Turbulence Intensity Adjustment Comparison Tool

CFARS Site Suitability Subgroup User's Guide

https://github.com/CFARS/site_suitability_tool 05/01/2021

USER'S GUIDE TABLE OF CONTENTS

		Page #
1.0	GENERAL INFORMATION	1
1.1	TURBULENCE INTENSITY ADJUSTMENT TESTS TOOL OVERVIEW	1
1.2	License	
1.3	CONTRIBUTORS AND POINTS OF CONTACT	2
1.4	ORGANIZATION OF THE USER GUIDE	
1.5	ACRONYMS AND ABBREVIATIONS	
2.0	GETTING STARTED	3
2.1	SETTING UP PYTHON	
	1.1 Confirm you have python 3 installed	
2.	1.2 If you do not have python installed, please install python 3 for macOS or Windows	4
	DOWNLOAD OR UPDATE THE REPOSITORY	
2.3	INSTALL REQUIREMENTS	5
3.0	PREPARING DATA AND CONFIGURATION	6
3.1	Preparing your data	6
3.2	PREPARING YOUR CONFIGURATION FILE	
3.	2.1 Match Your Columns	6
3.	2.2 Minimum required data columns	7
3.	2.3 Additional columns to label in the configuration filefile	8
3.	2.4 Site Metadata	9
3.	2.5 Filtering Metadata	9
4.0	RUNNING THE TOOL	10
4.1	Arguments	10
4.2	COMMAND IN TERMINAL/COMMAND PROMPT	10
5.0	INTERPRETING THE OUTPUT	10
5.1	AVAILABLE ADJUSTMENT METHODS FOR TESTING	10
5.2	STRUCTURE OF THE OUTPUT FILE	
5.2	ADJUSTMENT METHODS	13
5.3	EXTRAPOLATED TI	13
5.4	CAMPAIGN LENGTH SENSITIVITY / MODEL SENSITIVITY TO "TIME TO TRAIN"	

1.0 GENERAL INFORMATION

1.0 GENERAL INFORMATION

1.1 Turbulence Intensity Adjustment Comparison Tool Overview

The purpose of this software is to support the research initiative of the CFARS Site Suitability Subgroup. Using a "closed-data" approach, a standardized testing and analysis methodology is useful to investigate a variety of turbulence intensity adjustment techniques. This tool is designed so that users can extract results to submit to the research initiative without having to share data.

The secondary purpose of this tool is to act as an evaluation of various TI adjustment techniques on a user's data set. Please note this code is open source and in development.

1.2 License

The software includes a permissive BSD 3-Clause License.

Copyright © 2020, CFARS All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- 1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- 2. Re distributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
- 3. Neither the name of the copyright holder nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

HIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

1.3 Contributors and Points of Contact

For 2021, the point of contact for this software is Alexandra Arntsen <u>aea@nrgsystems.com</u> The maintenance of this software and point of contact beyond 2021 is yet to be determined.

Programming contributors include:

Alexandra Arntsen, NRG Systems Nikhil Kondabala, Apex Andrew Black, Vaisala Barrett Goudeau, NRG Systems Nigel Swytink-Binnema, Nergica Nicolas Jolin, Nergica

The CFARS Site Suitability Subgroup has collectively advised on and decided on tool functionality and content. Methods of adjusting turbulence intensity have not been created by the subgroup, but rather this software incorporates pre-existing methods for testing an analysis.

1.4 Organization of the User Guide

This manual is divided into 5 major sections, listed and defined below.

- 1.0 General Information
- 2.0 Getting Started: Describing the necessary set up of your system to run the python program.
- 3.0 *Preparing Data and Configuration:* Requirements for your input data and how to fill out the configuration template.
- 4.0 Running the Tool: How to execute the python script locally.
- 5.0 *Interpreting the Output*: How to navigate and interpret the output file.

1.5 Acronyms and Abbreviations

TACT - Turbulence Intensity Adjustment Comparison Tool TI - Turbulence intensity RSD - Remote sensing device

2.0 GETTING STARTED

2.1 Setting up python

The turbulence intensity adjustment tests tool is a python program. To extract results from your data set, you will need to have python installed on your local machine. Furthermore, this software is compatible with python 3.

2.1.1 Confirm you have python 3 installed.

For MacOS open terminal and type:

>>> python

```
aearntsen — python — 80×5

(python3env) Alexandras-MacBook-Pro:~ aearntsen$ python

Python 3.8.2 (default, Mar 26 2020, 10:43:30)

[Clang 4.0.1 (tags/RELEASE_401/final)] :: Anaconda, Inc. on darwin

Type "help", "copyright", "credits" or "license" for more information.

>>>
```

If the system reports a version of python 3 or higher, you should be ready to proceed.

However, if your system reports a version of python 2 (as seen below), you will need to update to python 3.

```
aearntsen — python — 84×11

Last login: Mon Jan 18 16:52:05 on ttys001
(base) Alexandras—MacBook—Pro:~ aearntsen$ python
Python 2.7.16 |Anaconda, Inc.| (default, Mar 14 2019, 16:24:02)
[GCC 4.2.1 Compatible Clang 4.0.1 (tags/RELEASE_401/final)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>>
```

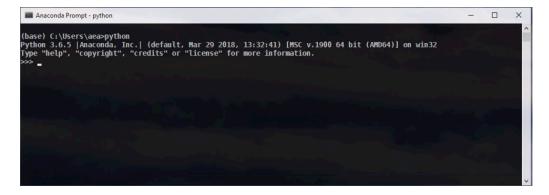
in terminal window type:

>>> conda install python=\$pythonversion\$

***NOTE: If in terminal "python is not recognized as an internal or external command", you need to add the installed python executable to your environment variable system path

For Windows open anaconda prompt and type:

>>> python



If the system reports a version of python 3 or higher, you should be ready to proceed.

However, if your system reports a version of python 2, you will need to update to python 3. Do update type:

>>> conda install python=\$pythonversion\$

2.1.2 If you do not have python installed, please install python 3 for macOS or Windows

https://www.anaconda.com/products.individual

2.2 Download or update the repository

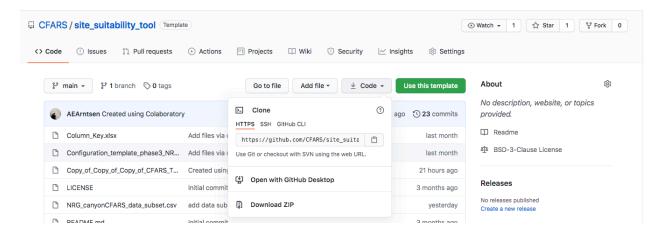
Once you confirm your python is set up correctly, you need to be sure to have a local copy of the tool repository.

https://github.com/CFARS/site suitability tool

If you are familiar with git, you can **clone the repository** above by in terminal or command line navigating to a directory of choice and typing

>>> git clone https://github.com/CFARS/site suitability tool.git

Another option is to **download the repository.** To do this, click the button in the repository website and then choose Download Zip as seen in the screen shot below.



2.3 Install requirements

The anaconda build of python comes with most of the packages you will need to run this tool, but to be sure your version of python is configured correctly, you can use the requirements file in the repository.

In terminal or command prompt, navigate to the site_suitability_tool directory you created by either downloading or cloning the repository. This directory contains all of the files from the repository, including a file named 'requirements.txt'.

Next, type

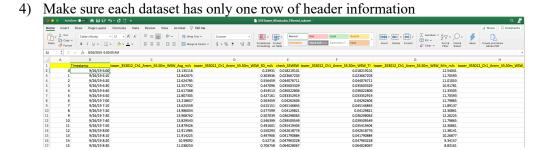
>>> pip install -r requirements.txt

3.0 PREPARING DATA AND CONFIGURATION

3.1 Preparing your data

The tool processes individual data sets separately, so each dataset should be a separate .csv file with a unique name.

- 1) Each data set should contain a reference anemometer collocated with a remote sensing device measuring at the same height. There should also be a second anemometer also measuring at this height.
- 2) Time stamp format should be month/day/year hour:minute:second Example: 07/15/20 14:50:00
- 3) Save each project data set to the directory that contains the tool



3.2 Preparing your configuration file

Each dataset requires a corresponding configuration file to be interpretable by the tool. Follow the below steps to setting up configuration files for each of your datasets.

In the repository directory you will find a file named 'Configuration_template.xlsx'. The first step is to create a version of this file for each one of your unique datasets with the dataset name included in the name of its configuration file. For example, if I have 2 datasets, 'NRG_project1.csv' and 'NRG_project2.csv' I will have 2 configuration files (copies of the template) named 'configuration_NRG_project1.xlsx' and 'configuration_NRG_project2.xlsx'.



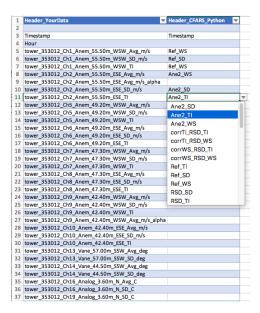
3.2.1 Match Your Columns.

It is important that we ensure column names are interpreted correctly by the tool. The first step in configuration is to label your data columns with the standardized column names.

In column A, copy and paste vertically the single row header information from your project file.



Next, label all the essential columns with the dropdown selection in column B of the template. *Note:* It is fine to have more columns in the input data that you label in the configuration file. You do not need to label every column in your input data set with a CFARS python header.



3.2.2 Minimum required data columns.

Note: the tool will not run without these minimum config file entries. It is designed to give feedback when there are missing data columns in the input file and or columns that should be labeled and accounted for but are not.

The minimum required columns to run the tool successfully:

Variable	Column Abbreviation	Notes
Timestamp	Timestamp	
Reference cup average wind speed	Ref_WS	
Reference cup wind speed standard	Ref_SD	
deviation		
Reference cup turbulence intensity	Ref_TI	
Redundant cup average wind speed	Ane2_WS	Same height as reference
Redundant cup wind speed standard	Ane2_SD	Same height as reference
deviation		
Redundant cup turbulence intensity	Ane2_TI	Same height as reference
Remote sensing device average wind	RSD_WS	Same height as reference
speed		
Remote sensing device wind speed	RSD_SD	Same height as reference;
standard deviation		Called wind speed dispersion in
		windcube data
Remote sensing device turbulence	RSD_TI	Same height as reference; Pre-
intensity		calculate in .csv if you are
		missing this TI =
		RSD_SD/RSD_WS
Remote sensing device direction	RSD_Direction	Use nearest height to the
		reference height; required to
		calculate TKE

3.2.3 Additional columns to label in the configuration file.

Some of these columns aid additional analysis, some additional corrections. See sections 5 & 6 for more information.

Optional data for further TI adjustment and analysis (add as many as data allows):

Additional comparison heights

Additional Data	Column Abbreviations	Notes
First Additional comparison height between cup and RSD	Ane_WS_Ht1, Ane_SD_Ht1, Ane_TI_Ht1, RSD_WS_Ht1, RSD_SD_Ht1, RSD_TI_Ht1	Optional, please include if your data happens to have other heights to compare
Second Additional comparison height between cup and RSD	Ane_WS_Ht2, Ane_SD_Ht2, Ane_TI_Ht2, RSD_WS_Ht2, RSD_SD_Ht2, RSD_TI_Ht2	Optional, please include if your data happens to have other heights to compare
Third Additional comparison height between cup and RSD	Ane_WS_Ht3, Ane_SD_Ht3, Ane_TI_Ht3, RSD_WS_Ht3, RSD_SD_Ht3, RSD_TI_Ht3	Optional, please include if your data happens to have other heights to compare
Fourth Additional comparison height between cup and RSD	Ane_WS_Ht4, Ane_SD_Ht4, Ane_TI_Ht4, RSD_WS_Ht4, RSD_SD_Ht4, RSD_TI_Ht4	Optional, please include if your data happens to have other heights to compare

Additional RSD direction columns for TKE calculations at additional comparison heights

Additional Data	Column Abbreviations	Notes
RSD direction average for first additional	RSD_Direction_Ht1	Optional, please include if your
comparison height		data happens to have other
		heights to compare
RSD direction average for second additional	RSD_Direction_Ht2	Optional, please include if your
comparison height		data happens to have other
		heights to compare
RSD direction average for second additional	RSD_Ditection_Ht3	Optional, please include if your
comparison height		data happens to have other
		heights to compare
RSD direction average for second additional	RSD_Direction_Ht4	Optional, please include if your
comparison height		data happens to have other
		heights to compare

Alpha calculation heights

Additional Data	Column Abbreviations	Notes

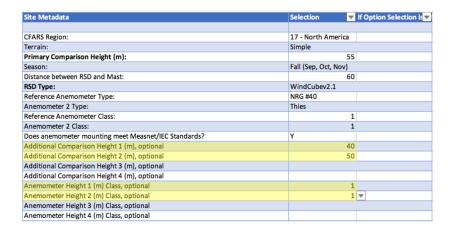
3.2.4 Site Metadata

Fill out the site metadata information in columns E & F

At a minimum, rows 3-13 in column E need a response in the configuration file.



If you add additional heights for comparison, you will need to also select these additional observation heights in the site metadata.



3.2.5 Filtering Metadata

Fill out information about filtering of the dataset in columns I & J. Filtering the data ahead of time is important to ensuring a true measure of adjustment method performance.

Key Cleaning Flags/Filters Applied	▼ Selection	▼ Notes	~
Tower Distortion using Conservative Method?	Υ		
Anemometer Icing?	Υ	Yes icing was removed	
Wind Speed Outliers?	N		
Sensor Degradation?	N		
RSD Low Data Availability Filter?	Υ		
RSD Low CNR Filter?	Υ		
Is the RSD wind speed corrected?	N		
Application of site-specific transfer function (slope/offset)	N		

Additional Notes:

- 1) Mark the version of firmware if your dataset contains ZX data
- 2) Refer the example files in the repository: 'Example_project_subset.csv' & its corresponding configuration file 'configuration Example project subset.xlsx'

4.0 RUNNING THE TOOL

4.1 Arguments

-in	location of input data file
-config	location of configuration file
-globalModel	specify the specific global model to test (only works with .pkl file model generated from a
	previous run). If this argument is not used, TACT will test the generic global model.
-rtd	location of windcube rtd files (only necessary if running WiSE* and LTERRA)
-res	location of results file
-saveModel	global model derived from the data set should be saved as user output. Tool will not generate a global output model if this argument is not used.
-timetestFlag	performs length of data collection analysis (note: this takes a several minutes to run)

^{*}Running the WiSE method requires and NDA

4.2 Command in terminal/command prompt

In terminal or command type a command that follows the following structure:

>>> python TIAT.py -in NRG_project1_data.csv -config configuration_NRG_project1.xlsx -rtd /Users/aearntsen/CFARS/NRG_project1/RTD -res /Users/aearntsen/CFARS/NRG_project1/out_NRG_project1.xlsx --timetestFlag

5.0 INTERPRETING THE OUTPUT

5.1 Available adjustment methods for testing

Adjustment Abbreviation	Description

SS-SF	Site-Specific Simple Filtered (regression correction iwith filter)
SS-S	Site-Specific Simple (regression correction)
SS-SS	Site-Specific Simple Stability (regression correction by TKE class)
SS-Match2	Site-Specific PDF transfer function
SS-WS	Site-Specific Wind speed Adjustment
SS-WS-Std	Site-Specific WS and Std. Dev. Adjustment
SS-LTERRA-MLa	Site-Specific LTERRA Machine Learning Version 1
SS-LTERRA-MLb	Site-Specific LTERRA Machine Learning Version 2
SS-LTERRA-MLc	Site-Specific LTERRA Machine Learning Version 3
G-LTERRA-ML	Global LTERRA Machine Learning
G-LTERRA-1Hz	Global LTERRA 1 Hz Adjusted
G-Sa	Global Simple
G-SFa	Global Simple with Filter
G-SFc	Global Simple with Filter Version 2 (site data incorporated)
G-C	Global 10-minute Std Deviation adjustment
WiSE	WiSE Method

5.2 Structure of the Output file

The output file is generated in the form of an excel workbook.

Tab Name	Description	
Baseline Results	Regression statistics for raw data; Regression stats by TKE class or by shear	
	exponent classification; Bin counts overall and by classification; length and	
	date of campaign, TKE and shear exponent classification breakdown	
Metadata	Sensor details; comparison heights; version of the tool; time of execution;	
	adjustments applied	
Sensitivity2TestLengthA	Adjusting the test train split (random % of data)	
Sensitivity2TestLengthB	Adjusting the test train split (incrementally adding days to training set)	
Sensitivity2TestLengthC	Adjusting the test train split (3 month sliding window)	
[Adjustment Key]	Separate tab for the results of each adjustment method	
TI-Extrap	Comparing Extrapolated TI to RSD TI	
K-S Tests	Statistical Similarity Tests	
Extrapolation Metadata	Metadata used for TI-Extrap	

Baseline Results

There is a table of regression results for the full data set (labeled "Full Comparison") as well as a table of regression results for the data parsed by stability class as defined by both TKE calculated from the RSD as well as by shear exponent (alpha) as calculated by the anemometer data and RSD if possible.

The number of observations in each bin (both 1 m/s and 0.5 m/s bin width) for the full data set as well as number of observations in each bin for the train and test data sets. The bin counts are also calculated for the different TKE classes and alpha classes.

The length of the test as well as start date and end date are recorded. The number of observations in each TKE and alpha class is recorded to categorize the "type of site".

Sensitivity2TestLengthA: Model generation sensitivity A

Tests the sensitivity of regression statistics to the % of the data set evaluated. The same sensitivity is also tested for a subset of adjustment methods, currently: SS-S, SS-SF, SS-WS-Std, SS-Match, SS-Match2, G-SFa.

Sensitivity2TestLengthB: Model generation sensitivity B

Tests the sensitivity of regression statistics to the number of days evaluated. The same sensitivity is also tested for a subset of adjustment methods, currently: SS-S, SS-SF, SS-WS-Std, SS-Match, SS-Match2, G-SFa in which the number of days to train the model varies.

Sensitivity2TestLengthC: Model generation sensitivity C

Tests the sensitivity of regression statistics to the season, so will evaluate a sliding window of 6-week periods, sliding by 1-week increments. The same sensitivity is also tested for a subset of adjustment methods, currently: SS-S, SS-SF, SS-WS-Std, SS-Match, SS-Match2, G-SFa.

Results for Each Adjustment Tab

Each Adjustment method has results recorded in a separate tab. Regression statistics for the "corrected" RSD TI are shown in the first table of the tab starting in column A, row 1. There is a separate table for regression results for the five TKE classes as well as the stability classes defined by both a calculation of alpha by the anemometer data and the RSD data.

Next, bin specific results are recorded. The first table in this section of results records an evaluation of mean TI at each 1 m/s bin interval including:

- 1) RSD TI (RSD TI)
- 2) Reference anemometer TI (Ref TI)
- 3) "Corrected" RSD TI (corrTI RSD TI)
- 4) secondary anemometer TI (Ane2 TI)
- 5) Error between RSD & Ref, corrected RSD & Ref, secondary anemometer & Ref (TI error RSD Ref, TI error corrTI RSD Ref, TI error Ane2 Ref)
- 6) RMSE between RSD & Ref, corrected RSD & Ref, secondary anemometer & Ref (TI rmse RSD Ref, TI rmse corrTI RSD Ref, TI rmse Ane2 Ref)
- 7) Difference between RSD & Ref, corrected RSD & Ref, secondary anemometer & Ref (TI_diff_RSD_Ref, TI_diff_corrTI_RSD_Ref, TI_diff_Ane2_Ref)
- * 1-7 for each stability class as defined by TKE

The second table in this section of results records an evaluation of mean TI at each 0.5 m/s bin interval including. The structure of the previous table remains.

The third table records an evaluation of the standard deviation of 10-minute TI at each 1 m/s bin interval including. The structure of the previous table remains.

The fourth table records an evaluation of the standard deviation of 10-minute TI at each 0.5 m/s bin interval including. The structure of the previous table remains.

5.2 Adjustment Methods

The following section provides a more in-depth description of each adjustment method testing in the tool.

5.3 Extrapolated TI

In

5.4 Campaign Length Sensitivity/ Model Sensitivity to "time to train"