NERC
NORTH AMERICAN ELECTRIC
RELIABILITY CORPORATION

Toward Ensuring Reliability: Reliability Performance Metrics

A NERC Staff White Paper

December 2007

Table of Contents

Executive Summary

From its beginnings in 1968, NERC has provided guidance and recommendations for the reliable planning and operation of the North American bulk power system. Today, NERC is the organization responsible for ensuring the reliability of that system through its legal authority to develop and enforce reliability standards as well as its other programs and activities.

The North American Electric Reliability Corporation's (NERC) mission is to ensure the bulk power system in North America is reliable. To achieve this objective, NERC develops and enforces reliability standards; monitors the bulk power system; assesses and reports on future reliability and adequacy; analyzes events for lessons learned; evaluates owners, operators, and users for reliability preparedness; and offers education and certification programs to industry personnel. NERC is a non-profit, industry self-regulatory organization that relies on the diverse and collective expertise of industry participants that comprise its various committees and sub-groups. It is subject to oversight by governmental authorities in Canada and the United States (U.S.).

In order for NERC's programs to be successful, it is important to track their influence on the reliability of the bulk power system. By defining various metrics and indices, it is possible to use amassed historical data to track the success of various initiatives and develop leading indicators and root causes of unreliable system performance based on past events. Until now, the industry has lacked an organized way to establish and track these metrics and indices.

The North American electric industry is at a critical juncture: the aging infrastructure and workforce will require billions of dollars of investment over the next decade; momentous changes in energy policy have made reliability standards mandatory; concerns about global warming have increased the likelihood of more environmental regulation; demand for electricity continues to rise faster than new supply capacity and transmission; and the availability of low carbon resources like natural gas threatens to decline. All of these issues – and many more – influence the future of electric reliability. The guidance these metrics and indices will provide is needed and necessary.

Metrics are a system of parameters or ways of quantitative and periodic assessment of a process that is to be measured.

An **index** is a single number calculated from an array of values or of quantities.

Leading indicators are indicators which tend to change before the general activity, and so may be used as a predictor.

Benchmark is a mark or a point of reference by which something is evaluated and measured.

This white paper proposes a plan to develop an advanced system for establishing those metrics, which will be used to:

- Measure:
    - Past and current reliability
    - Progress in ensuring reliability
    - Effectiveness of reliability standards and enforcement programs

- Identify
    - Factors that positively or negatively impact reliability
    - Reliability problems and solutions

Establishing reliability metrics requires three key elements, which this paper details and proposes for industry discussion:

- Selecting key metrics to quantify the reliability performance
- Identify reliability performance leading indicators from planning and operations data
- Assess the results to identify reliability performance trends

Metrics Objective

The goal of this white paper is to propose three main reliability indices, discuss Leading Indicators, investigate potential Adequate Level of Reliability metrics, and appointment of a Reliability Metrics Working Group. The white paper provides a conceptual framework for identifying performance data, leading indicators, and metrics that will transform data into meaningful information for judging the effectiveness of NERC's programs. Each chapter describes the proposed indices and their functions.

This document is part of a larger reliability metrics and benchmarking initiative of the Reliability Assessment & Performance Analysis program of NERC. The objectives of this NERC program are to develop guidelines for acceptable metrics, assess available metrics, maintain a performance metrics "dashboard" on NERC's Web site, and develop from analysis of these metrics appropriate reliability performance benchmarks.

This document also serves as a starting point for further benchmarking discussion. Section 809 (Reliability Benchmarking) of NERC's Rules of Procedure requires that NERC identify and track key reliability indicators as a means of benchmarking reliability performance and measuring reliability improvements.

Metrics Background

NERC initiated its work on reliability metrics in 2006 and established an initial Reliability Information Dashboard. Stakeholders were able to begin accessing the dashboard on the NERC website in early 2007 to view data that included the 10 most violated NERC Standards, the 10 entities cited for violating the greatest number of NERC Standards, and the capacity margin for each region. The dashboard is currently being revised and updated.

June 18, 2007 marked the beginning of mandatory and enforceable reliability standards in the United States, starting with 83 standards approved by the Federal Energy Regulatory Commission (FERC). Nearly 4,100 violations of the voluntary standards were self-reported prior to June 18. Notices of more than 300 violations of the standards were also received during the first month of standards enforceability.

How much risk is taken when the bulk power system is exposed to the 4,400 reported violations? Is the system becoming more or less reliable over time? How best to measure reliability management? The need for reliability performance metrics and benchmarking is ever increasing.

For assessment of the bulk power system reliability in this new era, NERC formed an internal Metrics & Benchmarking Team in July 2007 to develop key reliability indicators. In August 2007, the Team identified a set of 20 metrics. NERC management reviewed the team findings, and suggested two additional indices, which have been integrated into the current plans for development.

Key Terms and Data

Reliability
An electricity service level or the degree of performance of the bulk power system defined by accepted standards and other public criteria[1]. There are two basic, functional components of reliability: operating reliability and adequacy.

Operating Reliability
The ability of the electric system to withstand sudden disturbances such as electric short circuits or unanticipated loss of system components[2].

Adequacy
The ability of the electric system to supply the aggregate electrical demand and energy requirements of the end-use customers at all times, taking into account scheduled and reasonably expected unscheduled outages of system elements[3].

Standard Violations
A violation is a failure or inadequacy to meet a requirement of a reliability standard by an entity identified as responsible to comply with that requirement.

Risk
Each requirement within NERC Reliability Standards has been assigned a Violation Risk Factor (VRF) to provide clear, concise and comparative association between the violation of a requirement and the expected or potential impact of the violation on the reliability of the bulk power system. One of three defined levels of risk is assigned to each standards requirement: Lower Risk Factor, Medium Risk Factor, or High Risk Factor.

Severity
Violation Severity Levels (VSLs) are defined measurements of the degree to which a violator violated a requirement of a reliability standard. Whereas VRFs are determined pre-violation and indicate the relative potential impacts that violations of each standard requirement could pose to the reliability of the bulk power system, the VSLs are assessed post violation and are indicators of how severely the violator actually violated the standard(s) requirement(s) in question. Four VSLs have been designated for each standard requirement as: Lower, Moderate, High, or Severe.

---

[1] Not from NERC Glossary of Terms, derived for conceptual discussion in this paper only
[2] Definition of reliability http://www.nerc.com/~members/OC_PC/ALR/ as of December 12, 2007
[3] NERC Glossary of Terms ftp://www.nerc.com/pub/sys/all_updl/standards/rs/Glossary_02May07.pdf

Key Metrics

A reliability performance index is defined as a number showing the variation (increase or decrease) in reliability performance. Metrics are set of reliability measurements that quantify information, specifically units of performance, or reliability performance, such as operating reliability, adequacy, and compliance factors.

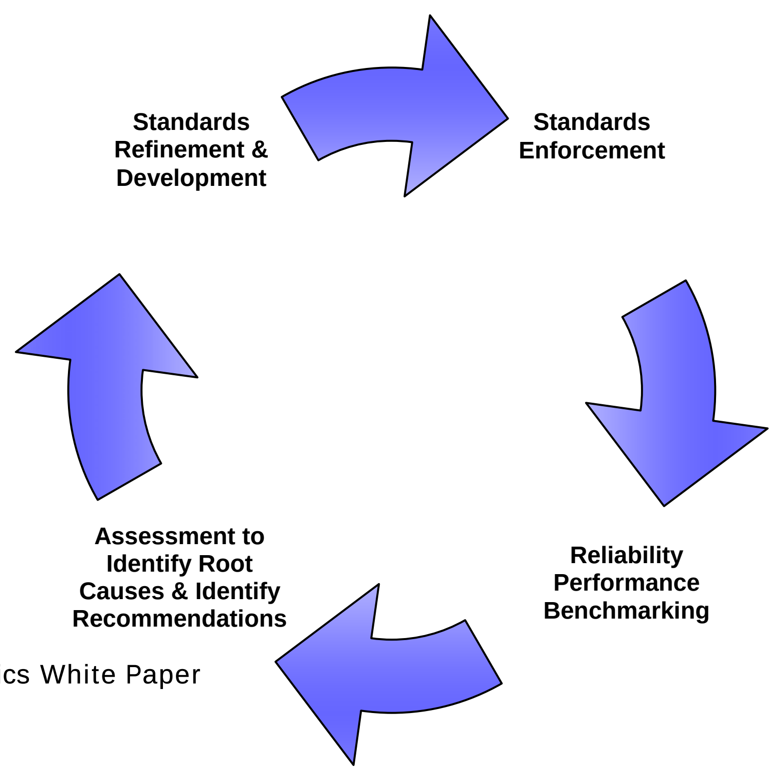
NERC initially proposes the development and use of three major indices as reliability performance metrics used to judge the relative performance of the bulk power system. The three indices are:

- **Reliability Performance Gap (RPG)**: designed to measure how far the system is from expected performance under contingencies (dynamic conditions).
- **Adequacy Gap (AG)**: designed to measure the capacity and energy shortage from expected adequacy level under steady state conditions.
- **Violation Index (VI)**: designed to measure the reliability improvement from compliance with NERC Reliability Standards.

These three indices are intended to capture and represent many complex reliability parameters into easy to understand reliability performance metrics. They provide clarity that reliability is NERC's strength and focus. NERC measures its success through the relative reliability improvements encapsulated in these metrics.

Through the reliability metrics/benchmarking cycle (Figure 1), NERC and its stakeholders can track performance and progress towards sustained reliability improvement. The event analysis results will be tied into standards refinement and development, be used in readiness evaluations, help identify training/education needs, and more. The defined metrics will also be used to measure effectiveness of reliability standards and NERC's compliance enforcement program.

Figure 1: Reliability Metrics/Benchmarking Cycle



Reliability Performance Gap (RPG)

The Reliability Performance Gap index (RPG) is designed to help measure reliability based on events (the ability to cope with contingencies) concerned with dynamic conditions.

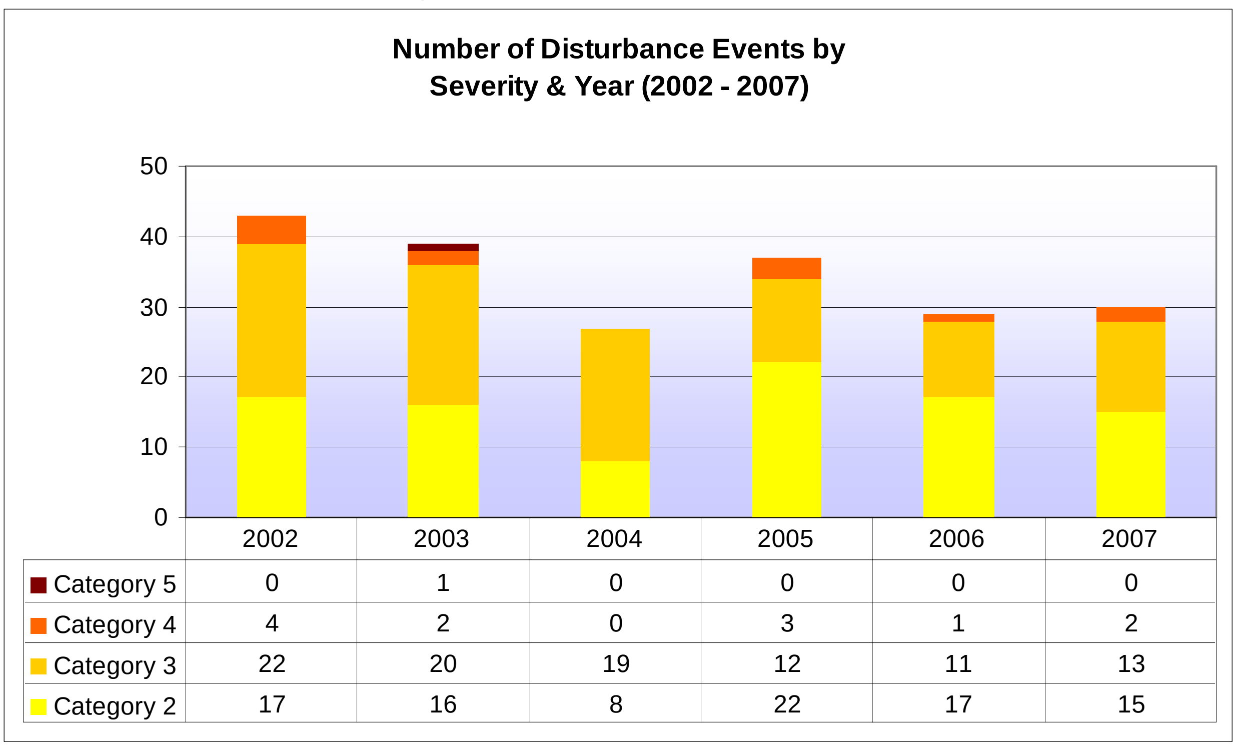
Sample Event Analysis Data

The number of disturbance events between 2002 and 2007 in NERC's Disturbance Analysis database can be complied by categories and trends. Results are shown in Table 1 and Figure 2. The events caused by factors other than the performance of the transmission system are not included. The event categories are defined in Appendix A, with Category 5 being the most severe.

Table 1 – Disturbance Event Trend

Categories	2002	2003	2004	2005	2006	2007
5	0	1	0	0	0	0
4	4	2	0	3	1	2
3	22	20	19	12	11	13
2	17	16	8	22	17	15

Figure 2 – Disturbance Event Trend

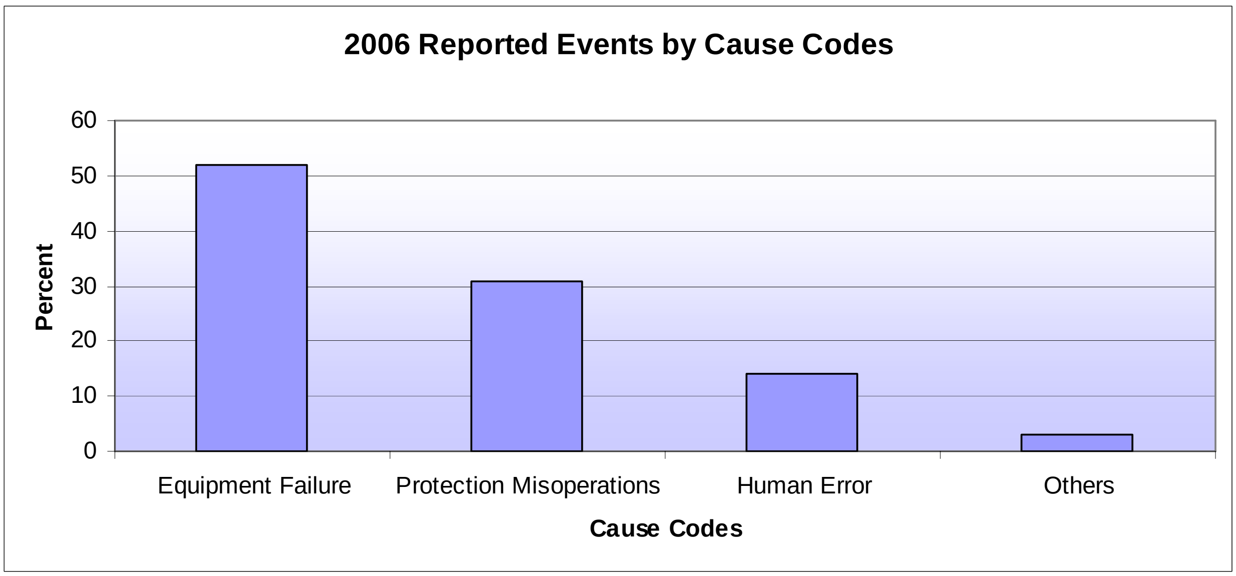


The RPG is defined as number of disturbance events in Categories 2 through 5 for a given period. It measures the change in system ability to cope with pre-studied contingencies and unforeseen disturbances. Ultimately, RPG indicates the most important measure of reliability standard and enforcement program effectiveness: the number of the events declining towards zero.

While the number of Category 2 events increased in 2005 and 2006, the number of events in Category 3 has continued the steady decline since 2002. The Category 4 events were up in 2005 and 2007 after declining in 2004 and 2006.

Of the 29 events in 2006, 15 (52%) occurred due to equipment failures, while system protection misoperations accounted for 31% and human errors accounted for 14%, as shown in Figure 3.

Figure 3 – Sample Event Root Cause Trend



NERC and the Regional Entities conduct compliance audits and event analyses. These programs are vital in identifying root causes and trends of past disturbances. With these findings, more resources can be devoted to proactively target the areas that experience and cause the greatest number of severe disturbance events.

NERC continues to broaden its effort to identify at-risk areas. Initiatives and strategies include a newly automated Event Tracking Tool and Alert System Process, designed to identify trends and disseminate findings and recommendations from the analysis of significant system events. NERC's goal is to reduce reliability performance gaps.

Adequacy Gap (AG)

The second index that NERC proposes is Adequacy Gap (AG) designed to measure capacity and energy deficiency (the ability to supply load) based on emergency events in steady-state conditions. The AG index is defined as number of capacity and energy emergency events for a study period. It is used to measure the change in system ability to supply load under steady-state conditions.

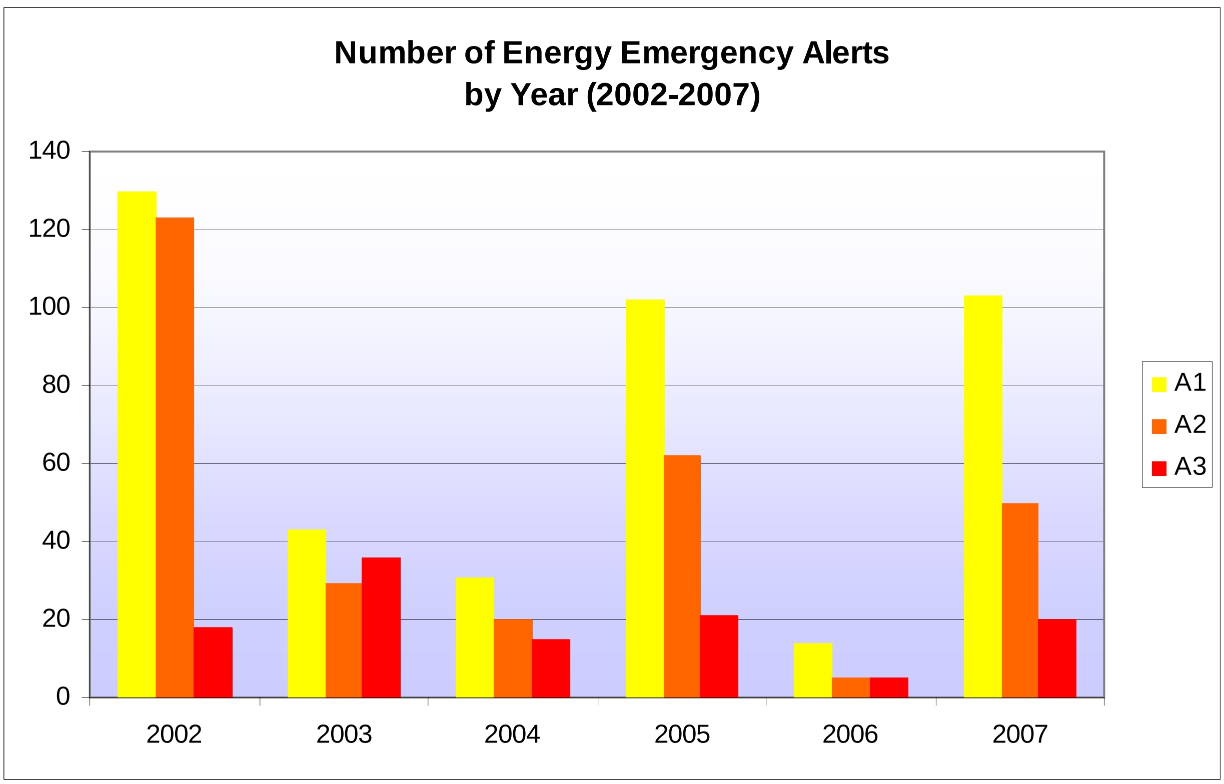
Sample AG Analysis using Real Event Data

The total number of capacity and energy emergency events between 2002 and 2007 in NERC's Reliability Coordinator Information System (RCIS) database can be sorted by categories and trends. Results are given in Table 2 and Figure 4. The emergency event categories A1, A2 and A3 are defined in Appendix A, with A3 being the most severe.

Table 2 – Adequacy Gap Trend

Categories	2002	2003	2004	2005	2006	2007
A3	18	36	15	21	5	20
A2	123	29	20	62	5	50
A1	130	43	31	102	14	103

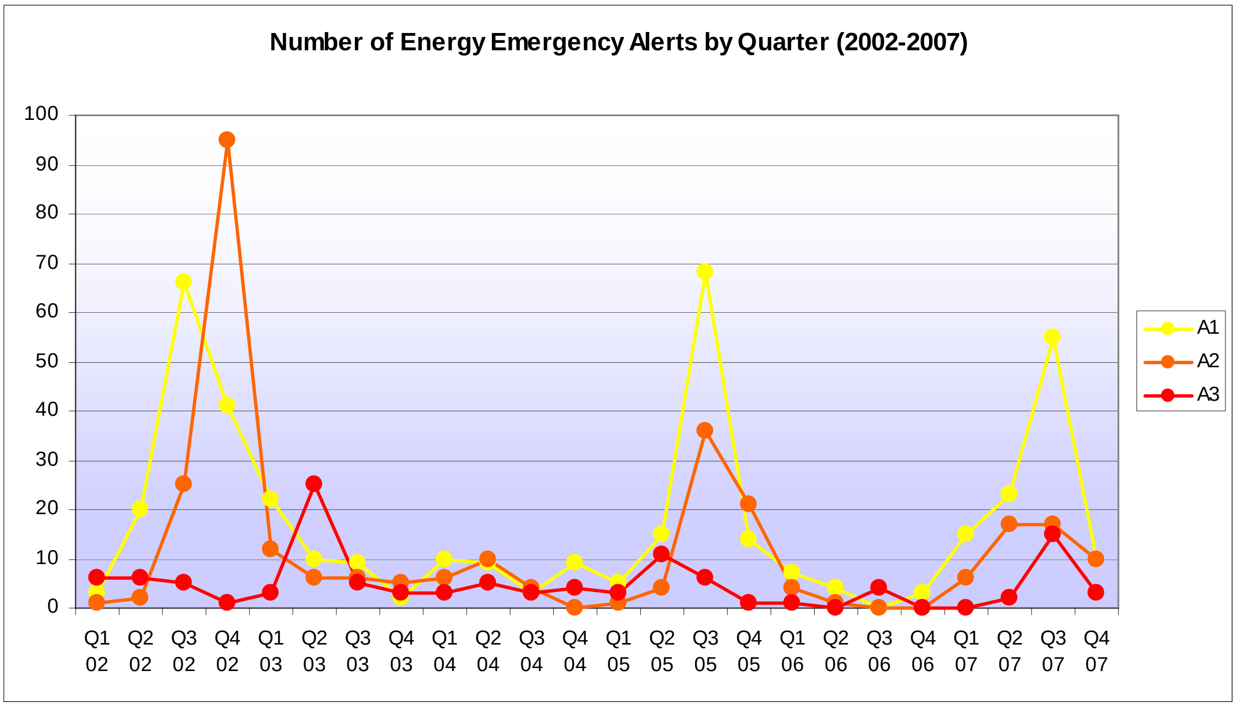
Figure 4 – Adequacy Gap Trend by Year



The 2002 – 2007 quarterly AG trend plots are provided in Figure 5. There is a seasonal pattern to the events over the year, with the 2006 summer weather particularly mild compared to other years. A clear periodicity is evident between the summer (Q3) and winter months (Q1 and Q4).

1Q – Winter (January, February, March)
2Q – Spring (April, May, June)
3Q – Summer (July, August, September)
4Q – Fall/Early Winter (October, November, December)

Figure 5 – Quarterly Adequacy Gap Trend



Event analysis has indicated that extreme weather, short-term load forecast errors and unplanned generation outages are main causes of the emergency events.

While this paper suggests how adequacy gaps may be formulated, more work is required to examine differences between planners' long-term assessment and actual adequacy performance. In order to produce a meaningful adequacy forecast, it is important to know how variations in actual generation availability, demand level and transmission constraints affect system adequacy performance so their impact and trend can be taken into account in adequacy assessment studies.

Violation Index (VI)

The third index NERC proposes is a metric measuring violations. The reliability performance data used in this metric would include data related to NERC's Violation Risk Factors and Violation Severity Levels. The key metric would be the Violation Index (VI) designed to measure improvement in the compliance with standards.

The VI for each entity is weighted based on each violation's:

- Violation Risk Factors (VRF)
- Violation Severity Levels (VSL)

A risk factor and severity level weighted-violation average can determine the change in reliability levels due to confirmed standard requirement violations. The weighting values can be derived by applying similar ratios developed in the NERC Sanction Guidelines[4] to assess the potential consequences of a particular violation. The sample weighting factors are included in Appendix B.

Violations of a higher risk factor requirement have a higher weighting value in the index than violations of lower risk factor requirements. The index decreases if the compliance improvement is achieved over a trending period.

For example, Table 3 lists the 3,419 confirmed violations with their Risk Factors and Severity Levels in NERC's compliance database (as of Oct. 19, 2007). Using the weighting factors in Appendix B, the VI is 187.27.

Table 3 – Number of Violations in second quarter of 2007

	Violation Severity Level			
Violation Risk Factors	**Lower**	**Moderate**	**High**	**Severe**
Lower	950	92	96	408
Medium	809	48	144	250
High	374	32	36	180

Since the second quarter of 2007 marked the beginning of mandatory and enforceable standards, the VI can be benchmarked with a base value of 100. All future VI values will be normalized based on this reference quarter.

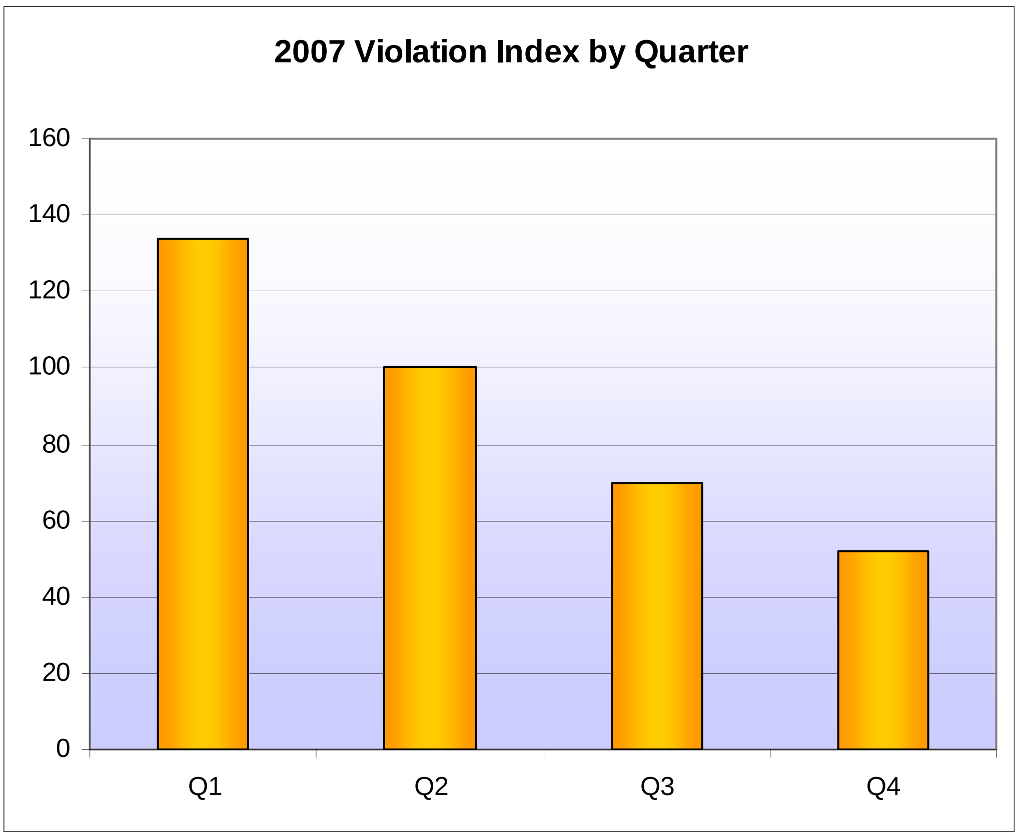
Other factors can also be included in the Violation Index calculation. The time horizons and size of an entity could be considered as additional weighting factors. Violations involving immediate

[4]Html address: ftp://www.nerc.com/pub/sys/all_updl/rop/Appendix4B-SanctionGuidelines.pdf

or real-time activities will have higher weighting factors than violations with longer time horizons. The more transmission equipment an entity owns/operates, the higher influence a violation could have on system reliability.

If the VI decreases over time, as shown in Figure 4, the Reliability Performance Gap index and Adequacy Gap index should become smaller as well. From event analysis results, root causes and recommendations will lead us to identified weak areas, worst performance and enabling recommendations for changes to Standards that should help minimize the occurrence of disturbance events.

Figure 4: Sample Graph Violation Index[5]



[5] For illustration only. Violations for Q1, Q3 and Q4 are sample data, not from the NERC's compliance database.

Leading Indicators

Reliability leading indicators forecast reliability changes. The indicators include the effects of human activities as well as the influence of equipment difficulty on the reliability performance.

The objective is to recognize and eliminate unreliable actions and at-risk conditions. If great attention is not given to the observation of unreliable performance, sooner or later disturbance events will become more severe. This is why not only events in Categories 2-5, but also near misses in Category 1 should be reported and analyzed. NERC is concerned about the number of unreliable situations leading to severe events.

When a near miss is observed, the root cause needs to be determined for this at-risk situation. As one observes, questions, and looks for unreliable situations, we can recognize and eliminate the causes of standard violations and disturbance events. The aim is to make problems go away so that they do not come back. NERC's goal is to have not one, but consecutive years with no events in Categories 2-5.

A good indicator measures how far reliability performance is from its goal and whether we are headed in the right direction. Choosing the right indicator is essential for effectively evaluating progress and should:

1. be relevant to the goal.
2. be easily understood.
3. be easily measured with regularly collected information.
4. provide meaningful information.

Table 4 lists proposed reliability and adequacy leading indicators and illustrates the relationship between indicators, objectives and the goal.

Table 4 – Potential Leading Indicators

Indicators for RPG	**Objective**	**Linkage to the Goal**
Percentage of the loss of a bulk power transmission component beyond recognized criteria, i.e., single phase line-to-ground fault with delayed clearing, line tripping due to growing trees, etc.	To minimize relay misoperations and tree contacts	Faulty protection systems and ineffective vegetation management are among root causes of blackouts.
Number of frequency excursions outside the Low or High Frequency Trigger Limits (FTLs) more than 5 minutes for a given period	To track real power balancing control performance	Degradation of regulation and reserve sharing capability decrease system recovery performance.
Minutes of an inter-area oscillation for a given period	To preserve stability of the system	A wide area oscillation could cause uncontrollable cascading outages.
Indicators for AG	**Objective**	**Linkage to the Goal**
Capacity margin up to 10 years	To address concerns with resource adequacy	Inadequate resources have negative impacts on reliability performance.
Demand assessment (demand side, load forecast and weather extremes) up to 10 years	Demand forecast is a fundamental element to transmission expansion planning	Effective use of demand side management improve reliability

Adequate Level of Reliability Indices

The next tier of metrics that NERC proposes for establishing Reliability Performance Benchmarks are metrics for determining an adequate level of reliability. The reliability performance data used for this metric could include, but is not limited to, data related to system limits and actual system conditions. The Adequate Level of Reliability Indices (ALRI) is designed to help measure reliability based on bulk power system characteristics, operations data and planning assessments.

An *Adequate Level of Reliability* is defined as when the bulk power system is planned and operated with the following System characteristics[2]:

1. Is controlled to stay within acceptable limits during normal conditions;
2. Performs acceptably after credible Contingencies;
3. Limits the impact and scope of instability and Cascading Outages when they occur;
4. Facilities are protected from unacceptable damage by operating them within Facility Ratings;
5. Integrity can be restored promptly if it is lost; and
6. Has the ability to supply the aggregate electric power and energy requirements of the electricity consumers at all times, taking into account scheduled and reasonably expected unscheduled outages of system components

To access and monitor the actual level and improvement of reliability performance and plans, NERC proposed the following metrics to measure each characteristic of the ALR and trend change patterns over operating and planning timeframes.

Characteristic #1 — The System is controlled to stay within acceptable limits during normal conditions

Metrics proposed:

- Number of frequency disturbance occurrences
- Percent of Reliability Coordinator real power reserve
- Number of emergency operations
- Number of System Operating Limits (SOL) and Interconnection Reliability Operating Limit (IROL) violations and near misses (less than 30 minutes)
- Number of Transmission Loading Relief Level 5 events

Characteristic #2 — The System performs acceptably after credible Contingencies

Metrics proposed:

- TADS event and outage data
- Disturbance event analysis data (Category 1 – 5)

Characteristic #3 — The System limits the impact and scope of instability and Cascading Outages when they occur

Metrics proposed:

- Disturbance event analysis data
- Transmission limitations for peak conditions from long term assessment

Characteristic #4 — The System's Facilities are protected from unacceptable damage by operating them within Facility Ratings

Metrics proposed:

- Disturbance event analysis data

Characteristic #5 — The System's integrity can be restored promptly if it is lost

Metrics proposed:

- Lost MW due to disturbance events
- Average restoration time following an event
- TADS: Mean Time To Repair (MTTR)

Characteristic #6 — The System has the ability to supply the aggregate electric power and energy requirements of the electricity consumers at all times, taking into account scheduled and reasonably expected unscheduled outages of system components

Metrics proposed:

- Capacity and Energy Emergency event data
- Capacity Margin up to 10 years
- Demand assessment data, including DSM trend

Next Steps

We propose the following steps and accompanying timeline shown in Table 4. Since data collection and analysis is the key to a valuable and successful metrics/benchmarking program, a group of industry experts with strong operations, planning and statistics background is required to advise and support the program needs.

A Reliability Metrics Working Group will be formed under the Planning Committee to advise and assist NERC staff in performing the following tasks:

1. Vet the white paper's concepts and incorporate comments received from stakeholders,
2. Develop general metrics of the Adequate Level of Reliability (ALR),
3. Define the measure, including formula or methodology for calculation, and
4. Identify data collection and reporting guidelines
5. Recommend metrics implementation plan

Table 4 Next Steps and Timelines

No.	Steps	Timeline
1	**Establish an initial dashboard.** Use the available data to establish an initial dashboard with event trends and compliance index.	December, 2007
2	**Appoint Reliability Metrics Working Group.** Form a working group with subject matter experts and vet the white paper's concepts and incorporate comments received from stakeholders,	December 2007 – February, 2008
3	**Define reliability metrics and methodology** Submit proposals and implementation plans for comment and suggestions to Planning and Operating Committees.	Starting 2008

Appendix A – Event Classifications

Events are categorized into two general classifications: disturbance events and capacity/energy emergency events.

Categories of Disturbance Events

Disturbance events are the disturbances that significantly affect the integrity of interconnected system operations. They are divided into 5 categories to take into account their different system impact.

Category 1 — An event results in any or combination of the following actions

- The loss of a bulk power transmission component beyond recognized criteria, i.e., single-phase line-to-ground fault with delayed clearing, line tripping due to growing trees, etc.
- A frequency below the Low Frequency Trigger Limit (FTL) more than 5 minutes
- A frequency above the High FTL more than 5 minutes
- An inter-area oscillation

Category 2 — An event results in any or combination of the following actions

- The loss of multiple bulk power transmission components
- SPS or RAS misoperation
- The loss of generation (between 1,000 and 2,000 MW in the Eastern Interconnection or Western Interconnection and between 500 MW and 1,000 MW in the ERCOT Interconnection)
- The loss of an entire generation station or 5 or more generators
- The loss of an entire switching station (all lines, 100 kV or above)
- The loss of dc converter station
- The occurrence of an islanding (weakly tied to the Interconnection in steady-state)

Category 3 — An event results in any or combination of the following actions

- The loss of generation (2,000 MW or more in the Eastern Interconnection or Western Interconnection and 1,000 MW or more in the ERCOT Interconnection)
- The loss of load (less than 1,000 MW)
- UFLS or UVLS operation

Category 4 — An event results in any or combination of the following actions

- The occurrence of an interconnected system separation or islanding
- The loss of load (1,000 to 9,999 MW)

Category 5 — An event results in any or combination of the following actions

- The occurrence of a blackout
- The loss of load (10,000 MW or more)

Categories of Capacity and Energy Emergency Events

The capacity and energy emergency events are divided into three main categories, exemplified in the three categories based on the Standard EOP-002-0 (Capacity and Energy Emergencies):

Category A1: All available resources in use

- Required operating reserves can not be sustained
- Non-firm wholesale energy sales have been curtailed

Category A2: Load management procedures in effect

- Public appeals to reduce demand
- Voltage reduction
- Interruption of non-firm end use loads in accordance with applicable contracts
- Demand-side management
- Utility load conservation measures

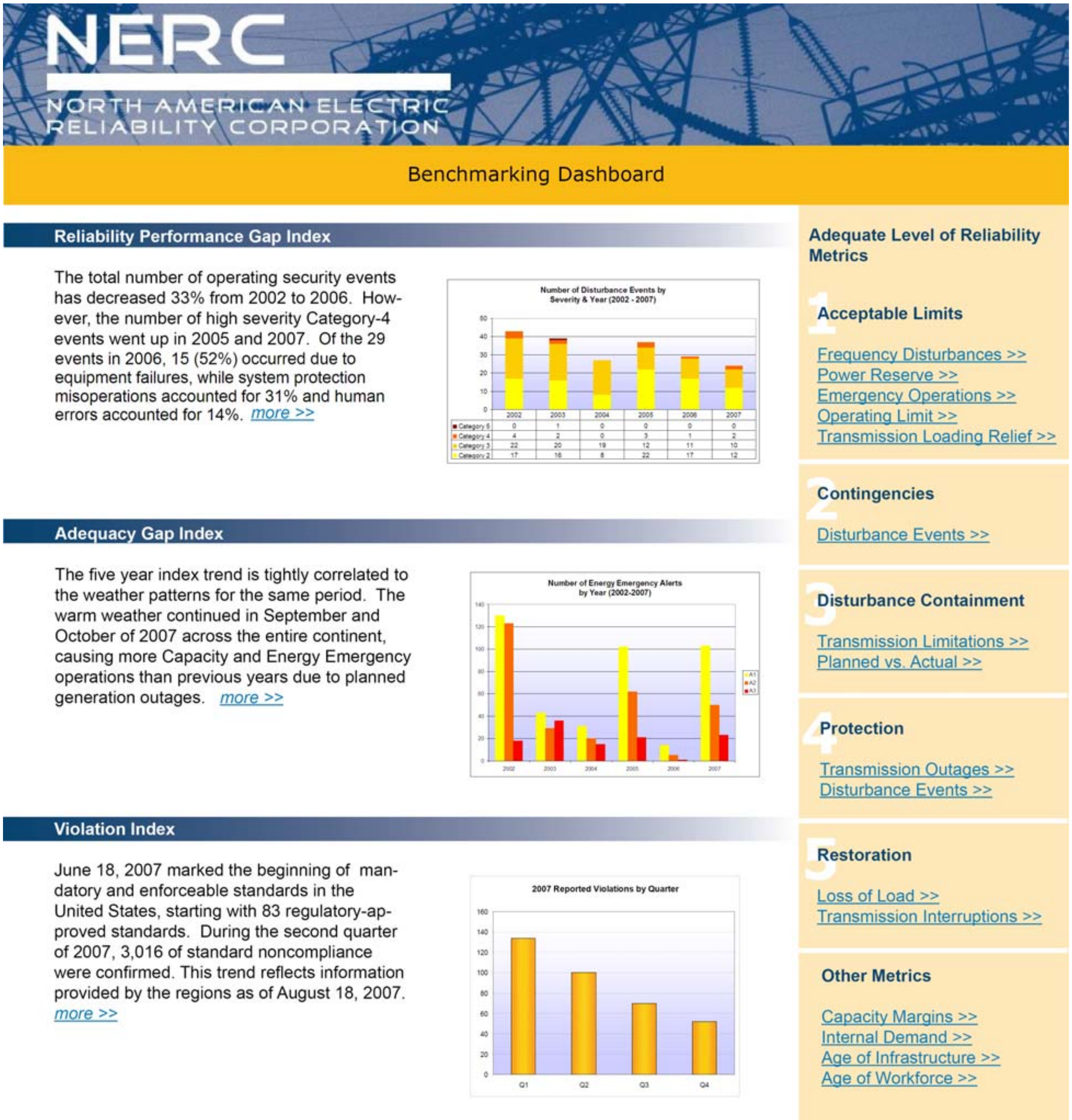
Category A3: Firm load interruption imminent or in progress

Appendix B – Weighting Factor Table

The following lists the weighting factors corresponding to combinations of Violation Risk Factors and Violation Severity Levels.

	Violation Severity Level			
Violation Risk Factors	**Lower**	**Moderate**	**High**	**Severe**
Lower	0.012987	0.025974	0.038961	0.06493
Medium	0.025974	0.051948	0.077922	0.12987
High	0.051948	0.103896	0.155844	0.25974

Appendix C – Display Design of Benchmarking Dashboard



NERC
NORTH AMERICAN ELECTRIC RELIABILITY CORPORATION

Benchmarking Dashboard

Reliability Performance Gap Index

The total number of operating security events has decreased 33% from 2002 to 2006. However, the number of high severity Category-4 events went up in 2005 and 2007. Of the 29 events in 2006, 15 (52%) occurred due to equipment failures, while system protection misoperations accounted for 31% and human errors accounted for 14%. more >>

	2002	2003	2004	2005	2006	2007
Category 5	0	1	0	0	0	0
Category 4	4	2	0	3	1	2
Category 3	22	20	19	12	11	10
Category 2	17	16	8	22	17	12

Adequacy Gap Index

The five year index trend is tightly correlated to the weather patterns for the same period. The warm weather continued in September and October of 2007 across the entire continent, causing more Capacity and Energy Emergency operations than previous years due to planned generation outages. more >>

Violation Index

June 18, 2007 marked the beginning of mandatory and enforceable standards in the United States, starting with 83 regulatory-approved standards. During the second quarter of 2007, 3,016 of standard noncompliance were confirmed. This trend reflects information provided by the regions as of August 18, 2007. more >>

Adequate Level of Reliability Metrics

1 Acceptable Limits

Frequency Disturbances >>
Power Reserve >>
Emergency Operations >>
Operating Limit >>
Transmission Loading Relief >>

2 Contingencies

Disturbance Events >>

3 Disturbance Containment

Transmission Limitations >>
Planned vs. Actual >>

4 Protection

Transmission Outages >>
Disturbance Events >>

5 Restoration

Loss of Load >>
Transmission Interruptions >>

Other Metrics

Capacity Margins >>
Internal Demand >>
Age of Infrastructure >>
Age of Workforce >>

NERC
NORTH AMERICAN ELECTRIC RELIABILITY CORPORATION

Benchmarking Dashboard

Adequacy Gap Index

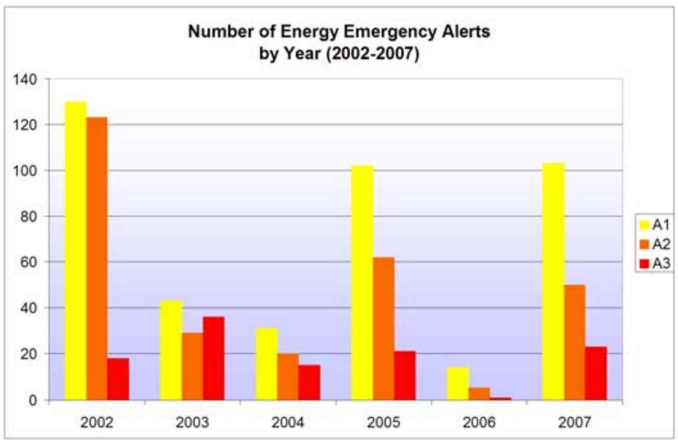


Chart View Options

Pull Down Menu

The five year index trend is tightly correlated to the weather patterns for the same period. The warm weather continued in September and October of 2007 across the entire continent, causing more Capacity and Energy Emergency operations than previous years due to planned generation outages.

About this Metric

Adequacy Gap (AG) is designed to measure capacity and energy deficiency (the ability to supply load) based on emergency events in steady-state conditions. The AG index is defined as number of capacity and energy emergency events for a study period. It is used to measure the change in system ability to supply load under steady-state conditions.

Event analysis has indicated that extreme weather, short-term load forecast errors and unplanned generation outages are main causes of the emergency events.