ALGORITHM THEORETICAL BASIS DOCUMENT: QUALITY FLAGS AND QUALITY METRICS FOR TIS DATA PRODUCTS

PREPARED BY:	**ORGANIZATION:**	**DATE:**
Derek Smith	FIU	07/29/2013
Stefan Metzger	FIU	07/29/2013

APPROVALS (Name):	**ORGANIZATION:**	**APPROVAL DATE:**
David Tazik	CCB PROJ SCI	09/01/2013
Hanne Buur	CCB DIR SE	09/16/2013

RELEASED BY (Name):	**ORGANIZATION:**	**RELEASE DATE:**
Stephen Craft	SE	09/17/2013

See Configuration Management System for Approval History.

CHANGE RECORD

REVISION	DATE	ECO #	DESCRIPTION OF CHANGE
A	09/17/2013	ECO-01303	Initial Release

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 DESCRIPTION

This document specifies how sensor flags and results from quality assurance and quality control (QA/QC) tests will be summarized to inform users of data quality.

1.1 Purpose

Each NEON DP has the ability to be flagged by various QA/QC tests and sensor flags. As such the need arises to assess these "quality flags" (QFs). QFs can either be attributed to individual measurements or a percent of measurements, e.g. when measurements become averaged for a time period. This document describes the theoretical background and entire algorithmic process used to produce information on the quality of NEON TIS data.

1.2 Scope

Information presented here relates to the data quality of NEON's TIS DPs. A final QF will be generated by NEON that determines whether a TIS DP is *valid* or *invalid*. This final quality flag will be based on an assessment of a DP's QF results. Any sensor specific details are specified in a sensor's specific algorithm theoretical based document (ATBD), in the algorithm implementation section.

2 RELATED DOCUMENTS AND ACRONYMS

2.1 Applicable Documents

Applicable documents contain information that shall be applied in the current document. Examples are higher level requirements documents, standards, rules and regulations.

AD [01]	NEON.DOC.001013	NEON Science Commissioning and Validation Plan

2.2 Reference Documents

Reference documents contain information complementing, explaining, detailing, or otherwise supporting the information included in the current document.

RD [01]	NEON.DOC.000008	NEON Acronym List
RD [02]	NEON.DOC.011009	FIU Dataflow and QA Plan
RD [03]	NEON.DOC.000243	NEON Glossary of Terms

2.3 Acronyms

Acronym	Explanation
ATBD	Algorithm Theoretical Basis Document
DP	Data Product
L0	Level 0

L1	Level 1
QA/QC	Quality Assurance/Quality Control
QF	Quality Flag
QM	Quality Metric
TIS	Terrestrial Instrument System

2.4 Verb Convention

"Shall" is used whenever a statement expresses a convention that is binding. The verbs "should" and "may" express non-mandatory provisions. "Will" is used to express a declaration of purpose on the part of the design activity.

3 DESCRIPTION OF DATA PRODUCTS

3.1 Reported Data Products

Each TIS Level 1 (L1) DP will be accompanied by a quality summary (Q_{sum}), which will summarize the QF results from the QA/QC analyses and sensor flags. Results of specific QFs as they relate to individual Level 0 (L0) DPs will be retained in a quality report (Q_{rpt}), but will not be presented unless requested. Detailed explainations of Q_{sum} and Q_{rpt} can be found in section 5.3.

3.2 Input Data Products

DP inputs will consist of any L0 DP that is used to create a given L1 DP.

3.3 Product Instances

This is dependent on the measurement in question and is detailed in a sensor's specific ATBD.

3.4 Temporal Resolution and Extent

The temporal resolution and extent of the QFs and QMs will be dependent on the sampling frequency of the measurement in question. Relevant information is detailed in a sensor's specific ATBD.

3.5 Spatial Resolution and Extent

The spatial resolution and extent of the QFs and QMs will be dependent on the location of the measurement in question. Relevant information is detailed in a sensor's specific ATBD.

4 SCIENTIFIC CONTEXT

Data quality has always been paramount to ecology; however, historically the need to automate its assessment has not been. This has primarily been due to the quantity of data that has been needed to answer research questions. Yet, there is an increasing need for "big data" to answer the research questions of tomorrow. As a result, there is a growing demand for automated approaches to assess data in order to replace time consuming manual assessments. Our goal was to develop an automated framework that allows the QA/QC and sensor flag results to be summarized in a transparent and easily interpretable way for TIS L1 DPs.

We derived techniques from standard timeseries analysis, which are widely used in eddy covariance research due to their need to routinely assess large data sets and scrutinize data quality. As such, we drew concepts from the following works, Mauder et al., 2013, Mauder, 2011, and Gockede et al., 2004. These approaches were designed to determine the quality of a data product that was produced using multiple other data products. Briefly, these schemes define criteria to rank the quality of a data product's inputs and then use that information to determine the quality of the final data product. A result is that these rank based approaches tend to be subjective and predominantly measurement specific. In addition, these approaches were not designed to be used for a multitude of measurements and as such are not easily transferable. Therefore, these approaches could not be directly transferred for NEON, but instead their concepts were drawn from to create our quality assessment scheme.

While, other quality flag systems for climatological data were examined, the sheer magnitude of the data that will be collected inhibits the use of many preexisting quality flag schemes. For example, many of NEON's TIS sensors collect data at 1 Hz and consequentially results many of the QA/QC and sensor flags are generated at the same rate. Thus, for flags generated at 1 Hz, there will be over 86,000 outcomes for each flag in a day and over 600,000 in a week. The ability to digest and interpret that quantity of data can quickly become overwhelming. Therefore, a framework was developed to objectively assess data quality, while remaining transparent and transferable to all of NEON's terrestrial sensors.

4.1 Theory of Algorithm

4.1.1 Quality Flags

Quality flags are generated by a number of QA/QC analyses as well as sensor flags. For example, QFs for L1 DPs include flags produced by plausibility, despiking, consistency, and sensor flags. While, L0 DPs flagged by some of these tests will result in a datum not being used to create a L1 DP, this is not true for all QFs. Information on which QFs will result in a datum being excluded when calculating a L1 DP is included in a sensor's specific ATBD. Each QF can be set to one of three states as shown in Eq. (1). For specific details on a QF please refer to their corresponding ATBDs.

$$QF = \begin{cases} 1 \text{ if the quality test failed} \\ 0 \text{ if the quality test passed} \\ -1 \text{ if NA i.e. not able to be run due to a lack of ancillary data} \end{cases}$$

Error! Bookmark not defined.Error! Bookmark not defined.(1)

4.2 Quality Metrics

Since each L1 DP is composed of multiple L0 DPs, any QFs that were applied to L0 DPs need to be summarized. Thus, a DP consisting of multiple observations will have three QMs associated with each QF. The three QMs will summarize as a percent of the total number of observations, used to create a DP, where a QF was set to 1, 0, and -1. QMs, will always be rounded half up to a whole percentage. QMs are defined as follows:

$$QM_{j,1} = \frac{\sum_{i=1}^{N}(QF_{j,i} \equiv 1)}{N} * 100$$

Error! Bookmark not defined.Error! Bookmark not defined.(2)

and

$$QM_{j,0} = \frac{\sum_{i=1}^{N}(QF_{j,i} \equiv 0)}{N} * 100$$

Error! Bookmark not defined.Error! Bookmark not defined.(3)

and

$$QM_{,-1} = \frac{\left|\sum_{i=1}^{N}(QF_{j,i} \equiv -1)\right|}{N} * 100$$

Error! Bookmark not defined.Error! Bookmark not defined.(4)

Where:

$QM_{j,1}$	= Quality metric associated with QF_j for the percent of tests set high
$QM_{j,0}$	= Quality metric associated with QF_j for the percent of tests set low
$QM_{j,-1}$	= Quality metric associated with QF_j for the percent of NA tests
QF_j	= Results of test, j, for a LL DP
i	= Running index over sample size
N	= Sample size

In reality one only needs to compute two of the three QMs and the third QM can be derived as follows:

$$100\% - \left(\mathrm{QM}_{j,1} + \mathrm{QM}_{j,-1}\right) = \mathrm{QM}_{j,0}$$ **Error! Bookmark not defined.Error! Bookmark not defined.**(5)

Table 1 illustrates how QMs for a L1 DP, where $\{QF_F | QF_j \in QF_F\}$ and $\{N | i \in N\}$, are determined according to Eq. (2), (3), and (4).

Observation	**QF_j**	…	**QF_F**
i=1	$1 \vee 0 \vee -1$	⋱	$1 \vee 0 \vee -1$
i=2	⋮	⋱	⋮
⋮	⋮	⋱	⋮
⋮	⋮	⋱	⋮
⋮	⋮	⋱	⋮
⋮	⋮	⋱	⋮
⋮	⋮	⋱	⋮
N	⋮	⋱	⋮

$QM_{j,1}$	**$QM_{j,0}$**	**$QM_{j,-1}$**		**$QM_{F,1}$**	**$QM_{F,0}$**	**$QM_{F,-1}$**
$\frac{\sum_{i=1}^{N}(QF_{j,i} \equiv 1)}{N} * 100$	$\frac{\sum_{i=1}^{N}(QF_{j,i} \equiv 0)}{N} * 100$	$\frac{\left\lvert\sum_{i=1}^{N}(QF_{j,i} \equiv -1)\right\rvert}{N} * 100$	…	$\frac{\sum_{i=1}^{N}(QF_{F,i} \equiv 1)}{N} * 100$	$\frac{\sum_{i=1}^{N}(QF_{F,i} \equiv 0)}{N} * 100$	$\frac{\left\lvert\sum_{i=1}^{N}(QF_{F,i} \equiv -1)\right\rvert}{N} * 100$

Table 1. Illustration of how QFs are summarized into QMs for a L1 DP.

5 ALGORITHM IMPLEMENTATION

5.1 α and β Quality Flags and Metrics

In order to assess the overall quality of a particular L1 DP, it is necessary to summarize the number of flagged observations among all L0 DPs that were used in its calculation. Thus, we define alpha (α) and beta (β) QFs and QMs, which will incorporate the outcomes of several QFs in order to assess the L1 DP's quality. The following defines how α and β QFs and QMs are determined.

α and β QFs will be calculated based on the outcomes of the QFs from QA/QC analyses as well as sensor flags. What QFs will be used to calculate α and β QFs is sensor-specific and must be specified in its corresponding ATBD. The calculation of α and β QFs is very similar except that QF_α will determine for a subset of QFs (defined in a sensor-specific ATBD) whether or not at least one QF was set to 1 for an observation. Likewise, QF_β will determine for a subset of QFs, whether or not at least one QF was set to -1 for an observation. The calculation of QF_α and QF_β are shown Eq. (6) and (7).

$$\mathrm{QF}_{\alpha} = \begin{matrix} 1 \text{ if } \sum_{\mathrm{i=j}}^{\mathrm{F}} (\mathrm{QF_i} \equiv 1) > 0 \\ 0 \text{ otherwise} \end{matrix}$$

Error! Bookmark not defined.Error! Bookmark not defined.(6)

and

$$\mathrm{QF}_{\beta} = \begin{matrix} 1 \text{ if } \sum_{\mathrm{i=j}}^{\mathrm{F}} (\mathrm{QF_i} \equiv -1) > 0 \\ 0 \text{ otherwise} \end{matrix}$$

Error! Bookmark not defined.Error! Bookmark not defined.(7)

QFs in Eq. (6) and (7) will be used to calculate α and β QMs. QM_α and QM_β will be calculated according to Eq. (2). Let it be noted that for a set of QFs, $\{\mathrm{QF_F} | \mathrm{QF_j} \in \mathrm{QF_F}\}$, QF_{α} and QF_{β} can both be set high for the same observation in the event one QF has been set high and another QF could not be computed due to a lack of ancillary data. Therefore, while the QMs for an individual QF will always sum to 100 %, QM_{α} and QM_{β} will always be less than or equal to 200 %.

Below, Table 2 summarizes how all QFs and QMs, including α and β QFs and QMs, are calculated for a particular DP.

Observation	QF_j	…	QF_F	QF_α	QF_β
i=1	$1 \vee 0 \vee -1$	⋱	$1 \vee 0 \vee -1$	If $(\sum_{i=f}^{F} \mathrm{QF}_i \equiv 1) > 0$ $\rightarrow \mathrm{QF}_\alpha = 1$ otherwise 0	If $(\sum_{i=f}^{F} \mathrm{QF}_i \equiv -1) > 0$ $\rightarrow \mathrm{QF}_\beta = 1$ otherwise 0
i=2	⋮	⋱	⋮	⋮	⋮
⋮	⋮	⋱	⋮	⋮	⋮
⋮	⋮	⋱	⋮	⋮	⋮
⋮	⋮	⋱	⋮	⋮	⋮
⋮	⋮	⋱	⋮	⋮	⋮
⋮	⋮	⋱	⋮	⋮	⋮
N	⋮	⋱	⋮	⋮	⋮

$QM_{j,1}$	$QM_{j,0}$	$QM_{j,-1}$		$QM_{F,1}$	$QM_{F,0}$	QM_{F-1}	QM_α	QM_β
Eq. (2)	Eq. (3)	Eq. (4)	…	Eq. (2)	Eq. (3)	Eq. (4)	Eq. (2)	Eq. (2)

Table 2. Illustration of how α and β QFs and QMs are determined for a DP with *N* observations and *F* QFs.

5.2 NEON Data Product Quality Assessment

When L0 DPs are used to generate a L1 DP, that L1 DP will be assessed by NEON on its data quality. A final QF, $\mathrm{QF_{NEON}}$, will be set depending on whether the DP has passed or failed NEON's assessment. The threshold for $\mathrm{QF_{NEON}}$ is based on the results of QM_α and QM_β, which establishes a limit for an acceptable amount of data that can either fail specific quality tests and or the tests could not be run due a lack of ancillary data. If a DP reaches or exceeds this threshold it is flagged as invalid data, i.e., $\mathrm{QF_{NEON}}$ = 1, and valid, i.e., $\mathrm{QF_{NEON}}$ = 0, if it does not. It is envisioned that the threshold for $\mathrm{QF_{NEON}}$ will change over time and may vary for different sensors. Here we present $\mathrm{QF_{NEON}}$ as a two to one ratio of QM_β to QM_α with maximums of 20% for QM_β and 10% $QM_{\alpha,}$ Eq. (8). Figure 3 represents the ratio of QM_β to QM_α in a graphical form.

$$\mathrm{QF_{NEON}} = \begin{matrix} 1 \text{ if } \mathrm{QM}_\beta + (2 * \mathrm{QM}_\alpha) \geq 20 \\ 0 \text{ otherwise} \end{matrix}$$

Error! Bookmark not defined.Error! Bookmark not defined.(8)

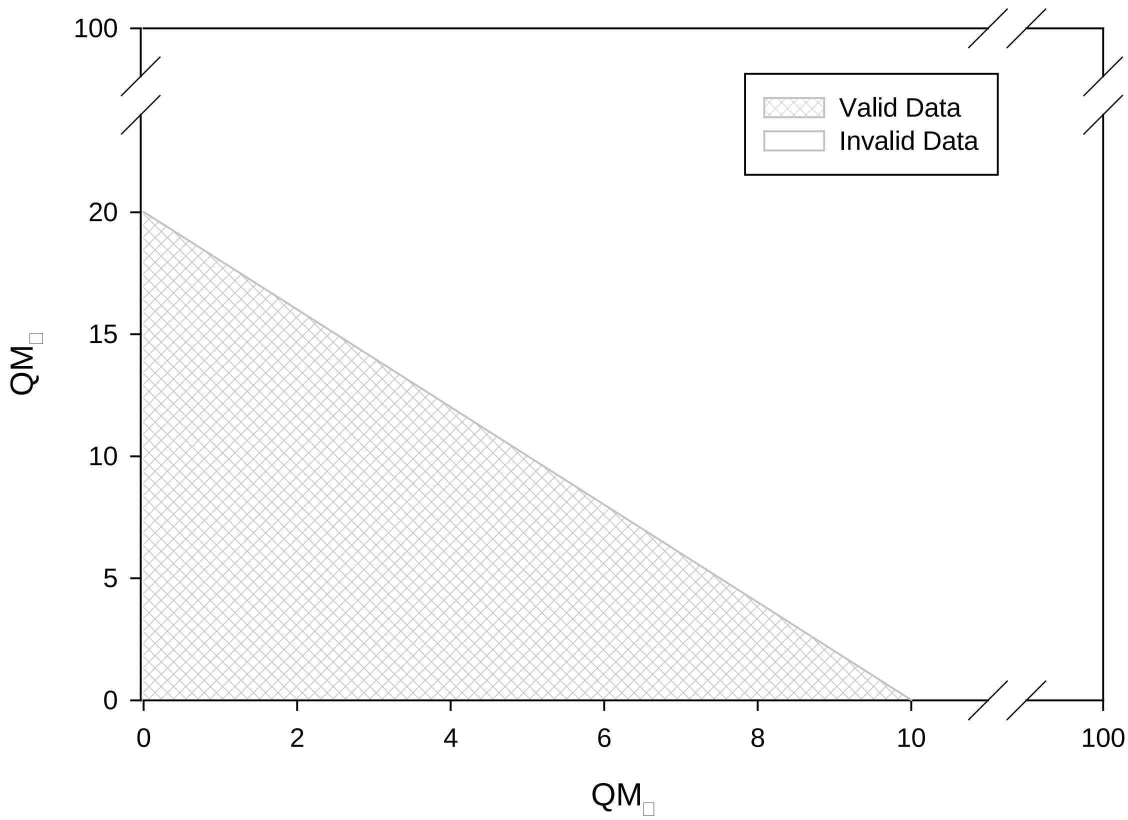

Figure 1. A graphical representation of QF_NEON, using Eq. 8, to determine whether a L1 DP is valid or not.

5.3 Quality Summary and Quality Report

Ultimately, each L1 DP will have the QF results from the QA/QC analyses and sensor tests presented in two separate schemes; a Q_{rpt} and a Q_{sum}. As previously stated, the Q_{rpt} will present the results of specific QFs as they relate to individual L0 DPs. For example, the Q_{rpt} for a thirty-minute temperature average, sampled at a rate of 1 Hz, allows the user to differentiate the 1800 outcomes of each QA/QC analysis and sensor test. The Q_{sum} will instead provide QMs for each QF. In addition, the Q_{sum} will include the final quality flag, $\mathrm{QF}_{\mathrm{NEON}}$, allowing users to quickly assess whether a DP is valid or not. A visual representation of QFs and QMs as they relate to the Q_{sum} and Q_{rpt} is shown in Figure 4. By presenting QA/QC and sensor tests in this manor, several levels of detail are retained on the quality of sensor data in order to facilitate data transparency and usability.

Q_{rpt}

Observation	QF_R	QF_δ	QF_s	QF_α	QF_β
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	0	0
5	0	0	0	0	0
6	0	0	0	0	0
7	0	0	0	0	0
8	0	0	1	1	0
9	1	0	1	1	0
10	1	0	0	1	0
11	0	0	0	0	0
12	0	0	0	0	0
13	0	0	0	0	0
14	0	0	0	0	0
15	0	-1	-1	0	1

QM_R (1)	QM_R (0)	QM_R (-1)	QM_δ (1)	QM_δ (0)	QM_δ (-1)	QM_s (1)	QM_s (0)	QM_s (-1)	QM_α	QM_β
13%	87%	0%	0%	93%	7%	13%	80%	7%	20%	7%

$QF_{NEON} = 1$

Q_{sum}

Figure 2. A QF and QM example for three QA/QC plausibility tests, range, delta, and step, and how they relate to the Qsum and Qrpt that will accompany a DP.

6 ALGORITHM VERIFICATION

Verification of the algorithms disclosed in this ATBD shall follow the procedures outlined in AD[01].

7 SCIENTIFIC AND EDUCATIONAL APPLICATIONS

While the framework presented here is applied only in the context of NEON DPs, it could also be easily transferred to other data sets. The quality of large data sets can often be difficult to interpret and the application of these algorithms is intended to balance information with accessibility. In addition, this framework was developed to cater to users with varying backgrounds and levels of expertise. It is our intention that the ideas presented here will be beneficial to others that are struggling to find meaningful techniques to assess the quality of their data.

8 FUTURE PLANS AND MODIFICATIONS

This ATBD will be version controlled, i.e. future developments might results in modifications to this ATBD, which will be documented accordingly.

9 APPENDIX

9.1 Variables

Variable	Explanation
i	Initial observation
N	Final observation
$\mathrm{QF_{NEON}}$	Final quality flag
QF_α	Alpha quality flag
QF_β	Beta quality flag
QM_α	Alpha quality metric
QM_β	Beta quality metric
$\mathrm{Q_{sum}}$	Quality summary
$\mathrm{Q_{rpt}}$	Quality report
$\mathrm{QF_j}$	Initial quality flag
$\mathrm{QF_F}$	Final quality flag
$\mathrm{QM_j}$	Initial quality metric
$\mathrm{QM_F}$	Final quality metric

10 BIBLIOGRAPHY

EPA. (2006). Data quality assessment: A reviewer's guide (QA/G-9R). U.S. EPA. pp. 61 [Online]. Avaliable: http://www.epa.gov/QUALITY/dqa.html [June, 2013].

Göckede, M., Rebmann, C., and T. Foken. (2004). A combination of quality assessment tools for eddy covariance measurements with footprint modeling for the characterization of complex sites. *Agricultural and Forest Meteorology*. 127, 175-188.

Mauder, M. and T. Foken. (2011). Documentation and instruction manual of the eddy-covariance software package TK3. pp. 60 [Online]. Available: http://opus.ub.uni-bayreuth.de/volltexte/2011/866/pdf/ARBERG046.pdf [June, 2013].

Mauder, M., Cuntz, M., Drüe, C., Graf, A., Rebmann, C., Schmid, H. P., Shmidt, M., and R. Steinbrecher. (2013). A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements. *Agricultural and Forest Meteorology*. 169, 122-135.

Wang, R. Y., Storey, V. C., and C. P. Firth. (1995). A framwork for analysis of data quality research. *IEEE Transactions on Knowledge and Data Engineering*. 7(4) 623-640.