

# EWX Flow Rules & Algorithms

## *Aggregation / Detection / Editing...*

### *Estimation*

Linear Interpolation <i>number of points</i>	<ul style="list-style-type: none"><li>• Uses gap check, zero reads, and variance validation to fill missing or extreme data with linearly interpolated data</li><li>• Decides which chunks need interpolation based on if they have gaps, low reads, or spikes/dips</li><li>• Linearly interpolates if the chunks are at least <b><i>number of points</i></b></li></ul>
Like Day Estimation <i>number like days</i> <i>max weeks to use</i>	<ul style="list-style-type: none"><li>• Fills gaps with average usage from like days using <b><i>number of like days</i></b></li><li>• Searches forwards or backwards up to <b><i>max weeks to use</i></b> to get like day usage data</li></ul>

## *Forecasting / Forwarding / Storage...*

### *Validation*

Interval Gap Check	<ul style="list-style-type: none"><li>• Checks for time gaps in interval usage data</li></ul>
Nonperiodic Zero Reads Check IDR <i>margin</i> <i>common threshold</i>	<ul style="list-style-type: none"><li>• Calculates proportion of reads less than (or equal to) <b><i>margin</i></b>, grouped by weekday and hour</li><li>• For each usage, if the proportion is less than <b><i>common threshold</i></b>, these reads are nonperiodic. Otherwise, they are periodic – meaning it is ok for usage on this day and time to be zero <b><i>common threshold</i></b> percent of the time</li></ul>
Energy Sum Validation <i>threshold</i>	<ul style="list-style-type: none"><li>• Checks periods based on heartbeats to determine if IDR sum and scalar values are similar</li><li>• Flags end of period if <math>sum(interval) \neq scalar \pm threshold\%</math></li><li>• Calculates percent diff b/t IDR sum &amp; scalar as <math>(IDR\ sum - scalar)/scalar * 100</math>, flags if <math>abs(percent\ diff) &gt; threshold</math></li></ul>

<p>Variance Validation</p> <p><i>window</i> <i>center (T/F)</i> <i>num of std dev's</i></p>	<ul style="list-style-type: none"> <li>• Calculates rolling mean &amp; sd of (<i>use – rolling mean</i>) for time period of size <b>window</b> seconds</li> <li>• Rolling window is either centered if <b>center</b> is true</li> <li>• <b>num of std dev's</b> throws a flag for reads which are this number of rolling sd's above (spikes) or below (dips) the rolling mean of (<i>use – rolling mean</i>)</li> </ul>
<p>Outlier Validation</p> <p><i>dip spike delta threshold</i> <i>dip spike hours</i> <i>dip spike lookback</i> <i>dip spike max outlier time</i> <i>dip spike min change</i> <i>periodicity num med abs dev</i> <i>periodicity num occur</i> <i>periodicity window</i></p>	<ul style="list-style-type: none"> <li>• Checks for periodicity or dip spike outliers based on if the daily usage is seasonal or not</li> <li>• <b>Periodicity</b>: calculates rolling mean w/ default window and centers usage. Groups centered data by day of week &amp; hour.</li> <li>• Calculated median and median average deviation per group</li> <li>• Flags usage if:  <math display="block">  \text{centered use} - \text{group med}   &gt; \text{num MAD} * \text{group MAD}</math> where group med and group MAD based on <b>window</b> </li> <li>• If num of flags per unique date are &gt; <b>num occur</b>, check this date</li> <li>• <b>Dip Spike</b>: calculates rolling mean for <b>hours</b> window &amp; min periods, if the percent change of rolling mean <b>hours</b> with starting roll mean as:  <math display="block">\text{delta roll mean} = (\text{roll mean}(\text{start} + \text{hours}) / (\text{roll mean}(\text{start})) - 1</math> for every time in data </li> <li>• If <i>delta roll mean</i> &gt; <b>delta threshold</b>, look backwards from a day ahead to <b>lookback</b> time before obs (index j) and look for a &lt;= 0 difference, this is the start of spike and roll mean here is right level</li> <li>• If nothing found, use roll mean of time – <b>lookback</b> as right level, check to see if usage within <b>max outlier time</b> if &lt;= right level</li> <li>• Otherwise we say it isn't a spike</li> <li>• If <i>delta roll mean</i> &lt; - <b>delta threshold</b>, repeat process</li> </ul>
<p>Ramp Up Ramp Down Validation</p> <p><i>compare num days</i> <i>percent threshold</i> <i>usage difference threshold</i></p>	<ul style="list-style-type: none"> <li>• Compare avg daily use between consecutive periods of <b>compare num days</b> days</li> <li>• <math>\text{percent variation} = (\text{first period ADU} - \text{last period ADU}) * 100 / \text{first period ADU}</math>, ramp up if &gt; <b>percent threshold</b>, ramp down if less</li> <li>• <math>\text{usage variation} = \text{first period ADU} - \text{last period ADU}</math>, same rule applies compared to <b>usage difference threshold</b></li> </ul>
<p>Load Factor IDR Validation</p> <p><i>load factor threshold</i> <i>peak threshold</i></p>	<ul style="list-style-type: none"> <li>• Flags usage if load factor is below <b>load factor threshold</b> and peak is above <b>peak threshold</b></li> <li>• Calculates <math>\text{load} = \text{usage} * 3600 / \text{heartbeat}</math>, use sum, count and max (as peak)</li> <li>• <math>\text{load factor} = \text{sum} / \text{peak} * \text{count}</math></li> </ul>