Supporting Information

Reliable data from low cost ozone sensors in a hierarchical network

Georgia Miskell[1], Kyle Alberti[2], Brandon Feenstra[4], Geoff S Henshaw[2], Vasileios Papapostolou[4], Hamesh Patel[2], Andrea Polidori[4], Jennifer A Salmond[3], Lena Weissert[1,3], David E Williams[1,*]

1. School of Chemical Sciences and MacDiarmid Institute for Advanced Materials and Nanotechnology, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
2. Aeroqual Ltd, 460 Rosebank Road, Avondale, Auckland 1026, New Zealand
3. School of Environment, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
4. South Coast Air Quality Management District, 21865 Copley Drive, Diamond Bar, CA 91765, USA

1. Diurnal variation of ozone concentration
2. The KS test
3. Regulatory observations compared with three different proxies
4. Co-located sensor control charts using the closest other regulatory station as the proxy
5. Summary results for the application of the framework on the sensor data, using closest other regulatory station as proxy.
6. Map showing the movement of “buddy” sensors from regulatory sites to test sites

1. Diurnal variation of ozone concentration

Figure S1 illustrates the strong diurnal variation of ozone concentration in the geographical region of the study. It also illustrates the seasonal effect, with the daily maximum ozone concentration being significantly higher in July than in January.

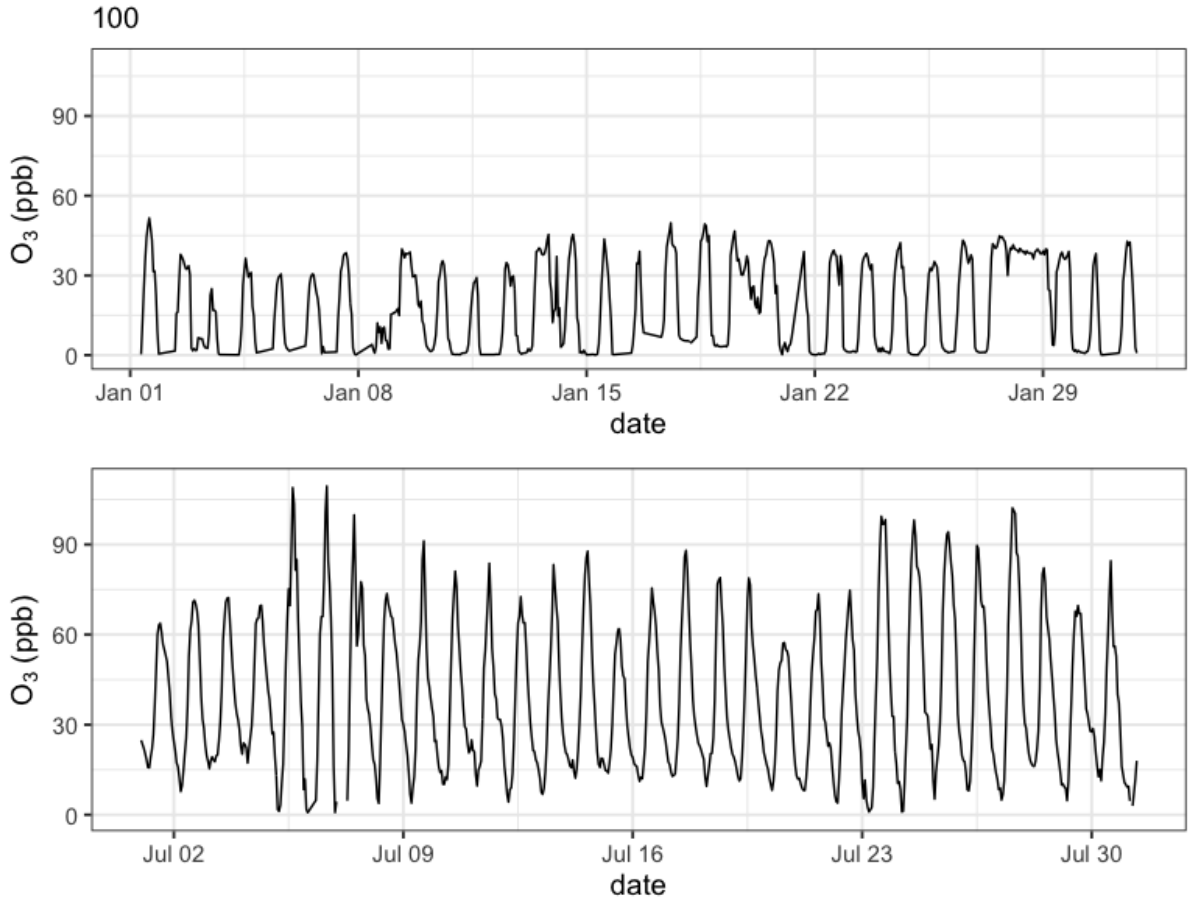


Figure S1. Time series illustrating the diurnal variation of ozone concentration.

2. The KS test

The Empirical Cumulative Distribution (ECD) for a sample with t_w observations and variable *X* has the cumulative probability:

$$\mathbb{P}(x > X) = \hat{F}_{t_w}(x) = \frac{1}{t_w + 1} \sum_{i=1}^{t_w} \begin{cases} 1, & x > x_i \\ 0, & x \leq x_i \end{cases}$$

Assumptions are x_i are *iid* random variables with a common underlying ECD, *F(X).* The KS statistic uses the ECD over a specified time window, t_w. The test hypothesizes

the two samples will have similar distributions by comparing the maximum absolute distance (supremum function, sup) between the cumulative probability curves on the *y*-axis for the same *x*-axis values. The KS statistic, *D,* for two samples with *m, t_w* observations and common variable *X* is:

$$D_{m,t_w} = \frac{sup}{x} \left|\hat{F}_m(x) - \hat{F}_{t_w}(x)\right|$$

KS statistics are translatable into *p*-values (p_{KS}) based on sample size and probability thresholds, with the null hypothesis, H_0, being the two samples could come from the same distribution. Assumptions are the two samples are independent, observations are independent and both come from continuous distributions.

2. Regulatory observations compared with three different proxies

Three proxies were tested against the regulatory data to examine their suitability as a proxy. The three proxies tested were: nearest regulatory station data not co-located with the instrument data under test, local network median (including regulatory and instrument data), and regulatory station with the most similar AADT within 5 km to the test location. The figures show the assessment of proxies by correlation of proxy-corrected regulatory station data with the original regulatory station data

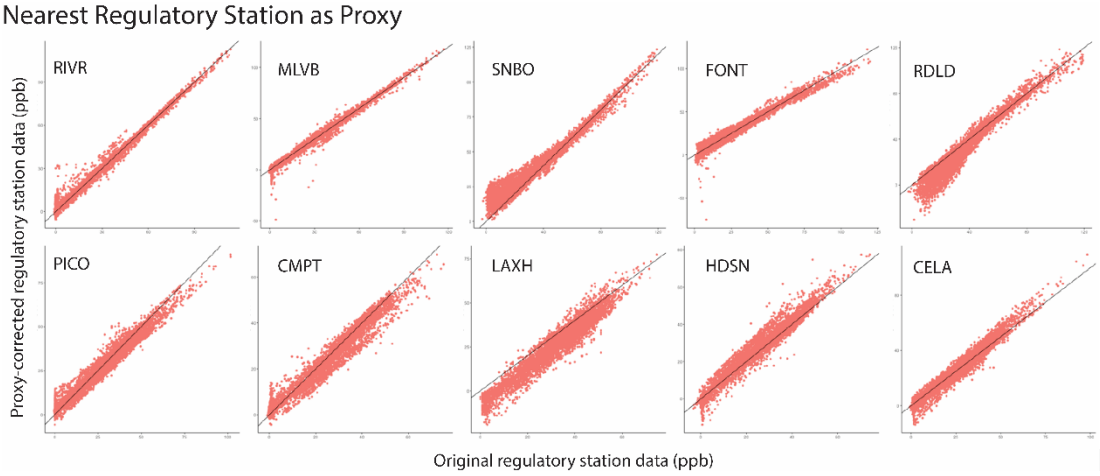


Figure S2: closest proximity proxy

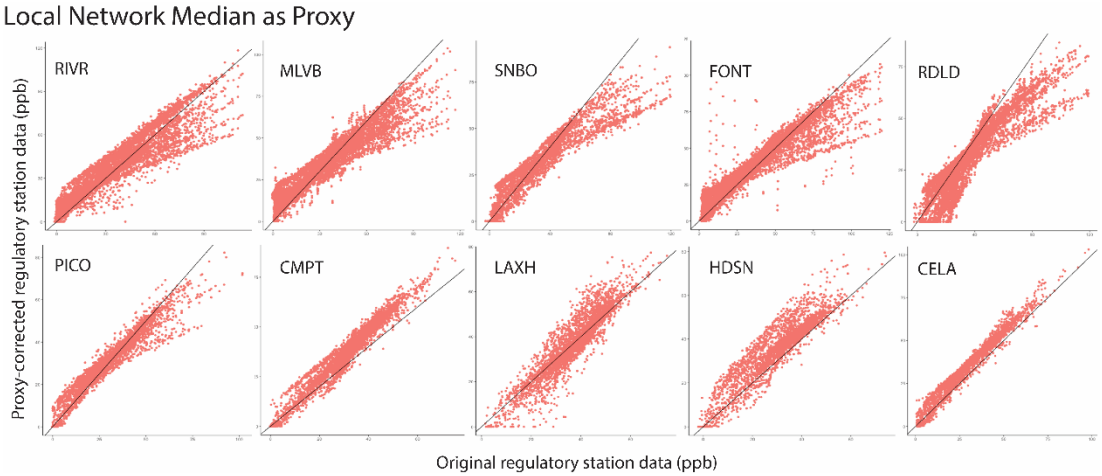


Figure S3: median of local low-cost sensor network as proxy

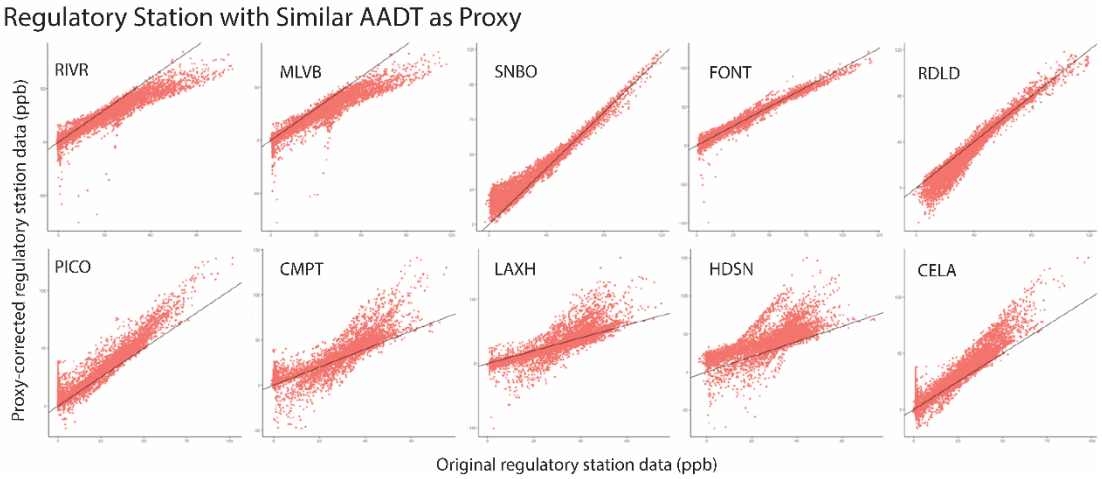
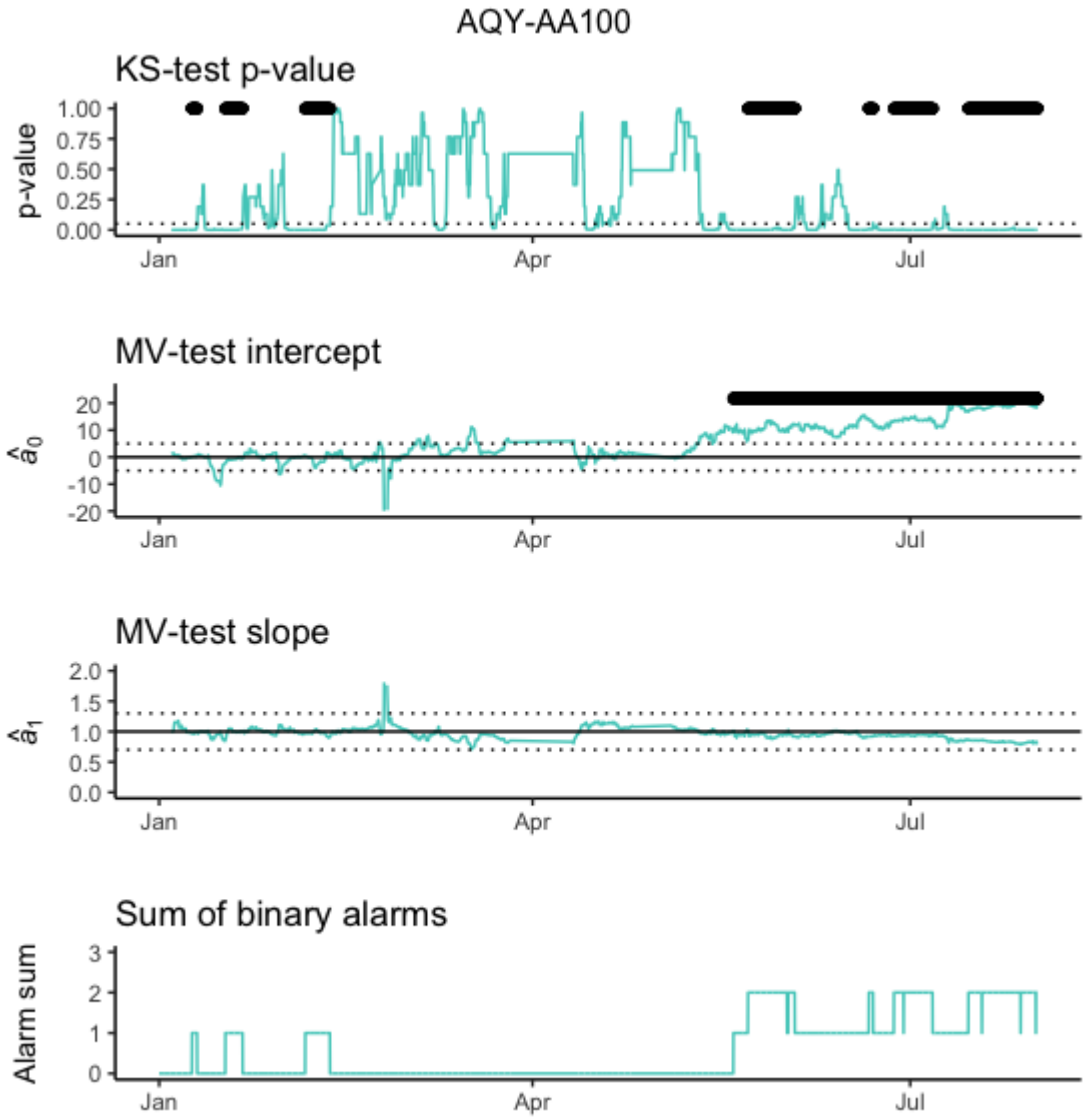
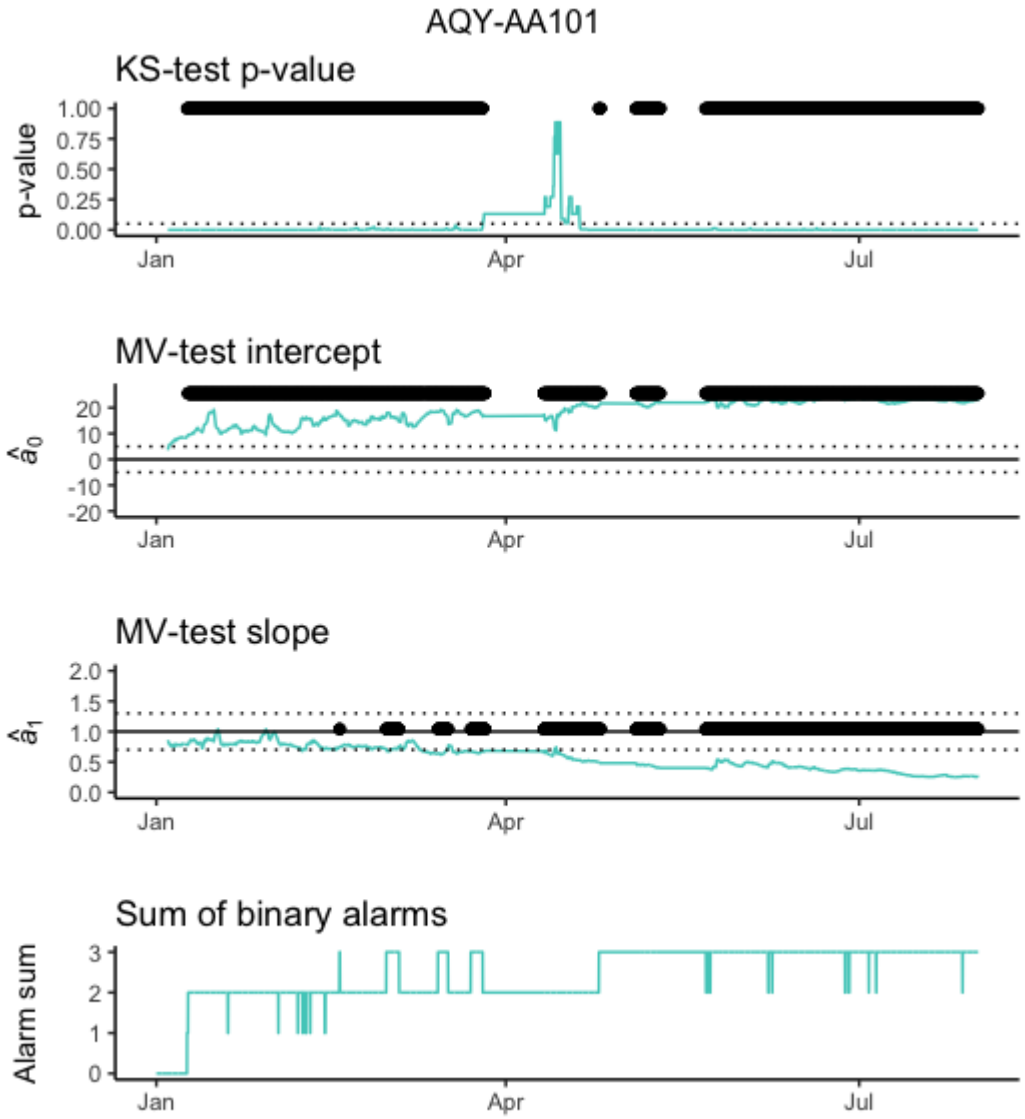


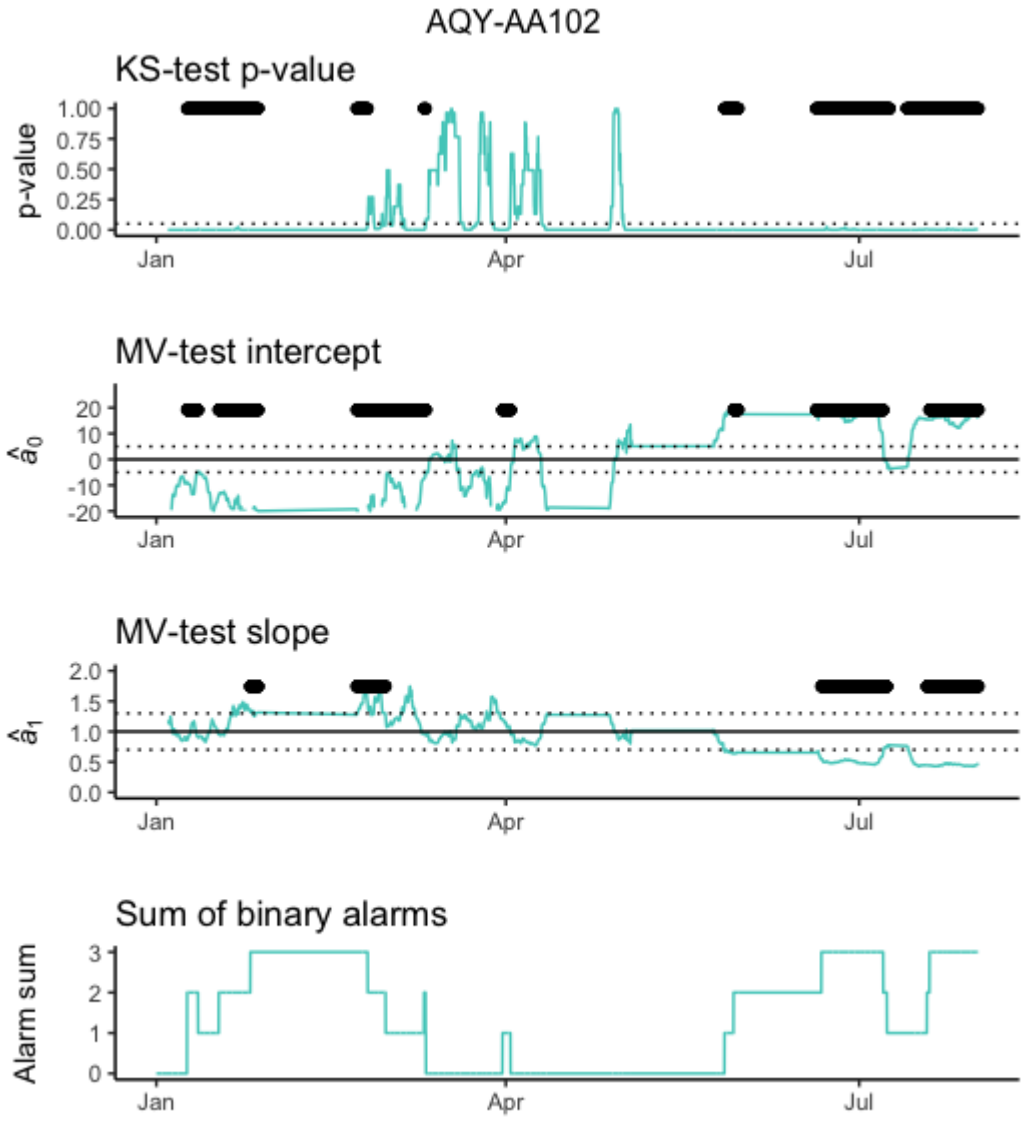
Figure S4: Regulatory station with similar Annual Average Density of Traffic as proxy

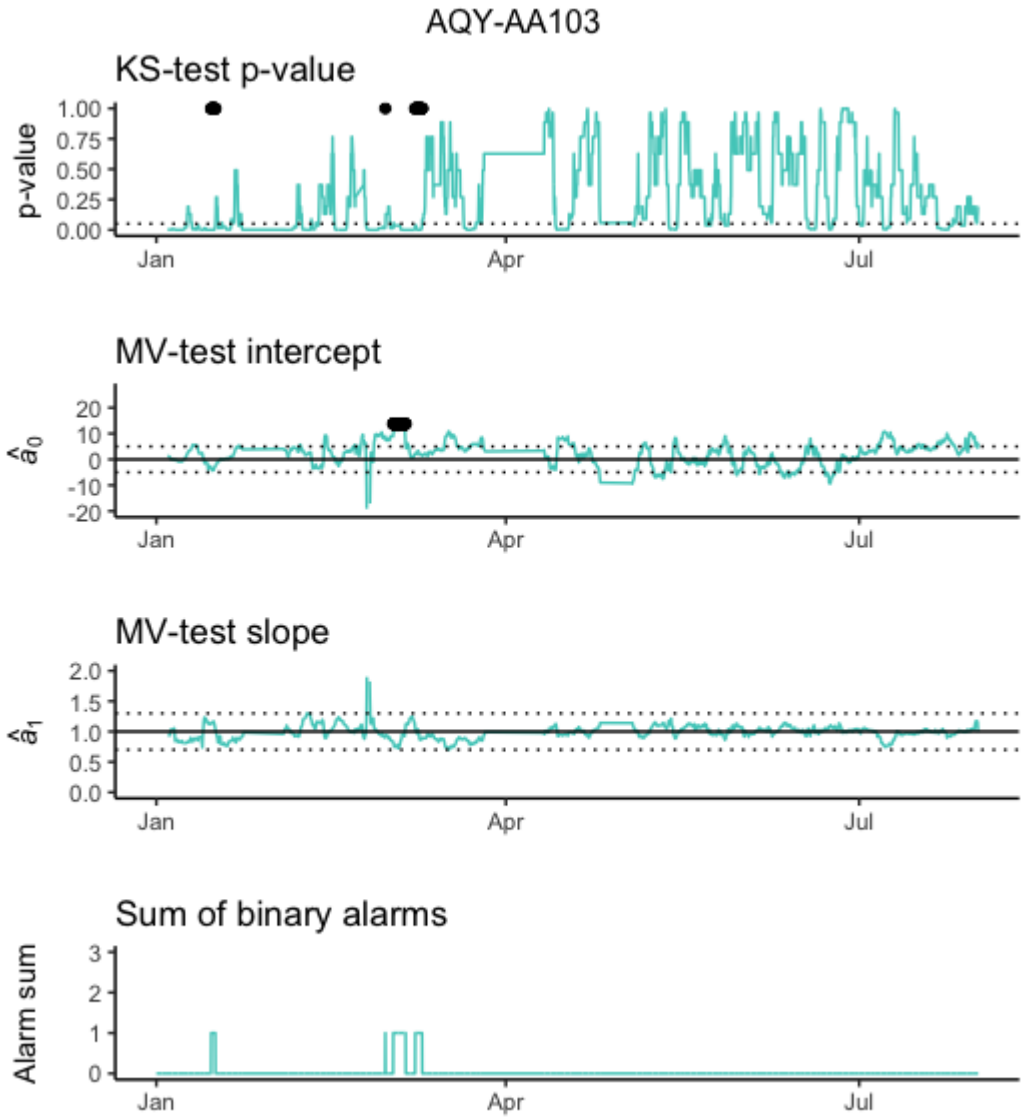
3. Co-located sensor control charts using the closest other regulatory station as the proxy

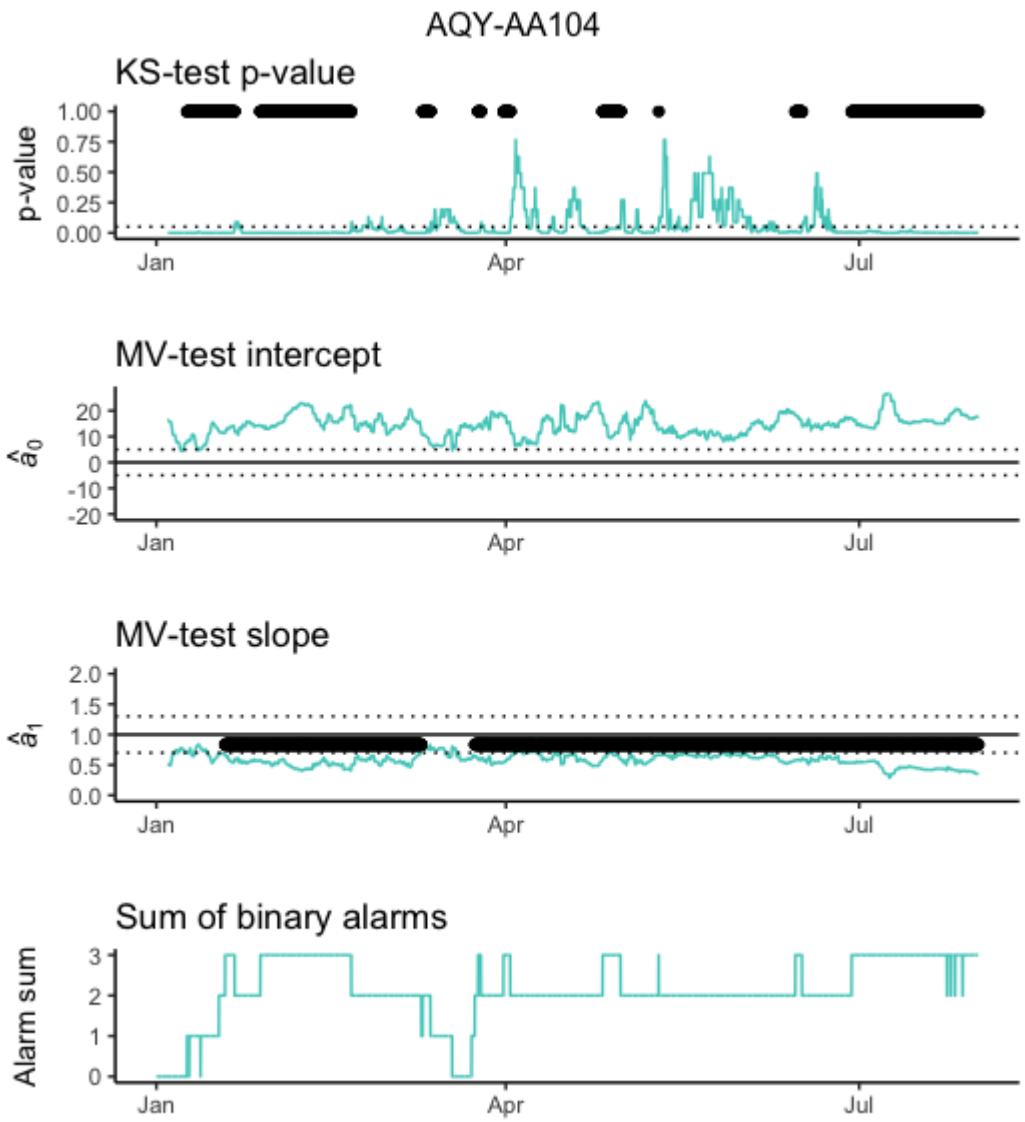
The four plots derived by comparison of the sensor data at each site with the closest proximity reference analyser as proxy show the KS-test p-value (p_{KS}) and the MV-test intercept ($\hat{a}_0$) and slope ($\hat{a}_1$), with dashed lines representing the defined thresholds and the large black points where the test result is outside of the threshold for the defined length, t_f. The bottom plot (in black) shows the sum of the black points – the test alarms - over time. Where this value is one or above, the correction algorithm from the MV test is applied to the sensor instrument data to derive the framework-corrected result.

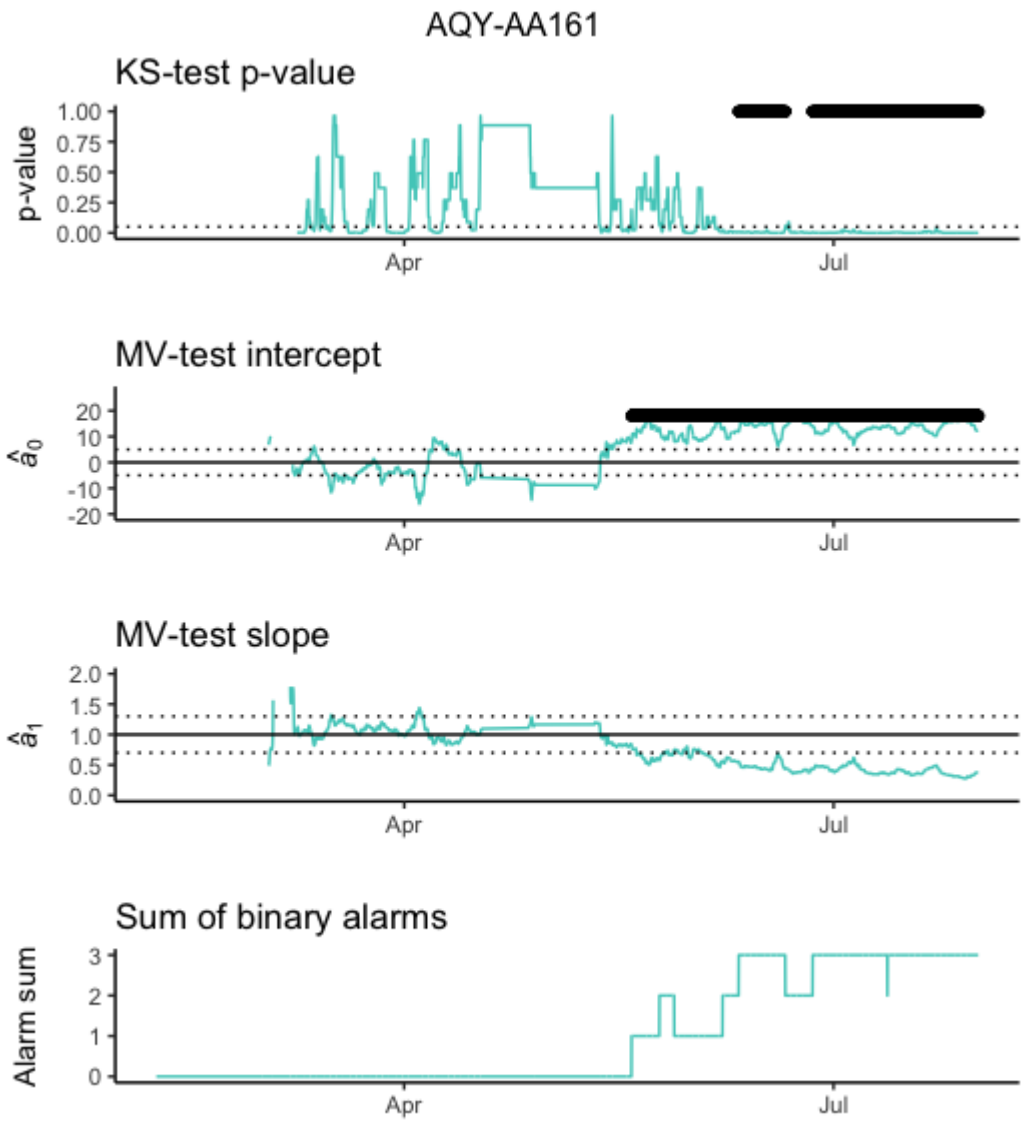

AQY-AA100
KS-test p-value
p-value
1.00
0.75
0.50
0.25
0.00
Jan
Apr
Jul
MV-test intercept
$\hat{a}_0$
20
10
0
-10
-20
Jan
Apr
Jul
MV-test slope
$\hat{a}_1$
2.0
1.5
1.0
0.5
0.0
Jan
Apr
Jul
Sum of binary alarms
Alarm sum
3
2
1
0
Jan
Apr
Jul

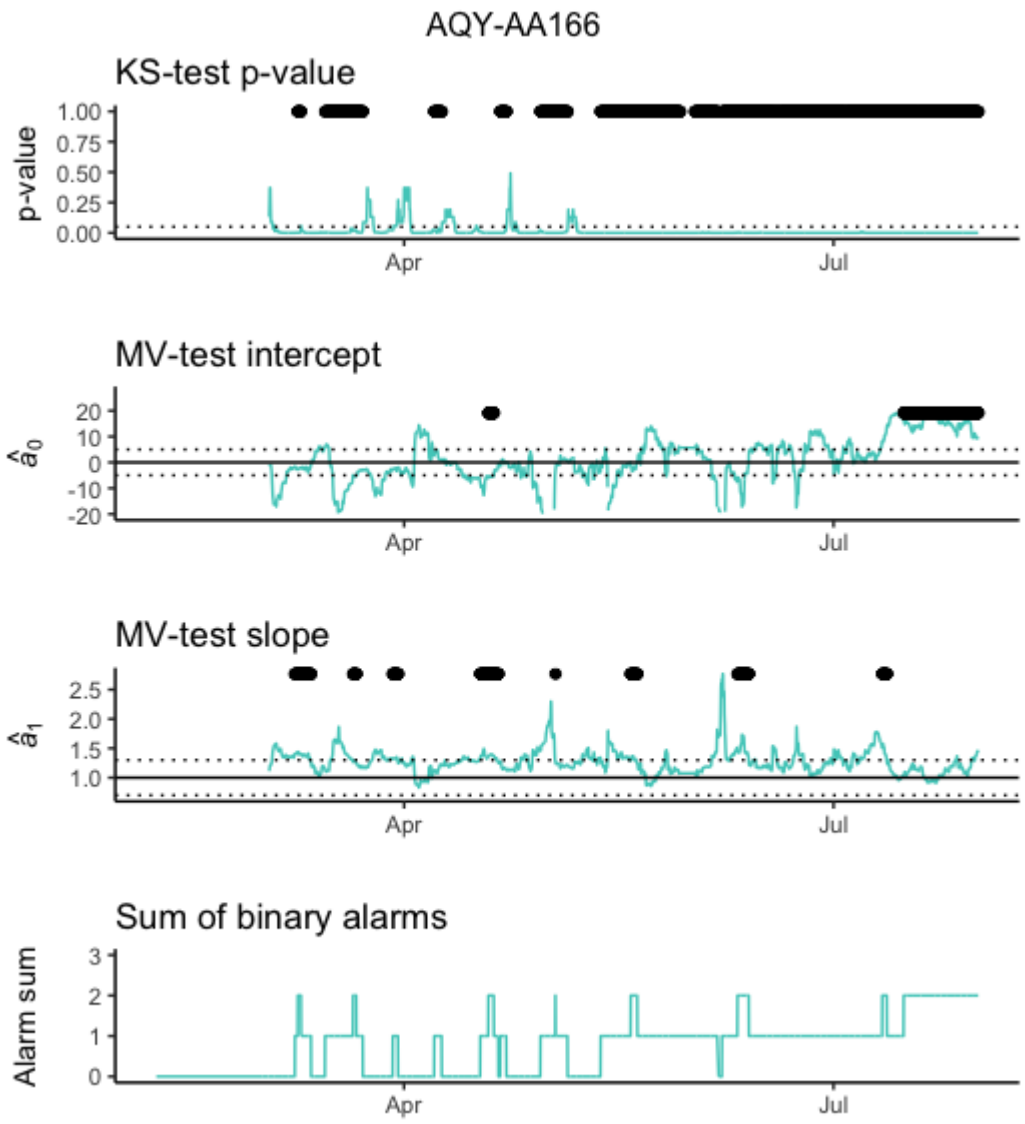

AQY-AA101
KS-test p-value
p-value
1.00
0.75
0.50
0.25
0.00
Jan
Apr
Jul
MV-test intercept
$\hat{a}_0$
20
10
0
-10
-20
Jan
Apr
Jul
MV-test slope
$\hat{a}_1$
2.0
1.5
1.0
0.5
0.0
Jan
Apr
Jul
Sum of binary alarms
Alarm sum
3
2
1
0
Jan
Apr
Jul

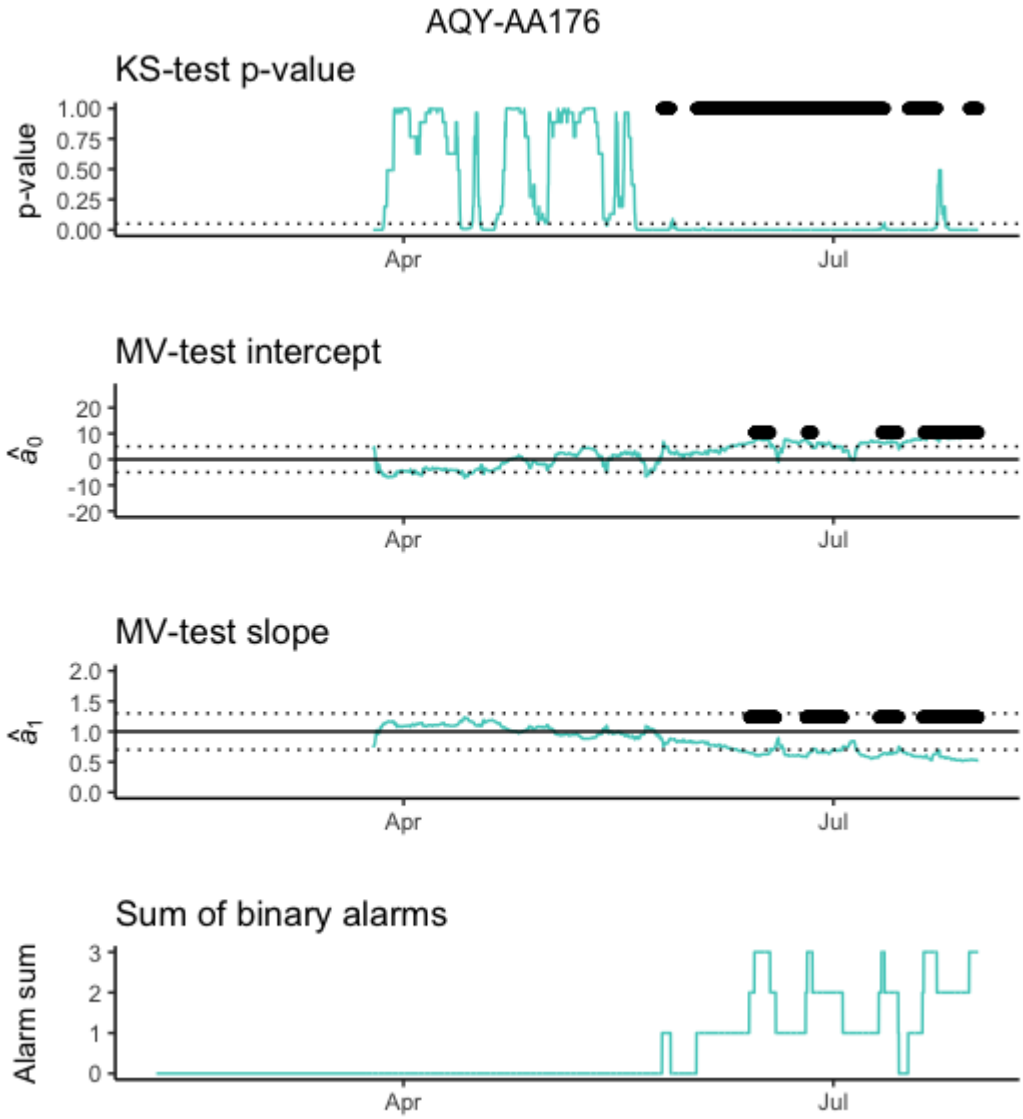

AQY-AA102
KS-test p-value
p-value
1.00
0.75
0.50
0.25
0.00
Jan
Apr
Jul
MV-test intercept
$\hat{a}_0$
20
10
0
-10
-20
Jan
Apr
Jul
MV-test slope
$\hat{a}_1$
2.0
1.5
1.0
0.5
0.0
Jan
Apr
Jul
Sum of binary alarms
Alarm sum
3
2
1
0
Jan
Apr
Jul

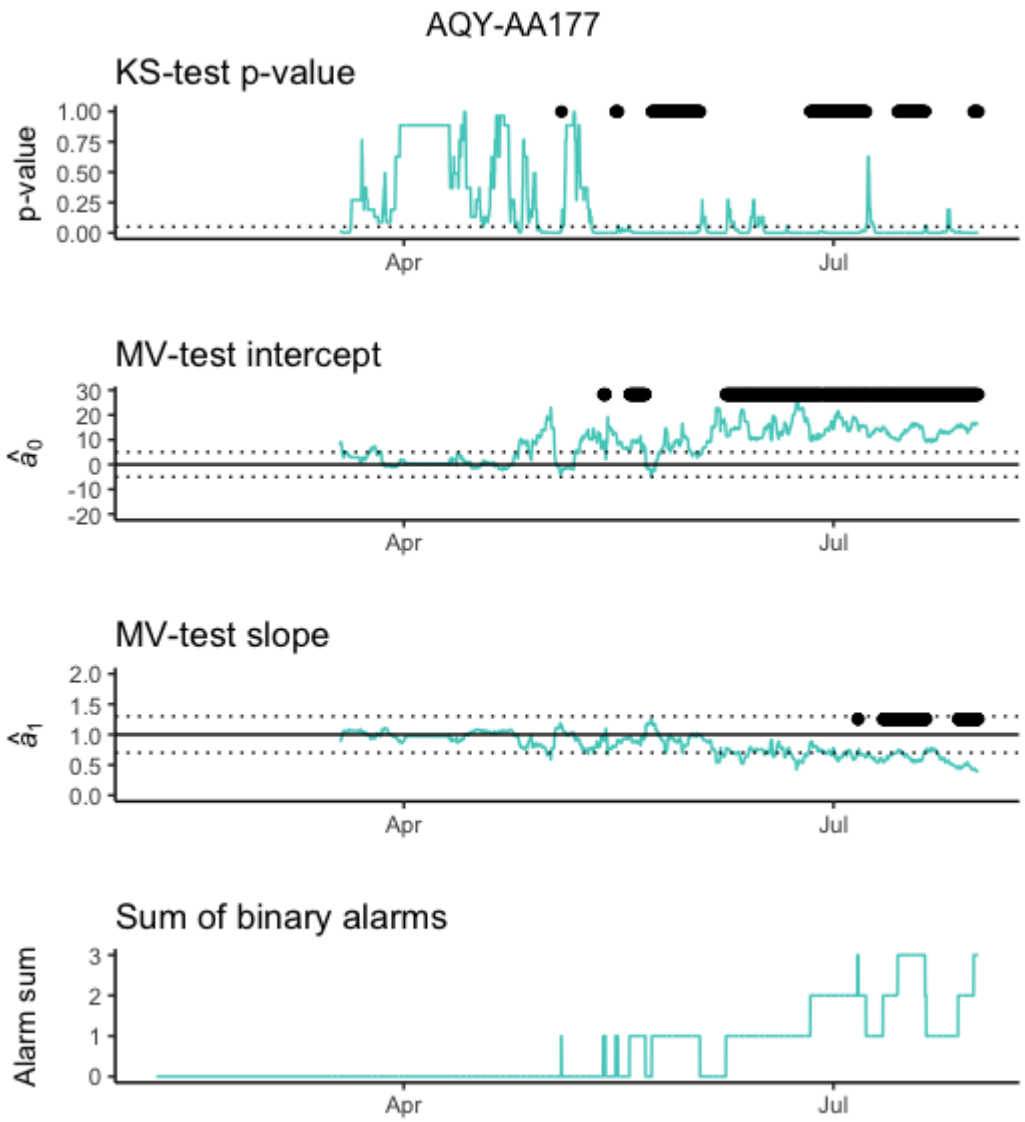

AQY-AA103
KS-test p-value
p-value
1.00
0.75
0.50
0.25
0.00
Jan
Apr
Jul
MV-test intercept
$\hat{a}_0$
20
10
0
-10
-20
Jan
Apr
Jul
MV-test slope
$\hat{a}_1$
2.0
1.5
1.0
0.5
0.0
Jan
Apr
Jul
Sum of binary alarms
Alarm sum
3
2
1
0
Jan
Apr
Jul


AQY-AA104
KS-test p-value
p-value
1.00
0.75
0.50
0.25
0.00
Jan
Apr
Jul
MV-test intercept
$\hat{a}_0$
20
10
0
-10
-20
Jan
Apr
Jul
MV-test slope
$\hat{a}_1$
2.0
1.5
1.0
0.5
0.0
Jan
Apr
Jul
Sum of binary alarms
Alarm sum
3
2
1
0
Jan
Apr
Jul


AQY-AA161
KS-test p-value
p-value
1.00
0.75
0.50
0.25
0.00
Apr
Jul
MV-test intercept
$\hat{a}_0$
20
10
0
-10
-20
Apr
Jul
MV-test slope
$\hat{a}_1$
2.0
1.5
1.0
0.5
0.0
Apr
Jul
Sum of binary alarms
Alarm sum
3
2
1
0
Apr
Jul


AQY-AA166
KS-test p-value
p-value
1.00
0.75
0.50
0.25
0.00
Apr
Jul
MV-test intercept
$\hat{a}_0$
20
10
0
-10
-20
Apr
Jul
MV-test slope
$\hat{a}_1$
2.5
2.0
1.5
1.0
Apr
Jul
Sum of binary alarms
Alarm sum
3
2
1
0
Apr
Jul


AQY-AA176
KS-test p-value
p-value
1.00
0.75
0.50
0.25
0.00
Apr
Jul
MV-test intercept
$\hat{a}_0$
20
10
0
-10
-20
Apr
Jul
MV-test slope
$\hat{a}_1$
2.0
1.5
1.0
0.5
0.0
Apr
Jul
Sum of binary alarms
Alarm sum
3
2
1
0
Apr
Jul


AQY-AA177
KS-test p-value
p-value
1.00
0.75
0.50
0.25
0.00
Apr
Jul
MV-test intercept
$\hat{a}_0$
30
20
10
0
-10
-20
Apr
Jul
MV-test slope
$\hat{a}_1$
2.0
1.5
1.0
0.5
0.0
Apr
Jul
Sum of binary alarms
Alarm sum
3
2
1
0
Apr
Jul

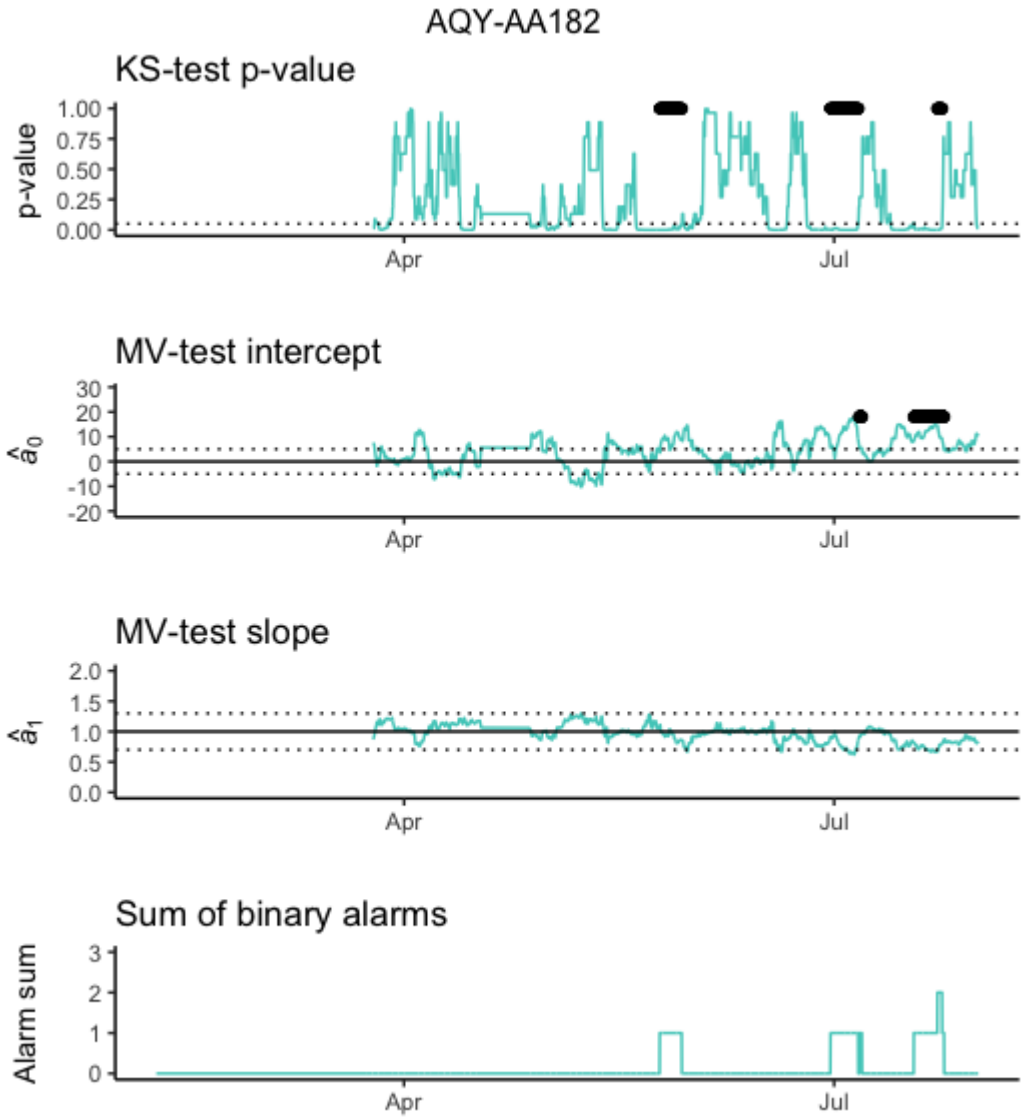


Figure S5: Control charts (four plots for each sensor) for the ten sensors co-located at reference sites, with the closest proximity other reference site as proxy.

4. Summary results for the application of the framework on the sensor data, using closest other regulatory station as proxy.

Site ID	Sensor ID	KS p_{KS} alarm %	MV $\hat{a}_0$ %	MV $\hat{a}_1$ %	Framework-corrected %
RIVR	100	13	0	0	0
MLVB	101	13	0	0	0
SNBO	102	17	39	19	22
FONT	103	9	17	0	0
RDLD	104	17	38	13	18
PICO	161	11	3	0	0.3
CMPT	166	5	0	0	0
LAXH	176	35	21	0	17
HDSN	177	5	0	0	0
CELA	182	11	2	0	0.3

The table shows the percentage of the total test time recorded as an alarm by each of the individual tests, and the percentage of the total time that the data were corrected using the framework described in the main text

5. Map showing the movement of “buddy” sensors from regulatory sites to test sites

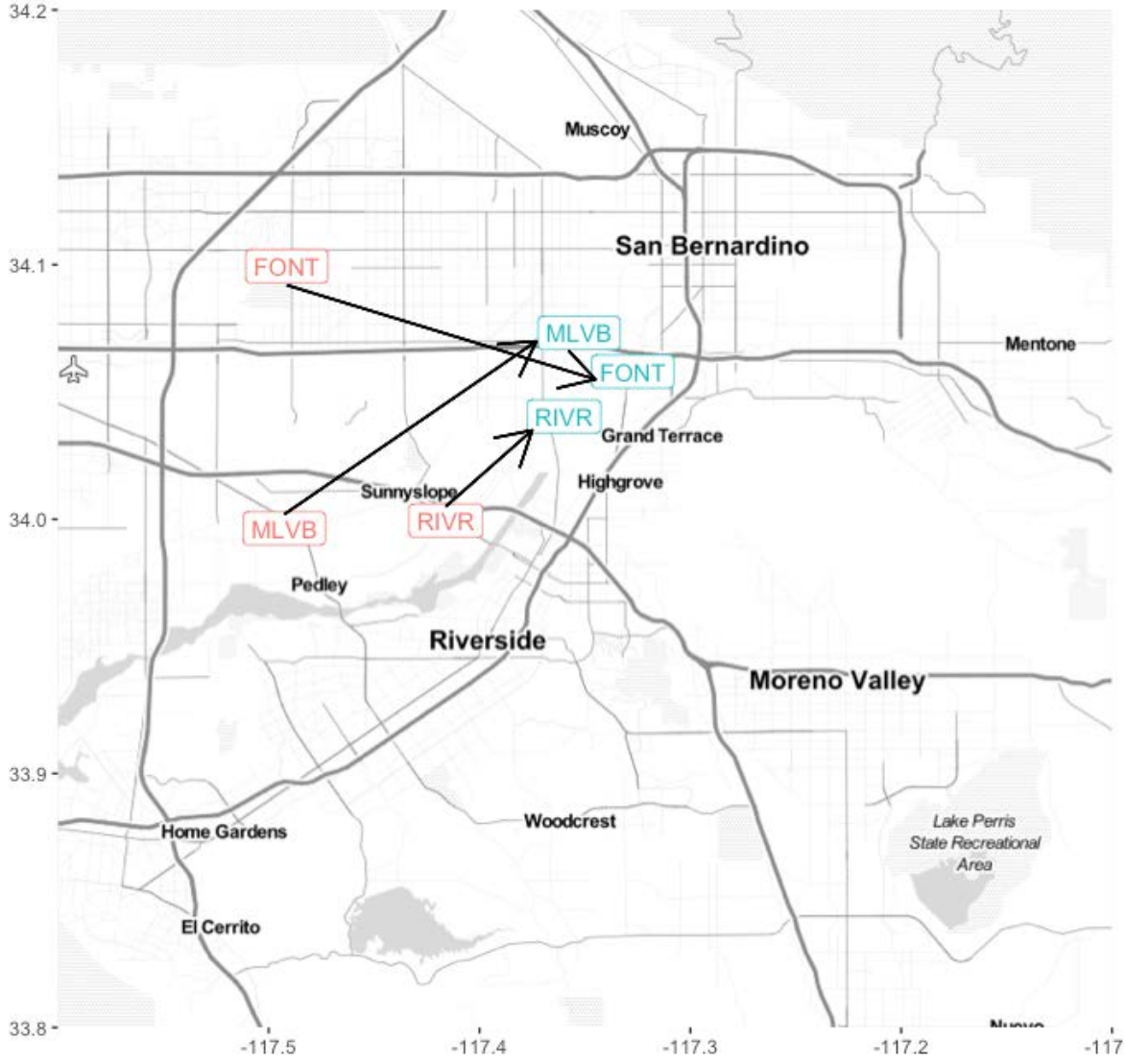


Figure S6: Map showing the movement of the “buddy” sensors (main text, section 3.4). Names refer to the initial co-located regulatory station. The blue locations where the arrows point to are the locations to which the “buddy” sensors were moved, to check the sensor mounted there. The pink locations are the locations of the regulatory sites where the “buddy” sensors were calibrated and from which they were moved.