasting LFE and WFE in the next time step (t+1). The forecasted values of LFE, WFE and SCE are used in turn to forecast system imbalance at time (t+1). In application, observed values at the current time step (t) are entered into the DBN model to generate forecasted values for LFE, WFE and SCE at the next time step (t+1). Forecast for future system imbalance at (t+n) can be generated using the hidden Markov model as needed.

As a proof of concept, the model shown in red in Figure 7 has been implemented and tested. The focus was put on these variables because they have the greatest impact on system imbalance and should serve well as an initial test of the feasibility of this modeling approach.

### 4.3 Preprocessing the Data

The state-spaces for BN nodes were derived from a historical time series of hourly observations on these components in the BPA system. The data set provided consists of continuous variables. While it is possible to build a BN model from continuous variables, the algorithms are much more complicated and the usual practice is to discretize the data. The process results in a histogram of the data and consists of dividing the data into discrete intervals that are non-overlapping and mutually exclusive. Data falling into each of these categories is given state names that correspond to the variable states defined in the BN model. The resulting discretization is simply a bar-plot of the frequency counts of observations in each bin. This can be done either using the R programming language, or in some cases it was done using the GeNIe® software (<http://genie.sis.pitt.edu>) that was used to implement the BN model. The GeNIe® program has user friendly utilities for discretizing data and viewing histograms and pie-charts of the resulting distributions. A screen shot showing the discretization of the temperature data using GeNIe is shown in Figure 8.

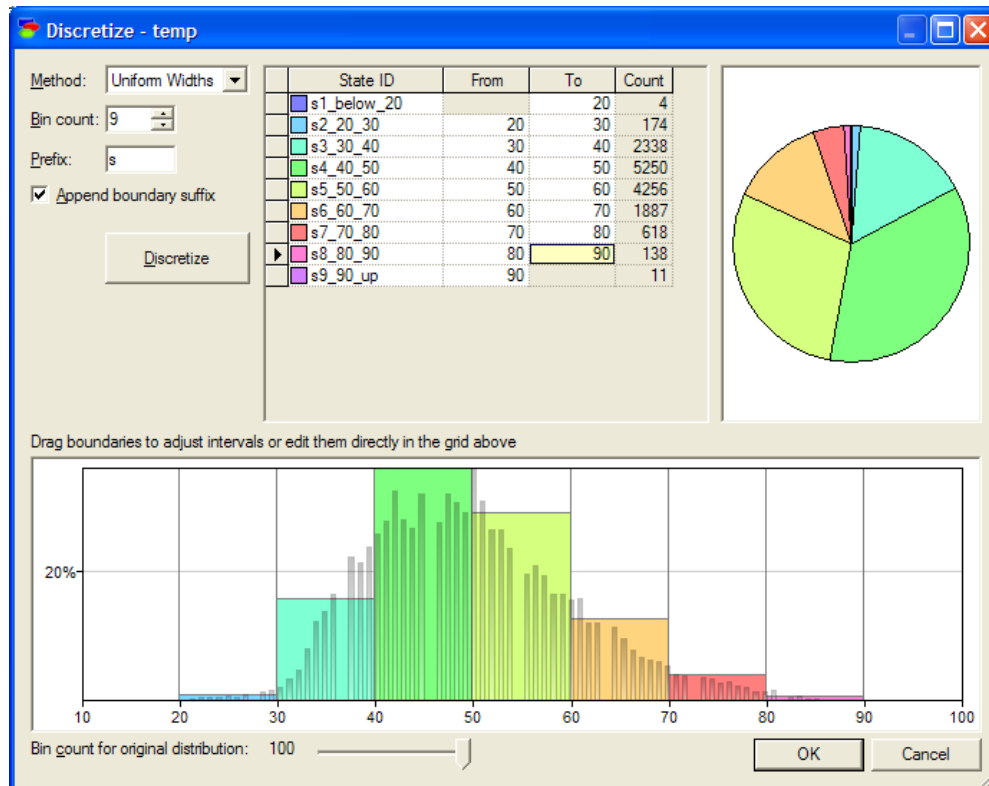


Figure 8. Screen shot of discretized temperature using GeNIe software.

### 4.4 Learning Probabilities from the Data

The discretized data is the basis for computing the conditional probability tables (CPT) for the BN model. For the root nodes (nodes without parents), the prior probability of being in a particular state is just the relative frequency for which the state occurs. The conditional probabilities are learned in an analogous fashion. For each combination of parent states, the proportion of times the child is in each of its states is determined. This becomes the basis for the learned probabilities. While this is the basic principle, the algorithm is slightly more complex because of the need to account for combinations of parent states that did not occur in the data set. These are accounted for in GeNIe software by using the expectation maximization (EM) algorithm [9].

Once the parameters were learned using the BN shown in red in Figure 7, the probability distributions were then used to build a DBN as in Figure 9. System imbalance depends on WFE and LFE. They each depend on themselves in the previous time period, as shown by the looping arrow with a “1”. In addition, Temperature at time t is used to predict LFE at time t+1. The time slices in the temporal plate of GeNIe is set to 4, thus one observation will produce predictions for three periods in the future. The model can be exercised with any number of time slices.

To make the four time slice model even more explicit we built a model in which the variables were duplicated for each of the four time slices. This is shown in Figure 10. Each color in the figure represents a single time slice. One can clearly see that this DBN is a first order Markov model because the probabilities in a given time slice depend only on the previous time period; i.e., conditionally probability arcs connect successive time periods. This DBN also has the property of stationarity, meaning the conditional distributions within time slices are the same across time periods.

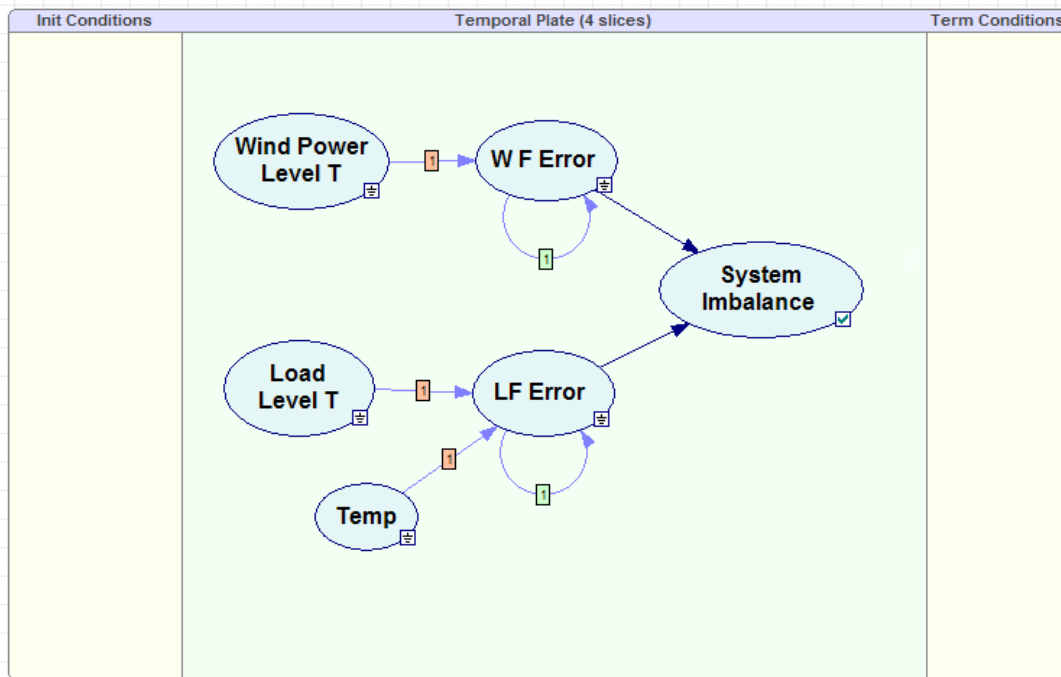


Figure 9. Screen shot of DBN for predicting system imbalance as implemented in GeNIe.

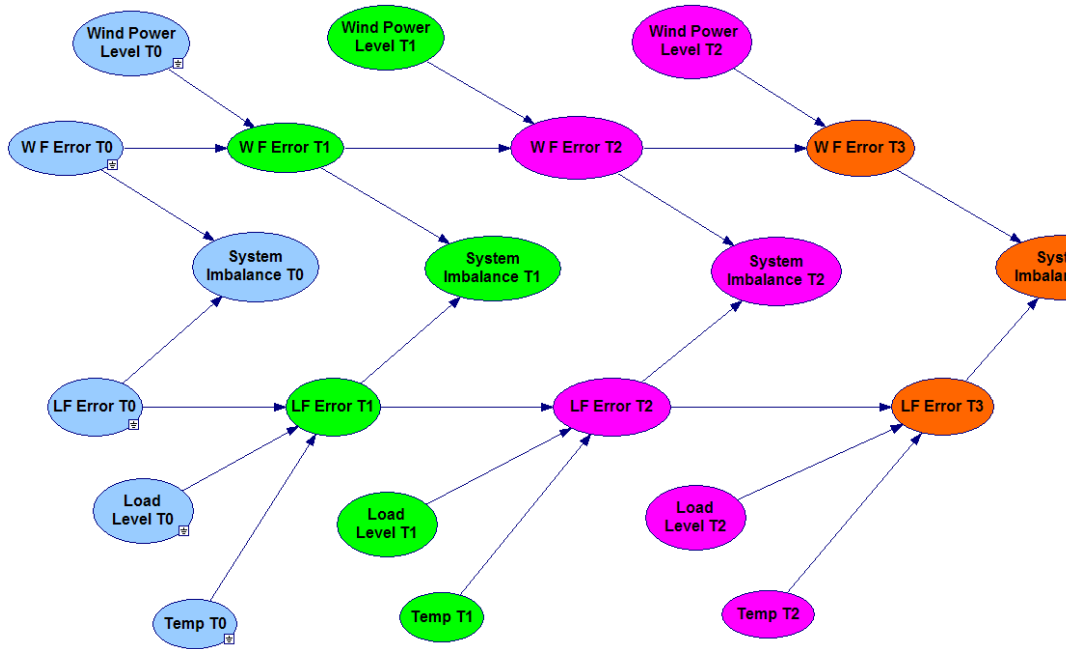


Figure 10. First order dynamic Bayes net used to predict system imbalance.

#### 4.5 Generation of Forecast Data

Table 1 shows the initial eleven records used to run the GeNIe DBN model shown in Figure 9. The complete data set consisted of 15,379 records from October 2006 to June 2008. For each hour, the data for time (t) were entered into the model as evidence and the model provided a forecast of system imbalance for three subsequent time periods. Figure 11 is a snapshot of exercising GeNIe for one time slice. A program was written in C++ to read evidence for time (t) and generate predictions for time (t+1), (t+2), and (t+3).

Table 1. Sample of data used to run the BN model

day	year	hour	load.t	loaderr.t	loaderr.t1	wind.t	winderr.t	winderr.t1	sys.t	sys.t1	temp.t
30	2006	10	MW4500	MW0	MW0	MW600	MW0	MW100	MW0	MW0	s05_50_60
30	2006	11	MW5000	MW0	MW0	MW600	MW100	MW_100	MW0	MW0	s05_50_60
30	2006	12	MW5000	MW0	MW_200	MW600	MW_100	MW0	MW0	MW_200	s05_50_60
30	2006	13	MW4500	MW_200	MW200	MW600	MW0	MW_100	MW_200	MW200	s05_50_60
30	2006	14	MW5000	MW200	MW400	MW400	MW_100	MW0	MW200	MW200	s05_50_60
30	2006	15	MW5500	MW400	MW200	MW400	MW0	MW100	MW200	MW200	s05_50_60
30	2006	16	MW5500	MW200	MW0	MW600	MW100	MW0	MW200	MW0	s05_50_60
30	2006	17	MW5500	MW0	MW0	MW600	MW0	MW100	MW0	MW0	s05_50_60
30	2006	18	MW5500	MW0	MW_200	MW600	MW100	MW0	MW0	MW_200	s05_50_60
30	2006	19	MW5500	MW_200	MW0	MW600	MW0	MW0	MW_200	MW_200	s05_50_60
30	2006	20	MW5500	MW0	MW0	MW600	MW0	MW0	MW_200	MW0	s05_50_60

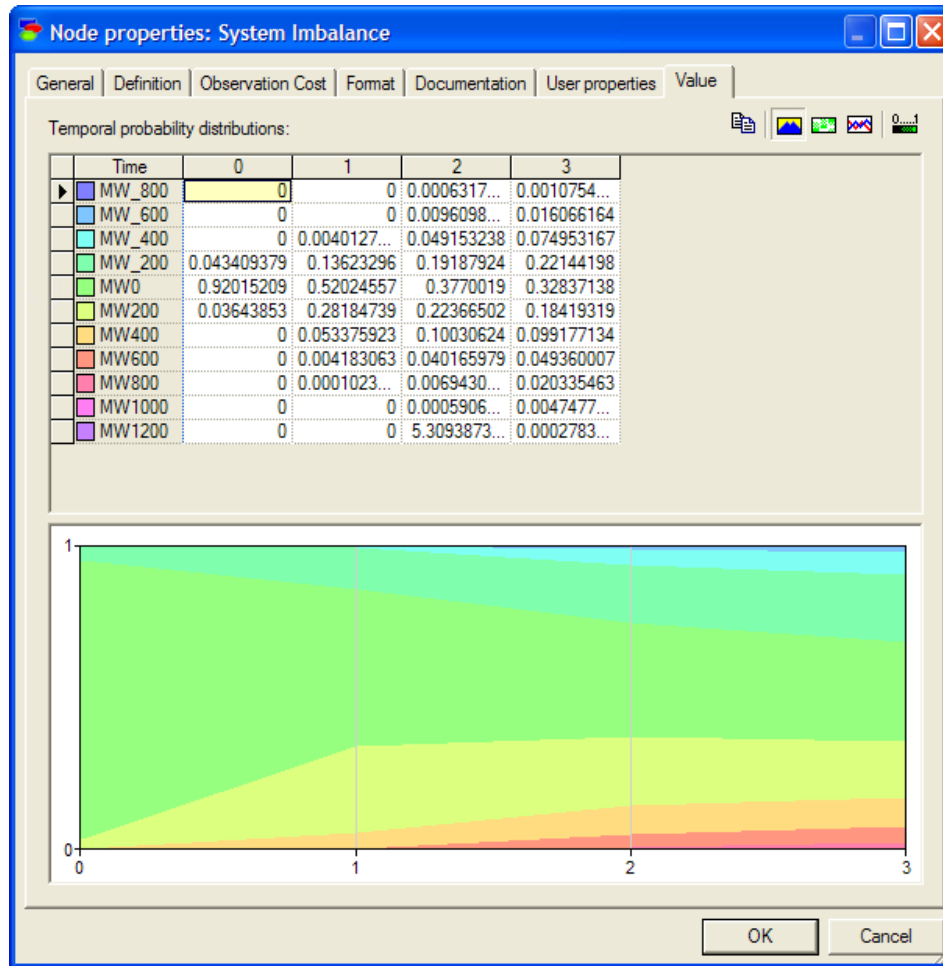


Figure 11. Output generated by GeNIe for predicting system imbalance.

In Figure 11, each color band represents a system imbalance state, such as MW0 (0 MW imbalance), MW\_200 (-200 MW imbalance), etc. The width of the color band at a particular hour (hour 1, 2 and 3) is the predicted probability of the corresponding state at that hour. Therefore, the sum of the width of all color bands is equal to “1”.

#### 4.6 Bayesian Network Model Output

An example output of the BN model shown in Figure 9 is the probability distribution of system imbalance in future time steps. Because hourly data were used to generate the model, prediction results have a time step of one hour. Figure 12 and Figure 13 show the predicted probability distribution of system imbalance in the next 1 and 2 hours, respectively. If the system is assumed to have a 500 MW upward regulating reserve and 700 MW downward regulating reserve, then system imbalance lower than -500 MW and higher than 700 MW indicates a tail event. In Figure 12 and Figure 13 the bar at -600 MW represents the interval between -500 MW and -700 MW, and the bar at 800 MW represents the interval between 700 MW to 900 MW. Therefore, the probability of a tail event is calculated by accumulating the probabilities including and beyond these two bars.

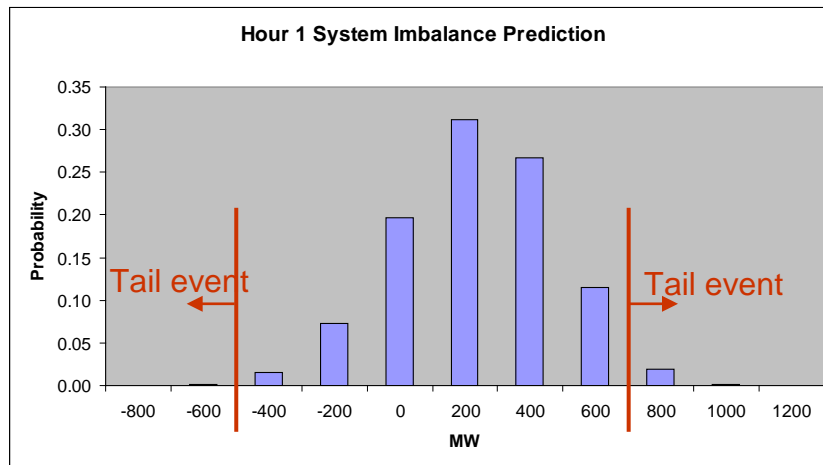


Figure 12. Prediction of system imbalance for the next hour.

Figure 12 shows that during the next operation hour, the probability of generation shortage in the provided example is 0.16%, and the probability of over generating is 1.95%. The most likely state of system imbalance is 200 MW, representing the interval between 100 MW and 300 MW.

Figure 13 shows that during the second operation hour, the probability of generation shortage is 0.08%, and the probability of over generating is 6.95%. The most likely state of system imbalance is 0 MW, representing the interval between -100 WM and 100 MW.

The actual system imbalance observed in the system for this example was: Hour 0 (current hour) = 132 MW, Hour 1 = 333 MW, Hour 2 = -162 MW.

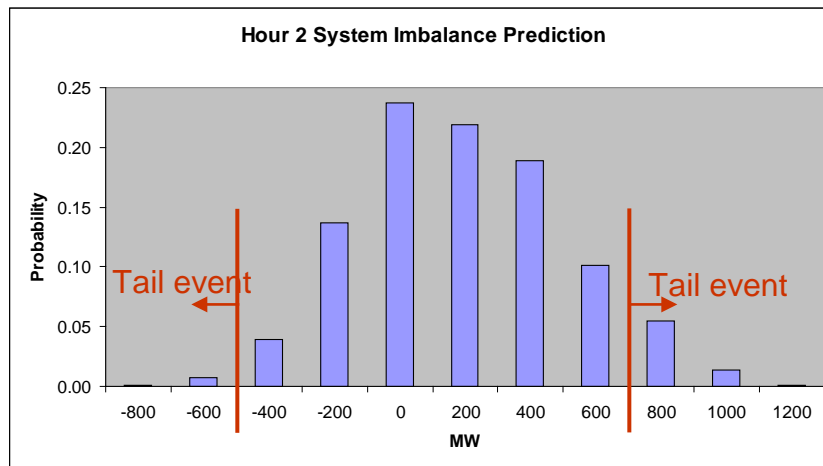


Figure 13. Prediction of system imbalance for the second hour.

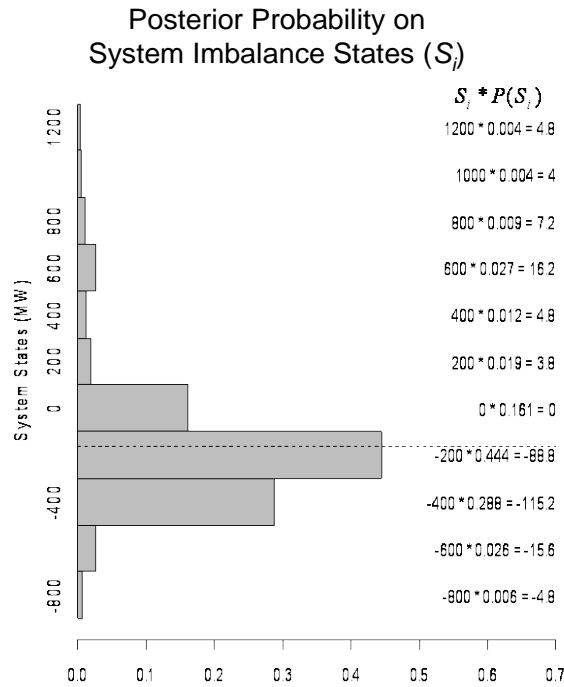
## 4.7 Validation Studies

### 4.7.1 Comparison with Naïve Persistence Forecasts

Validation of the BN forecasts on system imbalance were done in comparison to the naïve persistence (NP) forecast model. The NP model uses the observed system imbalance at time (t)



as the forecast for future times ( $t+n$ ;  $n \geq 1$ ). The NP model provides only a point estimate forecast without any measure of uncertainty. The BN model provides the Bayesian posterior probability distribution of the forecasted system imbalance conditioned on all system components in the model, and thus explicitly provides a measure of uncertainty. Before forecasts from the BN and NP models can be directly compared, the BN forecast needs to be converted to a point-estimate. BN point estimates are computed as probability-weighted averages of state-space interval mid-points. The algorithm for this conversion is diagrammed in Figure 14.



$$\sum_i S_i P(S_i) = -183.4 \text{ MW}$$

**Figure 14. Example demonstrating method for converting BN model forecasts from posterior probability distribution on system imbalance to point-estimates.**

Direct comparisons of the BN and NP model forecast accuracy are shown in Figure 15 and Figure 16 for 1-, 2-, and 3-hour forecasts as mean absolute error (MAE). Because the significance of system imbalance being negative or positive may be quite different in a power system, these two types of cases were compared separately in Figure 15 and Figure 16. Errors were computed by subtracting the observed system imbalance from the NP and BN forecasted system imbalances, respectively. Smaller MAE values indicate more accurate forecasts.

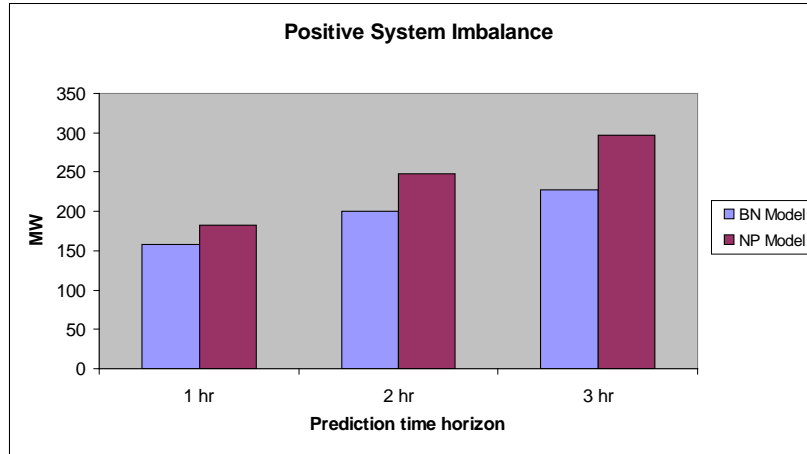


Figure 15. MAE comparison between BN and NP models forecasting at one, two and three hours: positive system imbalance cases

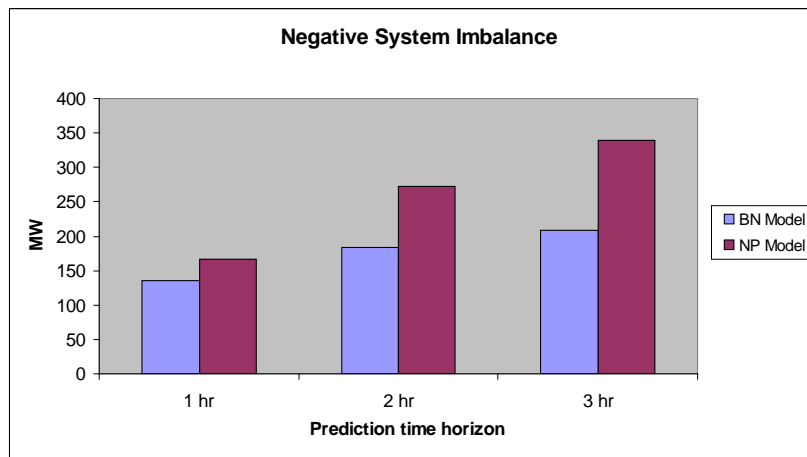


Figure 16. MAE comparison between BN and NP models forecasting at one, two and three hours: negative system imbalance cases

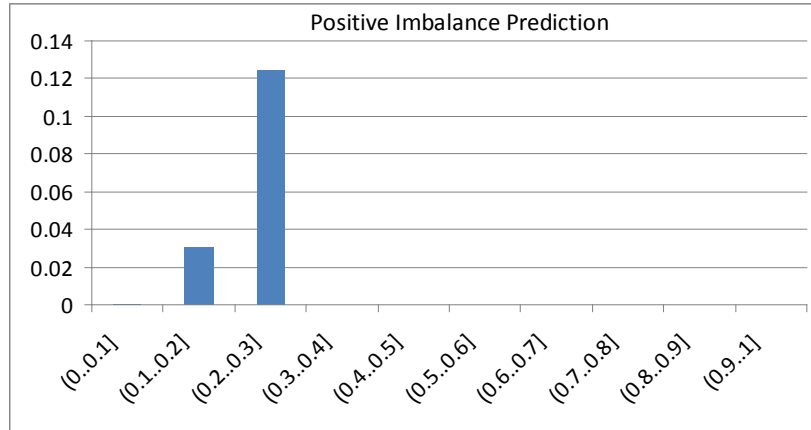
The BN model showed consistently improved accuracy over NP model on the 1-, 2- and 3-hour forecasts. The average improvements over NP model forecast results are 16%, 26% and 31% for 1-, 2- and 3-hour forecasts, respectively.

#### 4.7.2 Prediction of Tail Event Probability

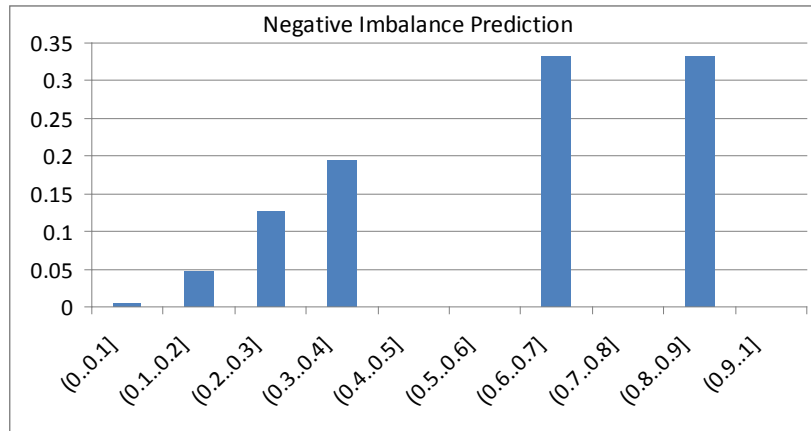
Probability of tail event can be calculated from the BN model output, as described previously. To validate the prediction results, all cases are grouped based on the probability of tail event predicted by the BN model, such as 0~0.1, 0.1~0.2, etc. Then in each group, the number of cases when tail events were observed (regulating capacity is insufficient) is divided by the total number of cases in that group. The results are deemed as the actual probability that tail events occurred. They are shown in Figure 17 and Figure 18 for positive imbalance and negative imbalance cases, respectively.

Figure 17 and Figure 18 show that the observed probability of tail events does not match very well with the probability predicted by the BN model, which demands improvements of the model. There could be a high false alarm rate, with the predicted probability always higher than

the actual probability observed. Nevertheless, there is a significant correlation between the two, which is more obvious in Figure 18. It does show that when the BN model predicts a high probability of tail event, the chance of a tail event actually occurring is also high.



**Figure 17. Observed probability of tail events vs. predicted probability by the BN model: positive system imbalance cases.**



**Figure 18. Observed probability of tail events vs. predicted probability by the BN model: negative system imbalance cases.**

## ***4.8 Potential Improvements on the Bayesian Network Model***

### **4.8.1 Improve Prediction Accuracy**

The current BN model has a time step of one hour. In real-time operations, predictions with higher time resolution are sometimes preferable because large system imbalance can be missed in forecasts if the forecasts are averages over long time intervals. Therefore, a BN model with 5 to 15 minute time interval could be constructed for the use in real time.

Another potential improvement is to use second or third order Markov process to see the trends of data series. For example, in the current model (first order Markov process), LFE at time (t+1) is affected by the temperature at time (t), and is irrelevant to temperature at and before time (t-1). If a second order Markov process is used, LFE at (t) would be affected by both the temperature

at time (t) and (t-1). The approach may be able to improve the prediction accuracy of the BN model.

Various techniques can also be explored and tested in dealing with the issue of insufficient data when forming the transition matrices between different system states.

On the other hand, dimension of the transition matrices for the model increases linearly with the time resolution and the order of Markov model. Lack of sufficient data will also become more challenging. These problems need to be taken care of appropriately to improve the model.

#### **4.8.2 Adapt the Prediction Model to Changed System Composition**

The need in balancing capacity depends on multiple factors and is driven mostly by the variability and forecast errors of load and variable generation resources. Composition of system load and generating resources affects their variability. As configuration of the control area changes, and penetration of wind power in the system increases, the statistical relations between system parameters and system imbalance in the BN model will certainly change accordingly. Improvement in forecast techniques will also change the transition probabilities between state variables in the BN model. To maintain a balancing reserve prediction model that represents the system accurately, the BN model needs to be adaptive. Historical load data, wind data, wind power data, and other necessary inputs to the model need to be saved. Forecast data together with the newly accumulated historical data can be used to modify the balancing reserve prediction model. In other words, after a certain period, the proposed tool can be trained with new sets of historical and simulated data that reflect the effect of new forecast efforts and changed system compositions.

## V. Framework of a Decision Support Tool for Within-hour Balance of the System

The framework of a decision support tool for within-hour balancing operations is proposed in this section. The tool will be able to determine the amount of balancing capacity needed in each hour, the chance of tail events and provide dispatch suggestions.

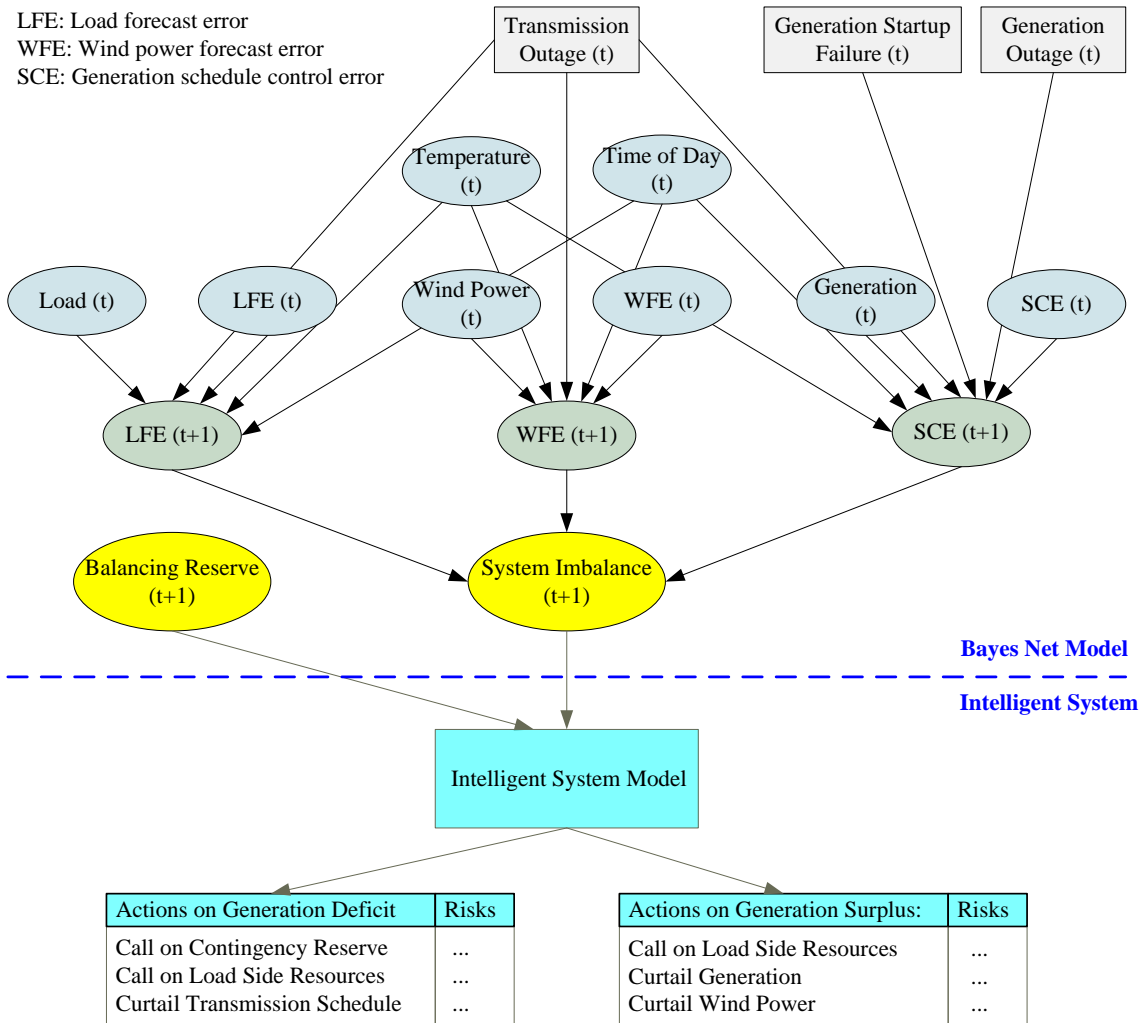
Based on studies of historical data, the statistical relationship between system imbalance and various influential factors, such as wind power, load, generation status, time of the day, and ambient air temperature can be modeled using a Bayesian network model that was introduced in Section IV. This model can be further developed using more BPA system data (involving more input variables) and made adaptive to changing system compositions (e.g., higher wind capacity). Load forecast, wind power forecast, ramp prediction, and other operational information generated using existing operation tools can be incorporated into the model. Prediction of balancing reserve requirements (system imbalance) for a number of operation hours will be the output of the model.

The tool will also contain an *intelligent system* model, using system imbalance prediction and available reserves as inputs, in addition to supervisory control and data acquisition (SCADA) measurements, to suggest dispatcher actions. BPA operational procedures and rules will be followed when making these suggestions, with meeting reliability standards and regulatory requirements as objectives. Since the delivery of balancing reserves is constrained by transmission capability and may affect voltage profile of the system, the impact of suggested actions on major transmission paths and buses need to be assessed. Risks and costs of taking these actions or no actions can also be formulated and evaluated. Intelligent system techniques, such as rule-based expert system, fuzzy logic, and artificial neural networks, can be explored in the development of this model.

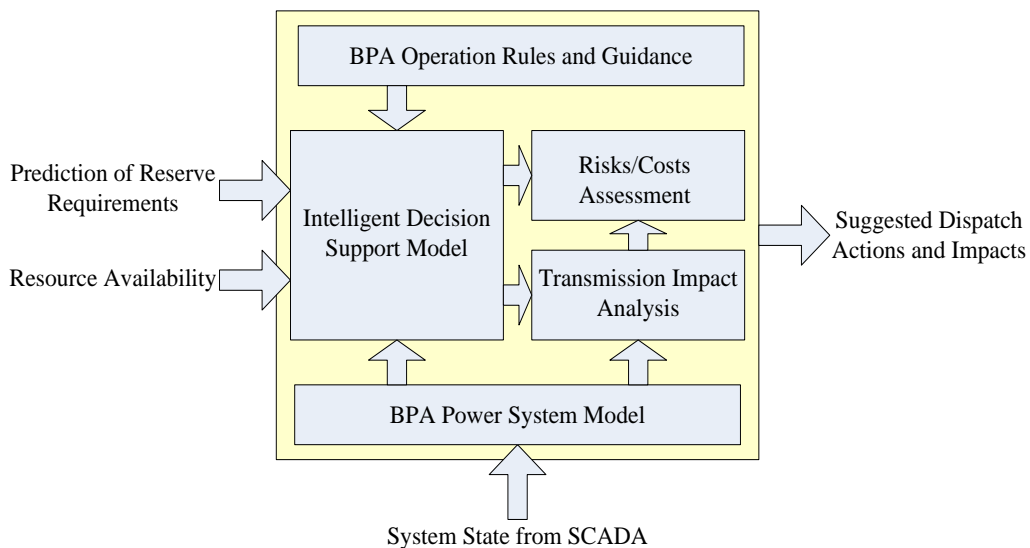
Development of the tool will need data collection, interviews and discussions with BPA engineers and operators, algorithm development, program coding, offline testing and verification, and feedback collection.

The decision support tool based on Bayesian network model is illustrated in Figure 19.

A block diagram of the Intelligent System model in the decision support tool is shown in Figure 20.



**Figure 19. Decision support tool for balancing operations based on Bayesian network model.**



**Figure 20. Block diagram of the intelligent system model in the decision support tool.**

## **VI. Conclusion**

This report summarizes the factors that affect balancing operations in the power system. A methodology based on the analysis of historical data is developed to identify distribution patterns of tail events (large system imbalance). The report also demonstrates that the relations between influential factors and system imbalance can be established using a Bayesian network model. It is shown that such models can provide better estimates of balancing reserve requirements in real-time operations than a persistence model, which assumes system imbalance in the future will be the same as the present time interval. The framework for a decision support tool is described, which includes the Bayesian network model and an intelligent system model. The tool will be able to determine the chance of tail events and provide dispatch suggestions.

## VII. References

- [1] ERCOT Operations Report on the EECF Event of February 26, 2008, [Online]. Available: [http://www.ercot.com/meetings/ros/keydocs/2008/0313/07.ERCOT\\_OPERATIONS\\_REPORT\\_EECP022608\\_public.doc](http://www.ercot.com/meetings/ros/keydocs/2008/0313/07.ERCOT_OPERATIONS_REPORT_EECP022608_public.doc)
- [2] ERCOT Event on February 26, 2008: Lessons Learned, [Online]. Available: [http://www.nrel.gov/wind/systemsintegration/pdfs/2008/ela\\_ercot\\_event.pdf](http://www.nrel.gov/wind/systemsintegration/pdfs/2008/ela_ercot_event.pdf)
- [3] ERCOT Protocols – Section 6: Ancillary Services, [Online]. Available: [http://www.ercot.com/content/mktrules/protocols/current/06-110109.doc#\\_Toc244479688](http://www.ercot.com/content/mktrules/protocols/current/06-110109.doc#_Toc244479688)
- [4] ERCOT Methodologies for Determining Ancillary Service Requirements, [Online]. Available: [http://www.ercot.com/content/meetings/tac/keydocs/2009/1105/12.ROS.ERCOT\\_Methodologies\\_for\\_Determining\\_Ancillary\\_Servic.doc](http://www.ercot.com/content/meetings/tac/keydocs/2009/1105/12.ROS.ERCOT_Methodologies_for_Determining_Ancillary_Servic.doc)
- [5] Shuai Lu, Yuri V. Makarov, Alan J. Brothers, Craig A. McKinstry, Shuangshuang Jin, and John H. Pease, *Prediction of Power System Balancing Requirement and Tail Event*, IEEE Transmission and Distribution Conference 2010, New Orleans, LA, April 19-22, 2010. Accepted.
- [6] Shuai Lu and Yuri Makarov, “Operational Impacts of Wind Energy Resources in the Bonneville Power Administration Control Area – Phase I Report”, Pacific Northwest National Laboratory, Richland, WA, Tech. Rep. PNNL-17558, May 2008.
- [7] Yuri Makarov, Shuai Lu, Bart McManus, John Pease, “The Future Impact of Wind on BPA Power System Ancillary Services”, *IEEE Transmission and Distribution Conference 2008*, Chicago, April 2008.
- [8] Jensen, Finn V. and Thomas D. Nielsen, 2007, *Bayesian Networks and Decision Graphs*, Second Edition, Springer.
- [9] Neapolitan, Richard E. 2004, *Learning Bayesian Networks*, Pearson Prentice Hall, Upper Saddle River, N.J.
- [10] Russell, Stuart and Peter Norvig, 2003 *Artificial Intelligence: A Modern Approach*, Second Edition, Prentice Hall Pearson Education, Inc, Upper Saddle River, New Jersey.