

OMI air-quality monitoring over the Middle East – Supplementary Material

Michael P. Barkley¹, Gonzalo Gonzalez Abad², Thomas P. Kurosu³, Robert Spurr⁴,
Sara Torbatian⁵, and Christophe Lerot⁶

¹Earth Observation Science Group, Department of Physics and Astronomy, University of Leicester,
UK.

²Atomic and Molecular Physics Division, Harvard-Smithsonian Center for Astrophysics,
Cambridge, Massachusetts, USA.

³NASA Jet Propulsion Laboratory, Pasadena, California, USA.

⁴RT Solutions Inc, Cambridge, Massachusetts, USA.

⁵Air Quality Meteorologist, Air Quality Control Company (AQCC), Tehran, Iran.

⁶Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium.

Correspondence to: Michael Barkley (mpb14@le.ac.uk)

1 Introduction

This supplementary material contains additional figures and tables to support the analysis described in the main text. It is structured as follows. Section 2 contains the supporting figures, whilst Section 3 contains tabulated statistics. Tables S2 to S4 showing the results of the sensitivity analysis conducted in Section 5 of the main text. The sensitivity tests are numbered in correspondence to their descriptions in the manuscript.

2 Ancillary Figures



Figure 1. Example of GoogleEarth imagery showing the ± 2 grid-cell spatial filtering mask (in light green) relative to the centre of the city of Kerman (Iran), as discussed in Section 3.2 of the main text. The mask corresponds to an approximate radial distance of about 10 km.

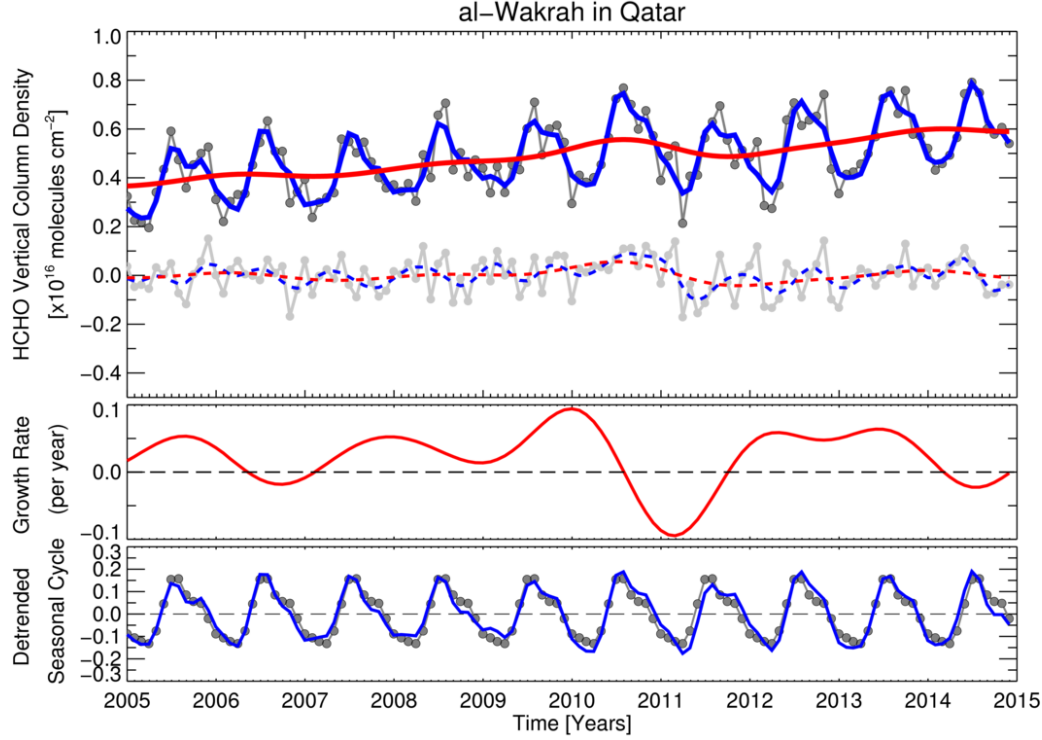


Figure 2. An example of a time series fit to observed HCHO data over al-Wakrah (Qatar), as outlined in Section 3.3. Top panel: the monthly OMI HCHO vertical columns are indicated by dark grey filled circles, whilst the light grey filled circles represent the fitting residual, which has been smoothed with a short-term 200-day filter (dashed blue line) and long-term 667-day filter (red dashed line). The solid red line is the long-term trend $F_T(t)$, given by the linear component of the fitted function $F(t)$ (equation 1) plus the residual filtered using the long-term filter. The solid blue line is the smoothed fitted curve $F_S(t)$ given by $F(t)$ plus the residual filtered using the short-term filter. Middle panel: The HCHO vertical column growth rate in 10^{16} molecules $\text{cm}^{-2} \text{yr}^{-1}$, which is the derivative of the long-term trend $F_T(t)$ shown in the top-panel. Bottom panel: the de-trended seasonal cycle $F_C(t)$ which is the difference between the long-term trend and the smoothed function fit (i.e. $F_S(t) - F_L(t)$). It represents the annual seasonal oscillation with any long-term trend removed. The dark grey filled circles are the fitted harmonic component of $F(t)$. At al-Wakrah (25.29°N , 51.61°E , population: 22712) there is a statistically significant large upward linear trend of $0.22 \pm 0.05 \times 10^{15}$ molecules $\text{cm}^{-2} \text{yr}^{-1}$. This corresponds to a linear growth of $4.61 \pm 0.94 \%$ yr^{-1} relative to the observed 2005–2014 median VCD. In this example, $|\omega/\sigma_\omega| = 4.9$ and the uncertainties of the trend (F_T) and smoothed curves (F_S) are about 4% and 6%, respectively. The median growth rate G is $6.56 \pm 5.61 \%$ yr^{-1} , whilst the mean seasonal amplitude is $3.05 \pm 0.36 \times 10^{15}$ molecules cm^{-2} (about $62.56 \pm 7.36\%$ relative to the median column). A similar analysis of the coincident time series of the HCHO AMF, cloud fraction, cloud-top pressure and number of samples, reveals no other significant trend. This indicates the upward growth in HCHO is not caused by a trend in any other retrieval parameter and is real at the 95% confidence level.

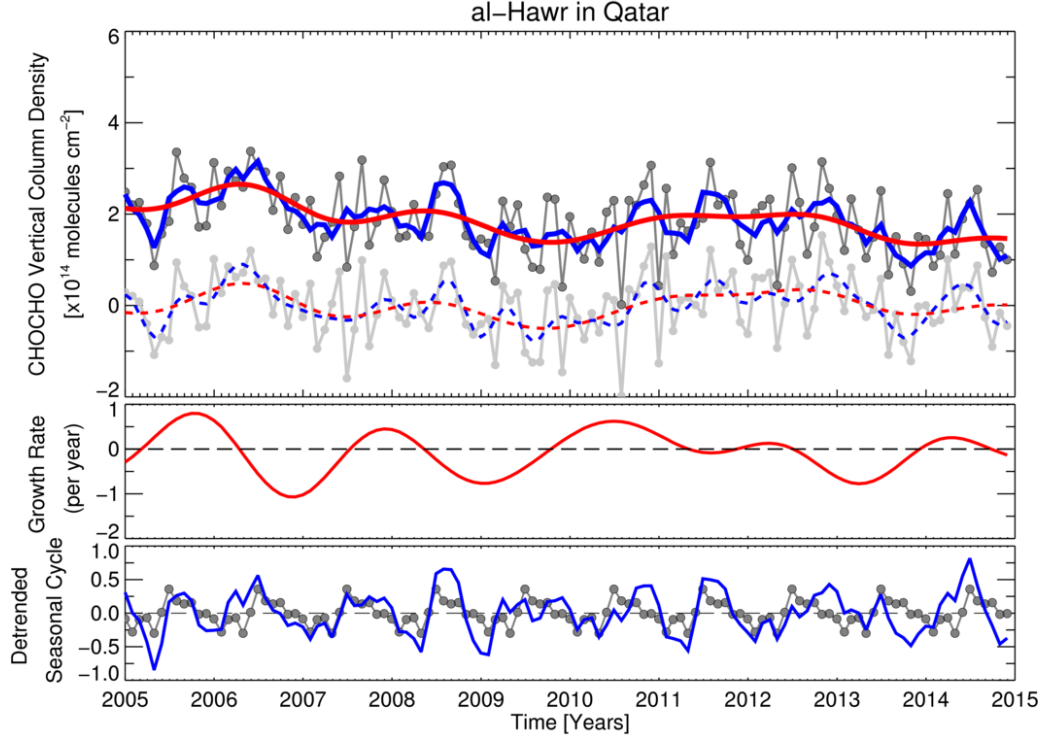


Figure 3. An example of a time series fit to observed CHOCHO data over al-Hawr (Qatar), as outlined in Section 3.3. Top panel: the monthly OMI CHOCHO vertical columns are indicated by dark grey filled circles, whilst the light grey filled circles represent the fitting residual, which has been smoothed with a short-term 200-day filter (dashed blue line) and long-term 667-day filter (red dashed line). The solid red line is the long-term trend $F_T(t)$, given by the linear component of the fitted function $F(t)$ (equation 1) plus the residual filtered using the long-term filter. The solid blue lines is the smoothed fitted curve $F_S(t)$ given by $F(t)$ plus the residual filtered using the short-term filter. Middle panel: The CHOCHO vertical column growth rate in 10^{14} molecules $\text{cm}^{-2} \text{yr}^{-1}$, which is the derivative of the long-term trend $F_T(t)$ shown in the top-panel. Bottom panel: the de-trended seasonal cycle $F_C(t)$ which is the difference between the long-term trend and the smoothed function fit (i.e. $F_S(t) - F_L(t)$). It represents the annual seasonal oscillation with any long-term trend removed. The dark grey filled circles are the fitted harmonic components of $F(t)$. At al-Hawr (25.69°N , 51.51°E , population: 2736) there is a statistically significant downward linear trend of $-0.83 \pm 0.23 \times 10^{13}$ molecules $\text{cm}^{-2} \text{yr}^{-1}$. This corresponds to a linear decrease of $-4.58 \pm 1.26 \%$ yr^{-1} relative to the observed 2005–2014 median VCD. In this example, $|\omega/\sigma_\omega| = 2.0$ and the uncertainties of the trend (F_T) and smoothed curves (F_S) are about 10% and 16%, respectively. The median growth rate G is $-0.89 \pm 13.88 \%$ yr^{-1} , whilst the mean seasonal amplitude is $9.50 \pm 3.35 \times 10^{13}$ molecules cm^{-2} (about $52.12 \pm 18.41\%$ relative to the median column). A similar analysis of the coincident time series of the CHOCHO AMF, cloud fraction, cloud-top pressure and number of samples, reveals no other significant trend. This indicates the downward trend in CHOCHO is not caused by a trend in any other retrieval parameter and is real at the 95% confidence level.

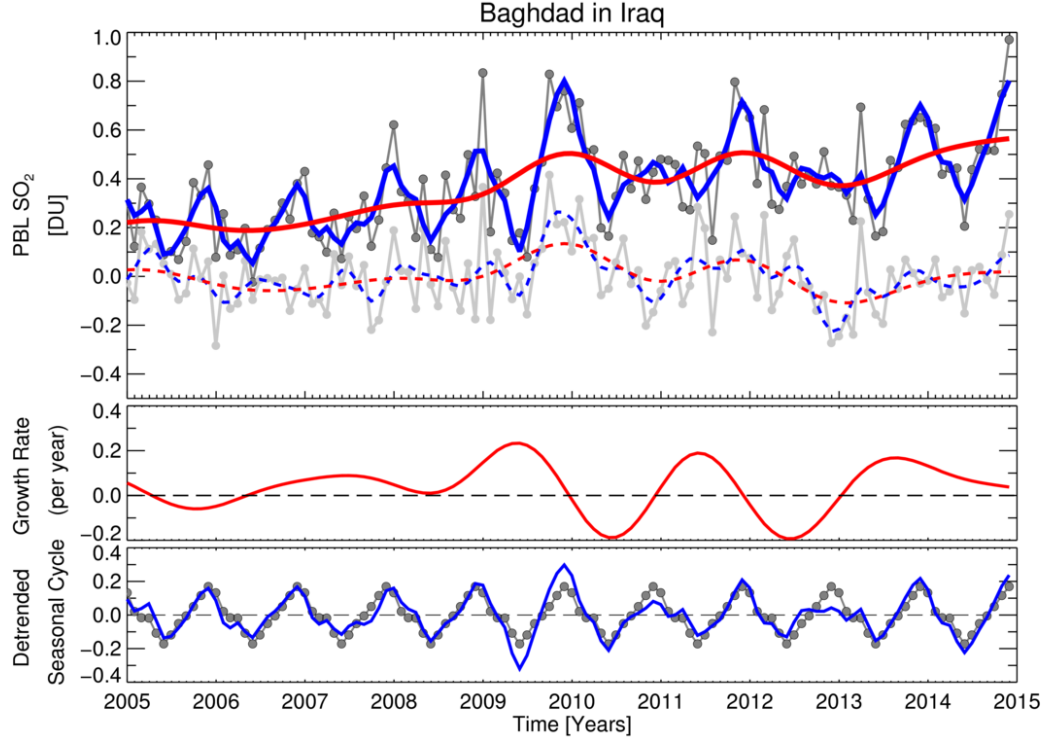


Figure 4. An example of a time series fit to observed SO_2 data over Baghdad (Iraq), as outlined in Section 3.3. Top panel: the monthly OMI SO_2 vertical columns are indicated by dark grey filled circles, whilst the light grey filled circles represent the fitting residual, which has been smoothed with a short-term 200 day filter (dashed blue line) and long-term 667 day filter (red dashed line). The solid red line is the long-term trend $F_T(t)$, given by the linear component of the fitted function $F(t)$ (equation 1) plus the residual filtered using the long-term filter. The solid blue line is the smoothed fitted curve $F_S(t)$ given by $F(t)$ plus the residual filtered using the short-term filter. Middle panel: The SO_2 vertical column growth rate in DU yr^{-1} , which is the derivative of the long-term trend $F_T(t)$ shown in the top-panel. Bottom panel: the de-trended seasonal cycle $F_C(t)$ which is the difference between the long-term trend and the smoothed function fit (i.e. $F_S(t) - F_L(t)$). It represents the annual seasonal oscillation with any long-term trend removed. The dark grey filled circles are the fitted harmonic components of $F(t)$. At Baghdad (33.34°N , 44.39°E , population: 5511490) there is a statistically significant large upward linear trend of $0.035 \pm 0.008 \text{ DU yr}^{-1}$. This corresponds to a linear growth of $9.78 \pm 2.25 \% \text{ yr}^{-1}$ relative to the observed 2005–2014 median VCD. In this example, $|\omega/\sigma_\omega| = 4.4$ and the uncertainties of the trend (F_T) and smoothed curves (F_S) are about 9% and 15%, respectively. The median growth rate G is $13.18 \pm 13.37 \% \text{ yr}^{-1}$, whilst the mean seasonal amplitude is $0.37 \pm 0.06 \text{ DU}$ (about $102.78 \pm 17.78\%$ relative to the median column). A similar analysis of the coincident time series of the SO_2 cloud fraction, cloud-top pressure and number of samples, reveals no other significant trend. This indicates the upward growth in SO_2 is not caused by a trend in any other retrieval parameter and is real at the 95% confidence level.

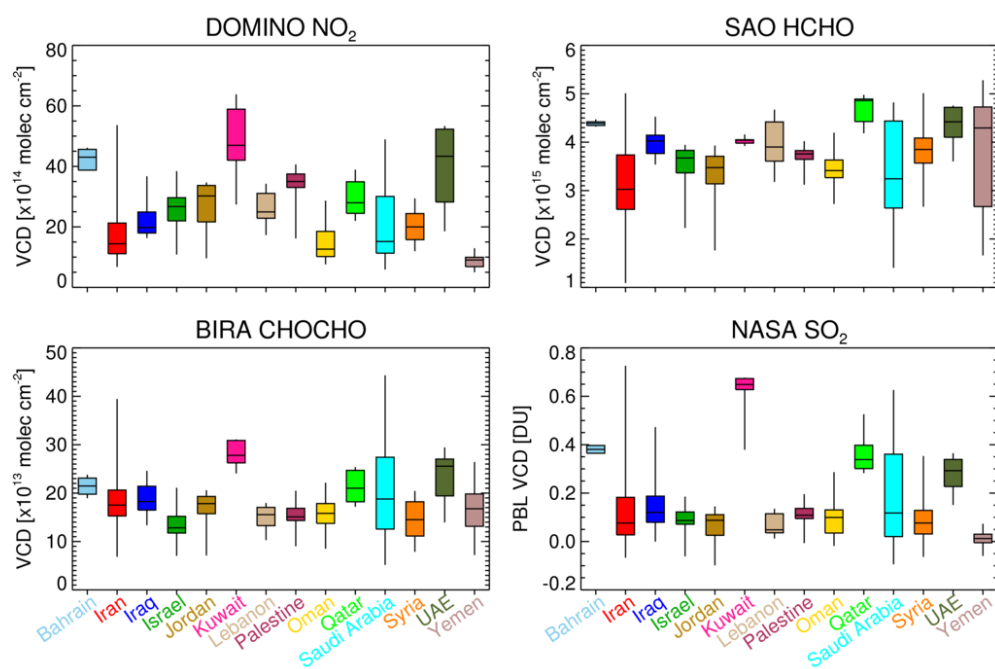


Figure 5. Box-and-whisker plots showing the minimum, 25th percentile, median, 75th percentile, and maximum values of the observed median vertical columns, for urban targets categorised by country.

3 Statistical Summaries

Table 3. The difference between the linear trend (in % yr⁻¹) and inferred growth rates (in % yr⁻¹) per species and target category.

Species	Urban	Refinery	Ports	Plants
NO ₂				
Maximum	4.84	3.18	2.49	4.21
Minimum	-5.29	-1.56	-0.95	-4.90
Median	-0.18	-0.09	-0.11	0.78
HCHO				
Maximum	3.42	1.33	1.10	3.42
Minimum	-2.46	0.00	-0.70	-2.34
Median	-0.06	0.90	0.08	-0.93
SO ₂				
Maximum	31.50	5.66	6.03	77.06
Minimum	-21.63	-2.44	-0.36	-16.63
Median	0.21	-0.97	2.84	0.33
CHOCHO				
Maximum	-3.69	-	-	3.14
Minimum	-3.69	-	-	3.14
Median	-3.69	-	-	3.14

Table 4. Statistical summary of the DOMINO NO₂ vertical column density (VCD) analysis. Tests 1–6 are defined as follows: (1) construction of each 10 year time series using mask of ± 4 grid-cells (~ 20 km radius around each target), instead of the default ± 2 grid-cells (~ 10 km radius around each target), (2) use of cloud fraction filter of 40%, (3) use of unaffected OMI detector rows 5 to 23 only, (4) increased smoothing of gridded maps (spatial filter of $0.35^\circ \times 0.35^\circ$ with a $2\text{-}\sigma$ width), (5) no filtering for outliers in the time series analysis, (6) focus on only locations with $>500,000$ people using a spatial mask of ± 16 grid-cells (~ 80 km radius around each target). Column headers are: n=number of locations with trends, n_{same} number of same locations with trends as found in default analysis, n_{new}= number of new locations trend found in the this test, and n_{miss} is the number of locations from the original analysis that have not been detected.

Test	Urban Targets (default=198)				Refinery default=17				Oil Ports default=6				Power Plants default=57			
	n	n _{same}	n _{new}	n _{miss}	n	n _{same}	n _{new}	n _{miss}	n	n _{same}	n _{new}	n _{miss}	n	n _{same}	n _{new}	n _{miss}
Test 1	220	170	50	28	17	14	3	3	6	5	1	1	69	51	18	6
Test 2	165	159	6	39	13	13	0	4	6	6	0	0	49	43	6	14
Test 3	95	57	38	141	6	4	2	13	5	3	2	3	30	14	16	43
Test 4	207	194	13	4	15	15	0	2	6	6	0	0	59	56	3	1
Test 5	199	172	27	26	16	15	1	2	6	6	0	0	58	52	6	5
Test 6	22	12	10	186	–	–	–	–	–	–	–	–	–	–	–	–

Table 5. Statistical summary of the SAO HCHO vertical column density (VCD) analysis. Tests 1–6 are defined as follows: ((1) construction of each 10 year time series using mask of ± 4 grid-cells (~ 20 km radius around each target), instead of the default ± 2 grid-cells (~ 10 km radius around each target), (2) use of cloud fraction filter of 40%, (3) use of unaffected OMI detector rows 5 to 23 only, (4) increased smoothing of gridded maps (spatial filter of $0.35^\circ \times 0.35^\circ$ with a $2\text{-}\sigma$ width), (5) no filtering for outliers in the time series analysis, (6) focus on only locations with $>500,000$ people using a spatial mask of ± 16 grid-cells (~ 80 km radius around each target).. Column headers are: n=number of locations with trends, n_{same} number of same locations with trends as found in default analysis, n_{new}= number of new locations trend found in the this test, and n_{miss} is the number of locations from the original analysis that have not been detected.

Test	Urban Targets (default=34)				Refinery default=6				Oil Ports default=4				Power Plants default=26			
	n	n _{same}	n _{new}	n _{miss}	n	n _{same}	n _{new}	n _{miss}	n	n _{same}	n _{new}	n _{miss}	n	n _{same}	n _{new}	n _{miss}
Test 1	46	29	17	5	7	6	1	0	8	4	4	0	34	21	13	5
Test 2	34	27	7	7	3	3	0	3	4	3	1	1	27	22	5	4
Test 3	19	7	12	27	4	3	1	3	1	1	0	3	14	10	4	16
Test 4	43	33	10	1	6	6	0	0	5	4	1	0	33	26	7	0
Test 5	33	31	2	3	6	6	0	0	4	4	0	0	29	25	4	1
Test 6	9	3	6	31	–	–	–	–	–	–	–	–	–	–	–	–

Table 6. Statistical summary of the NASA SO₂ vertical column density (VCD) analysis, tests 1–6 are defined as follows: (1) construction of each 10 year time series using mask of ± 4 grid-cells (~ 20 km radius around each target), instead of the default ± 2 grid-cells (~ 10 km radius around each target), (2) use of cloud fraction filter of 40%, (3) use of unaffected OMI detector rows 5 to 23 only, (4) increased smoothing of gridded maps (spatial filter of $0.35^\circ \times 0.35^\circ$ with a $2\text{-}\sigma$ width), (5) no filtering for outliers in the time series analysis, (6) focus on only locations with $>500,000$ people using a spatial mask of ± 16 grid-cells (~ 80 km radius around each target).. Note that test 2 (different cloud fraction filtering) was not applied to the SO₂ data. Column headers are: n=number of locations with trends, n_{same} number of same locations with trends as found in default analysis, n_{new}= number of new locations trend found in the this test, and n_{miss} is the number of locations from the original analysis that have not been detected.

Test	Urban Targets (default=18)				Refinery (default=3)				Oil Ports (default=2)				Power Plants (default=9)			
	n	n _{same}	n _{new}	n _{miss}	n	n _{same}	n _{new}	n _{miss}	n	n _{same}	n _{new}	n _{miss}	n	n _{same}	n _{new}	n _{miss}
Test 1	18	9	9	9	3	3	0	0	3	1	2	1	9	6	3	3
Test 3	13	8	5	10	1	1	0	2	1	1	0	1	10	5	5	4
Test 4	17	17	0	1	3	3	0	0	3	2	1	0	9	8	1	1
Test 5	14	11	3	7	2	2	0	1	1	1	0	1	4	3	1	6
Test 6	3	1	2	17	–	–	–	–	–	–	–	–	–	–	–	–

Table 7. Statistical summary of the BIRA CHOCHO vertical column density (VCD) analysis, tests 1–5 are defined as follows: (1) construction of each 10 year time series using mask of ± 4 grid-cells (~ 20 km radius around each target), instead of the default ± 2 grid-cells (~ 10 km radius around each target), (2) use of cloud fraction filter of 40%, (3) use of unaffected OMI detector rows 5 to 23 only, (4) increased smoothing of gridded maps (spatial filter of $0.35^\circ \times 0.35^\circ$ with a $2\text{-}\sigma$ width), (5) no filtering for outliers in the time series analysis, (6) focus on only locations with $>500,000$ people using a spatial mask of ± 16 grid-cells (~ 80 km radius around each target). Note that test 2 (different cloud fraction filtering) was not applied to the SO₂ data. Column headers are: n=number of locations with trends, n_{same} number of same locations with trends as found in default analysis, n_{new}= number of new locations trend found in the this test, and n_{miss} is the number of locations from the original analysis that have not been detected.

Test	Urban Targets (default=1)				Refinery (default=0)				Oil Ports (default=0)				Power Plants (default=1)			
	n	n _{same}	n _{new}	n _{miss}	n	n _{same}	n _{new}	n _{miss}	n	n _{same}	n _{new}	n _{miss}	n	n _{same}	n _{new}	n _{miss}
Test 1	2	1	1	0	1	0	1	0	0	0	0	1	1	1	0	0
Test 3	2	0	2	1	0	0	0	0	0	0	0	0	0	0	0	1
Test 4	2	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0
Test 5	1	0	1	0	0	0	0	0	0	0	0	0	3	1	2	0
Test 6	0	0	0	0	–	–	–	–	–	–	–	–	–	–	–	–