

PERMANENT STANDARDS WORKING GROUP

APPENDIX C-VEE

REQUIREMENTS FOR VALIDATING, EDITING, AND ESTIMATING MONTHLY AND INTERVAL DATA IN DIRECT ACCESS

July 29, 1998

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

VOTING RESULTS (Plenary Meeting on July 13, 1998)

For: 19	Against: 1	Abstain: 1
ABB	ENRON	FirstPoint
Applied Metering Technologies		
California Energy Commission		
CPUC-ORA		
Cellnet		
EPRI		
eT Communications		
GE		
ITRON		
LADWP		
NERTEC		
PG&E		
Phaser		
SCE		
Schlumberger		
SDG&E		
Sierra Pacific Power		
Southern Company		
TeCom		

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Section A: CA Interval Data VEE Rules (Revision 2.0)

1. Introduction

This section defines the interval data validation, editing, and estimation techniques required to participate in the California market as an MDMA.

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 MDMA and MSP business practices should ensure that the meter is programmed correctly for the required revenue data and that the MDMA 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 MDMA verifies that the meter is programmed correctly.

All validation checks must be run. Failure of one check does not preclude the MDMA 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

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

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Check	Purpose
Time check of meter reading device/system	Check for the following: <ul style="list-style-type: none"> • time drift of meter reading device/system outside CPUC standard
Meter identification check	Check for the following: <ul style="list-style-type: none"> • meter ID was reported correctly • meter has not been changed out • data is being reported for the correct meter
Time check of meter	Check for the following: <ul style="list-style-type: none"> • time drift of meter clock outside CPUC standard
Pulse Overflow check	Check for the following: <ul style="list-style-type: none"> • improper scaling factor in meter • improperly sized transformer • hardware problem
Test Mode check	Check for the following: <ul style="list-style-type: none"> • Data collected when meter was in test mode that represents test load rather than actual usage
Sum check	Check for the following in combination meter/recorder installations: <ul style="list-style-type: none"> • Crossed channels between meter and recorder • pulse relay problems <p>Check for the following for all installations:</p> <ul style="list-style-type: none"> • invalid PT and CT ratios, • invalid meter constants
Spike check	Check for the following for all installations: <ul style="list-style-type: none"> • transmission error • spike resulting from meter test. <p>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.</p>
kVARh check (for kWh data only if corresponding kVARh data available)	Check for the following: <ul style="list-style-type: none"> • kWh channels are correctly mapped to kVARh channel • meter is operating correctly

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Check	Purpose
High/Low Usage check	Check for the following in all installations: <ul style="list-style-type: none"> • dropped phases • inaccurate meter constants • energy diversion • fast/slow meters Also check for the following in combination meter/recorder installations: erratic pulse input to recorder

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 ensures that the collection device is synchronized to the national time standard before data collection begins.

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 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.)

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

- 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

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 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 to 75 minutes, prorate the data using one of the algorithms in Attachment C-VEE-A A, Section 1.
- 2) If the meter time was off by more than 75 minutes, the data must be estimated.

3.4. Either as data is collected or prior to publishing on MDMA 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

If the MDMA determines the meter was in test mode, the MDMA must ensure that the customer is not billed for the test load. If no actual customer usage data is available for the time in which the meter was in test mode, zero usage is reported for that period; this data is valid. If the meter is inadvertently left in test mode, the data must be estimated if actual usage data is not available from the meter .

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

3.4.3. Perform Sum Check

The sum check is performed to 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 \leq two meter multipliers, the data passes the sum check. (meter multiplier = $CTR \times VTR$, where CTR is current transformer ratio and VTR is voltage transformer ratio)
- If difference is $>$ two 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 2.1 for more information.
 - (c) Redo the sum check, taking into account missing or incomplete intervals. See Attachment C-VEE-A, Section 2.2 for more information.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

- (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.
- 3) If the sum check failure cannot be resolved, the data must be estimated.

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 do the Spike Check

- 1) For each 24-hour period, identify the highest and third 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

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

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, subtract the third highest peak from the highest peak and divide by the third highest peak.

Pass/Fail Criteria

- If $(\text{highest peak} - \text{third highest peak}) / \text{third highest peak} \leq 1.8$, the interval passes the spike check.
- If $(\text{highest peak} - \text{third highest peak}) / \text{third highest peak} > 1.8$, 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

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

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 4 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 4 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

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

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 $|(\text{historical daily average} - \text{this month's daily average})| \leq 0.5 * \text{historical daily average}$, the data passes the high/low usage check.
- If $| (\text{historical daily average} - \text{this month's daily average}) | > 0.5 * \text{historical daily average}$, all data in the billing month fails the high/low usage check.

If data fails the High/Low Usage Check

- 1) 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

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

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, optionally rerun validation checks to ensure reasonableness of estimates.

3.7. Record Keeping Requirements

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 MDMA must record the estimation algorithm used. Interval data estimation algorithms include:

- less than 2 hours (4.1)
- greater than 2 hours (4.2) – not scaled based on usage
- greater than 2 hours (4.2) – scaled based on usage
- time-drifted intervals prorated (Attachment C-VEE-A, Section 1.1)
- time-drifted intervals prorated (Attachment C-VEE-A, Section 1.2)
- intervals adjusted (4.3)
- intervals manually estimated (4.2.8)

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

- intervals estimated due to meter interval programmed incorrectly (4.4)
- load profile template used (4.2.7)

3.8. Irregular Usage Customers

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 MDMA 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 MDMA verifies that the data represents the actual customer usage, OR
- 2) MDMA is notified by the customer's ESP or previous MDMA of the irregular usage pattern.

The data used to identify an irregular usage customer could be data collected by the MDMA, or historical data provided by the previous ESP or MDMA. An MDMA 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 MDMA must notify both the customer's ESP and UDC of the customer's irregular usage status and what modified checks will be performed.

The goal of the modified checks is to automate the manual procedures the MDMA would perform to verify that this is the customer's normal usage pattern. An MDMA may use a variation of the spike check or high/low usage check based on the actual usage pattern. Note that the MDMA 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 MDMA must determine whether a customer is a permanent or temporary irregular usage customer. Temporary irregular usage customers

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

must be reviewed annually to determine if they are still irregular usage customers or should be returned to the normal checks.

4. Interval Data Estimation Rules

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 the average of selected reference days to estimate the data.

Rules and definitions for selecting reference days for estimation:

- “Same weekdays” are defined as the same day of week as the day that needs estimation. In the case of holidays, “same weekdays” are holidays.
- “Like days” are defined as the same daytype (i.e., weekday, weekend, or holiday) as the day that needs estimation.
- A standard list of holidays will be used, regardless of the tariff or service territory of the meter. The “legal” definitions of the holiday is used; if the holiday falls on a Sunday, the “legal” holiday is the following Monday. Otherwise the “legal” holiday is the same date as the actual holiday. The holidays used are the following:
 - New Years Day
 - Presidents Day
 - Memorial Day
 - Independence Day
 - Labor Day

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

- Veterans Day
 - Thanksgiving Day
 - Christmas Day
-
- Only “valid” intervals can be used for estimation. Valid intervals are defined as those that have passed all validation checks or have been verified. Estimated intervals cannot be used for estimation.
 - Data from days with a power failure cannot be used for estimation. Power failures can cause irregular usage patterns, resulting in data that is not typical for the customer.
 - Valid intervals from “partial” days can optionally be used for estimation. “Partial” days are defined as those containing estimated data or those on which data collection began at some time other than midnight.
 - Historical data up to 90 days prior to the day needing estimation and from the current billing period may be used for estimation.
 - Reference days are chosen to be the closest chronologically to the data needing estimation, regardless of seasonal crossover. This may include days after the day requiring estimation. If two potential reference days are equidistant from the day requiring estimation, use the earlier day first. For example, if June 2, 1998 needed estimation and the billing period was from June 1 to June 30, the reference days used would be May 19, May 26, and June 9, provided they contained valid data.

4.2.1. Develop a daily profile

- 1) Find the three “same weekdays” with valid data closest in time to the day with the data needing estimation based on the rules listed in the previous section. If the day with data needing estimation is a holiday, the “same weekdays” are holidays, not the same day of week. If not enough historical holidays exist in the current billing period or previous 90-day period, use Sundays. Calculate the average daily profile using the three selected days.
- 2) If only two same weekdays are available from the current billing period and 90 days of historical data, calculate the average daily profile using the two selected days.
- 3) If only one same weekday is available from the current billing period and 90 days of historical data, use it as the daily profile.
- 4) If no same weekdays are available in the current billing period and 90 days of historical data, look for the three “like” days that are closest chronologically to the day with intervals needing

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

estimation. For example, if the intervals needing estimation were on Tuesday, use Monday, Wednesday, and Thursday. Only use weekdays with weekdays; only use weekends with weekends; only use holidays or Sundays with holidays. Calculate the average daily profile using the three selected days.

- 5) If only two like days are available from the current billing period and 90 days of historical data, calculate the average daily profile using the two selected days.
- 6) If only one like day is available from the current billing period and 90 days of historical data, use it as the daily profile.

4.2.2. Use the daily profile to estimate the required data

- 1) Estimate the data needing estimation by applying the appropriate intervals from the average daily profile to fill the missing intervals.
- 2) If start and stop meter readings are available and known to be good, they may optionally be used to scale the estimated interval data. See Attachment C-VEE-A, Section 3 for more information.

4.2.3. If there are no similar days or like days, use the load profile for the customer's class to estimate the data. Use this month's usage, if available, to scale the load profile. If this month's usage is not available, use last month's usage or last year's usage, whichever is determined to be more reasonable, to scale the load profile. Refer to the UDCs' load profile documentation for more information on applying load profiles. Note that the load profiles are for hourly data. If 15 minute data is required, assume a flat load throughout the hour (i.e., each 15-minute interval would have 1/4th the hourly usage).

4.2.4. If there is no historical data that can be used, the data must be estimated manually and the process and assumptions documented.

4.3. Correcting Data Problems Attributable to Metering Problems

If on investigation the cause of the data problem is determined to be a problem with the meter or meter installation and the data can be corrected by scaling the intervals and meter readings, the MDMA will be notified of:

- 1) The time period requiring correction.
- 2) The scaling factor to be applied to each interval in that period.

Examples of these situations include a meter running slow, a meter running fast, one or two phases dropped, etc.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

When the data is posted, it is marked as estimated if it had not been previously posted, and marked as adjusted if it had previously been posted

4.4. Interval in meter doesn't match tariff or settlement requirements

If the meter programming and the MDMA requirements are inconsistent, the data is calculated as follows:

- 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 ($10 + 0.5 \cdot 20$) and 40 kWh ($0.5 \cdot 20 + 30$). 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.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Section B: CA Monthly Data VEE Rules(Revision 1.1)

1. Introduction

This document defines the data validation, editing, and estimation techniques required to participate in the California market as an MDMA for monthly data. Monthly data includes consumption, demand, and Time-of-Use (TOU) consumption and demand..

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 not to reflect actual usage.

The following checks are required for monthly data validation for kWh and kW data. Similar checks would apply to kVARh and kVAR data if those values are required.

Check	Purpose
Time check of meter reading device/system (applies to devices/systems collecting TOU data only)	Check for the following: <ul style="list-style-type: none"> • time drift of meter reading device/system outside CPUC standard
Time check of meter (applies to meters collecting TOU data only)	Check for the following: <ul style="list-style-type: none"> • time drift of meter clock outside CPUC standard
High/low usage check	Check for the following: <ul style="list-style-type: none"> • misread • fast/slow meter • broken meter • incorrect multipliers • energy diversion • dropped phases
High/low demand check (applies to demand readings only)	Check for the following: <ul style="list-style-type: none"> • misread • fast/slow meter • broken meter • incorrect multipliers • energy diversion • dropped phases

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Check	Purpose
Time-of-use check (applies to TOU data only)	Check for the following: <ul style="list-style-type: none"> • misread • fast/slow meter • broken meter • incorrect multipliers • energy diversion • dropped phases
Zero consumption for active meters	Check for the following: <ul style="list-style-type: none"> • energy diversion • meter doesn't register
Usage for inactive meters	Check for the following: unauthorized usage at a site for which there is no customer with financial responsibility
Number of dials on meter	Check for the following: <ul style="list-style-type: none"> • wrong meter • misread
Dial decimal quantity	Check for the following: <ul style="list-style-type: none"> • wrong meter • misread
Meter identification	Check for the following: <ul style="list-style-type: none"> • that the meter ID was reported correctly • the meter has not been changed out • the data is being reported for the correct meter

3. Rules for Monthly Data Validation Check

Some monthly data validation checks must be done at the time of meter reading, and other checks could be done anytime after the meter is read until the data is posted. For example, the meter identification check must be performed at the time of the meter reading, while the high/low usage check can be performed in a handheld system as the meter is read, or back in a host system. All checks are not applicable to all types of data. The following table summarizes which checks must be done for each type of data, and provides a recommended sequence.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Check	Must be done at time of meter read?	Consumption	Demand	TOU Consumption	TOU Demand
Time in meter	Yes	n/a	n/a	3	3
High/Low usage	No	3	n/a	4	n/a
High/Low demand	No	n/a	3	n/a	4
TOU usage	No	n/a	n/a	6	n/a
Zero consumption for active meters	No	4	n/a	5	n/a
Number of dials	Yes	2	n/a	2	n/a
Number of demand decimal places	Yes	n/a	2	n/a	2
Meter ID	Yes	1	1	1	1

Most of the checks and estimation algorithms are based on historical data for the same customer and the same site. In areas with wide fluctuations in weather, this may not provide the best data for residential customers, as residential usage patterns vary much more with changes in weather than larger customers. A separate set of High/Low usage validation check and estimation rules are provided based on day-before usage of similar customers in the same geographic area.

3.1. Time check of Meter Reading Device/System

This check only applies for meter reading devices and systems collecting Time-of-Use data. Time check of meter reading device/system ensures that the collection device is synchronized to a national time standard before data collection begins

3.2. Time tolerance check of meter

The time tolerance check is only required if the meter is collecting Time-of-Use (TOU) data. It verifies that the meter's time is correct, and that TOU data represents the appropriate time periods. Note that depending on the communication technology used, network latency must be taken into account.

- 3.2.1. If time in meter is within +/- 3 minutes of the time standard, the data has passed Time Tolerance check. Note that if the meter is within +/- 3 minutes of the standard, the time in the meter can optionally be corrected.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

- 3.2.2. If time in meter is off > 3 minutes but <= 55 minutes, the data passes the Time Tolerance check. The data does not need to be estimated, but the MDMA must record the fact that the meter's time was off by this amount in case there is a later question about the data. The meter time may optionally be reset.
- 3.2.3. If time in meter is off > 55 minutes, the data fails the time tolerance check and must be estimated. The time in the meter must be reset. If the meter fails the time tolerance check after being reset for three consecutive months, the meter must be physically inspected/tested.

3.3. High/Low Usage

The High/Low Usage check validates cumulative consumption (kWh). Two methods are provided - one based on historical data, and one based on previous day data from similar customers. An MDMA may implement either check, depending on weather characteristics and density of meter population served by the MDMA. The second check requires a minimum density of meter population to be statistically accurate; this still needs to be determined.

3.3.1. Method based on historical data

- 3.3.1.1. Calculate the average daily usage (ADU) for the present billing period.

For example, if all constants and factors have already been applied to the reads, the ADU could be calculated by:

$$\text{ADU} = (\text{current billing read} - \text{previous billing read}) / (\# \text{ days between billing reads})$$

If the previous billing read were on June 1, and the present billing read is on June 30, there would be 29 days between billing reads.

- 3.3.1.2. Calculate the historical ADU

- 3.3.1.2.1. If there is not at least one month (minimum 27 days) of historical billing data available for the same customer and site, this check is not performed.

- 3.3.1.2.2. If data for the same customer and site is available, calculate the ADU for the same billing period last year. Use this as the historical billing ADU. One way to determine which billing period last year is "the same" is to choose the mid-point of this year's billing

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

period, and find the billing period last year that included the same date. For example, if the billing period this year was from April 13 to May 13, the mid-point would be April 28. If the billing period has an even number of days, use the day after the middle as the mid-point. For example, if the billing period was from June 1 to June 30, the mid-point would be June 16. Another way to determine which billing period is “the same” is to find the billing period last year with its read date in the same calendar month as the read date of the data being validated.

- 3.3.1.2.3. If there is no data from a year ago but there is data for the last billing period (minimum 27 days), calculate the ADU for the last billing period and use this as the historical ADU.
- 3.3.1.3. Compare the present billing period ADU with the historical billing period ADU. If the present billing period ADU is between 40% and 200% (inclusive) of the historical ADU, the data passes this check. (Note that some systems may convert ADU to a billing period usage to perform the check.) Optional trend factors that take into account peer group usage based on demographics, climactic areas, and customer class may be applied to the ADU to refine the High/Low comparison check. Sample trend factor calculations are to be provided at a later date.
- 3.3.1.4. If the present billing period ADU is not within 40% to 200% (inclusive) of the historical billing period ADU, the data fails this check. Optionally re-read the meter.
 1. If the reread is essentially at the same time as the original read and a different value is obtained, assume the first reading was a misread and perform the check again with the new reading. If the same reading was obtained, assume the meter reading is correct; the data failed the High/Low Usage check but is verified. OR

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

2. If the reread is not at the same time as the original read, recompute the average daily usage using the new reading. If the new ADU is within +/- 20% of the previous ADU, the data fails the High/Low usage check but is verified. If the new ADU is not within +/- 20% of the previous ADU, the data fails the High/Low usage check.

- 3.3.1.5. Data that fails the High/Low usage check and has not been verified may be investigated and manually verified if justification is found; otherwise the data must be estimated.

3.3.2. Method based on previous day usage of similar customers

Note that this method requires a certain density of customer data for residential customers in the same geographic area, where weather patterns are typically consistent throughout the geographic area.

- 3.3.2.1. The following steps are performed at the end of each meter reading cycle day for each geographical area in order to validate and estimate usage the following day:

- 3.3.2.1.1. At the end of the reading day, for each good meter read (open account, billed, between 27-33 days & ADU =< 100), perform the following calculations to determine an ADU for the billing period:

1. Calculate ADU (= KWH/days in billing period)
2. Add ADU to Sum of Current ADU
3. Calculate ADU squared
4. Add ADU squared to Sum of Current ADU squared
5. Add 1 to total meters
6. Calculate last month's ADU
7. Calculate current ADU times last month's ADU
8. Calculate last month's ADU squared

- 3.3.2.1.2. Determine which range of usage (high, medium or low) the current ADU should be grouped with by comparing current ADU to yesterday's ADU Low and High Range Factors (Reference 3.3.2.1.3 for ADU low and high range factor calculation methodology)

1. If the current ADU is less than yesterday's ADU low range factor:
 - Add current ADU to Sum of current low ADU
 - Add last month's ADU to Sum of last month's

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

low ADU

- Add current ADU times last month's ADU to Sum of current low ADU times last month's ADU
- Add current ADU squared to Sum of current low ADU squared
- Add last month's ADU squared to Sum of last month's low ADU squared
- Add 1 to total low meters

2. If the current ADU is not less than the ADU low range factor from yesterday and is less than the ADU high range factor, add the figures to the medium range following same format is in 3.3.2.1.2. step 1.
3. Otherwise, add the current ADU to the ADU high range following the same format in 3.3.2.1.2. step 1.

3.3.2.1.3. Calculate an aggregated ADU for current data for each geographic area

3.3.2.1.3.1. Sum together the ADU values for each geographic area

3.3.2.1.3.2. Calculate the mean for the total ADU (=Sum of Current ADU / total meters)

3.3.2.1.3.3. Calculate the standard deviation for the total

3.3.2.1.3.4. Calculate the current ADU low and high range factors:

- ADU Low Range Factor = mean - .43 Standard Deviation. If ADU Low range factor is less than the total current mean * .5, the ADU low range factor becomes the mean half.
- ADU High Range Factor = mean + .43 Standard Deviation
- NOTE: By determining the low & high factors, the Medium Range = (mean - .43 Standard Deviation) to (mean + .43 Standard Deviation)

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

3.3.2.1.4. For each of the three ranges determined above (low, medium, and high), calculate a percent of change of monthly usage for each geographic area.

3.3.2.1.4.1. After each meter's current billing period's (ADU) is grouped in 3 Ranges (Low, Medium, and High) as specified in 3.3.2.1.2, the following data are summed up by ADU range and area:

- Number of customers
- Sum of all last month's ADU
- Sum of all current month's ADU
- Sum of {each last month's ADU times current month's ADU}
- Sum of {all last month's ADU squared} i.e., Square all ADU, then sum them.
- Sum of {all current month's ADU squared}

From the data above modified ADU mean factors and standard deviation factors are determined for each range as follows:

Modified Mean Factor:

Sum of {last month ADU times current month ADU} divided by the sum of {all last month's ADU squared}

Modified Standard Deviation Factor:

Step 1: (Sum of {all current month's ADU squared} minus (Mean squared times sum of {all last month's ADU squared}) divided by (Total Meters minus 1).

Step 2: Take square root of Step 1

3.3.2.1.5. Calculate high and low range factors.

Calculate high and low range factors (HRF and LRF) for each of the 3 usage ranges within a geographic area. The mean is used to calculate estimated reads, and the high and low range factors are used in this validation check. 2.8 and 3.5 are used in the below example to represent the range deviation factor and will allow for an appropriate meter read error rate.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

This factor can be changed to control the error rate.

High Range Factor Formula:

$$\text{HRF} = 1 + \{(2.8 \times \text{Modified Standard Deviation} \times \text{Number of meters}) / (\text{Sum of Current month's ADU})\}$$

Low Range Factor Formula:

$$\text{LRF} = 1 - \{(3.5 \times \text{Modified Standard Deviation} \times \text{Number of meters}) / (\text{Sum of Current month's ADU})\}$$

3.3.2.2. As each meter is read, perform the following using the values calculated from the previous meter reading days' data.

- 3.3.2.2.1. Determine the usage from the preceding billing month and the preceding billing reading for the customer and site.
- 3.3.2.2.2. Calculate low limit for this month's usage by multiplying the preceding month's usage by the low range factor determined above.
- 3.3.2.2.3. Calculate high limit for this month's usage by multiplying the preceding month's usage by the high range factor determined above.
- 3.3.2.2.4. If the current usage is between the low and high limit calculated in the previous two steps, the data passes the High/Low check.

The following is a representation of how the High and Low Range Factors are used to validate meter usage:

Sample High/Low Usage Check:

Customer's previous usage = 400 kWh

High Range Value: 400kwh X 1.115 HRF* = 446

Low Range Value: 400kwh X .885 LRF* = 354

Usage values falling between 354 and 446 are accepted. Usage values outside this range fail the check.

* HRF/LRF = High and Low Range Factors, see description above.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

- 3.3.2.2.5. If the current usage is outside the low and high limit, the data fails the High/Low check. Optionally re-read the meter.
1. If the reread is essentially at the same time as the original read and a different value is obtained, assume the first reading was a misread and perform the check again with the new reading. If the same reading was obtained, assume the meter reading is correct; the data failed the High/Low Usage check but is verified. OR
 2. If the reread is not at the same time as the original read, re-compute the average daily usage using the new reading. If the new ADU is within +/- 20% of the previous ADU, the data fails the High/Low usage check but is verified. If the new ADU is not within +/- 20% of the previous ADU, the data fails the High/Low usage check.
- 3.3.2.2.6. Data that fails the High/Low usage check and has not been verified may be investigated and manually verified if justification is found; otherwise the data must be estimated.

3.4. High/Low Demand

The High/Low Demand Check compares the demand against historical data as a reasonableness check.

- 3.4.1. Determine the peak demand for this billing period.
- 3.4.2. Determine the historical peak demand.
- 3.4.2.1. If there is not at least one month (at least 27 days) of historical billing demand data available, skip this check.
 - 3.4.2.2. If demand data for this customer and site is available for the same billing period last year, use that as the historical peak demand. (Refer to section 3.3.1.2.2 to determine same billing period last year.)
 - 3.4.2.3. If demand data is not available for the same billing period last year but there is demand data for the last billing period, use the peak demand from the preceding billing month as the historical peak demand.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

- 3.4.3. Compare the present peak demand with the historical peak demand. If the present peak demand is between 40% and 200% of the historical peak demand, the data passes this check.
- 3.4.4. If the present peak demand is not within 40% to 200% of the historical peak demand, the data fails the High/low demand check. Optionally reread the meter if the same data is still available in the meter. For example, the same data would still be available if demand reset has not yet been performed or meter stores preceding billing period data.
1. If the reread is essentially at the same time as the original read and a different value is obtained, assume the first reading was a misread and perform the check again with the new reading. If the same reading was obtained, assume the meter reading is correct; the data failed the High/Low Demand check but is verified. OR
 2. If the reread is not at the same time as the original read and the new demand value is within +/- 20% of the previous demand value, the data fails the High/Low demand check but is verified. If the re-read results in the same value, the data fails the High/Low demand check but is verified.
- 3.4.5. Data that fails the High/Low Demand check and has not been verified may be investigated and manually verified if justification is found. Otherwise the data must be estimated.

3.5. TOU Usage

The TOU usage check compares the sum of the kWh meter readings for all periods against the current season total kWh meter reading. Note that this check must be done in whatever units are read from the meter. For example, if the meter provides kWh, the kWh values must be summed and compared. If the meter provides pulses, the pulse values must be summed and compared.

- 3.5.1. For the current billing period, calculate the total kWh by summing all the periods, including all seasons.
- 3.5.2. Compare the calculated total kWh with the current total kWh read from the meter. If they are within +/- the number of periods (active or inactive) summed together, the data passes the check. If they are not, the data fails the TOU Usage check and must be estimated. (Note: some TOU rates may include more periods in one season than another, causing "inactive" periods. For example, a summer season may have three periods, and a winter season only two. The period that appears only during the summer is "inactive" during the winter season.)

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

For example, assume there were two periods, peak and off-peak, and a season change occurred during the month.

Period	Previous Season	Current Season
Peak	50	150
Off-Peak	100	200

Determine the valid range by calculating the sum of the periods +/- the number of periods, or $(50 + 100 + 150 + 200) +/- 2$. If the current total kWh read from the meter was between 488 and 502 inclusive, the data would pass the check. If the current total kWh read from the meter was less than 488 or greater than 502, the data would fail the check.

If the meter is programmed to provide readings for all but one of the periods, this test is modified to verify the sum of the periods with readings is \leq the total kWh.

3.6. Zero Consumption for Active Meters

The Zero Consumption checks for zero usage during the billing month.

- 3.6.1. If the meter is an active meter (i.e., is associated with a customer who has financial responsibility), calculate the usage for the present billing month.
- 3.6.2. If the usage is greater than 0, the data passes the zero consumption check.
- 3.6.3. If the usage is 0, the data failed the zero consumption check. Optionally verify the meter reading by re-reading the meter and/or testing the meter. If the reread is the same and the usage is still 0, the data failed the Zero Consumption check but is verified. If a new, different meter reading is obtained, run all the checks again using the new data
- 3.6.4. Data that fails the zero consumption check may be manually investigated and verified if justification is found (for example, a building or equipment that is only used seasonally). If the data is not validated or verified, it must be estimated.

3.7. Usage for Inactive Meters

An inactive meter is one for which there is no customer with financial responsibility. This does not apply to non-UDC MDMAs and is not a required check.

3.8. Number of Dials on Meter

This check applies to cumulative consumption only. It checks that the number of "dials" (digits) reported in the read is consistent with the number of dials (or

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

digits) on the meter display. This check is performed for both remote and local reads if supported by the meter reading technology. If the meter reading technology doesn't support this check, it is not performed.

- 3.8.1. Determine the number of digits in the meter reading.
- 3.8.2. If the number of digits in the meter reading is consistent with the number of digits/dials on the meter, the data passes this check.
- 3.8.3. If the number of digits is not consistent, re-read the meter to verify that the correct meter is being read and that it has the correct number of digits/dials. If the re-read provides the same values, the meter reading failed the Number of Dials on Meter check but is verified. The situation must be investigated and records must be corrected. If the re-read produces different values, perform the check again with the new values.

3.9. Meter Read Demand Decimal Quantity Difference

The Meter Read Demand Decimal Quantity Difference check verifies that the number of demand decimal places displayed on the meter is correct. Note this check is only performed for on-site meter reads, and is not performed for remote meter reads.

- 3.9.1. When the meter is read on-site, the meter reader compares the number of decimal places displayed by the meter with the number of decimal places expected. If they are the same, the reading passes the Meter Read Dial Decimal Quantity check.
- 3.9.2. If they are not the same, re-read the meter to verify that the correct meter is being read and that it has the correct number of decimal places. If the re-read provides the same values, the meter reading failed the Meter Read Dial Decimal Quantity Difference check but is verified. The situation must be investigated and records must be corrected. The meter may need to be re-programmed. If the re-read produces different values, perform the check again with the new values.

3.10. Meter Identification

There are two types of Meter Identification checks depending on how the meter is read - Internal Meter Identification check and External Meter Identification check. The following table summarizes when each check is required:

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Meter Reading Method	Perform External ID Check?	Perform Internal ID Check?
Remote	No	Yes
Optical Port	Yes	Yes
Manual	Yes	No

1. If the meter is read remotely or via its optical port, the Internal Meter Identification check is performed. This compares the meter's internal identification (often its serial number) with the identification expected by the meter reading system. If they match, the data passes this check. If they don't match, the MDMA must investigate why the meter is different than indicated by their records and resolve the inconsistency
2. If the meter is read locally (including via an optical port), the External Meter Identification Check is performed. This compares the Meter ID on the meter nameplate with the Meter ID expected by the meter reading system. If they match, the data passes this check. If they don't match, the MDMA must investigate why the meter is different than indicated by their records and resolve the inconsistency.

3.11. Irregular Usage Monthly Data

An irregular usage customer is one whose usage pattern at a specific location does not follow normal usage patterns and consistently fails the High/Low usage or zero consumption for active meters checks. A customer may be determined to be an irregular usage customer by an MDMA if the MDMA verifies that:

1. the customer data fails the standard validation check for three consecutive months, and that the data represents the actual customer usage, OR
2. the MDMA is notified by the customer's previous MDMA.

The data used to determine a customer is an irregular usage customer could be data collected by the MDMA, or historical data provided by the previous ESP or MDMA. An ESP may notify the MDMA that a customer is a potential irregular usage customer based on conversations with the customer, triggering an inspection of the data. If a customer is determined to be an irregular usage customer, the MDMA may optionally omit the check the customer normally fails. For example, if an irregular usage customer typically fails the High/Low usage check but not the zero consumption check, the zero consumption check must still be performed but the high/low usage check may be omitted. The MDMA must notify both the customer's ESP and UDC of the customer's irregular usage status and what checks will not be performed.

For some customers, irregular usage patterns are symptomatic of the business and will always be present, such as agricultural or seasonal customers. For other

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

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 MDMA must determine whether a customer is a permanent or temporary irregular usage customers. Temporary irregular usage customers must be reviewed at least annually to determine if they are still irregular usage customers or should be returned to the normal checks.

4. Monthly Data Estimation Rules

Note that the MDMA must record the estimation algorithm used for each data element that is estimated. The MDMA must retain this information for the same period required for raw and validated data (3 years). Monthly data estimation algorithms include:

- Estimation based on previous year's data
- Estimation based on preceding billing period's data (≥ 27 days)
- Estimation based on similar customers
- Estimated demand based on average load
- Other estimation method (MDMA must document when this is used)

4.1. Estimating Usage

Two methods to estimate usage are provided. They are similar to the two methods of performing the High/Low Usage check. The first is based on historical usage for the same customer and site; the second is based on historical usage for the same customer and site combined with a factor based on present usage of customers of the same class and same geographic area. The number of decimal places included in ADU calculations must be sufficient so that significant rounding errors do not occur. The recommended value is 2 decimal places. Final estimated usage is truncated to an integer.

4.1.1. Method 1 - Based on Historical Usage

4.1.1.1. Calculate ADU to be applied

- 4.1.1.1.1.** If billing data is available from the same customer and same site for the same billing period last year and it is not estimated, calculate the ADU for the same billing period last year and use this value as the ADU. Refer to section 3.3.1.2.2 to determine the same billing period last year. Optional trend factors that take into account peer group usage based on demographics, climactic areas, and customer class may be applied to the ADU to provide a more accurate estimation. Sample trend factor calculations are to be provided at a later date.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

In this case, the estimation algorithm is estimation based on previous year's data.

- 4.1.1.1.2. If there is no data from the previous year but there is a full preceding billing month (at least 27 days) calculate the ADU for the preceding billing month and use this value for the ADU. In this case, the estimation algorithm is estimation based on preceding billing period's data.
- 4.1.1.1.3. If neither of the previous two options are available, data must be estimated based on any available data, such as similar customers, load profiles, average usage for the customer class, meter reads since last billing read, other historical data, etc. In this case, the estimation algorithm is other estimation algorithm. The MDMA must document how the data is estimated.

- 4.1.1.2. Calculate the number of days since the last good meter reading within the current billing cycle to the end of this billing period. If the meter is read monthly, this would typically be last month's billing meter reading. If the meter is read more frequently, this could be more recent than last month's billing reading.
- 4.1.1.3. Multiply the ADU (including any constants or factors) by the number of days since the last good reading. If necessary, divide this value by a meter constant or other factor to convert it to the same units reported in the meter reading. Truncate the value to an integer, and add the truncated value to the last good reading to obtain an estimated reading. This is the estimated meter reading. Mark the reading as being estimated using the appropriate algorithm.

4.1.2. Method 2 - Based on historical usage and similar customers

- 4.1.2.1. For the residential meter population (i.e., same geographic area and customer class), utilize the following determinants as determined in 3.3.2:
 - ADU Low Range Factor (3.3.2.1.3.4)
 - ADU High Range Factor (3.3.2.1.3.4)
 - Low Range Modified Mean Factor (3.3.2.1.4.1)
 - Medium Range Modified Mean Factor (3.3.2.1.4.1)
 - High Range Modified Mean Factor (3.3.2.1.4.1)

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

- 4.1.2.2. Calculate the ADU from last month's billing period for that customer.
- 4.1.2.3. Calculate the modified ADU for a specific meter by multiplying last month's ADU (from step 4.1.2.2) by yesterday's medium range modified mean factor (above) for that geographical area.
- 4.1.2.4. Determine if the modified ADU is in yesterday's low, medium, or high range.
 - If the modified ADU is less than the ADU low range factor, yesterday's low range modified mean factor is used to calculate estimated ADU in the succeeding steps.
 - If the modified ADU is equal to or greater than ADU low range factor but less than the ADU high range factor, yesterday's medium range modified mean factor is used to calculate estimated ADU in the succeeding steps.
 - If the modified ADU is greater than or equal to the ADU high range factor, yesterday's high range modified mean factor is used to calculate estimated ADU in the succeeding steps.
- 4.1.2.5. Multiply the prior ADU by the modified mean factor determined in 4.1.2.4. This becomes the new estimated ADU.
- 4.1.2.6. Continue with steps 4.1.1.2, and 4.1.1.3 using the estimated ADU calculated in the preceding step.

4.2. Estimating Demand

- 4.2.1. If demand data is available from the same customer and same site for the same billing period last year and it is not estimated, use that demand as the estimated demand. Refer to section 3.3.1.2.2 to determine the same billing period last year. In this case, the estimation algorithm is estimation based on previous year's data.
- 4.2.2. If there is no demand data from the previous year but there is demand data from a full preceding billing month (at least 27 days), use the preceding month's demand as the estimated demand. In this case, the estimation algorithm is estimation based on preceding billing period's data
- 4.2.3. If neither of the above two options are available, calculate the average demand for the billing period. This is done by dividing the actual or estimated usage by the number of hours in the billing period. Use this value as the estimated demand. For example, if the billing period is 30 days, divide the usage for the billing period by 720 (the number of hours

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

in 30 days)

4.3. Estimating TOU Usage

For missing TOU usage data, each period must be estimated separately, using historical data from the same TOU period (defined by time frames) and season as the data requiring estimation. Optional trend factors that take into account peer group usage based on demographics, climactic areas, and customer class may be applied to the ADU to provide a more accurate estimation. Sample trend factor calculations are to be provided at a later date. The number of decimal places included in ADU calculations must be sufficient so that significant rounding errors do not occur. The recommended value is 2 decimal places. Final estimated usage is truncated to an integer.

4.3.1. For each period requiring estimation, the following steps are performed. Note that if there is a season change during the time period requiring estimation, each season needs to be done separately. If season crossover occurs in the month requiring estimation, reference data could be selected from the last month with crossover between the same seasons, or the last full month of the season. There are two cases to consider. For example:

1. If season changes occur on October 1 and May 1, and the billing month April 15 to May 15 (including the season crossover) requires estimation, reference data for the winter period may be chosen from the billing period that contained the October 1 crossover, or from the preceding billing month. Reference data for summer could be chosen from the billing period that contained the October 1 crossover, or from the last full summer month. If the reference month selected does not contain the same seasons as the month requiring estimation, an appropriate month containing the correct seasons should be selected.
2. If season changes occur on October 1 and May 1, and the billing month May 15 to June 15 requires estimation, reference data may be chosen from the billing period that contained the last full month of summer data, or from the summer portion of the preceding billing month.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

- 4.3.1.1. Calculate ADU to be applied
 - 4.3.1.1.1. If billing data is available from the same customer, same site, and same TOU period for the same billing period last year, and last year's data is not estimated, calculate each period's ADU for the same billing period last year and use the values as each period's ADU. Refer to section 3.3.1.2.2 to determine the same billing period last year. For example, if the billing period this year was from April 13 to May 13, the mid-point would be April 28. If a season change occurred during the month, use data from the appropriate season as reference data. Optionally, use data from the month before or after the same billing period last year to get at least one week's worth of data for the season. In this case, the estimation algorithm is estimation based on previous year's data.
 - 4.3.1.1.2. If there is no data from the previous year but there is at least one month's data (minimum 27 days) available from the preceding billing month and it is not estimated, calculate each period's ADU for the preceding billing month and use the values for each period's ADU. If a season change occurred during the month, use data from the appropriate season as reference data. Optionally, use data from the month before or after the same billing period last year to get at least one week's worth of data for the season. In this case, the estimation algorithm is estimation based on preceding billing period's data.
 - 4.3.1.1.3. If there is less than one week's historical data available, each period's data must be estimated by other methods based on any available data, such as similar customers, load profiles, average usage for the customer class, meter reads since last billing read, other historical data, etc. In this case, the estimation algorithm must be documented. The estimation algorithm is other estimation algorithm.
- 4.3.1.2. Calculate the number of days since the last good meter reading within the current billing cycle to the end of this billing period for each period requiring estimation. If the meter is read monthly, this would typically be last month's billing reading. If

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

the meter is read more frequently, this could be more recent than last month's billing reading.

- 4.3.1.3. For each period, multiply the ADU for that period by the number of days requiring estimation. This is the estimated usage by period. Sum the periods to derive the total estimated usage for the billing period. Mark the reading as being estimated using the appropriate algorithm.

4.4. Estimating TOU Demand Data

For missing TOU demand data, estimate each period required for billing separately. Note that if there is a season change during the time period requiring estimation, each season needs to be done separately.

- 4.4.1. For each period requiring estimation, the following steps are performed. Note that if there is a season change during the time period requiring estimation, each season needs to be done separately. If season crossover occurs in the month requiring estimation, reference data could be selected from the last month with crossover between the same seasons, or the last full month of the season. There are two cases to consider. For example:
 1. If season changes occur on October 1 and May 1, and the billing month April 15 to May 15 (including the season crossover) requires estimation, reference data for the winter period may be chosen from the billing period that contained the October 1 crossover, or from the preceding billing month. Reference data for summer could be chosen from the billing period that contained the October 1 crossover, or from the last full summer month. If the reference month selected does not contain the same seasons as the month requiring estimation, an appropriate month containing the correct seasons should be selected.
 2. If season changes occur on October 1 and May 1, and the billing month May 15 to June 15 requires estimation, reference data may be chosen from the billing period that contained the last full month of summer data, or from the summer portion of the preceding billing month.
 - 4.4.1.1. If demand data is available from the same customer and same site for the same billing period last year and it is not estimated, use that demand as the estimated demand. Refer to section 3.3.1.2.2 to determine the same billing period last year. In this case, the estimation algorithm is estimation based on previous year's data.
 - 4.4.1.2. If there is no demand data from the previous year but there is demand data from a full preceding billing month (at least 27

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

days), use the preceding month's demand as the estimated demand. In this case, the estimation algorithm is estimation based on preceding billing period's data.

- 4.4.1.3. If neither of the above two options are available, calculate the average demand for the billing period. This is done by dividing the actual or estimated usage by the number of hours in the billing period. Use this value as the estimated demand. For example, if the billing period is 30 days, divide the usage for the billing period by 720 (the number of hours in 30 days).

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-A: Interval Data VEE Technical Methods (Revision 2.0)

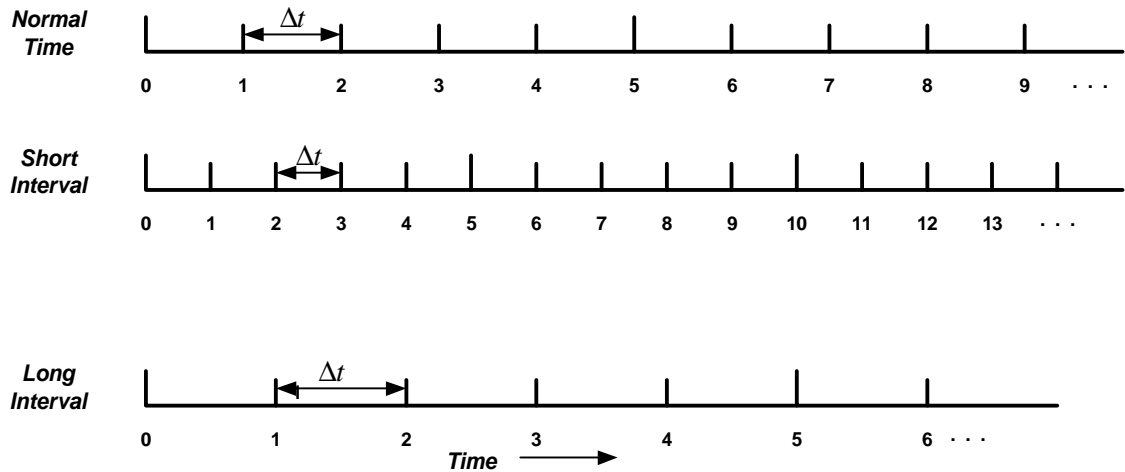
This attachment provides technical information for Section A (CA Interval Data VEE Rules) of Appendix C-VEE.

1.1. Prorating Time Drifted Data

Two options are provided for correcting the data when the actual number of intervals retrieved from the meter does not equal the expected number of intervals based upon the elapsed time.

1.1.1. Option 1 for Prorating Time Drifted Data

This section describes how to normalize interval data when the clock in the meter does not agree with the clock in the computer reading the meter. This phenomenon is called Clock Drift. Clock drift can be both a negative or positive value, depending upon whether the real time (at the computer) is greater than [negative drift] or less than [positive drift] than the clock in the meter. This is illustrated below.



For each of the illustrations shown above, the actual interval of measurement is different. We assume that the meter and the computer systems are synchronized at some time, T_0 {for example, the last meter read} and that the meter is now being read at a read time, T_r . For each case above, assume that the meter reading system reads at the same time. The elapsed time is given by

$$\text{Elapsed Time} = \Delta T_e = T_r - T_0$$

The elapsed time for the shorter interval case is given by

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-A: Interval Data VEE Technical Methods (Revision 2.0)

$$\text{Short Interval Elapsed Time} = \Delta T_{es} = T_{ms} - T_0$$

where T_{ms} is the time that the meter clock gives when the meter is read. The elapsed time for the longer interval case is given by

$$\text{Longer Interval Elapsed Time} = \Delta T_{el} = T_{ms} - T_0$$

where T_{ml} is the time that the meter clock gives when the meter is read. Note that $T_{ms} > T_0$, i.e. the meter clock is running faster therefore clocking more intervals and more elapsed time. And that $T_{ml} < T_0$, i.e., the meter clock is running slower, hence fewer intervals and a shorter elapsed time.

In each case, the internal clock in the meter is registering that the interval length is the length that is specified for the meter, namely Δt or 15 minutes for CA interval meters. However, since the clock is running faster for the short interval case and slower for the long interval case, we must adjust the values such that the correct usage is obtained for each meter. The total drift time for the clock can be calculated for each case as follows:

$$\text{Total Drift Time For Short Intervals } TD_s = |\Delta T_e - \Delta T_{es}| = |T_r - T_{ms}|$$

and

$$\text{Total Drift Time For Long Intervals } TD_l = |\Delta T_e - \Delta T_{el}| = |T_r - T_{ml}|$$

Note that TD_s would be negative had we not taken the absolute value.

The actual number of intervals for each case can be calculated using the elapsed time measured by the meter and dividing it by the preset meter interval, Δt or 15 minutes. Or when truncated to an integer:

$$\text{Expected Number of Intervals, } N = \Delta T_e / \Delta t$$

$$\text{Number of Short Intervals } N_s = \Delta T_{es} / \Delta t$$

$$\text{Number of Long Intervals } N_l = \Delta T_{el} / \Delta t$$

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-A: Interval Data VEE Technical Methods (Revision 2.0)

Two cases are considered.

Case # 1 is when $N = N_s = N_l$, i.e. the clock drift is small enough that no additional interval is generated. For this situation, we could elect to do nothing since for a thirty day reading schedule for 15 minute interval data, there will be about 2880 intervals and a time drift error of less than 0.04% in each interval. The actual drift time for each interval, ignoring any fractional intervals at the end of the read period can be calculated as

$$\Delta DT_s = TD_s / N$$

and for the long interval case,

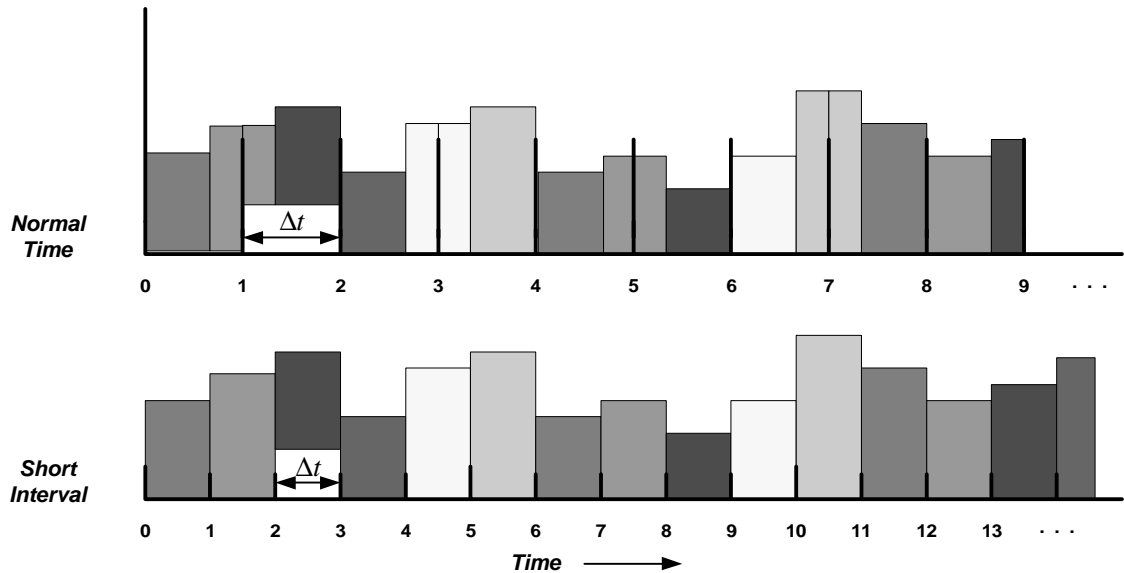
$$\Delta DT_l = TD_l / N$$

This result could be used in combination with the actual interval width to increase or decrease the usage in the particular interval.

Case #2 involves a situation when $N \neq N_s$ or $N \neq N_l$ and hence there are a fewer or greater number of intervals that expected. Here we suggest that a procedure be adopted that distributes the measured interval values into intervals that have the correct length, namely Δt . Thus we increase or decrease the interval length to correspond to actual interval length that should have been in the meter based on the drift, $\Delta t + \Delta DT_l$ or $\Delta t - \Delta DT_s$. Thus the interval distribution that is shown in the above figure would now have proper interval lengths in minutes. The procedure for distribution this data is fairly straightforward. We simply insert the real 15-minute interval grid on the actual interval grid taken from the meter and adjusted by the above technique. Next we divide each of the obtained usage values at the corresponding 15-minute grid points.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-A: Interval Data VEE Technical Methods (Revision 2.0)



This figure shows how to allocate each of the respective values. For example in the first corrected interval, the usage would equal the sum of the usage in the incorrect time interval case plus the fraction to the left of the corresponding interval in the correct zone. For example, if the drift per interval was found to be 4 minutes for short interval case, then the time grid would be 0, 11, 22, 33, 44, 55, 66, 77, etc. The amount taken from the second short interval would be the usage in interval 2 multiplied by 4/11 or 34%. For the case of the third correct interval, there would be components from interval 3, 4, and 5 of the incorrect interval usage. From the third interval, it would be 3/11 of the usage, from the fourth interval it would be 100% of the usage, and from the fifth interval it would be 1/11 of the usage. This scheme should be easy to implement. The one caution is that additional manipulation of the usage data will be required if the interval values are stored as **kW** for the interval, rather than **kWh**. But this should not be a problem.

1.1.2. Option 2 for Prorating Time Drifted Data

The objective of this algorithm is to create the expected number of intervals while preserving the usage recorded in the actual number of intervals.

When the Actual number > Expected Number...

- 1) Truncate the extra intervals.
- 2) The total usage will decrease by the amount recorded in the truncated intervals. In order to preserve the recorded usage, the usage in the

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-A: Interval Data VEE Technical Methods (Revision 2.0)

remaining intervals will need to be scaled up as described below.

When the Actual number < Expected Number...

- 1) Interpolate using the last good interval to create the expected number of intervals.
- 2) The total usage will increase by the amount in the interpolated intervals. In order to preserve the recorded usage, the usage in each interval will need to be scaled down as described below.

To scale the usage in the truncated or interpolated intervals...

- 1) Calculate the recorded usage.

Recorded Usage = total usage for the period. It can be derived from the sum of the usage in the “actual” intervals or from good meter readings.

- 2) Calculate the pre-scaled usage.

Pre-scaled Usage = The total usage in the intervals after they have been truncated or interpolated to the expected number of intervals.

- 3) Calculate the scaling factor.

Scaling Factor = (Recorded Usage/Pre-scaled usage).

- 4) Multiply the usage in each interval by the scaling factor to create corrected usage. The sum of the usage in the scaled intervals should now be equal to the recorded usage.
- 5) Flag all intervals for the period as estimated.

1.2. Sum Check Failure Troubleshooting Techniques

The objective of the sum check is to compare the energy use recorded by the meter to the energy use recorded by the pulse recorder over the same time period. Due to data collection methods, often the period represented by the meter reads does not correspond exactly to the period represented by the interval data. For example, the period of data collection may span from 5/1/98 01:12 AM to 6/1/98 01:22 AM, with the meter readings corresponding exactly to this time period. With 15-minute interval data, the interval data for this same period of data collection would begin at 5/1/98 01:00 AM and end at 6/1/98 01:15 AM. The difference of 12 minutes from the start meter reading and 7 minutes from the end meter reading could be the source of error in the failure of the sum check.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-A: Interval Data VEE Technical Methods (Revision 2.0)

1.2.1. Account for Start and End Time Differences

The following technique enables the MDMA to resolve sum check failures by taking into account time differences between the meter readings and the interval data.

Redo the sum check, taking into account the differences in time 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:

- 1) Calculate a prorated start meter reading to be used in this check by doing the following:
 - (a) Calculate the percentage of an interval that has elapsed between the start time of the first interval and the time of the start meter reading. For example, if the meter was read at 3:30 PM, the first interval in an hourly interval data stream would start at 3:00 PM. The percentage of time elapsed is $(30 \text{ min.}/60 \text{ min.}) = 50\%$.
 - (b) Multiply the usage from the first interval by the percentage from the previous step. For example, if the usage in the first interval is 240 kWh, the percentage usage is $(240 * 0.50) = 120 \text{ kWh}$.
 - (c) Determine how many meter increments are represented by the percentage usage from the previous step. For a meter multiplier of one, the usage is equal to the number of meter increments, so 120 kWh is equal to 120 meter increments. For a meter multiplier of 80, 120 kWh is equal to 1 meter increment (i.e., 120 divided by 80 and rounded down to the nearest integer).
 - (d) Calculate a prorated start meter reading by subtracting the number of meter increments from the previous step from the actual start meter reading. For example, if the start meter reading is 55555, and the number of meter increments is equal to 120, the prorated start meter reading would be $(55555 - 120) = 55435$.
- 2) Calculate an allowable margin of error to be used in this check by doing the following:
 - (a) Calculate the percentage of an interval that has elapsed between the end time of the last interval and the time of the stop meter reading. For example, if the meter was read at 11:15 AM, the last interval in an hourly interval data stream would start at 11:00 AM. The percentage of time elapsed is $(15 \text{ min.}/60 \text{ min.}) = 25\%$.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-A: Interval Data VEE Technical Methods (Revision 2.0)

- (b) Multiply the usage from the last interval by the percentage from the previous step. For example, if the usage in the last interval is 120 kWh, the percentage usage is $(120 \times 0.25) = 30$ kWh.
- (c) Determine how many meter increments are represented by the percentage usage from the previous step. For a meter multiplier of one, the usage is equal to the number of meter increments, so 30 kWh is equal to 30 meter increments. For a meter multiplier of 80, 30 kWh would result in .375 meter increments.
- (d) Calculate the allowable margin of error by adding 2 to the value calculated in the previous step.
- (e) Redo the sum check using the prorated start and original stop meter readings and the allowable margin of error instead of the two multipliers.

1.2.2. Account for Missing or Incomplete Intervals

With some metering and data collection technologies, it is possible for the meter or cumulative usage register to reflect accurate usage even when the interval data is missing or incomplete. The following technique enables the MDMA to resolve the sum check failure for those intervals that were successfully collected.

If some intervals are missing or incomplete, redo the sum check after scaling the difference between the adjusted start read and the stop read by the percentage of good intervals:

- 1) Count the number of good intervals in the data stream.
- 2) Calculate the percentage of good intervals by dividing the count from the previous step by the number of intervals elapsed between start time and stop time.
- 3) Multiply the percentage by the difference between the start reading and the stop reading. (Note that you may use the actual start and stop readings or the prorated start and stop readings from 1.2.1 in this step.)
- 4) Compare the new difference with the sum of the usage in the good intervals. Note that the values must be in the same units for the comparison.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-A: Interval Data VEE Technical Methods (Revision 2.0)

- 5) If the difference is \leq allowable margin from 1.2.1, the good intervals pass the sum check. The missing or incomplete intervals need to be estimated.

1.3. Scaling estimated data using good meter readings

If start and stop meter readings are available and are known to be good, they may optionally be used to scale the estimated interval data as follows:

- 1) Determine the total usage for the time period based upon the meter readings.

$$\text{Total Usage} = ((\text{Stop Reading} - \text{Start Reading}) * \text{Meter Multiplier})$$

- 2) Sum together the valid intervals.
- 3) Subtract the sum of the valid intervals from the total usage to determine the total estimated usage.

$$\text{Total Estimated Usage} = \text{Total Usage} - \text{Sum of Valid Intervals}$$

- 4) Sum together the previously estimated intervals.
- 5) Calculate the scaling factor by dividing the total estimated usage by the sum of the estimated intervals.

$$\text{Scaling Factor} = \text{Total Estimated Usage} / \text{Sum of Estimated Intervals}$$

- 6) Multiply each estimated interval by the scaling factor.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-B: Monthly Data VEE Technical Methods

This attachment provides an example of validating and estimating usage based on previous day usage of similar customers. The text is a duplicate of the corresponding sections in the CA Monthly VEE Rules. The numbers to the left of the text correspond to the numbers on the spreadsheet at the end of the attachment.

3.3.2. Method based on previous day usage of similar customers

Note that this method requires a certain density of customer data for residential customers in the same geographic area, where weather patterns are typically consistent throughout the geographic area.

3.3.2.1. The following steps are performed at the end of each meter reading cycle day for each geographical area in order to validate and estimate usage the following day:

→
Represents the numerical association on the attached spreadsheet calculations. Used for mapping only.

3.3.2.1.1. At the end of the reading day, for each good meter read (open account, billed, between 27-33 days & ADU =< 100), perform the following calculations to determine an ADU for the billing period:

- 1 → 1. Calculate ADU (= KWH/days in billing period)
- 2 → 2. Add ADU to Sum of Current ADU
- 3 → 3. Calculate ADU squared
- 4 → 4. Add ADU squared to Sum of Current ADU squared
- 5 → 5. Add 1 to total meters

- 6 → 6. Calculate last month's ADU
- 7 → 7. Calculate current ADU times last month's ADU
- 8 → 8. Calculate last month's ADU squared

3.3.2.1.2. Determine which range of usage (high, medium or low) the current ADU should be grouped with by comparing current ADU to yesterday's ADU Low and High Range Factors (Reference 3.3.2.1.3 for ADU low and high range factor calculation methodology)

1. If the current ADU is less than yesterday's ADU low range factor:
 - Add current ADU to Sum of current low ADU
 - Add last month's ADU to Sum of last month's low ADU
 - Add current ADU times last month's ADU to Sum of current low ADU times last month's ADU

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-B: Monthly Data VEE Technical Methods

- Add current ADU squared to Sum of current low ADU squared
 - Add last month's ADU squared to Sum of last month's low ADU squared
 - Add 1 to total low meters
2. If the current ADU is not less than the ADU low range factor from yesterday and is less than the ADU high range factor, add the figures to the medium range following same format as in 3.3.2.1.2. step 1.
3. Otherwise, add the current ADU to the ADU high range following the same format as in 3.3.2.1.2. step 1.
- 3.3.2.1.3. Calculate an aggregated ADU for current data for each geographic area
- 7 → 3.3.2.1.3.1. Sum together the ADU values for each geographic area
- 8 → 3.3.2.1.3.2. Calculate the mean for the total ADU (=Sum of Current ADU / total meters)
- 9 → 3.3.2.1.3.3. Calculate the standard deviation for the total
- 3.3.2.1.3.4. Calculate the current ADU low and high range factors:
- 10 → • ADU Low Range Factor = mean - .43 Standard Deviation. If ADU Low range factor is less than the total current mean * .5, the ADU low range factor becomes the mean half.
- 11 → • ADU High Range Factor = mean + .43 Standard Deviation
- 12 → • NOTE: By determining the low & high factors, the Medium Range = (mean - .43 Standard Deviation) to (mean + .43 Standard Deviation)
- 3.3.2.1.4. For each of the three ranges determined above (low, medium, and high), calculate a percent of change of monthly usage for each geographic area.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-B: Monthly Data VEE Technical Methods

3.3.2.1.4.1. After each meter's current billing period's (ADU) is grouped in 3 Ranges (Low, Medium, and High) as specified in 3.3.2.1.2, the following data are summed up by ADU range and area:

- 13 → • Number of customers
- 14 → • Sum of all last month's ADU
- 15 → • Sum of all current month's ADU
- 16 → • Sum of {each last month's ADU times current month's ADU}
- 17 → • Sum of {all last month's ADU squared} i.e., Square all ADU, then sum them.
- 18 → • Sum of {all current month's ADU squared}

From the data above modified ADU mean factors and standard deviation factors are determined for each range as follows:

Modified Mean Factor:

- 19 → Sum of {last month ADU times current month ADU} divided by the sum of {all last month's ADU squared}

Modified Standard Deviation Factor:

- 20 → Step 1: (Sum of {all current month's ADU squared} minus (Mean squared times sum of {all last month's ADU squared})) divided by (Total Meters minus 1).

Step 2: Take square root of Step 1

3.3.2.1.5. Calculate high and low range factors.

Calculate high and low range factors (HRF and LRF) for each of the 3 usage ranges within a geographic area. The mean is used to calculate estimated reads, and the high and low range factors are used in this validation check. 2.8 and 3.5 are used in the below example to represent the range deviation factor and will allow for an appropriate meter read error rate. This factor can be changed to control the error rate.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-B: Monthly Data VEE Technical Methods

21 → High Range Factor Formula:
$$\text{HRF} = 1 + \{(2.8 \times \text{Modified Standard Deviation} \times \text{Number of meters}) / (\text{Sum of Current month's ADU})\}$$

22 → Low Range Factor Formula:
$$\text{LRF} = 1 - \{(3.5 \times \text{Modified Standard Deviation} \times \text{Number of meters}) / (\text{Sum of Current month's ADU})\}$$

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-B: Monthly Data VEE Technical Methods

3.3.2.2. As each meter is read, perform the following using the values calculated from the previous meter reading days' data.

- 23 → 3.3.2.2.1. Determine the usage from the preceding billing month and the preceding billing reading for the customer and site.
- 24 → 3.3.2.2.2. Calculate low limit for this month's usage by multiplying the preceding month's usage by the low range factor determined above.
- 25 → 3.3.2.2.3. Calculate high limit for this month's usage by multiplying the preceding month's usage by the high range factor determined above.
- 26 → 3.3.2.2.4. If the current usage is between the low and high limit calculated in the previous two steps, the data passes the High/Low check.

4.1. Estimating Usage

4.1.2. Method 2 - Based on historical usage and similar customers

4.1.2.1. For the residential meter population (i.e., same geographic area and customer class), utilize the following determinants as determined in 3.3.2:

- ADU Low Range Factor (3.3.2.1.3.4)
- ADU High Range Factor (3.3.2.1.3.4)
- Low Range Modified Mean Factor (3.3.2.1.4.1)
- Medium Range Modified Mean Factor (3.3.2.1.4.1)
- High Range Modified Mean Factor (3.3.2.1.4.1)

27 → 4.1.2.2. Calculate the ADU from last month's billing period for that customer.

28 → 4.1.2.3. Calculate the modified ADU for a specific meter by multiplying last month's ADU (from step 4.1.2.2) by yesterday's medium range modified mean factor (above) for that geographical area.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-B: Monthly Data VEE Technical Methods

- 29 → 4.1.2.4. Determine if the modified ADU is in yesterday's low, medium, or high range.
- If the modified ADU is less than the ADU low range factor, yesterday's low range modified mean factor is used to calculate estimated ADU in the succeeding steps.
 - If the modified ADU is equal to or greater than ADU low range factor but less than the ADU high range factor, yesterday's medium range modified mean factor is used to calculate estimated ADU in the succeeding steps.
 - If the modified ADU is greater than or equal to the ADU high range factor, yesterday's high range modified mean factor is used to calculate estimated ADU in the succeeding steps.
- 30 → 4.1.2.5. Multiply the prior ADU by the modified mean factor determined in 4.1.2.4. This becomes the new estimated ADU.
- 4.1.2.6. Continue with steps 4.1.1.2, and 4.1.1.3 using the estimated ADU calculated in the preceding step.

PSWG APPENDIX C-VEE: Requirements for Validating, Editing, and Estimating Monthly and Interval Data in Direct Access

Attachment C-VEE-B: Monthly Data VEE Technical Methods

Spreadsheet Calculations:

Meter Number	Current Usage	Current ADU	Last Month Usage	Last Month ADU	Meter Reading	Current ADU * Last Month ADU	Current ADU Squared	Last Month ADU Squared	Range Grouping
1	125	4.17	135	4.50	3731	18.75	17.36	20.25	Low
2	275	9.17	250	8.33	7475	76.39	84.03	69.44	High
3	114	3.80	142	4.73	3697	17.99	14.44	22.40	Low
4	178	5.93	165	5.50	4888	32.63	35.20	30.25	Medium
5	254	8.47	260	8.67	7358	73.38	71.68	75.11	High
6	299	9.97	189	6.30	6835	62.79	99.33	39.69	High
7	178	5.93	190	6.33	5278	37.58	35.20	40.11	Medium
8	158	5.27	136	4.53	4176	23.88	27.74	20.55	Medium
9	165	5.50	140	4.67	4329	25.67	30.25	21.78	Medium
10	235	7.83	255	8.50	7033	66.58	61.36	72.25	High
11	218	7.27	235	7.83	6500	56.92	52.80	61.36	High
12	110	3.67	119	3.97	3286	14.54	13.44	15.73	Low
13	98	3.27	105	3.50	2912	11.43	10.67	12.25	Low
14	85	2.83	129	4.30	3117	12.18	8.03	18.49	Low
15	169	5.63	110	3.67	3913	20.66	31.73	13.44	Medium
16	200	6.67	168	5.60	5221	37.33	44.44	31.36	High
17	147	4.90	178	5.93	4688	29.07	24.01	35.20	Low
18	165	5.50	201	6.70	5281	36.85	30.25	44.89	Medium
19	180	6.00	135	4.50	4446	27	36.00	20.25	Medium
Sum		111.77		108.07		681.63	727.99	664.82	

Assume yesterday's ADU low range factor was 4.95 and the ADU high range factor was 6.4.

Mean 8 5.88
 Std Dev 9 1.98
 ADU Low Range Factor 10 5.03
 ADU High Range Factor 11 6.73

Ranges	Low	High
Low	0	5.03
Medium	5.03	6.73
High	6.73	or higher

	13 Meters	15 Curr ADU	14 Last mo ADU	16 Curr * Last Mo ADU	18 Curr Mo sqrd	17 Last Mo ADU Sqd
Low	6	22.63	26.93	103.97	87.95	124.33
Medium	7	39.77	35.90	204.26	226.38	191.27
High	6	49.37	45.23	373.40	413.66	349.22

	19 Modified Mean	20 Mod Std Deviation	22 Low Range Factor	21 High Range Factor
Low	0.84	0.45	0.58	1.33
Medium	1.07	1.17	0.28	1.58
High	1.07	1.70	0.28	1.58

Example validation routine for meter read the following day using the determinants calculated above:

Customer's Last Month's Usage = 200 kWh 23 ADU = 6.67 Customer's usage falls in Medium Range

High Range Value: 200 kWh X 1.58 = 316 25
 Low Range Value: 200 kWh X 0.28 = 56 24

Usage falling between 56 and 316 is acceptable. Usage outside of this range fails the high/low usage check. 26

Example estimation routine for meter read the following day using the determinants calculated above:

Customer's Previous Usage = 200 kWh 27 ADU = 6.67

Modified ADU = 6.67 X 1.07 (Medium Range Modified Mean Factor)
28 = 7.14 Modified ADU

Determine where 7.14 is grouped according to the ADU low and high range factors (5.03 & 6.73 respectively) 29
 7.14 > 6.73, therefore, the high range modified mean factor of 1.07 is used.

ADU Estimated usage = last month's ADU (6.67) X 1.07 = 7.14, rounded to 7
 Total usage = ADU Estimated usage X number of days in the billing period. 30