GADS Database management project – Task details

Revision history

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Revision number** | **Issued on** | **Created / Modified by** | **Checked by** | **Approved by** |
| Draft |  | ER |  |  |

# Introduction

The Generating Availability Data System (GADS) is a data reporting standard for utility generators begun in 1982. GADS Wind is the reporting format for wind generators. Wind farms with an installed capacity of 75 MW of more and commissioned after January 1st, 2005 are required to report GADS data.

WEICan receives funding from NRCan (Natural Resources Canada) to collect and analyse GADS data. WEICan’s role is to collect annual GADS data from project participants and to produce annual reports for each wind farm. The annual reports compare the farm’s performance to a national average on several factors. CanWEA (Canadian Wind Energy Association) facilitate the process of gathering data.

# Aims of this project

1. Set up a version control system to manage project code.
2. Automate the processing and analysis of GADS data files.
3. Develop a database management system that allows for managing data categories. This included adding data, checking for duplicates, flagging invalid data, filling in missing data, reimporting data files etc
4. Developing a front-end for the database that interfaces with the database. It should allow for generating reports on the whole dataset and with filters applied.
5. Project documentation including software architecture and error handling

# Relevant standards

GADS Wind Turbine Generation – Data reporting instructions – January 1 2018 – Version 2

GADS Wind Turbine Generation – Data reporting instructions – January 1 2017 – Version 1.2

The 2017 version is the most current standard. Some farm operators submit data according to older standards such as:

GADS Wind Turbine Generation – Data reporting instructions – January 1 2016

GADS Wind Turbine Generation – Data reporting instructions – January 2011 – Version 1.1.0

Data Reporting Instructions

# GADS project workflow

This section briefly describes how data flows between the project participants and WEICan.

1. Wind farm owners sign an NDA (Non-disclosure agreement) with WEICan. ([Link to example NDA](file:///\\weicserver\Docs\Projects\5032%20-%20GADS\Reports\Report%202015%20Data\NRCan\Appendix%202%20-%20CanWEA%202016%20GADS%20Package.pdf)).
2. Once every year (at minimum), the wind farm operator submits GADS data for each of their wind farms to WEICan. Farm data is grouped according to parameters such as turbine manufacturer, construction year etc. For the NERC classification of groups and sub-groups, see the 2017 GADS wind standard, Page 3.

The process of generating this GADS data depends on several factors such as the owner’s SCADA system, access to data and mapping turbine manufacturer fault codes to GADS fault codes. Compiling this data sometimes produces errors or leaves out critical information.

1. After farm owners send GADS data to WEICan, WEICan checks the data for formatting errors and also checks the plausibility of reported values. NERC’s GADS standard was revised in 2015 and in 2017. This adds some complexity to reported data as not all farm owners submit data according to the latest standard.
2. WEICan then calculates GADS parameters for each farm. The results of the per farm calculation is used to calculate the same scores averaged across all farms. Note that some calculations are averaged across all farms, some across all turbines and some per megawatt.
3. The calculation results are used to generate per-farm reports. The per-farm report shows the performance of the farm on the GADS parameters. It also compares the performance of the farm to the national average. National data is anonymised and averaged so that no identifiable information is retained.
4. Per farm reports are submitted to individual farm owners. A separate report showing a national average is submitted to CanWEA and NRCan.



# Current setup

The current setup uses a series of Excel (VBA) macros to import participant’s data and calculate results. With the current setup, significant amounts of manual work are required. Data must be in a fixed format and any deviations will produce errors or incorrect results. Managing the database is also done manually. Modifying data is difficult and can result in duplicate entries. Some inputs are hard-coded and identifying them is difficult. The VBA code uses fixed cell references which causes two problems. One, code is difficult to read because it is not immediately clear what data is in a particular cell. Two, because cell references are fixed (static), changing the position of data causes calculation errors.

A brief description of the current GADS code can be found at [this link](file:///W:\Projects\5032%20-%20GADS\2015%20GADS\GADS%202016%20Operating%20Manual.docx).

# Participants’ data submission requirements

In order to participate in this program, after signing the NDA, data files are to be submitted to WEICan in accordance with the most recent revision of the GADS Wind Turbine Generation Reporting Instructions. The most up to date version can be found on the NERC website at: <http://www.nerc.com/pa/RAPA/gads/Pages/default.aspx>.

The main data to be submitted includes power generation data and records of downtime. Each outage occurrence has its duration assigned to one of 132 numerical GADS codes by the wind farm operator. Outage hours and occurrences are split into 3 categories: forced, maintenance and planned. A more detailed explanation is available in the NERC reporting instructions. When transferring data from a SCADA system to GADS, contributors should take care to report to the most suitable code as the quality of each user’s reporting is directly reflected in report quality. If there is uncertainty about which category an outage should be assigned to WEICan can be contacted for support. A list of common submission errors found by WEICan can be found in appendix A.

When submitting data, 3 files must be generated; sub-group, performance and component. The contents of each file are outlined in the NERC reporting instructions. The CanWEA program is currently utilizing monthly data, thus, 3 files per month are required, 36 for the year. The sub-group file does not change from month to month and can therefore be copied 12 times for simplicity. The file should be reviewed regularly to ensure that the data is up to date in case changes at the reporting wind farms. Example files are available upon request from WEICan. WEICan has worked with some contributors to provide tools to compile data and can provide assistance for a nominal fee.

Files should be prefixed with the corresponding numerical month prior to submission as follows:

|  |
| --- |
| 1 Outage.csv |
| 1 Performance.csv |
| 1 Subgroup.csv |
| 2 Outage.csv |
| 2 Performance.csv |
| 2 Subgroup.csv |
| 3 … |

Correctly naming the submitted files is critical to proper data interpretation by the report generation program.

Following data submission by all participants, reports with content similar to the sample included in this information package will be distributed. One report summarizing the past year will be provided for each wind farm. Reports for each group or subgroup as well as more in-depth monthly reports can be discussed with WEICan.

WEICan is committed to data privacy while still providing a useful benchmarking tool. Feedback on report formatting and data presentation is encouraged.

Example reports & relevant files

|  |  |
| --- | --- |
| GADS Project directory on server | \\weicserver\Docs\Projects\5032 - GADS |
| GADS Code directory on server (no spaces in path) | \\weicserver\Docs\Projects\5032\_GADS\_Source\_Code |
| NERC presentation on GADS as applicable to wind generators | \\weicserver\Docs\Projects\5032 - GADS\NERC Standards\2017\Files from NERC website\Wind Training\_Wind Binder.pdf |
| 2017 GADS Wind turbine reporting instructions | \\weicserver\Docs\Projects\5032 - GADS\NERC Standards\2017\wind\_DRI\_DL\_GADS Wind DRI-2016 Master.pdf |
| WEICan GADS 2015 Report | \\weicserver\Docs\Projects\5032 - GADS\2015 GADS\Reports\PDF\WEICAN DEWIND.pdf |
| WEICAN GADS 2016 Report | \\weicserver\Docs\Projects\5032 - GADS\2016 GADS\Reports\PDF\WEICAN DEWIND.pdf |
| 2016 GADS data submitted by participants (contains errors!) | \\weicserver\Docs\Projects\5032 - GADS\2016 GADS\Data |
| 2015 GADS data submitted by participants (contains errors!) | [\\weicserver\Docs\Projects\5032 - GADS\2015 GADS\Raw Data\2015](file:///\\weicserver\Docs\Projects\5032%20-%20GADS\2015%20GADS\Raw%20Data\2015) |
| 2017 GADS data submitted by participants (contains errors!) | \\weicserver\Docs\Projects\5032 - GADS\2017 GADS\Data |
| Examples of 2016 GADS data after formatting and errors removed (Files named GADS 2016 – XXXXX) | \\weicserver\Docs\Projects\5032 - GADS\2016 GADS |
| Existing Excel macro (Open file, navigate to “Readme” sheet, right click on Calculate button and then access VB code as shown in screenshots below) | \\weicserver\Docs\Projects\5032 - GADS\2016 GADS\Master 2016.xlsm |

# Project phases

Broadly, this project will consist of two development tasks:

1. Development and management of the database
2. Development of the front-end interface

The project will be split into two phases: The first assumes that data files have no errors and will focus on developing a working database and implementing accurate calculations. Once the database is correctly implemented and the calculations are verified to be correct, the second phase can begin. The second phase focuses on handling errors in the input data files such as missing data or implausible values.

|  |  |  |  |
| --- | --- | --- | --- |
| **Phase** |  | **Task** | Summary |
|  |  | **Project setup** |  |
|  | Understanding GADS project, existing work |  |
|  | Setting up version management system (revision control system) |  |
|  | Define modules, overall function, coding blocks |  |
|  | Define list of database parameters and parameter data types and ranges | Example: Net Actual Generation (NAG); Datatype – Float; Range -1e8 to +1e8 |
|  |  | **Data import** |  |
|  | Reading participants data files | Assume no errors in file |
|  | Standard classification files | Including region |
|  | Missing data checks | Offer opportunity to add missing data such as number of turbines. Flag additions. |
|  | Data conditioning & plausibility checks | Are data such as hours reported plausible give number of turbines? See Appendix J of 2018 GADS. |
|  | Adding to database | Check that data is not already included in database, avoid duplicates |
|  | Importing existing data and handling changing report formats | Existing data is already in different formats. Adding / subtracting / remapping data points must be possible in case reporting standard changes in the future. |
|  |  | **Database development & management** |  |
|  | Data editing |  |
|  | Editing data in database, autofill missing data |  |
|  | Calculations |  |
|  | Grouping data by year & reporting format |  |
|  | Data export (as CSV file) according to filters |  |
|  | Database improvement: handling improperly formatted files |  |
|  |  | **Front-end development** |  |
|  | Report generation |  |
|  | Generate test reports using 2015 GADS data and compare results against existing reports |  |
|  | Generate reports using filters | Filter by turbine age, filter by geographic location etc |
|  | Verify report data by manual calculation |  |
|  |  | **Documentation** |  |
|  | Document all modules, functions with arguments, program flow | Do not document calculation steps and equations |
|  | Document error messages and error handling |  |
|  | Managing report versions / revisions |  |

# Specific project requirements

*(incomplete)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *File name* | *Number of columns*  *(GADS 2011)* | *Number of columns*  *(GADS 2016)* | *Number of columns*  *(GADS 2017)* | *Number of columns*  *(GADS 2018)* |
|  | *Outages file* | *14* | *20* | *19 (6 optional)* | *19* |
|  | *Performance* | *30* | *40* | *33* | *33* |
|  | *Sub-group* | *21* | *22* | *40* | *40* |
|  | *Plants* | *2* | *2* | *-* | *-* |
|  | *Groups* | *20* | *20* | *-* | *-* |

Use reference tables in Appendix F of 2018 GADS standard.

# Project Background

The text below is copied from another file and gives a brief overview of WEICan’s role in the GADS project.

**Background**

At the end of 2015, Canada had commissioned more than 11,000 MW of wind-power capacity, and even more wind farms will be coming online in the coming years. Although the wind industry supplies approximately 5 percent of Canada’s electricity needs, to date it has not benefited from a broad and consistent understanding with respect to its performance, especially when compared to data that is available for traditional forms of energy.

While some sources of information do exist with regards to wind-energy data, these tend to conflict with each other and typically underestimate basic details such as tracking current installed wind-energy capacity. With the current Federal and Provincial focus on climate change, it is increasingly important to have access to reliable and consistently produced baseline data for renewable energy, in particular wind energy, which is one of the fastest growing sources of new electricity in Canada. Moreover, as wind parks transition from construction to operation and maintenance, the need for comparative statistics also increases. Power generation is a competitive business, and when electricity prices are low, wind energy has extremely low margins due to lack of fuel costs. So it is critical to track all causes and occurrences of downtime to direct process improvements and forecast future expenditures.

**Project Description**

The Canadian Wind Energy Association (CanWEA), along with wind farm owners and operators throughout Canada, have recognized the need for standardized reporting to support wind industry internal benchmarking, preventative maintenance, and research.

In early 2014 CanWEA initiated a project to collect key performance indicators from wind turbine owners. CanWEA currently has non-disclosure agreements (NDAs) with wind turbine owners and operators that comprise 30% of the installed wind energy capacity in Canada. These wind turbine owners and operators agreed to adopt the existing Generating Availability Data System (GADS) format, which is the type of reporting used for conventional (i.e. not renewable) generating units.

**WEICan’s role in the project**

The Wind Energy Institute of Canada (WEICan) is proposing to manage this CanWEA project. WEICan’s technical staff will collect and aggregate from participating wind farms, following execution of an NDA between WEICAN and wind plant owners. WEICan will lead the analysis of the data on behalf of the various stakeholders based on direction provided, and help lead the research coordination for this work, as well as dissemination of data to stakeholders. Wind farm owners will receive a yearly report, prepared by WEICan, showing how their wind farms’ performances compare to the Canadian average. All individual wind farm data will be kept confidential, and representative and aggregated data will be used for public reports. WEICan will also attend meetings and conferences to present the data to various stakeholders. In addition, WEICan will provide aggregated reports and data sets for further research to NRCan as part of their primary deliverable to NRCan under this project scope of work. However, in order to receive such data, NRCan will be required to execute an NDA with WEICan.

**Scope of Work**

**The contractor must complete the following:**

1. Data collection and integration: WEICan will securely obtain individual wind farm production and availability data in GADS format from participating wind farm owners. WEICan will work with the wind farm operators and owners to ensure that the data is coded appropriately and will create a program for each wind farm that will allow the data to be analyzed. The data from the wind farms will be entered and individual reports will be generated. WEICan will verify data quality and consistency.
2. Research coordination: The large volume of data will provide many research opportunities. For example, one of the turbine owners, a university professor, or research institution may have a research question or may see value in a specific analysis of the data. WEICan will work with CanWEA, NRCan, and the owners/operators to establish research objectives. WEICan will coordinate the transfer and analysis of data, according to the NDAs.
3. Information dissemination: Data will be aggregated and anonymized, according to the NDA. Each wind farm, owner/operator who participates in this program will receive an individual report allowing comparison to other wind farms and year-to-year comparison within their own wind farm. An overall report with the aggregated and anonymized data will be provided to the broader wind industry. Reports will also be provided that can be distributed globally, for example, at conferences and on the CanWEA or NRCan websites. Lastly, NRCan will receive an aggregated dataset (protected via NDA), in order to undertake additional research internally.

**Schedule and Deliverables**

WEICan will collect, aggregate, analyze, and disseminate wind turbine data from wind turbine owners and operators from across Canada. The general task structure with 3 tasks listed in the following table is the same for each year.

|  |  |  |  |
| --- | --- | --- | --- |
| **Tasks** | | **Deliverables** | **Dates** |
| 1 | Data collection and integration | The contractor will produce a confidential database | December/January 2017, 2018, 2019, 2020, 2021 |
| 2 | Research Coordination | Research objectives will be confirmed | February 2017, 2018, 2019, 2020, 2021 |
| 3 | Information Dissemination | * Anonymized and aggregated data will be published in an annual report. * Datasets for further research and analysis will be provided to NRCan * Presentations will be prepared and presented at various forums, such as CanWEA or UVIG. | March 2017, 2018, 2019, 2020, 2021 |