# ARIZONA VALIDATING, EDITING, AND ESTIMATING (VEE) MONTHLY AND INTERVAL DATA

**1.      Introduction**

This section defines the interval data validation, editing, and estimation techniques required to participate as an MRSP.  Arizona VEE Rules will be reviewed six months from implementation for any changes, adjustments, or modifications.

**2.      Required Data Validation Checks**

Data validation checks are designed to identify things that can go wrong at the meter/recorder and cause the data collected to not reflect actual usage.

These rules apply to both kWh and kVARh data, depending on the data required by the meter's tariff(s).  If data is provided for informational purposes only (not used for billing purposes), validation is not required.  Data that has not gone through the validation process is raw data.

General MRSP's and MSP's business practices should ensure that the meter is programmed correctly for the required revenue data and that the MRSP's system is set up to accurately maintain information such as interval size, meter constants, and what quantity is recorded by what channel.  These VEE rules do not require or describe how the MRSP verifies that the meter is programmed correctly.

All validation checks must be run.  Failure of one check does not preclude the MRSP from performing other validation checks.

**Several words are used to describe the quality of interval data.**

- Raw data - data that has not gone through the VEE process

- Valid data - data that has gone through all required validation checks and either passed them all or been verified

- Verified data - data that failed at least one of the required validation checks but was determined to represent actual usage

- Estimated - data that has been calculated based on standard estimation rules because the raw data was not valid or available

- Adjusted – data that has been changed because of an incorrect pulse multiplier.

# ARIZONA VALIDATING, EDITING, AND ESTIMATING (VEE) MONTHLY AND INTERVAL DATA

The following validation checks are required for interval kWh and kVARh data.  They are described below in section 3 :

| Check | Purpose | Requirement |
|---|---|---|
| Time check of meter reading device/system | Check for the following:<br>• time drift of meter reading device/system outside the Applicable Regulatory Authority standard | Time of read |
| Meter identification check | Check for the following:<br>• meter ID was reported correctly<br>• meter has not been changed out<br>• data is being reported for the correct meter | Time of read |
| Time check of meter reading device system | Check for the following:<br>• time drift of meter clock outside the Applicable Regulatory Authority standard | Time of read |
| Pulse Overflow check | Check for the following:<br>• improper scaling factor in meter<br>• improperly sized transformer<br>• hardware problem | Time of read |
| Test Mode check | Check for the following:<br>• Data collected when meter was in test mode that represents test load rather than actual usage | Time of read |
| Sum check | Check for the following in combination meter/recorder installations:<br>• Crossed channels between meter and recorder<br>• pulse relay problems<br><br>Check for the following for all installations:<br>• invalid VT and CT ratios,<br>• invalid meter constants | Monthly (Time of read Optional) |
| Spike check | Check for the following for all installations:<br>• transmission error<br>• spike resulting from meter test.<br>Note that a spike can also occur after an outage - in this case the data is valid, but may or may not be used for peak billing depending on the tariff and company policy. | Time of Read (24-hour period) |
| kVARh check (for kWh data only if **corresponding** kVARh data available) | Check for the following:<br>• kWh channels are correctly mapped to kVARh channel<br>• meter is operating correctly | Monthly |

| Check | Purpose | Requirement |
|---|---|---|
| | | |
| High/Low Usage check | Check for the following in all installations:<br>• dropped phases<br>• inaccurate meter constants<br>• energy diversion<br>• fast/slow meters<br>Also check for the following in combination meter/recorder installations:<br>erratic pulse input to recorder | Monthly |

3. **Interval data collection and validation rules**

If interval data is read more often than required for billing, checks need to be performed at different times in the process. Some must be done as the data is read from the meter; some can be done anytime between when the data is collected from the meter and the end of the cycle, and others have to be done on a billing period basis at the end of the billing cycle. They are broken out that way in this description.

3.1. **Time check of meter reading device/system** will be performed by the MSP and MRSP once every 24 hours to ensure **that the collection device is synchronized to the national time** standard

3.2 **Collect data**

3.3. **As data is collected**

3.3.1. Check meter identification – verify that the meter's identification matches what is expected.

3.3.2. Perform Time Tolerance check on meter and data.

The time tolerance check is performed to minimize and correct meter clock drift and to minimize and correct the data problems associated with meter clock drift.

How to do Time Tolerance check on meter

To perform a time tolerance check on the meter, compare meter time to data collection device/system time. Note that depending on the communication technology used, network latency must be taken into account.

Pass/Fail Criteria
• If meter time is within 3 minutes of time standard, the meter passes the time tolerance check. (Note that if the meter time is

within the 3-minute tolerance, the meter time can optionally be corrected.)

- If meter time is off by more than 3 minutes, the meter time must be corrected. If the meter fails the time tolerance check for three consecutive months, the meter must be physically inspected/tested.

How to do Time Tolerance check on data (Interval Check)

To perform a time tolerance check on the data, compare the number of intervals retrieved from the meter to the number of intervals expected given the elapsed time.

Pass/Fail Criteria

- If the actual number of intervals is equal to the expected number, the data passes the time tolerance check.

- If the actual number of intervals differs from the expected number, the data fails the time tolerance (interval) check. The data to be corrected includes all intervals from the last time the meter time was determined to be good (i.e., within the 3-minute tolerance) and when it was discovered that the meter time was off by more than 3 minutes and the meter time was reset.

If data fails the Time Tolerance Check…

1) If the meter time was off by less than or equal 15 minutes, a point to point linear interpelation shall be performed.

2) If the meter time was off by more than 15 minutes, the data must be estimated. See estimation rules – 4.3

### 3.4. Either as data is collected or prior to publishing on MRSP Server

3.4.1. Perform Pulse Overflow check

Inspect each interval for this condition. If a pulse overflow occurs, the meter requires physical meter test/maintenance. Intervals with pulse overflows must be estimated.

3.4.2. Perform Test Mode check

Anytime the meter is placed in test mode and the actual data is not available, the data must be estimated (see Estimation Rules section 4.3).

3.4.3.   Perform Sum Check

The sum check is performed to on meters with encoded reads ensure that the difference between the energy use recorded in the intervals and the energy use recorded in the meter over the same time period is within an acceptable range.  This check may be done on either consumption or pulse data, provided the data scaling is consistent throughout the period.

How to do the Sum Check
1)   Calculate the energy use recorded in the intervals by summing the intervals between the start and stop meter readings.

2)   Calculate the energy use recorded by the meter by taking the difference between the start and stop readings accounting for possible rollover between start and stop readings.  For example, if the start reading was 99968 and the stop reading was 00294, and the meter reading rolls over at 99999, the difference would be 326.

3)   Compare the energy use recorded in the intervals to the energy use recorded by the meter.  Note that the values must be in the same units for the comparison.

Pass/Fail Criteria
*   If difference is <= 1.5 meter multipliers, the data passes the sum check.  (meter multiplier = CTR x VTR, where CTR is current transformer ratio and VTR is voltage transformer ratio)

*   If difference is > 1.5 meter multipliers, the data fails the sum check.

If data fails the Sum Check
1)   Several optional steps may be taken to resolve the sum check failure.

(a)   Reread the meter and redo the sum check from original start meter reading to new stop meter reading.

(b)   Redo the sum check, taking into account the differences between the time of the start read and the start of the first interval, and the time of the stop read and the end of the last interval.  See Attachment C-VEE-A, Section 1.2. for more information.

(c) Redo the sum check, taking into account missing or incomplete intervals. See Attachment C-VEE-A, Section 1.2.2 for more information.

(d) Additional checks may be performed, based on the technology used, to verify that the interval data is an accurate representation of usage as measured by the meter readings.

2) If sum check is not resolved, perform manual inspection of data.

(a) Verify meter and pulse multipliers. If a multiplier was incorrect, redo the sum check using the correct multipliers.

(b) Check for a meter change between the start and stop meter readings. If the meter was changed, redo the sum check for each meter independently.

(c) Manually inspect data. If the data seems reasonable, it can be considered verified.

(d) If the data does not seem reasonable, perform physical meter test/inspection. If meter tests OK, the data can be considered verified. If a problem is found with the meter, the data must be estimated. (Note: if the problem existed prior to this billing period, previously posted data must be adjusted and re-posted.) If unable to visit site and perform meter test prior to posting the data, the data must be estimated.

(e) If interval data is available but meter readings are not available, manually inspect the data. Data that seems reasonable (compared with historical data) can be considered verified. Any data that does not seem reasonable must be estimated. Pending – with resolution of  VEE Issue #3

3)

3.4.4. Perform Spike Check The spike check is performed to identify intervals with suspiciously high usage relative to the surrounding intervals. This check may be done on either

consumption or pulse data, provided the data scaling is consistent throughout the period.

How to  perform daily Spike Check

1)      For each 24-hour period, identify the highest and sixth highest peaks.  (Normally the 24-hour period is from midnight to midnight.  If the data is at the beginning of the span and doesn't start at midnight, use enough data from the next day of data to get 24 hours of data.  If the data is at the end of the span and doesn't stop at midnight, use enough data from the next to last day of data to get 24 hours.)

2)      If the highest peak is less than or equal to the spike check threshold of 10 pulses, skip the spike check.  (A spike check threshold is used to eliminate false spikes for meters with very low usage.)

3)      If the highest peak is greater than the spike check threshold of 10 pulses, divide the highest peak by the sixth highest peak

Pass/Fail Criteria

•       If highest peak divided by the Sixthhighest peak)is <= 1.8, the interval passes the spike check.

•       If highest peak–divided by the  sixth highest peakis  > 1.8, the interval fails the spike check.

If data fails the Spike Check

1)      Optionally reread the meter. If you get different value from reread, redo spike check.

2)      If value is the same on reread or you cannot reread the meter, perform manual inspection of data.

        (a)     Look for similar patterns on similar days.  If a similar pattern is found and this seems reasonable, the data can be considered verified.

        (b)     Optionally check with customer for unusual conditions at the time of the spike.  If a legitimate reason for spike is found, the data can be considered verified.

3)      If no similar pattern or legitimate reason for spike is found, the interval with the spike must be estimated.

4)      If there is a regular pattern of failing this check, the customer may be an irregular usage customer.  See section on Irregular Usage Customers for additional information.

3.4.5. If interval kVARh data is available, perform kVARh Check

The kVARh check is performed to identify intervals where reactive load (kVARh) is present and active load (kWh) is not, indicating a suspicious usage pattern and possible meter malfunction. This check is only required when both kWh and kVARh are used for billing. If kVARh data is available but not used for billing, the check is optional. This check may be done on either consumption or pulse data, provided the data scaling is consistent throughout the period.

How to do the kVARh Check
1)      If multiple kWh channels map to a single kVARh channel, or multiple kVARh channels map to single kWh channel, the appropriate channels must be totaled prior to this check.

2)      If there are any kWh intervals with zero consumption, check the corresponding kVARh interval.

Pass/Fail Criteria
*       If the corresponding kVARh interval is also zero or less than or equal to the kVARh check threshold of 10 pulses, the kWh data passes the kVARh check. (A kVARh check threshold is used to eliminate false errors for meters with very low usage.)

*       If the corresponding kVARh interval is greater than the kVARh check threshold of 10 pulses, the kWh interval fails the kVARh check.

If data fails the kVARh Check
1)      Several optional steps may be taken to resolve the kVARh failure.

(a)     Investigate to determine if this data represents actual customer usage, in which case the data can be considered verified.

(b)     If multiple kWh channels map to a single kVARh channel, investigate to determine if the problem can be directly traced to specific kWh channels. If this is the case, only data for those channels must be estimated. If the problem is not attributable to specific channels, all kWh channels need to be estimated.

2)      If no legitimate reason for the kVARh failure is found, the intervals with failures must be estimated.

3)      If there is a regular pattern of failing this check, the customer may be an irregular usage customer.  See section on Irregular Usage Customers for additional information.

### 3.5.    On the billing cycle for the meter

3.5.1.  High/Low Usage Check

This test must be performed on the data that has passed or been verified for previous checks, with no estimated values included.  This identifies metered usage that is suspiciously high or low relative to historical usage.  It may optionally also be performed on all data (valid and estimated) to provide a reasonableness check on the estimates derived using the standard estimation techniques.

This check must be done on consumption data, not pulses.

How to do the High/Low Usage Check
1)      If last year's data is available, calculate average daily usage for same billing month last year; use summed VEE or historical billing interval data if available, if not use VEE or historical billing usage (i.e., difference between register readings).

2)      If last year's data is not available, calculate average daily usage for the previous billing month; use summed VEE or historical billing interval data if available, if not use VEE or historical billing usage (i.e., difference between register readings).

3)      If last year's data and last month's data are not available, skip the high/low usage check.

4)      Calculate average daily usage for this billing month using either summed VEE data (if check includes estimated data) or sum of all intervals not requiring estimation (if check does not include estimated data).  If not all intervals are included in the sum, prorate the sum accordingly.

Pass/Fail Criteria
•      If |(historical daily average - this month's daily average)| <= 75 – 125% * historical daily average, the data passes the high/low usage check.

•      If |(historical daily average - this month's daily average)| > 75 – 125% * historical daily average, all data in the billing month fails the high/low usage check.

If data fails the High/Low Usage Check

1.        <u>Check to see if there has been a meter exchange</u>

2.        <u>Check the meter and pulse multiplier</u>

3.        Perform manual inspection of data.

    (a)     Look at recent history for the meter.  If monthly usage has been on a trend in the appropriate direction and this seems reasonable, the data can be verified.

    (b)     Optionally check with customer for changed usage patterns.  If changed usage patterns match change in data, the data can be verified.

    (c)     Check to see if some of the data looks reasonable; reasonable looking data can be verified.  For example, if a meter fails sometime during the month, the data at the beginning of the month may be OK, while the data after the meter failure may be obviously bad.

    (d)     If the data does not seem reasonable, perform physical meter test/inspection. If meter tests OK, the data can be verified.  If a problem is found with the meter, the data must be estimated.  (Note: if the problem existed prior to this billing period, previously posted data must be adjusted and re-posted.)  If unable to visit site and perform meter test prior to posting the data, the data must be estimated.

2)     If the data is investigated and found to be accurate, the data is verified.

3)     If the data fails high/low usage check, suspect data must be estimated.

4)     If there is a regular pattern of failing this check, the customer may be an irregular usage customer.  See section on Irregular Usage Customers for additional information.

**3.6.**    **After all validation checks have been performed and required data has been estimated, the MRSP must rerun all validation checks to ensure reasonableness of estimates.  If data fails validation after estimation, it needs to be manually verified for reasonableness.**

**3.7.**    **Record Keeping Requirements**

The data and the logs should be kept for a period of three years. If data failed one or more validation checks, the specific checks that the data failed must be recorded on an interval level, and:

1) If the data was manually verified, that information must be recorded on an interval level. Verified data is valid.

2) If the validation failure(s) were not resolved through accepted methods, the data must be estimated.

For each interval that is estimated, the MRSP must record the estimation algorithm used. Interval data estimation algorithms include:

- less than 2 hours (4.1)

- greater than 2 hours (4.2) – scaled based on usage

- intervals adjusted

- intervals manually estimated

- intervals estimated due to meter interval programmed incorrectly

### 3.8. Irregular Usage Customers Including Series of Spike Checks etc.

An irregular usage customer is one whose usage pattern does not follow normal usage patterns and consistently fails the spike check, kVARh check, or high/low usage check. An MRSP can identify a customer as an irregular usage customer if:

1) the customer data fails the standard validation check for three consecutive months and the MRSP verifies that the data represents the actual customer usage, OR

2) MRSP is notified by the customer's ESP or previous MRSP of the irregular usage pattern,

The data used to identify an irregular usage customer could be data collected by the MRSP, or historical data provided by the previous ESP or MRSP. An MRSP may modify the spike check and/or high/low usage check, and skip the kVARh check if an irregular usage customer consistently fails the check. The MRSP must notify both the customer's ESP and UDC of the customer's irregular usage status and what modified checks will be performed immediately.

The goal of the modified checks is to automate the manual procedures the MRSP would perform to verify that this is the customer's normal usage pattern. An MRSP may use a variation of the spike check or high/low usage check based on the actual usage pattern. Note that the MRSP may not skip the spike

check or high/low usage check.  If the data passes the modified check, the data is valid and does not need to be marked as verified.

- Examples of modifications for the spike check include modifying the spike check value (180%) or the pulse threshold value (10 pulses).

- Examples of modifications for the high/low usage check include changing the percentages (+/- 50%), using the year's average instead of one billing period's average, or comparing to the minimum and maximum values for the past year.

For some customers, irregular usage patterns are symptomatic of the business and will always be present, such as co-generation customers.  For other customers, irregular usage patterns may be a temporary condition, such as when a factory adds a second shift and fails the high/low usage check for the first 12 months.  The MRSP must determine whether a customer is a permanent or temporary irregular usage customer.  Temporary irregular usage customers must be reviewed annually to determine if they are still irregular usage customers or should be returned to the normal checks.

# ARIZONA VALIDATING, EDITING, AND ESTIMATING (VEE) MONTHLY AND INTERVAL DATA

4.     **Interval Data Estimation Rules**

A customers bill is considered estimated when 10% or more of intervals are estimated in a billing period.

Estimate intervals needing estimation using the following estimation rules

**4.1.   If section of data needing estimation is 2 hours or less in length, use point-to-point linear interpolation to estimate the data.  Intervals containing a power failure cannot be used as end points for interpolation.**

How to apply Point-to-Point Linear Interpolation
1)     If the section occurs in the middle of the data, the "first point" is the last valid interval before the section, and the "second point" is the first valid interval after the section.

2)     If the section occurs at the beginning of the span, use the last interval from the historical data as the first point if the historical data is available and valid.  Otherwise, use the second point (the first valid interval after the section) as the first point – this will cause the load to be estimated as a flat load.

3)     If the section occurs at the end of the span, use the first point (the last valid interval before the section) as the second point – this will cause the load to be estimated as a flat load.

**4.2    If the section of data needing estimation is more than 2 contiguous hours, use a scaling methodology to spread the kWh over selected reference days to estimate the data.**

**Determine the amount of kWh (X) for the section of data needing estimation.**
- If the dial reads are good for the bill cycle, use them to compute the total kWh for the bill cycle.  To compute the amount of kWh (X) for the section of data needing estimation:  X = total kWh for the billing cycle minus the kWh for the good intervals
- If the dial reads do not include enough kWh to include the missing portion of the interval data, then the kWh for the missing portion or the section needing estimation should be estimated.
  - If the timeframe of data needing estimation exceeds one week, then check with the UDC *(if applicable)* to be sure there wasn't an outage, and/or check with the MSP to find out what happened to the meter and check with the customer to be sure that they were operating normally during the time when the interval data is missing or needs estimation.

# ARIZONA VALIDATING, EDITING, AND ESTIMATING (VEE) MONTHLY AND INTERVAL DATA

- Estimate the kWh for the missing data or timeframe needing estimation using the average daily kWh from the current bill cycle if available or the average daily kWh for the billing cycle from the previous year, same bill month or the average daily kWh for the billing cycle from the same year, previous month or one of these values modified to account for current weather conditions and customer operating patterns.

### Determine the reference days.
- The goal is to select reference days whose load shapes will most closely resemble the timeframe needing estimation.
- Select days closest in time. If none are available that would be appropriate, try the same seasonal group of months. If none are available that would be appropriate, try the same time frame in the previous year.
- Seasonal months are:      Fall (September 15 through November 15)
  Winter (November 16 through February 15)
  Spring (February 16 through May 15)
  Summer (May 16 through September 14)
- Select "same weekdays" (Mondays for Mondays, Tuesdays for Tuesdays) when possible. If not enough are available, use "like days" (weekdays for weekdays, weekends for weekends).
- If a holiday falls in the estimation period and the customer observes that holiday, then use a holiday load shape as a reference day. If no holidays are available and the customer's holiday usage is similar to his Sunday usage, substitute a Sunday for the holiday. If the customer does not observe the holiday, use the appropriate "same weekday".
- If the estimation period exceeds one week, try to have at least one week of reference days.

### Allocate the kWh for the estimation period.
- If the reference period includes multiples of the "same weekdays", average them together to create an average Monday, Tuesday etc.
- Place the reference days into the portion of the file to be estimated.
- Sum the kWh (**Y**) of the estimated (reference days) portion of the file.
- Take the kWh calculated for the section needing estimation and divide it by the sum of the kWh in the reference days.      **Scale factor=X/Y**
- Multiply every interval in the estimation period by the scale factor.

**The maximum demand for the bill cycle should not be estimated, unless the entire cycle is estimated. In this case, use historical date to define the maximum demand.**

### 4.3 Correcting data problems attributable to metering problems

# ARIZONA VALIDATING, EDITING, AND ESTIMATING (VEE) MONTHLY AND INTERVAL DATA

## Rules for assigning the file status and individual interval status

If on investigation the cause of the data problem is determined to be a problem with the meter or meter installation use the following rules to correct the data by scaling or estimating the intervals or the meter reads.

**4.3.1** In the case of an incorrect pulse or meter constant being applied to the data, the MRSP will be notified of the following:

1. The time period requiring action.
2. The correct pulse or meter constant.

### Without prior knowledge of a problem

- For the original data file, if the intervals were sent prior to having any knowledge that the meter data was incorrect, the intervals would be marked as **Actual**.
- Any subsequent corrections will result in the new file being marked as **Corrected** and each affected interval marked as **Adjusted**.

### With prior knowledge of a problem

- For the original file, if there was prior knowledge that the meter interval was incorrect and corrections are made by applying the correct pulse constant to each of the affected intervals and the intervals would be marked as **Actual**.
- Any subsequent corrections will result in the new file being marked as **Corrected** and each affected interval being marked as **Adjusted**.

**4.3.2** In the case of the meter being programmed incorrectly, wired incorrectly causing the meter to run fast or slow or one or more phase being dropped, the MRSP will be notified of the following:

1. The time period requiring action.

### Without prior knowledge of a problem

- For the original file, if the intervals or were sent prior to having any knowledge that the meter intervals was incorrect, the data would be marked as **Actual**.
- For any subsequent corrections the new file will be marked as **Corrected** and each affected interval will be marked as **Estimated**.

# ARIZONA VALIDATING, EDITING, AND ESTIMATING (VEE) MONTHLY AND INTERVAL DATA

## With prior knowledge of a problem

- For the original file, if there was prior knowledge that the meter data was incorrect the file will be marked as **original** and each affected interval will be marked as **Estimated**.

- For the original file, if there was prior knowledge that the meter data was incorrect and corrections result in less than 10% of the total intervals being estimated, the file will be marked as Original and each affected interval marked as **Estimated**.

- **For any subsequent corrections which result in the sum of the original and subsequent estimated intervals the new file will be marked as Correct and each affected interval marked as Estimated.**

### 4.3.3 Interval in meter doesn't match tariff or settlement requirements

If the meter programming and the MRSP requirements are inconsistent, the data is calculated as follows and the meter program must be corrected immediately:

1) The meter is programmed to collect data at a smaller interval than required by its tariffs, and the meter's interval evenly divides into the interval required by the tariff – for example, the meter was programmed to collect 5-minute data, and the tariff requires 15-minute data. Sum the 5-minute intervals into 15-minute intervals on even 15-minute boundaries. If the data passed all the other validation checks, it is valid and does not need to be marked as estimated or verified.

2) The meter is programmed to collect data at a larger interval than required by the meter's tariff. For example, the meter is programmed to collect 60-minute intervals, but the tariff require 15-minute intervals. Prorate the data by assuming an even load distribution during the interval. In this example, the usage in the 60-minute interval would be divided by 4 to estimate the usage in a 15-minute interval. The data is marked as estimated. The meter must be reprogrammed to the correct interval.

3) The meter is programmed to collect data at a smaller interval than required by its tariff, but the meter's interval doesn't evenly divide into the interval required by the tariff. For example, the meter is programmed to collect 10-minute intervals, and the tariff requires 15-minute intervals. The data would be estimated and marked as estimated. To estimate data, all collected intervals that are contained within the

required reporting interval are included in the appropriate reporting interval.  Collected intervals that cross the boundaries of required reporting intervals are included proportionally in both reporting intervals. In this example, if there were three 10-minute intervals containing 10 kWh, 20 kWh, and 30 kWh, the corresponding estimated 15-minute intervals would contain 20 kWh (10 + 0.5*20) and 40 kWh (0.5*20 + 30).  This is similar to the prorating technique discussed in Attachment C-VEE-A Section 1.1.  The meter must be reprogrammed to the correct interval.

## 5.0    Re-posting Data

The MRSP will be required to re-post actual or corrected data regardless of the amount of usage or time that has elapsed since the posting of the original file.

It will be up to the discretion of the biller and/or settlement whether the data is used for rebilling.