PNNL-18769

U.S. DEPARTMENT OF ENERGY

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: Task 2 Report

S Lu
YV Makarov
CA McKinstry
AJ Brothers
S Jin

September 2009

Pacific Northwest
NATIONAL LABORATORY

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**

**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**
**online ordering: http://www.ntis.gov/ordering.htm**

This document was printed on recycled paper.

(8/00)

PNNL-18769

# Low Probability Tail Event Analysis and Mitigation in the BPA Control Area: Task 2 Report

S Lu, Ph.D.: Senior Research Engineer, Project Manager
YV Makarov, Ph.D.: Chief Scientist - Power Systems
CA McKinstry: Senior Research Scientist
AJ Brothers, Ph.D.: Senior Research Scientist
S Jin, Ph.D.: Research Engineer

September 2009

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

Pacific Northwest National Laboratory
Richland, Washington 99352

# SUMMARY

This report presents a methodology developed at the Pacific Northwest National Laboratory (PNNL) for the Bonneville Power Administration (BPA) for the prediction of power system balancing requirement and the probability of tail event (large imbalance between generation and load) in the BPA system. Maintaining sufficient balancing reserves to match the difference between hourly generation schedule and real-time variable load and intermittent resources becomes more and more challenging with the increasing penetration of intermittent energy sources. The presented methodology uses yearly distributions and hourly distributions of balancing requirement and tail events to provide a high level look at the issue and show to system operators those hours when problems are most likely to occur. For real-time prediction, a Bayes net model is constructed to model the statistical relationships between system imbalance and forecast errors, generation schedule control errors and other influential factors. The methodology will be able to provide reference information to system operators in determining the sufficiency of system balancing reserve and taking appropriate control actions.

# ACKNOWLEDGEMENT

The authors gratefully acknowledge the contributions of John Pease, BPA Project Manager, Technology Innovation Group, for his sincere advice and help to the team, as well as the following BPA colleagues for sharing their system operation 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 ............................................................................................. 1
II. YEARLY DISTRIBUTION OF TAIL EVENTS .................................................... 3
III. HOURLY DISTRIBUTION OF TAIL EVENTS ................................................... 5
IV. REAL-TIME PREDICTION OF BALANDING RESERVE REQUIREMENT AND TAIL EVENTS ........................................................................................................ 7
4.1 Bayes Net Models ........................................................................................ 7
4.2 Building a Bayes Net Model for the BPA Power System .................................... 7
4.3 Preprocessing the Data .................................................................................. 8
4.4 Learning Probabilities from the Data ............................................................... 9
4.5 Generation of Forecast Data ......................................................................... 11
4.6 Byes Net Model Output ................................................................................ 12
4.7 Validation Studies ........................................................................................ 14
4.7.1 Comparison with Naïve Persistence Forecasts ............................................. 14
4.7.2 Prediction of Tail Event Probability ............................................................ 15
4.8 Potential Improvements on the Bayes Net Model ............................................. 16
V. CONCLUSION ............................................................................................... 18
VI. REFERENCES .............................................................................................. 19

# FIGURES

Fig. 1. 2007 BPA system regulation capacity requirement and tail event distribution ...... 3
Fig. 2. 2010 BPA system regulation capacity requirement and tail event distribution ...... 3
Fig. 3. Average regulation capacity shortage of all tail events: distribution corresponding to each hour during a day ........................................................................................ 5
Fig. 4. Bayes net model for predicting system imbalance. ................................................ 8
Fig. 5. Screen shot of discretized temperature using GeNIe software .............................. 9
Fig. 6. Screen shot of DBN for predicting system imbalance as implemented in GeNIe. 10
Fig. 7 First order dynamic Bayes net used to predict system imbalance. ....................... 11
Fig. 8. Output generated by GeNIe for predicting system imbalance. ............................ 12
Fig. 9. Prediction of system imbalance for the next hour. .............................................. 13
Fig. 10. Prediction of system imbalance for the second hour .......................................... 13
Fig. 11. Example demonstrating method for converting BN model forecasts from posterior probability distribution on system imbalance to point-estimates ...................... 14
Fig. 12. Mean absolute prediction error comparison between BN and NP models forecasting at one, two and three hours: positive system imbalance cases....................... 15
Fig. 13. Mean absolute prediction error comparison between BN and NP models forecasting at one, two and three hours: negative system imbalance cases ...................... 15
Fig. 14. Observed probability of tail events vs. predicted probability by the BN model: positive system imbalance cases ........................................................................................ 16
Fig. 15. Observed probability of tail events vs. predicted probability by the BN model: negative system imbalance cases ....................................................................................... 16

# TABLES

Table 1. Sample of data used to run the BN model ........................................................ 11

# I. INTRODUCTION

*Tail event* refers to the situation in a power system when unfavorable forecast errors of load and wind are superposed onto fast load and wind ramps, or non-wind generators falling short of scheduled output, the imbalance between generation and load becomes very significant. This type of events occurs infrequently and appears on the tails of the distribution of system power imbalance; therefore, is referred to as tail events.

With the increasing penetration of intermittent energy sources in the system, including wind and solar, large imbalance is encountered more frequently than ever. Maintaining sufficient balancing reserves, both upward and downward, to match the difference between hourly generation schedule and real-time variable load and intermittent resources, becomes more and more challenging. This project developed the methodology at the Pacific Northwest National Laboratory for the Bonneville Power Administration (BPA) for the analysis and online prediction of system balancing requirement and the probability of tail event. The objective of this study is to provide reference information to system operators helping them be aware of those times a tail event is likely to occur and evaluate the capability of system to deal with the possible amount of imbalance in the next several operation hours and perform dispatches accordingly. Because BPA uses only regulating reserve to balance its system under normal conditions, balancing reserve and balancing requirement will be used interchangeably with regulating reserve and regulation requirement in the report. For systems also having a real-time dispatch/load following process, the balancing reserve discussed here is equivalent to the sum of load following and regulating reserves.

The methodology presented in this report contains three parts:

1. Yearly distributions showing the occurring frequency versus MW level of system imbalance;
2. Hourly distributions showing the regulation requirement and average MW level of tail events corresponding to each of the 24 hours of a day;
3. A model to predict the distribution of regulation requirement in real-time operation and the probability of tail event occurrence in each operation hour for the next several hours.

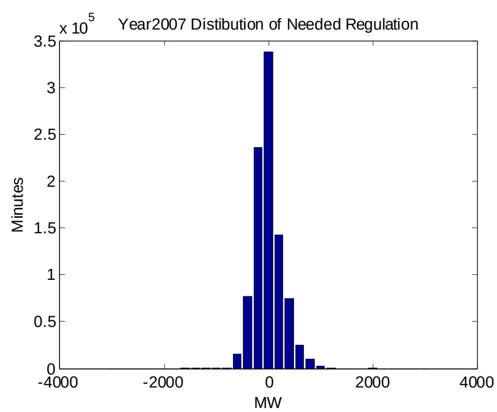
Part 1 and 2 uses the approach developed in [1] and [2] out of a previous study for BPA, providing a high level look at system balancing requirement and the frequency of tail events at various MW levels. In Part 3, the statistical relationships between system imbalance and forecast errors, generation schedule control errors and other influential factors such as weather, temperature, wind speed, etc., are modeled using an approach called Bayes net (BN). This model preserves the statistical characteristics obtained from system historical data and uses them as the basis for the prediction of future. It is similar to the process that an experienced system operator estimates what the difference between generation hourly schedule and the actual generation need would be in the system, based on the operating experience he/she has accumulated. The prediction given by the BN model is simply more quantitative. The model is expected to be able to help system operators in real-time by determining the sufficiency of regulating reserve of the system and suggesting appropriate control actions.

In the presented work, it is assumed that the balancing reserve is able to compensate for 99.5% of the system imbalance cases. A tail event is defined as when the balancing capacity needed is larger than the amount available. The study is focused on capacity requirements of the regulating reserve; however, other types of requirements, such as ramp rate and ramp duration [1], can also be analyzed in a similar fashion. All of the results shown in this paper were generated based on BPA 21-month historical and forecast data.

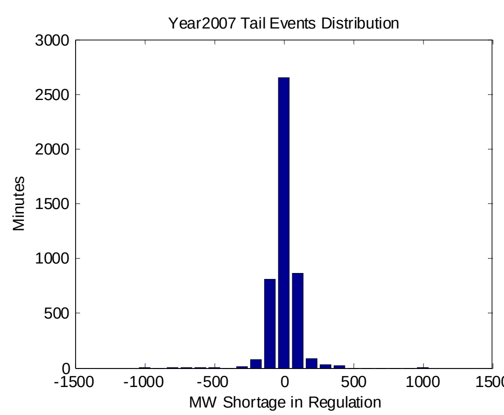
The report is organized as follows: In Section II and Section III, the yearly and hourly distribution plots, i.e., Part 1 and Part 2 of the methodology are described. In Section IV, the BN model, i.e., Part 3 of the methodology is introduced for the real-time prediction of regulation requirements and probability of tail events. Validation of the BN model prediction results were performed and are reported in Section IV as well. Section V concludes the report and is followed by references.

# II. YEARLY DISTRIBUTION OF TAIL EVENTS

Yearly distributions are based on the analysis of system regulation requirement, which is generated from BPA system load and wind data using the methodology developed in [1]. Available system regulating reserve is defined as the MW level that can cover 99.5% of the cases that have been simulated. Subtracting available regulating reserve from the regulation requirement, we get the MW shortage of the system. Fig. 1 and Fig. 2 show the distribution of the regulation capacity requirement and distribution of MW shortage in the BPA system in 2007 and 2010, respectively. The 2010 results were obtained based on forecasted load and wind data.

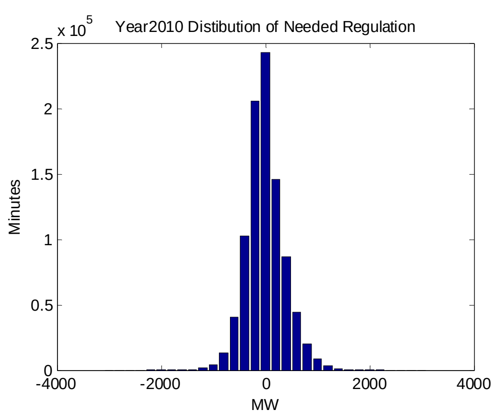



(a) Regulation

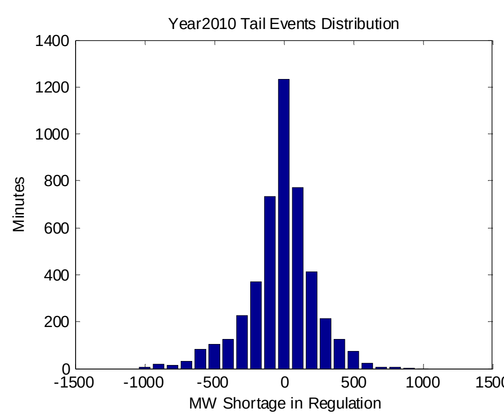



(b) Tail events

Fig. 1. 2007 BPA system regulation capacity requirement and tail event distribution




(a) Regulation




(b) Tail events

Fig. 2. 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 seen. For example, from plot (b) of Fig. 1 and Fig.2 it can be seen that in 2007, 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 a high level look at the regulation 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 regulating reserve.

# III. HOURLY DISTRIBUTION OF TAIL EVENTS

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

1. Data series of regulation capacity requirement are generated using the methodology developed in [1].
2. A time series representing the MW regulation capacity shortage, both regulation-up and regulation-down, is derived by subtracting the available regulating reserve from the data series of regulation capacity requirement.
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 regulation shortage.

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

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, as shown in Fig. 3.

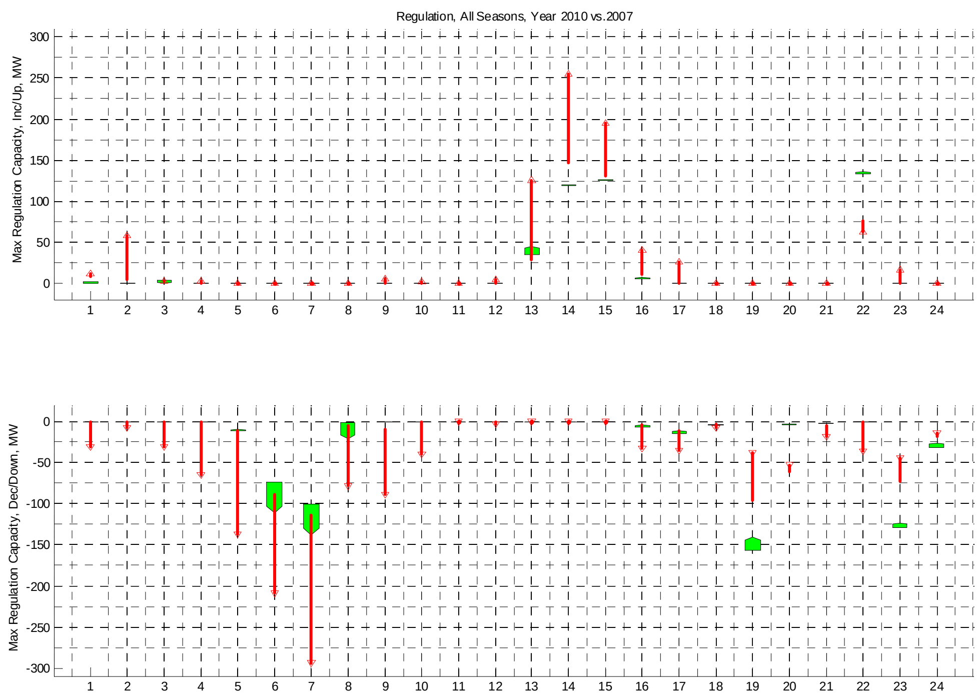



Fig. 3. Average regulation capacity shortage of all tail events: distribution corresponding to each hour during a day

In Fig. 3, the average tail event MW shortage 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. 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 BALANDING RESERVE REQUIREMENT AND TAIL EVENTS

## *4.1 Bayes Net Models*

Bayes net (also called Bayesian Network) models can be used to graphically represent the causal relationships amongst variables in which uncertainty is the predominant characteristic. The graphical representation consists of *nodes* and *directed links* or *arrows*. The nodes represent the variables and the arrows show the inter-dependencies between variables. Arrows point from *parent* nodes to the *child* nodes and show the direction of conditional dependence. The resulting structure of nodes and arrows forms a directed acyclic graph (DAG). Child nodes are conditionally dependent on their parents and are conditionally independent of their non-descendents given their parents.

Dynamic systems that change continuously through time, such as power system operations, can by modeled using a dynamic Bayesian Network (DBN) model where the state space of the system is modeled on successive time intervals. There are two simplifying assumptions that are typically used in constructing DBNs. One is stationarity—the probabilities 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 previous time periods. For a first order DBN they would depend only on the previous time period. Additional information on BN models can be found in [3], [4] and [5].

## *4.2 Building a Bayes Net Model for the BPA Power System*

A DBN model as a decision support tool for real-time operation of the BPA system was developed. The objective of this model, shown in Fig. 4, is to forecast the state of system imbalance (SI) in future time steps conditioned on the state of system components in the current time step. The uncertain nodes are depicted in the model as ovals. The model also extends from predicting SI to include decisions that might be made depending on the forecast system imbalance. These decisions consist of curtailment operations and are shown as rectangles. The model also identifies two relevant outcomes that result from system imbalance. These are line congestion and control performance standard (CPS) violations which are shown as hexagons.

As can be seen in Fig. 4, system imbalance is identified to have 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 the load and temperature at that time step and the wind/storm. LFE is strongly affected by diurnal and seasonal cycles, as well as meteorological events such as the passage of cold 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 at time (t). SCE is also predicted from relevant variables in a previous time step as shown in Fig. 4.

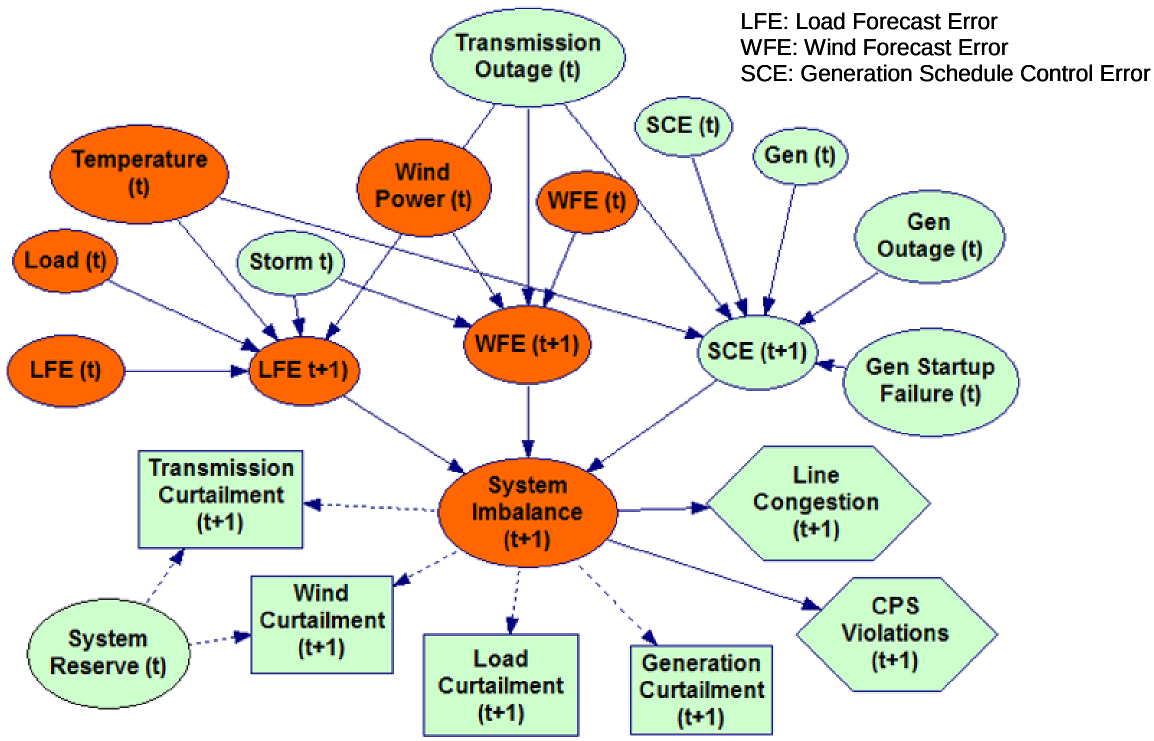



Fig. 4. Bayes net model for predicting system imbalance.

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 are deterministic in nature and are incorporated into the model to show their potential impact on LFE, WFE, SCE and subsequently system imbalance.

The conditional probability hierarchy shown in Fig. 4 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.

Till the time of writing this report, the model shown in red in Fig. 4 has been implemented. 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 Fig. 5.

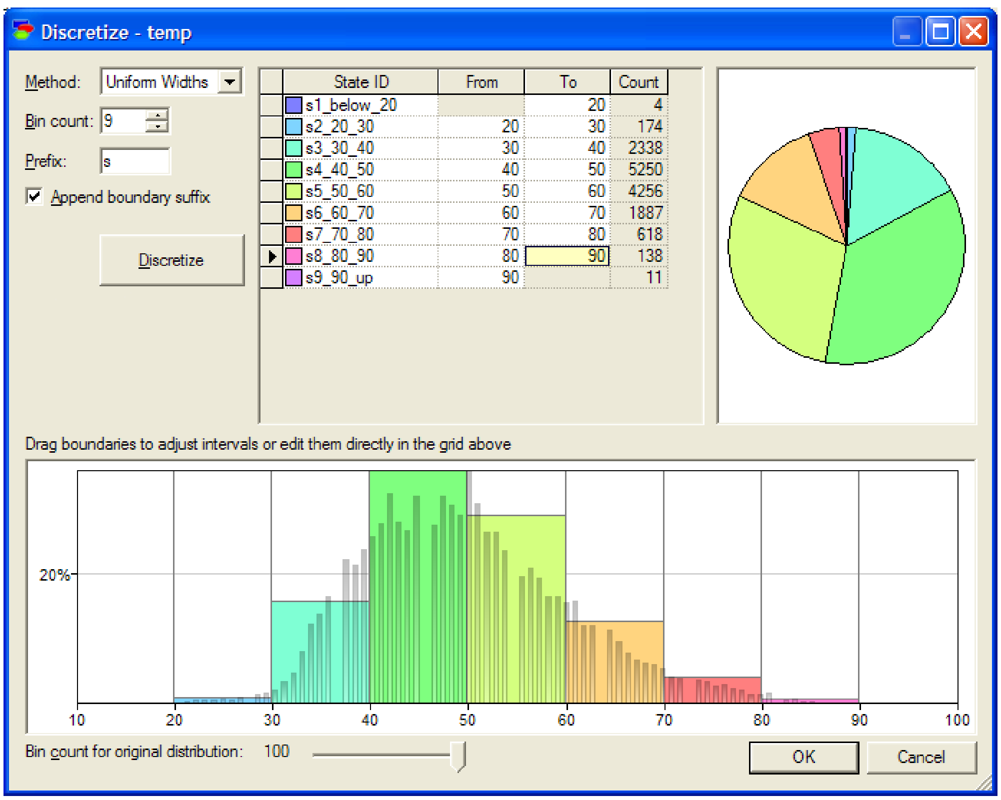



Fig. 5. Screen shot of discretized temperature using GeNIe software

### 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 [4].

Once the parameters were learned using the BN shown in red in Fig. 4, the probability distributions were then used to build a DBN as in Fig. 6. 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 time 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 Fig. 7. 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.

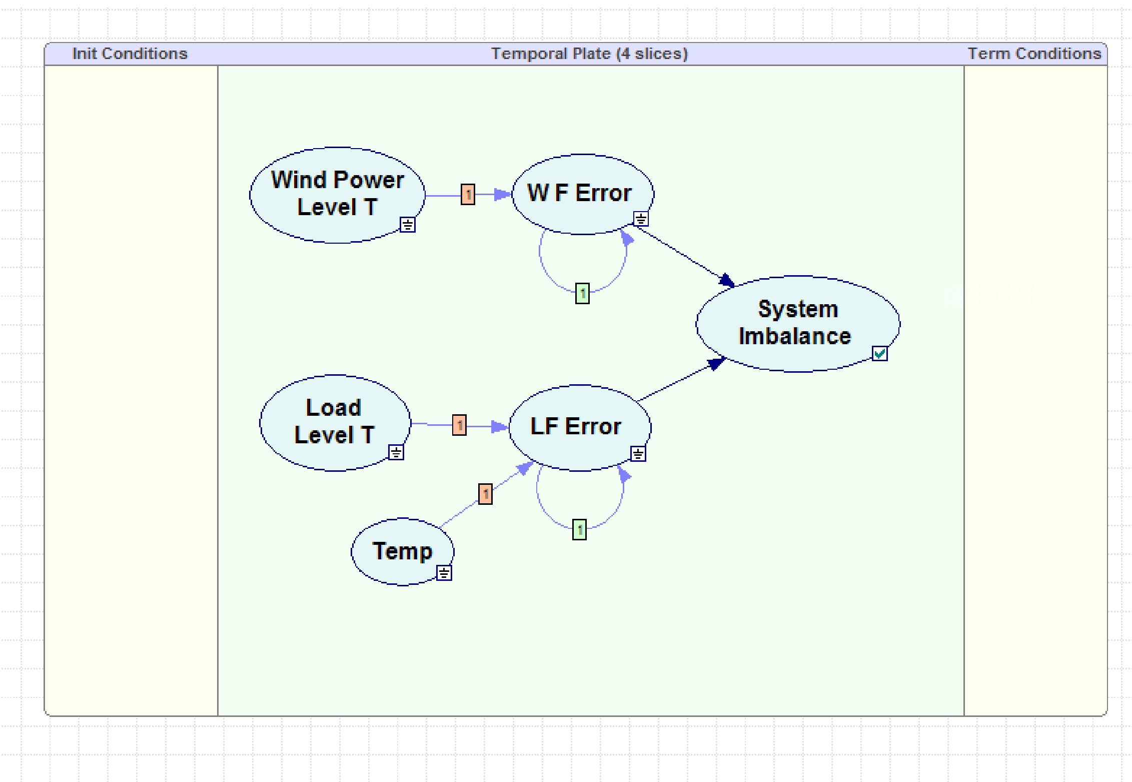



Fig. 6. Screen shot of DBN for predicting system imbalance as implemented in GeNIe.

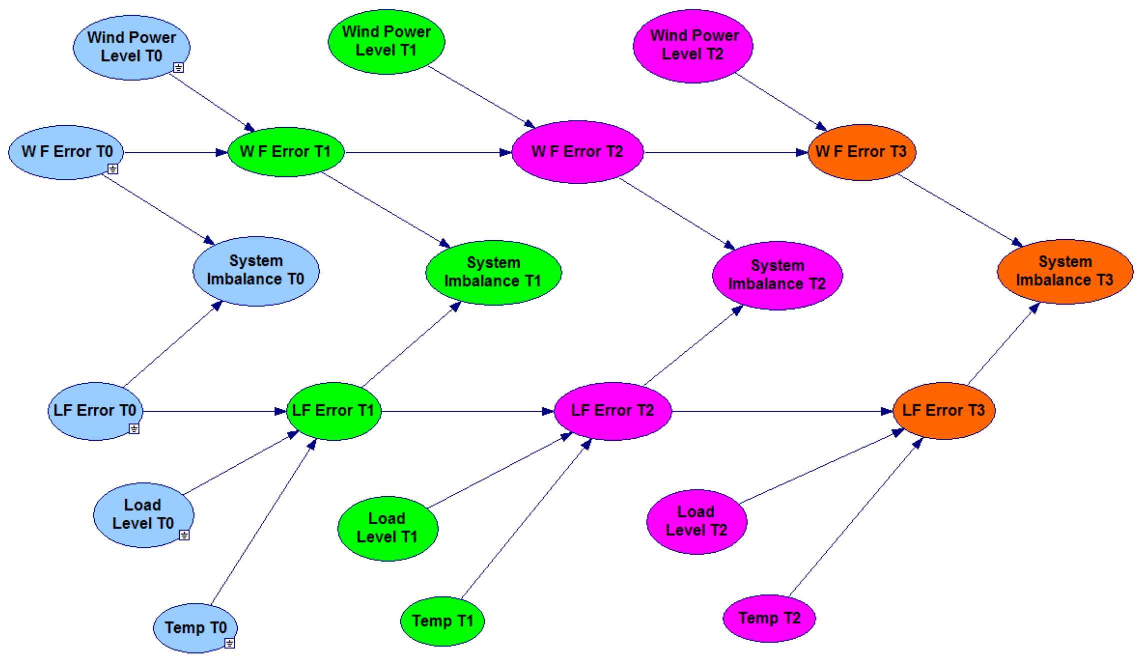



Fig. 7 First order dynamic Bayes net used to predict system imbalance.

### *4.5 Generation of Forecast Data*

Table 1 shows the initial eleven records used to run the GeNIe DBN model shown in Fig. 6. 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. Fig. 8 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 |

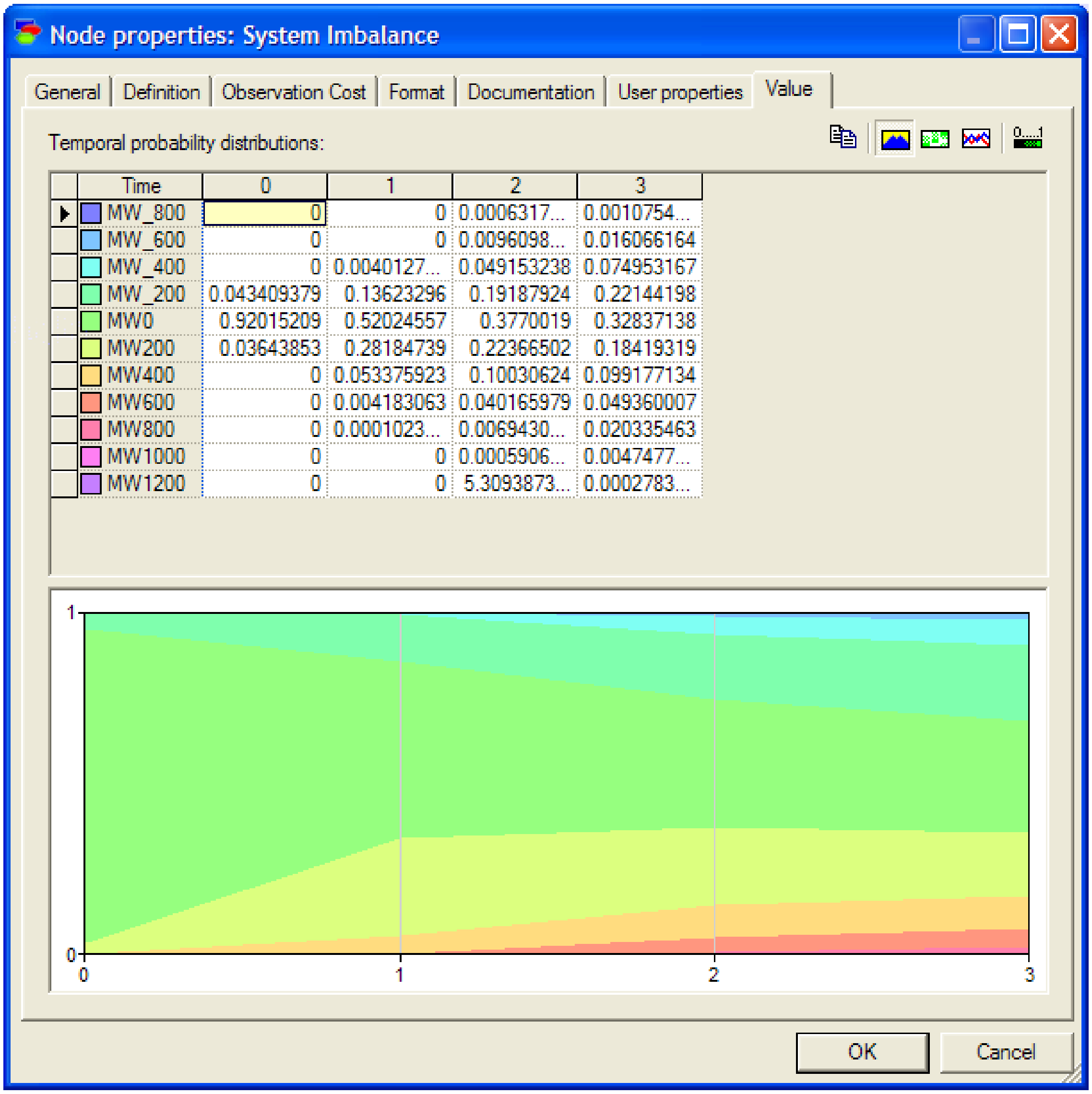



Fig. 8. Output generated by GeNIe for predicting system imbalance.

In Fig. 8, 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 Byes Net Model Output*

The output of the BN model shown in Fig. 6 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 1 hour. Fig. 9 and Fig. 10 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 Fig. 9 and Fig. 10 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.

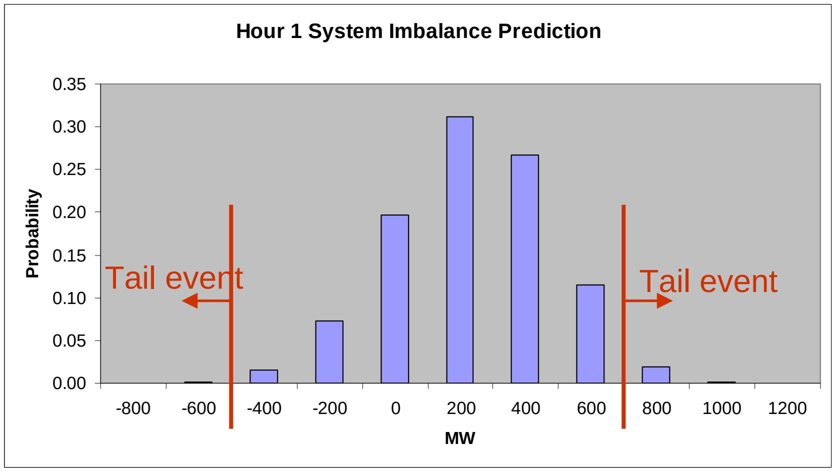



Fig. 9. Prediction of system imbalance for the next hour.

Fig. 9 shows that during the next operation hour, the probability of being short of generation is 0.16%, and the probability of being over generating is 1.95%. The most likely state of system imbalance is 200 MW, representing the interval between 100 MW and 300 MW.

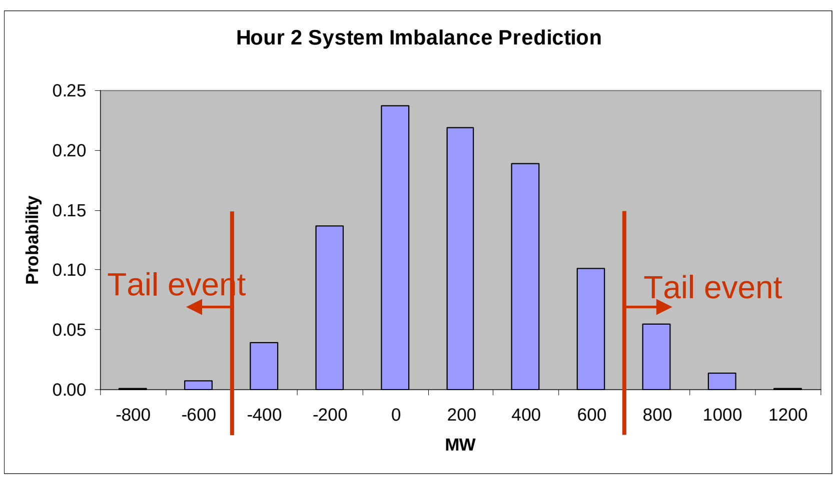



Fig. 10. Prediction of system imbalance for the second hour

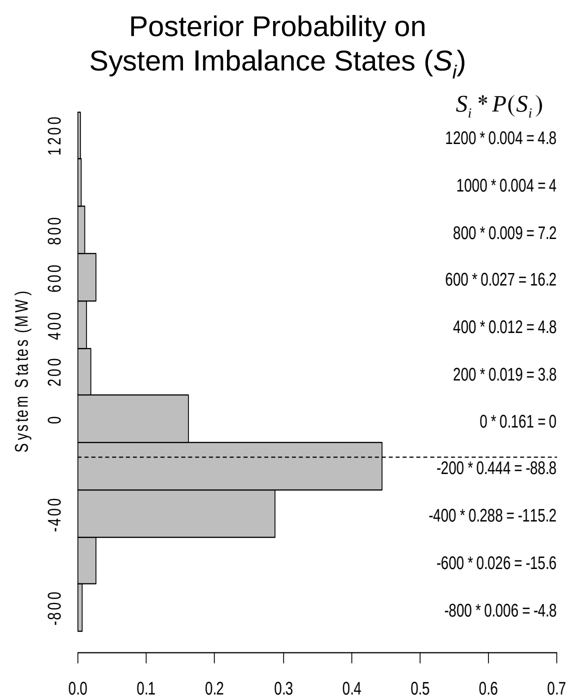
Fig. 10 shows that during the next operation hour, the probability of being short of generation is 0.08%, and the probability of being over generating is 6.95%. The most likely state of system imbalance is 0 MW, representing the interval between -100 WM and 100 MW.

The actual system imbalance is: Hour 0 (present hour) = 132 MW, Hour 1 = 333 MW, Hour 2 = -162 MW.

## *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≥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 diagramed in Fig. 11.




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

Fig. 11. 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 Fig. 12 and Fig. 13 for 1-, 2-, and 3-hour forecasts as mean absolute prediction error (MAPE). 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 Fig. 12 and Fig. 13. Errors were computed by subtracting the observed system imbalance from the NP and BN forecasted system imbalances, respectively. Smaller MAPE values indicate more accurate forecasts.

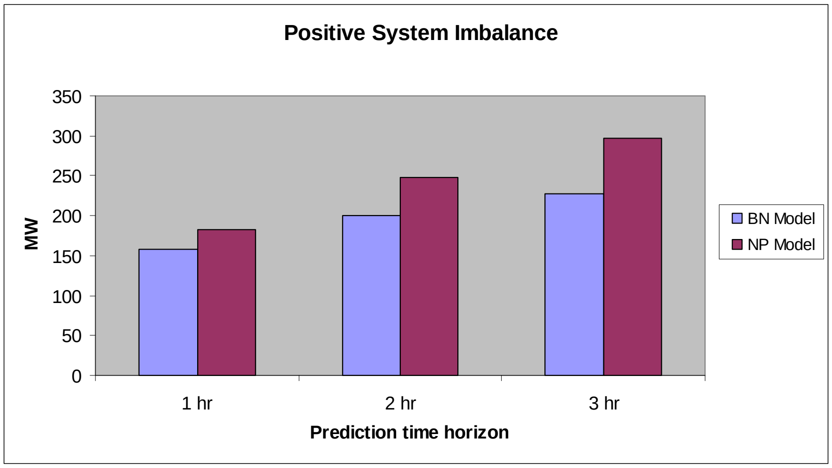



Fig. 12. Mean absolute prediction error comparison between BN and NP models forecasting at one, two and three hours: positive system imbalance cases

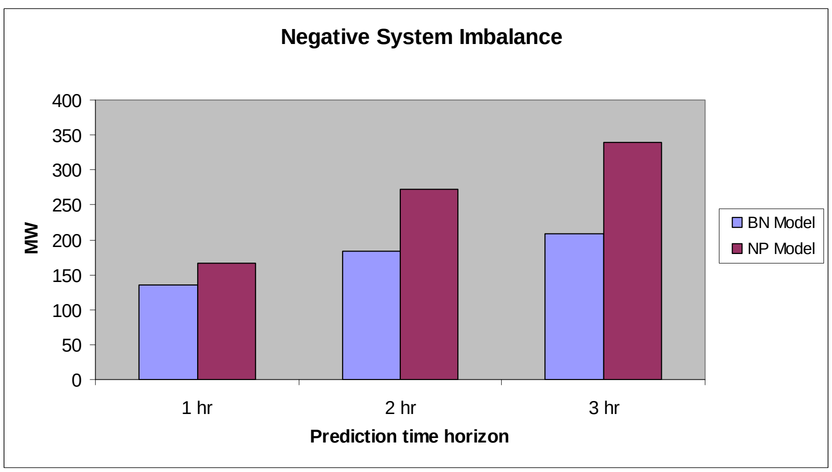



Fig. 13. Mean absolute prediction error comparison between BN and NP models forecasting at one, two and three hours: negative system imbalance cases

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 reserve 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 Fig. 14 and Fig. 15 for positive imbalance and negative imbalance cases, respectively.

Fig. 14 and Fig. 15 show that the observed probability of tail events does not match very well with the probability predicted by the BN model. There could be a high false alarm rate, with the predicted probability always higher than the actual observed. Nevertheless, there is a significant correlation between the two, which is more obvious in Fig. 15. 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.

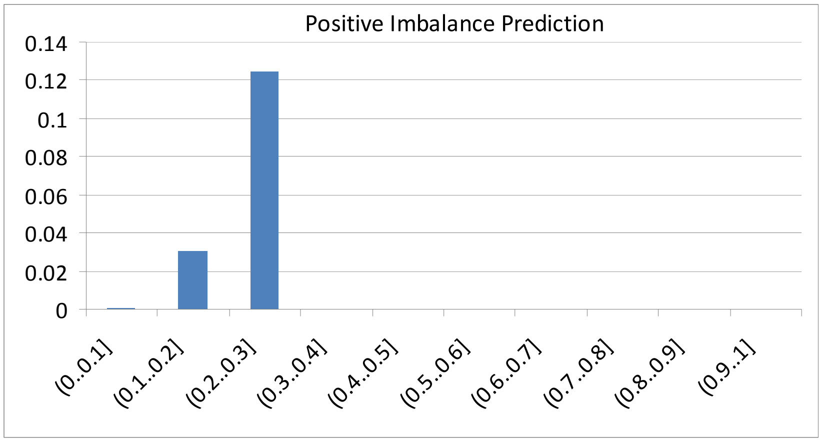



Fig. 14. Observed probability of tail events vs. predicted probability by the BN model: positive system imbalance cases

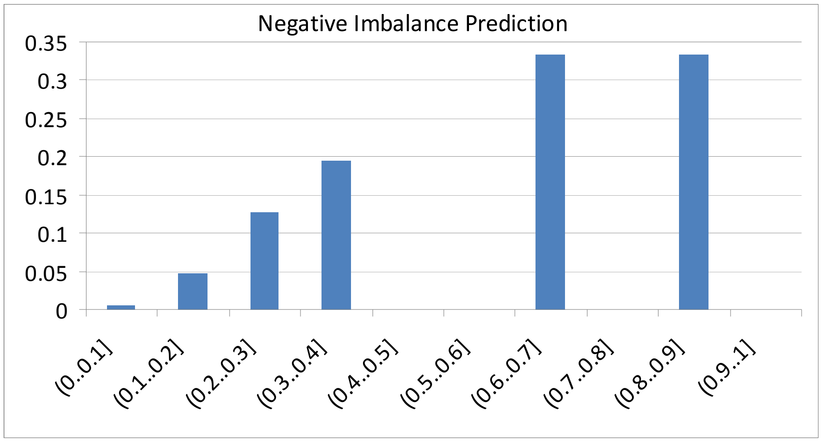



Fig. 15. Observed probability of tail events vs. predicted probability by the BN model: negative system imbalance cases

### *4.8 Potential Improvements on the Bayes Net Model*

The current BN model has a time step of 1 hour. In real-time operations, predictions with higher time resolution are usually 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 should 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.

# V. CONCLUSION

This report presents a methodology to analyze and predict the balancing reserve requirement and probability of tail event for power systems with intermittent resources. The methodology contains three parts:

1. Yearly distributions of balancing requirement and tail events, which show the level of system imbalance and the corresponding total period that can be seen in the system, providing a high level look at the issue and a quick estimate of risks associated with any particular reserve level.

2. Hourly distributions of balancing requirement and tail events, providing information on those hours when the sufficiency of balancing reserve should be carefully watched.

3. A model called Bayes net (BN), predicting the real-time need for regulating reserves of the power system and probability of tail event.

Accumulating system operating experience is a slow and long process for human system operators. The BN model adopted in the project study essentially "remembers" all cases that have happened in the past and uses this experience and current system status to generate an estimate of balancing requirements. It is similar to the process of an experienced system operator giving a prediction based on his own knowledge of the system, but in a more quantitative fashion. The model should be able to provide good reference information to system operators.

Based on the prediction of system imbalance, available reserves and system policy, the BN model can be extended to provide early warnings of load, wind or transmission curtailment, facilitating the coordination between different parties in the operation of the power system.

As penetration of intermittent sources increases in the system, the balancing reserve requirement is more variable. Determining whether the system will have sufficient reserve capacity in real-time operation becomes more difficult. The methodology presented in this report will help to address this problem.

# VI. REFERENCES


[1] 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.

[2] 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.

[3] Jensen, Finn V. and Thomas D. Nielsen, 2007, *Bayesian Networks and Decision Graphs*, Second Edition, Springer.

[4] Neapolitan, Richard E. 2004, *Learning Bayesian Networks*, Pearson Prentice Hall, Upper Saddle River, N.J.

[5] Russell, Stuart and Peter Norvig, 2003 *Artificial Intelligence: A Modern Approach,* Second Edition, Prentice Hall Pearson Education, Inc, Upper Saddle River, New Jersey.