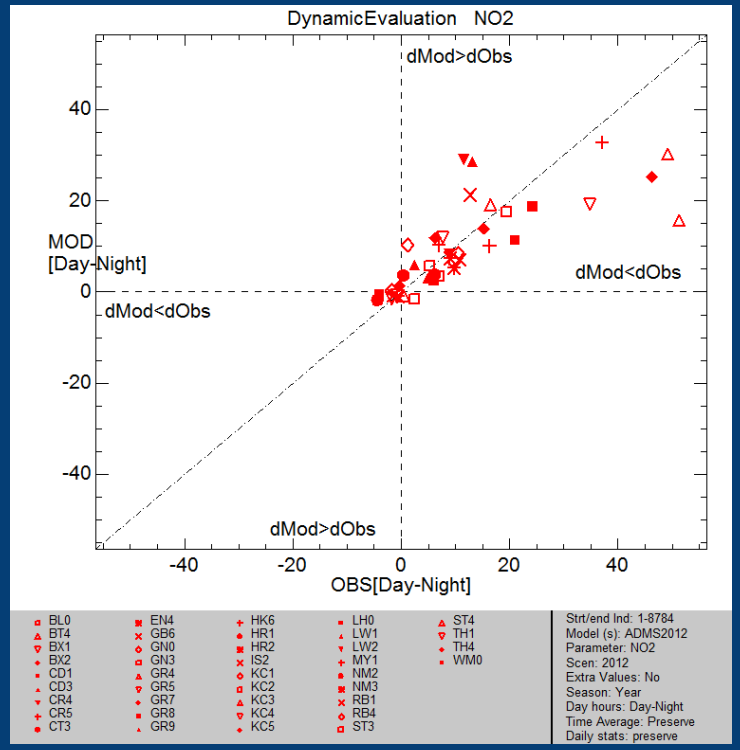
Comparing QA/QC metrics of two street-scale model configurations



Jenny Stocker,
Kate Johnson

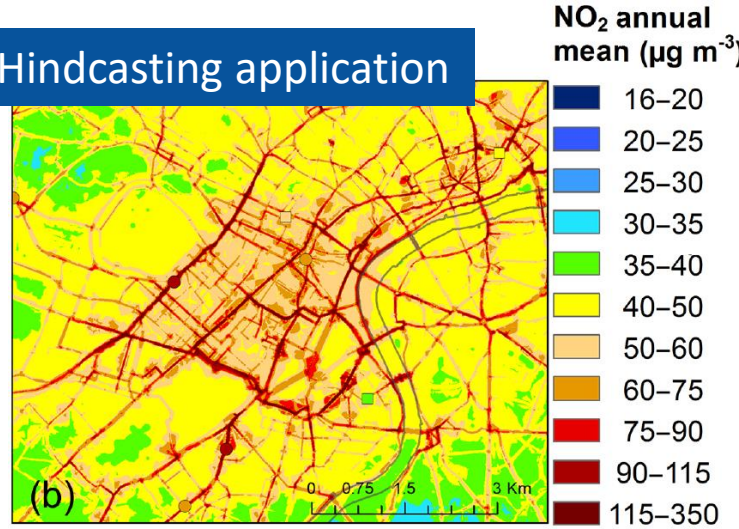
FAIRMODE

8th October 2021

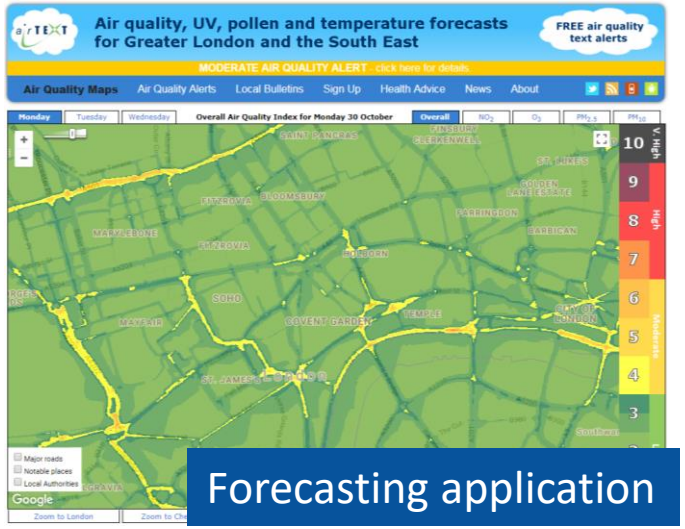
Online

Model descriptions

Hindcasting application



- Hourly modelled concentrations from CERC's quasi-Gaussian ADMS-Urban model
- Street-scale resolution allows evaluation at a range of site types
- **2 applications**



- **Research project** involving modelling AQ in London (2012)
- 45 sites (17 background, 28 traffic)
- Measured meteorology and long-range pollutant concentration data used as input, so **good model performance expected**

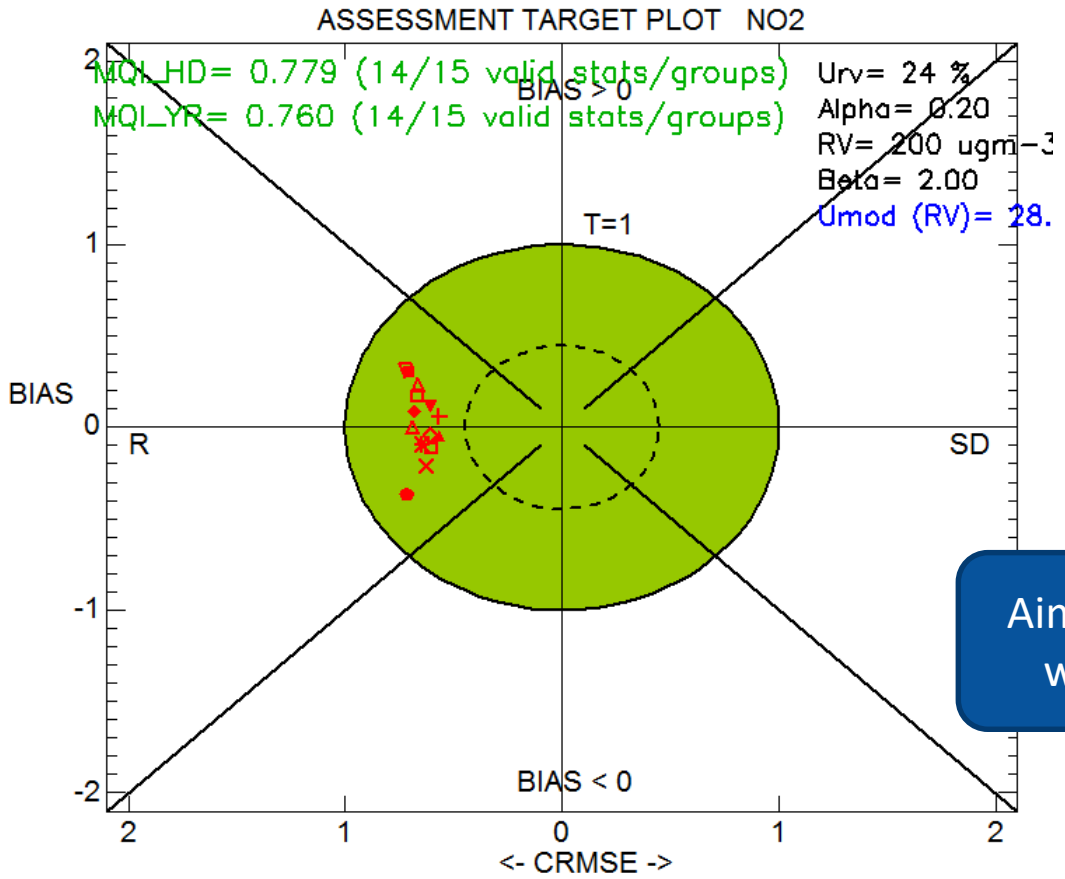
Hood, C., MacKenzie, I., Stocker, J., Johnson, K., Carruthers, D., Vieno, M. and Doherty, R., 2018. Air quality simulations for London using a coupled regional-to-local modelling system. Atmospheric Chemistry and Physics, 18(15), pp.11221-11245.

- **Air quality forecast results** from the airTEXT (www.airtext.info) system for London (2018)
- 84 sites (13 background, 52 traffic, 4 industrial)
- Forecast meteorology and forecast long-range pollutant concentration data (CAMS) used as input, so **model performance likely to be less good**

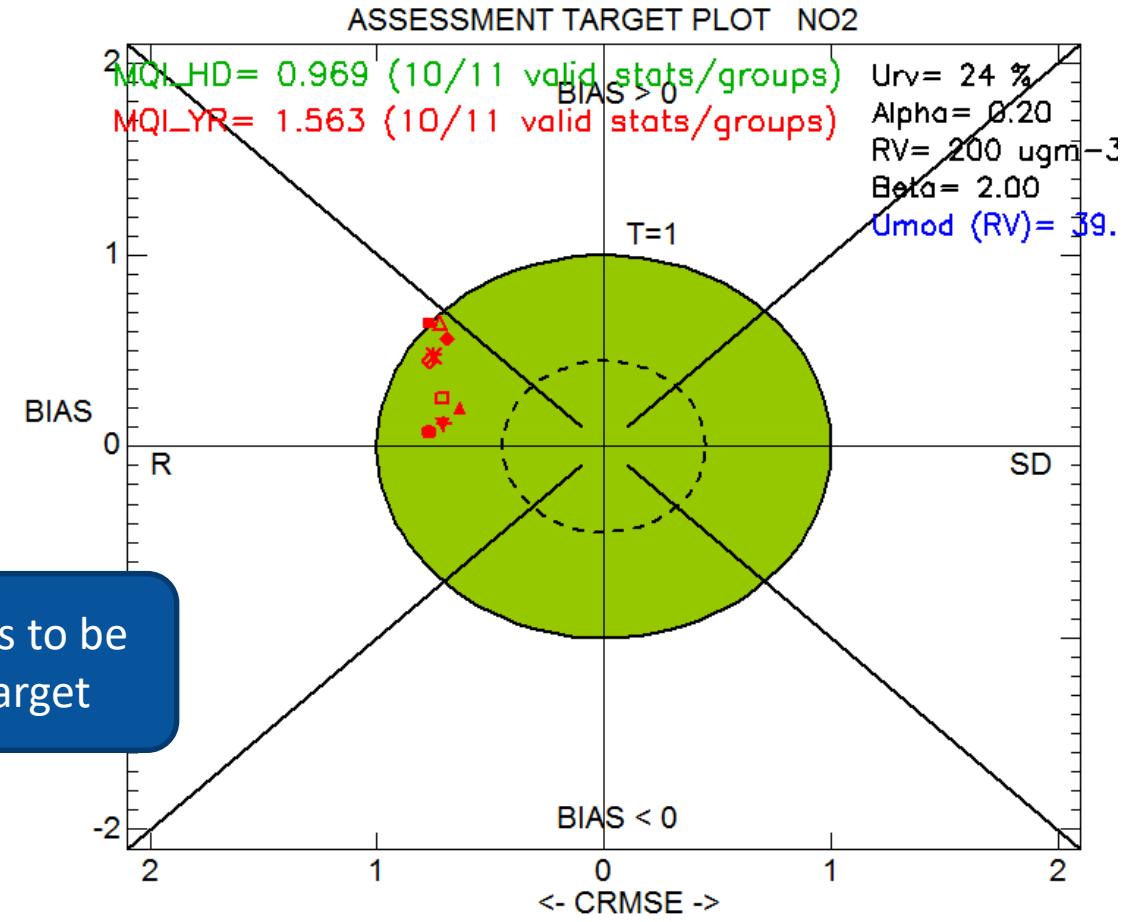
Stidworthy, A., Jackson, M., Johnson, K., Carruthers, D. and Stocker, J., 2018. Evaluation of local and regional air quality forecasts for London. International Journal of Environment and Pollution, 64(1-3), pp.178-191.

Compare performance: Background site target plot NO2

- Hindcast



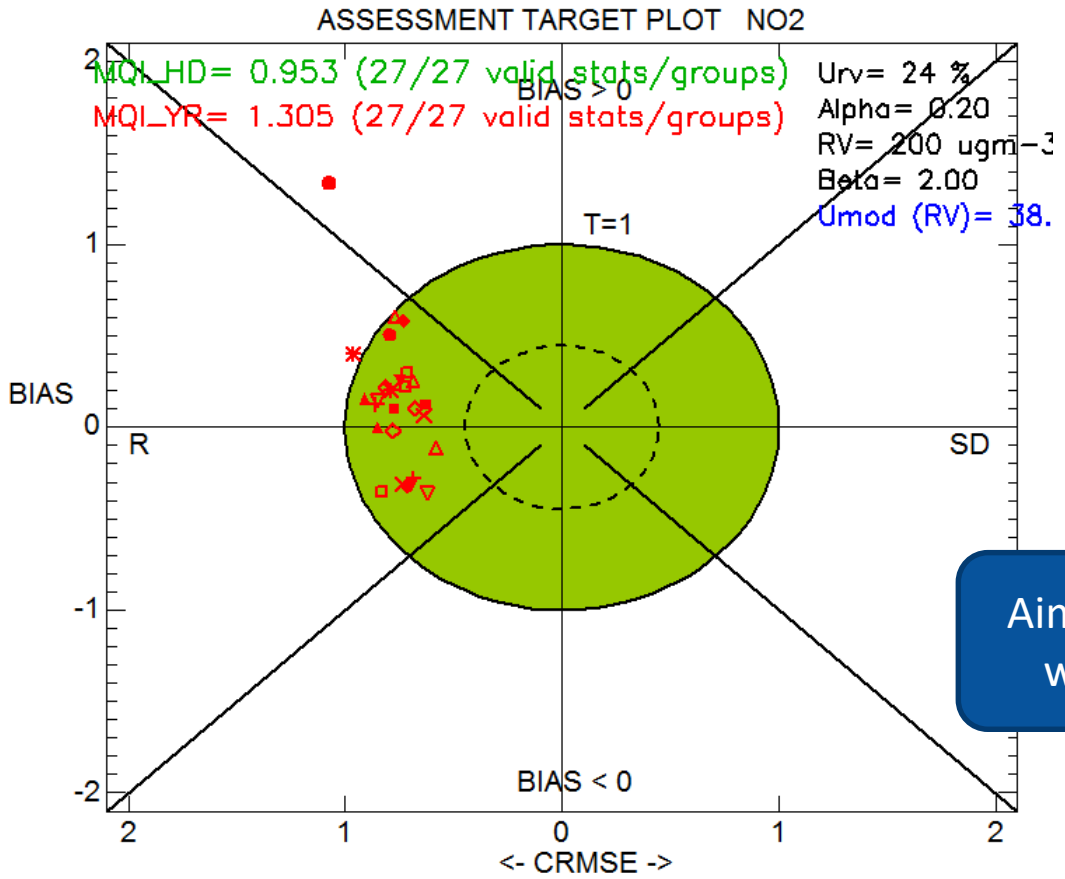
- Forecast



Aiming for points to be within green target

Compare performance: Traffic site target plot NO2

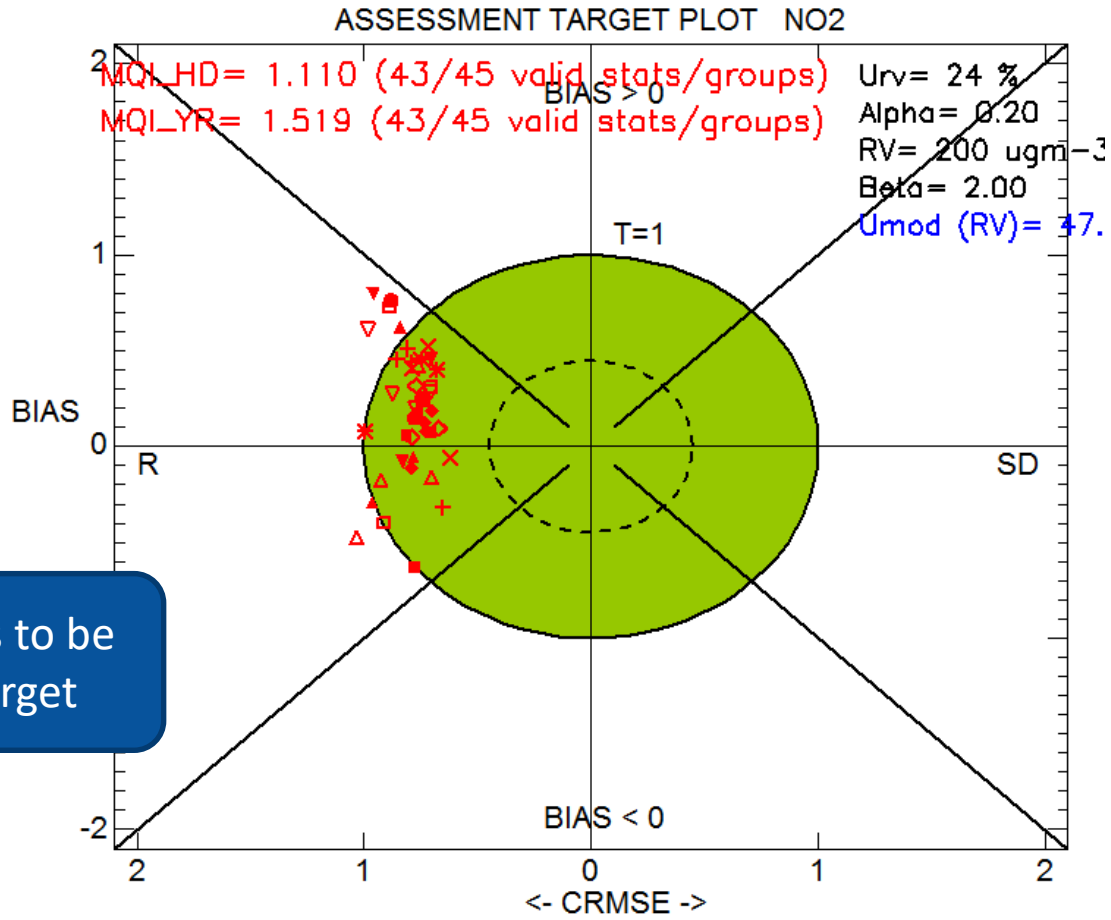
- Hindcast



BT4 CD1 CD3 CR4 CR5 EN4 GB6 GN0 GN3 GR5 GR7 GR8 GR9 HG1 HK6 HR2 IS2 KC2 KC3 KC4 KC5 LW2 MY1 NM2 RB4 ST4 TH4

Strt/end Ind: 1-8784
Model (s): ADMS2012
Parameter: NO2
Scen: 2012
Extra Values: No
Season: Year
Day hours: All 24h
Time Average: Preserve
Daily stats: preserve

- Forecast



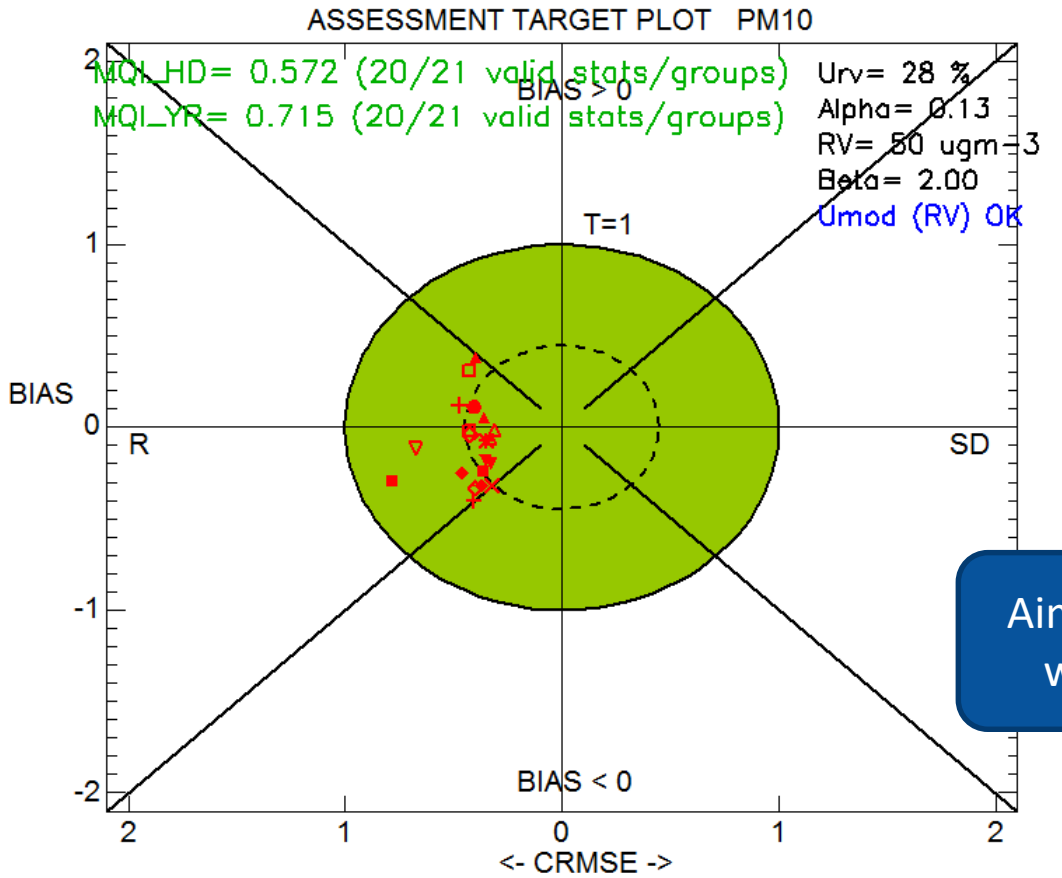
Aiming for points to be within green target

BG1 BG2 BT4 BT6 BX1 BX2 CD1 CR5 CT6 EA6 EI1 EN1 EN4 EN5 GB6 GN0 GN3 GN4 GR4 GR7 GR8 GR9 HG1 HI0 HK6 HR2 HV1 HV3 IS2 LW2 MY1 RB4 RI1 RI2 SK5 ST4 ST6 TH2 TH4 TK8 WA7 WA8 WM6

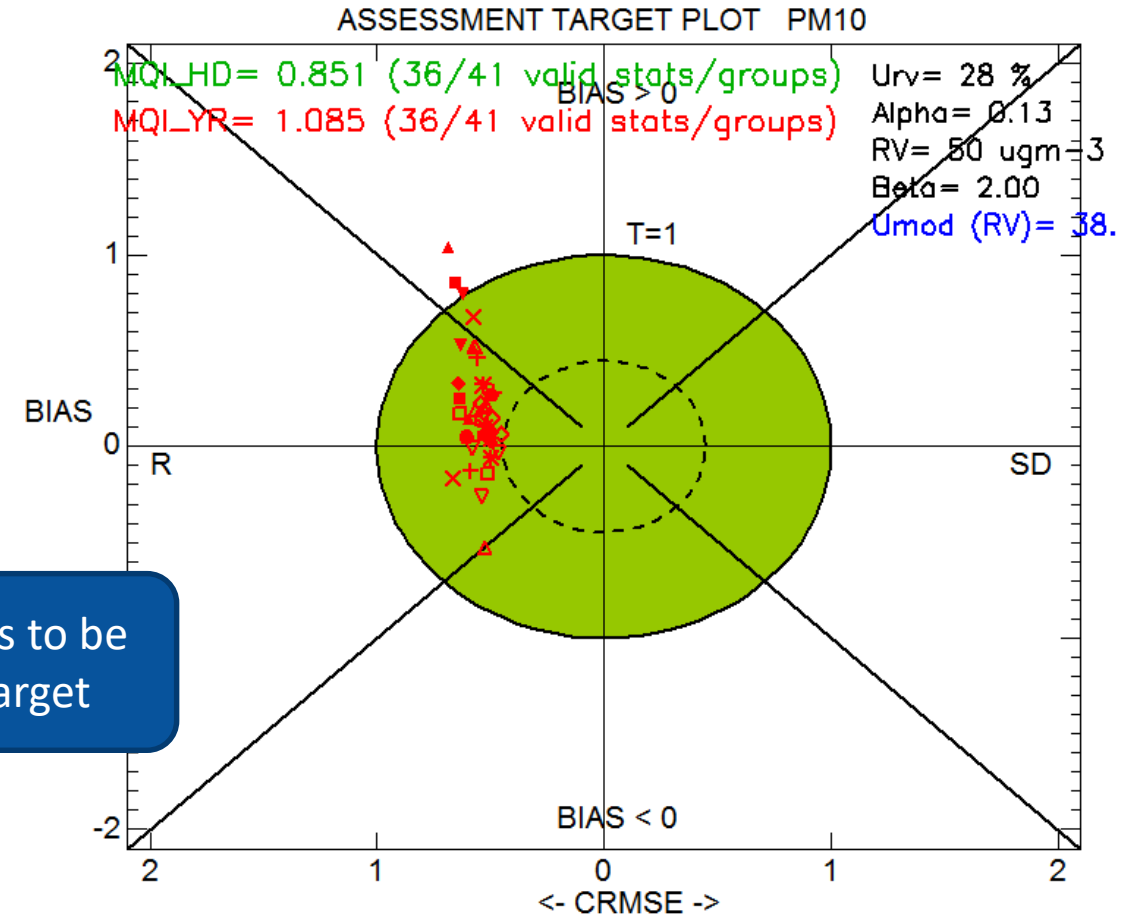
Strt/end Ind: 1-8760
Model (s): Day0
Parameter: NO2
Scen: 2018
Extra Values: No
Season: Year
Day hours: All 24h
Time Average: Preserve
Daily stats: preserve

Compare performance: Traffic site target plot PM10

- Hindcast



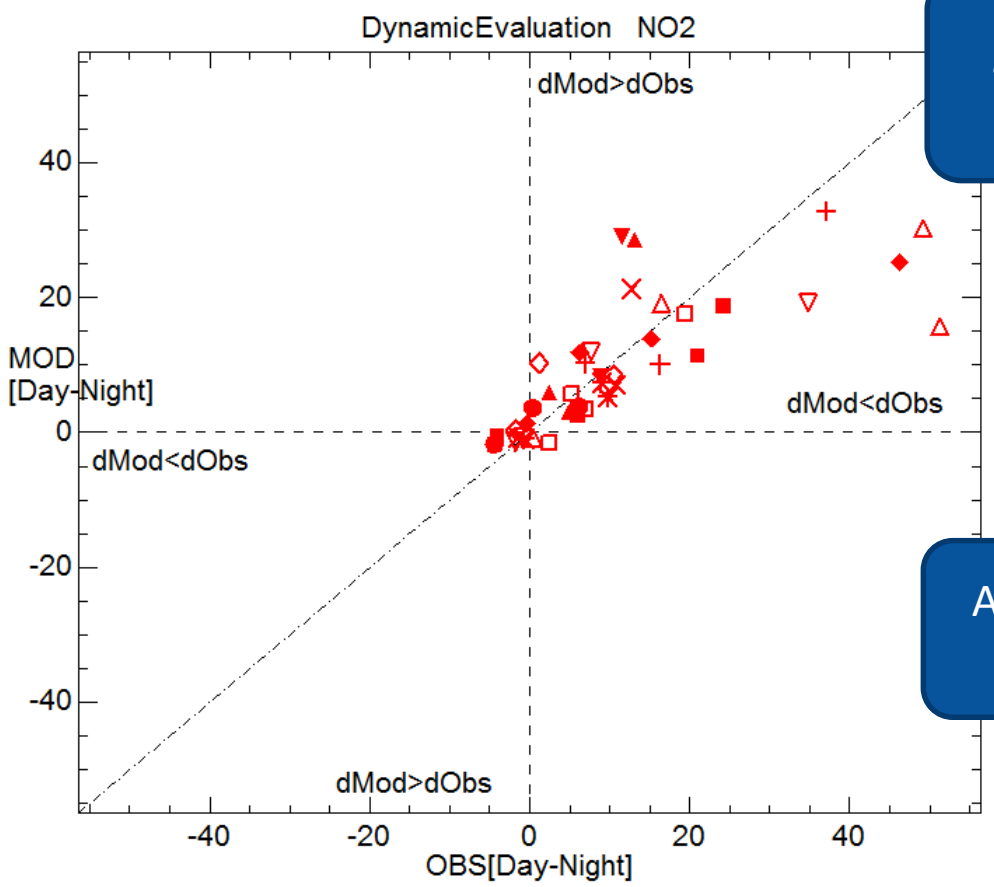
- Forecast



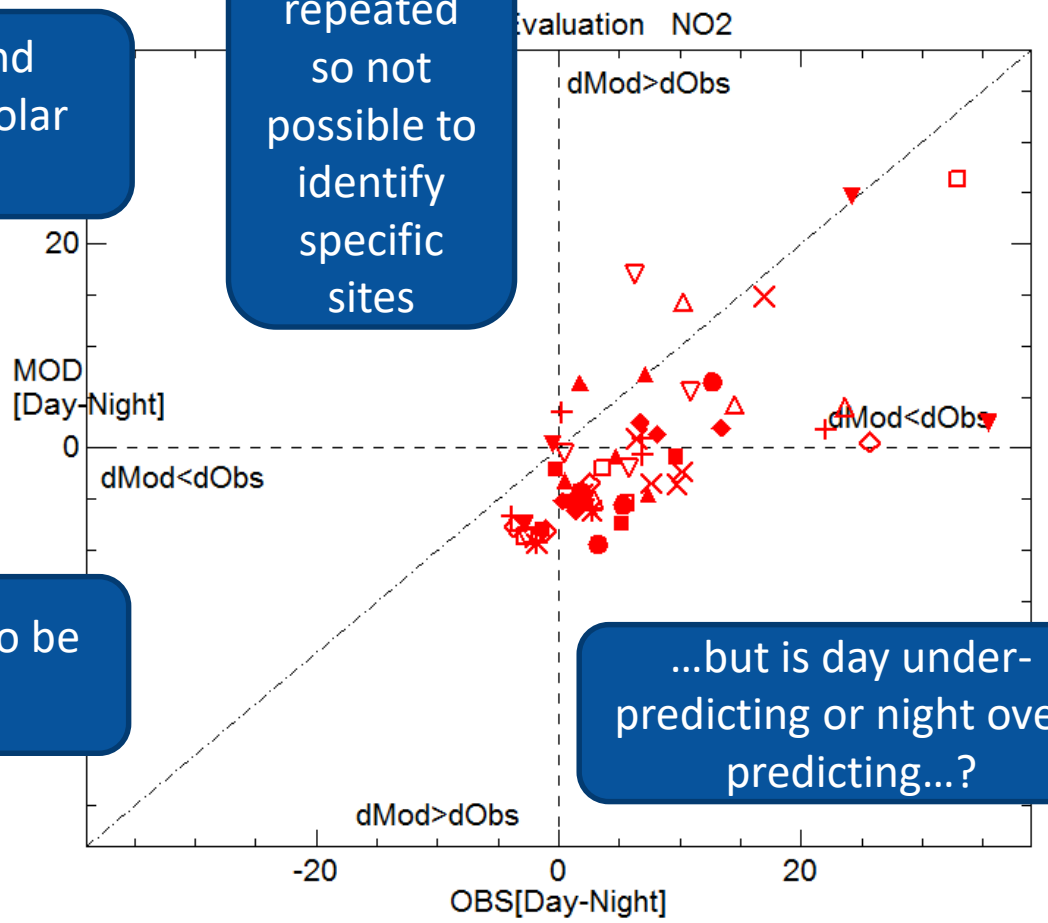
Aiming for points to be within green target

Compare performance: Day – Night NO2

- Hindcast



- Forecast



How are 'day' and 'night' defined? Solar elevation?

Symbols repeated so not possible to identify specific sites

Aiming for points to be on the 1:1 line

…but is day under-predicting or night over-predicting…?

BL0 BT4 BX1 BX2 CD1 CD3 CR4 CR5 CT3 EN4 GB6 GN0 GN3 GR4 GR5 GR7 GR8 GR9 HK6 HR1 HR2 IS2 KC1 KC2 KC3 KC4 KC5 LH0 LW1 LW2 MY1 NM2 NM3 RB1 RB4 ST3 ST4 TH1 TH4 WM0

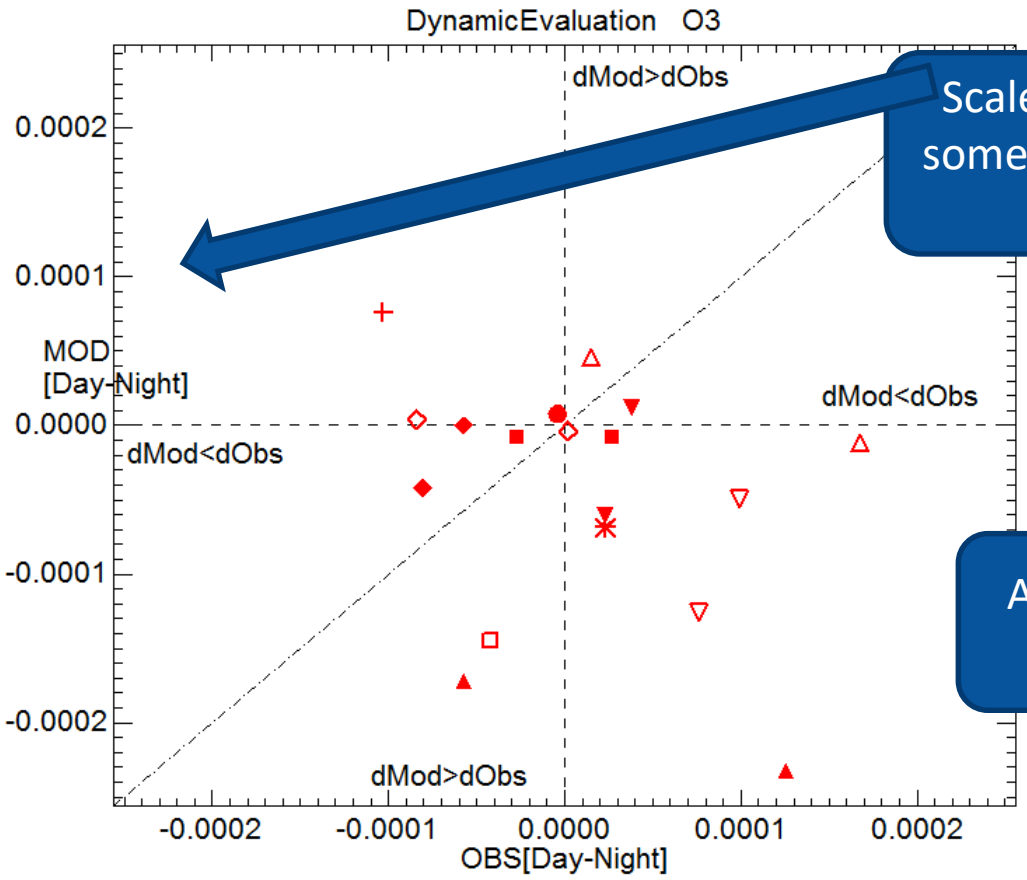
Strt/end Ind: 1-8784
Model (s): ADMS2012
Parameter: NO2
Scen: 2012
Extra Values: No
Season: Year
Day hours: Day-Night
Time Average: Preserve
Daily stats: preserve

BG2 BL0 BQ7 BT4 BT5 BT6 BX1 BX2 CD1 CR5 CT3 CT6 EA6 EA8 EI1 EN1 EN4 EN5 GB6 GN0 GN3 GN4 GR4 GR7 GR8 GR9 HG1 HI0 HK6 HR2 HV1 HV3 IS2 IS6 KC1 LB5 LB6 LW1 LW2 MY1 RB4 RI1 RI2 SK5 ST4

Strt/end Ind: 1-8760
Model (s): Day0
Parameter: NO2
Scen: 2018
Extra Values: No
Season: Year
Day hours: Day-Night
Time Average: Preserve
Daily stats: preserve

Compare performance: Day – Night O3

- Hindcast



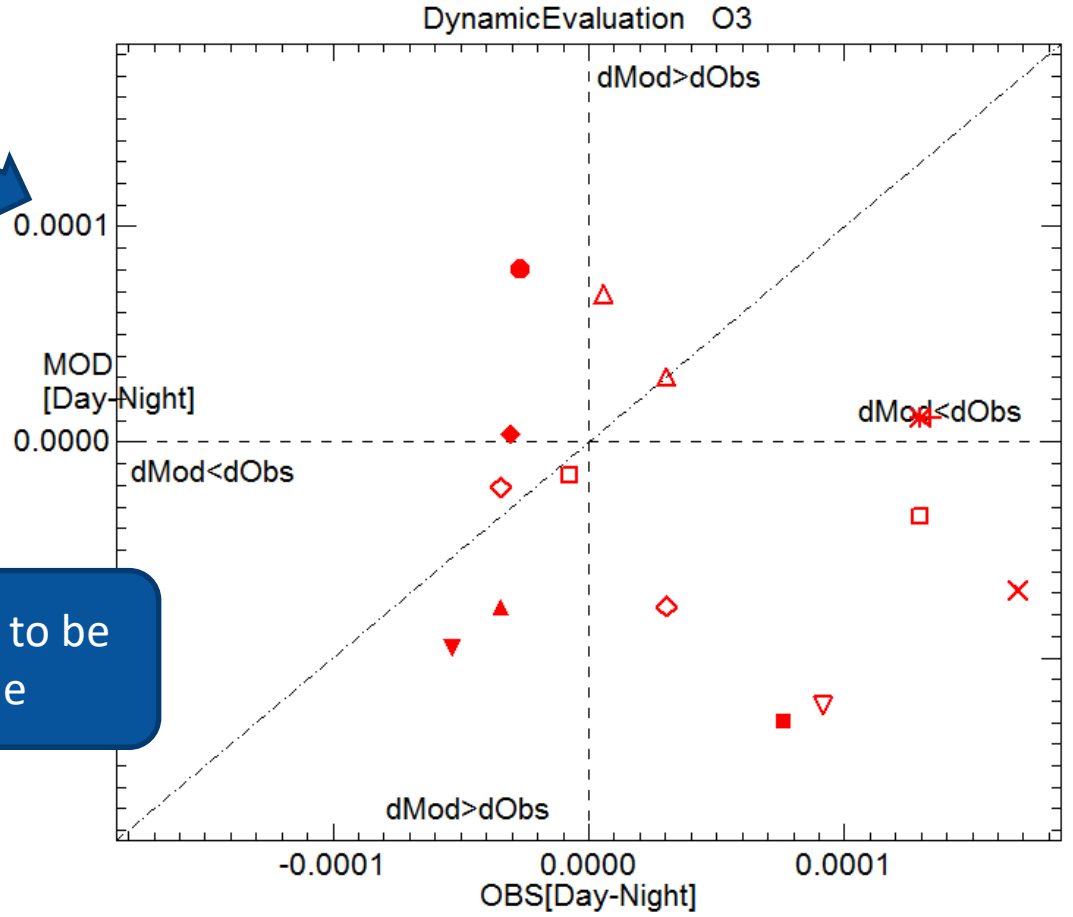
BL0 BX1 GB6 GN3 GR4 GR8 GR9 HK6 KC1 LH0 MY1 NM2 NM3 RB1 ST3 TH1 TH4 WM0

Strt/end Ind: 1-8784
Model (s): ADMS2012
Parameter: O3
Scen: 2012
Extra Values: No
Season: Year
Day hours: Day-Night
Time Average: 8h
Daily stats: Max

Scales very low, something seems wrong

Aiming for points to be on the 1:1 line

- Forecast

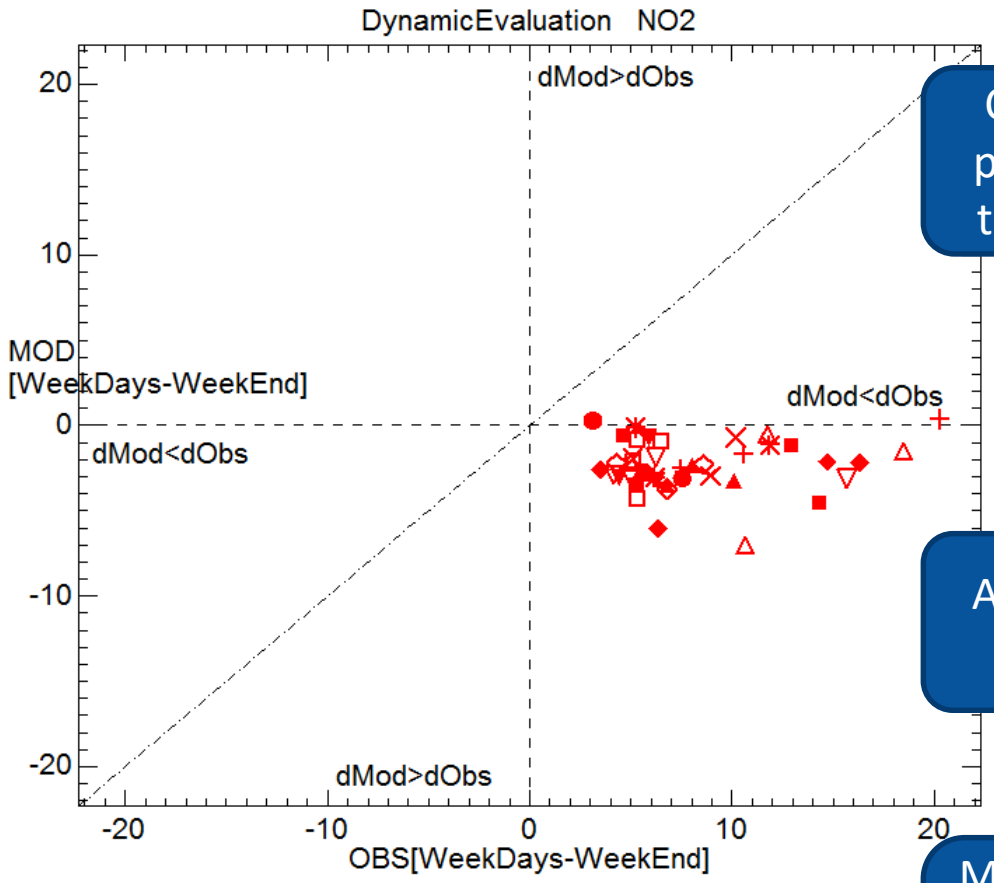


BL0 BQ7 BT4 BX1 GB6 GN3 GR4 GR8 GR9 HI0 HK6 KC1 MY1 RI2 TH4

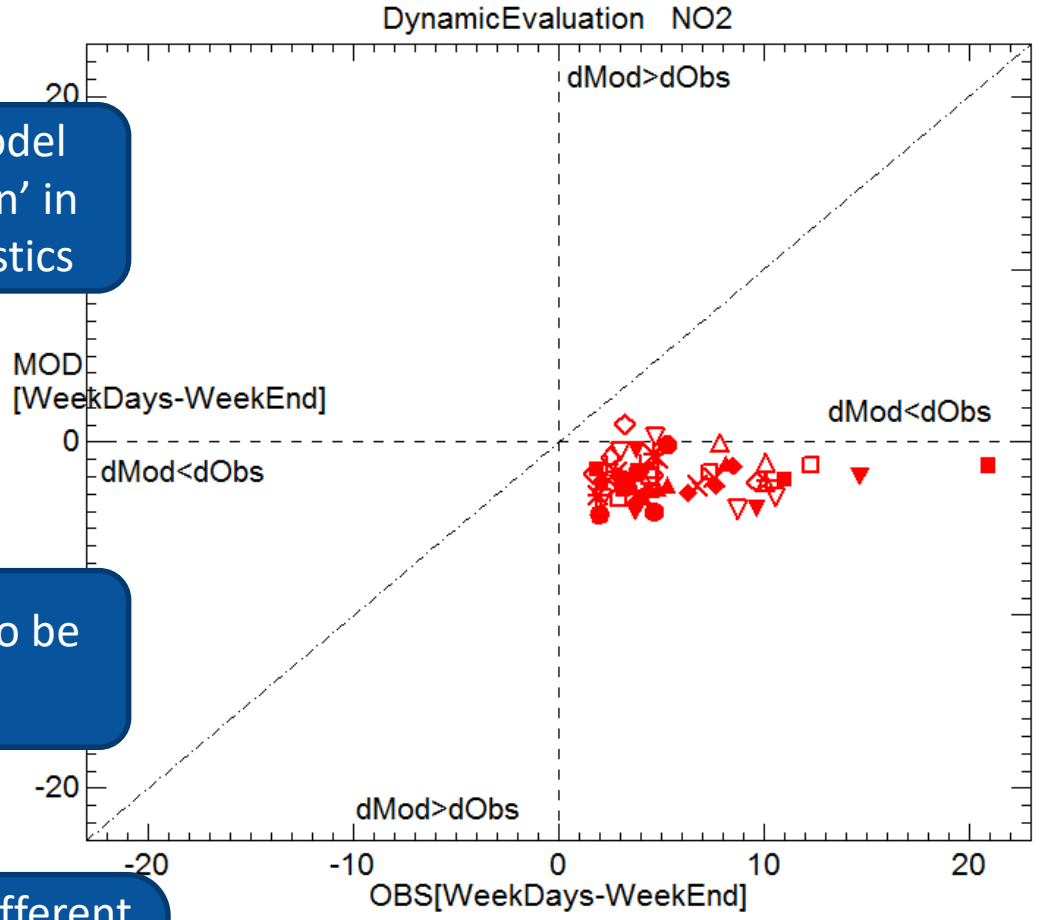
Strt/end Ind: 1-8760
Model (s): Day0
Parameter: O3
Scen: 2018
Extra Values: No
Season: Year
Day hours: Day-Night
Time Average: 8h
Daily stats: Max

Compare performance: Weekday / weekend NO2

- Hindcast



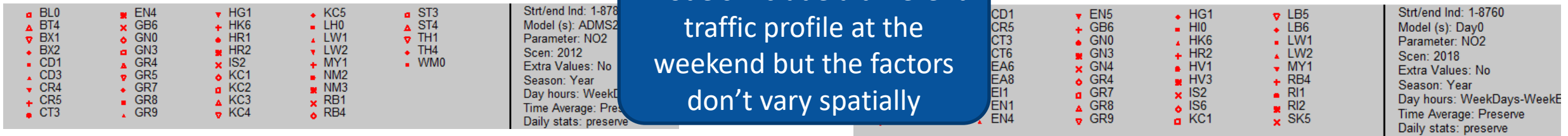
- Forecast



Corresponding model performance 'green' in the summary statistics

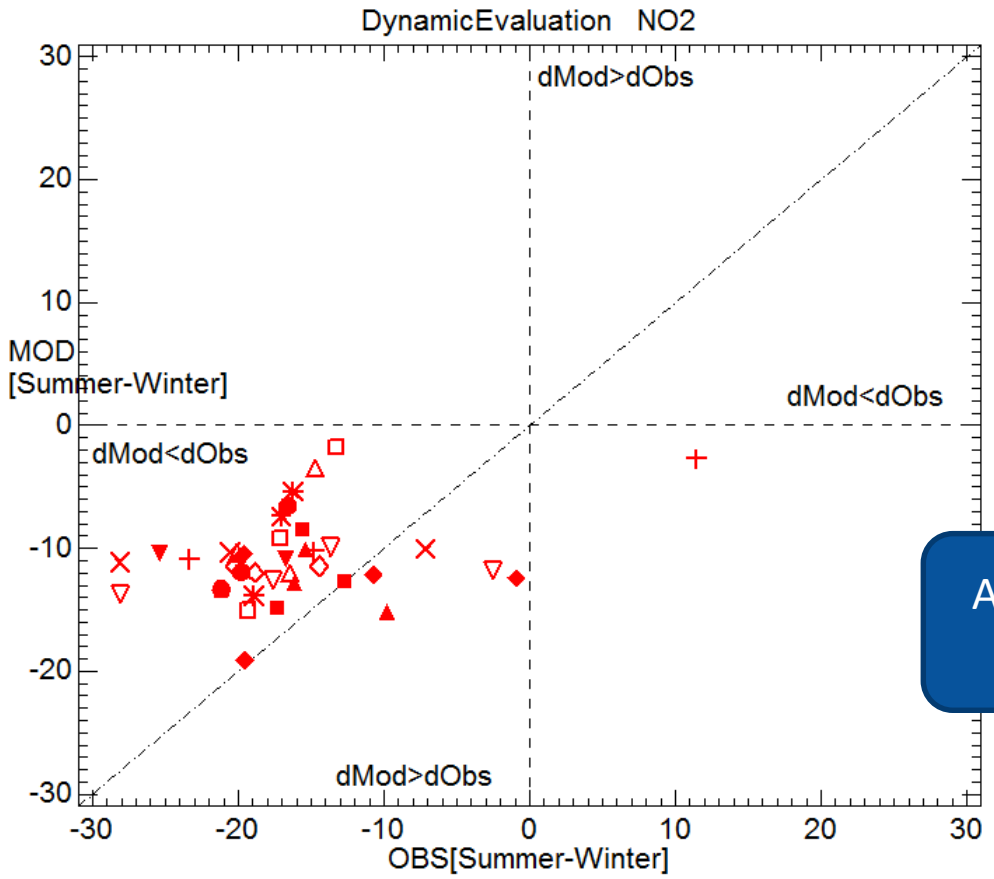
Aiming for points to be on the 1:1 line

Models include a different traffic profile at the weekend but the factors don't vary spatially



Compare performance: Summer – Winter NO2

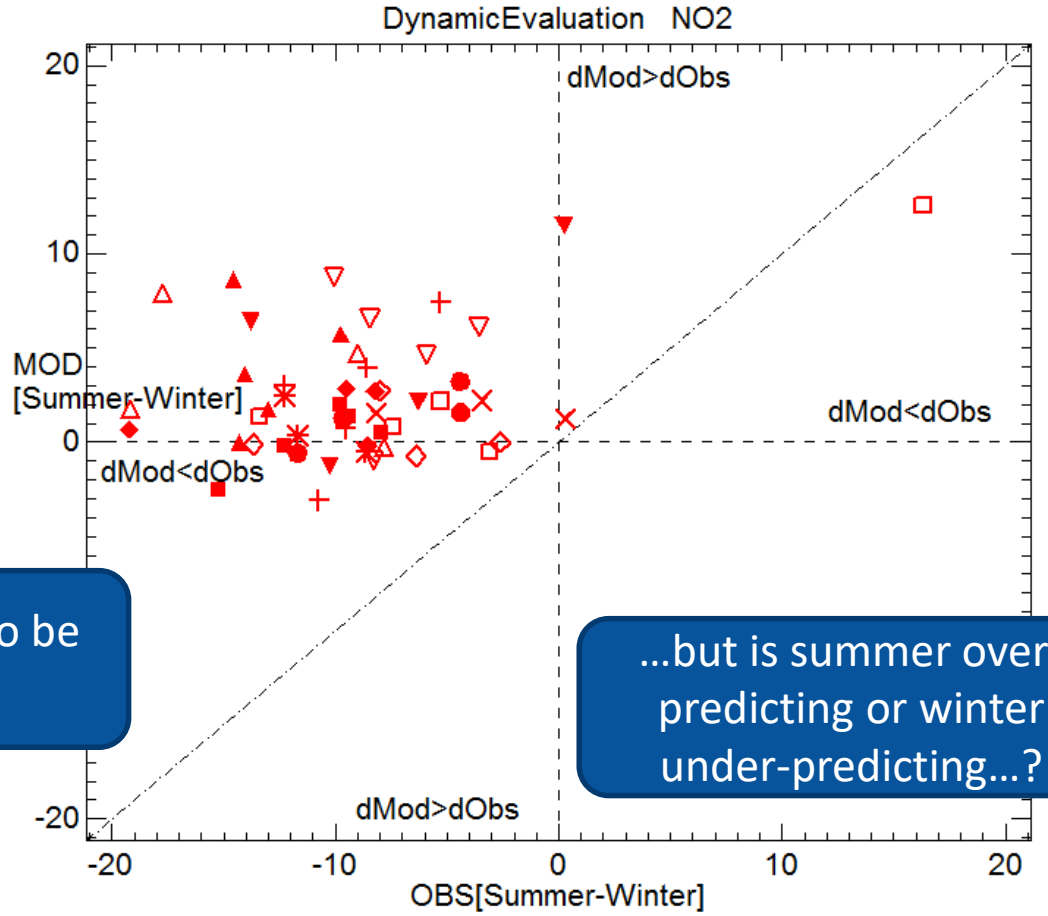
- Hindcast



BX1 BX2 CD1 CD3 CR4 CR5 CT3 EN4 GB6 GN0 GN3 GR4 GR5 GR7 GR9 HK6 HR1 HR2 IS2 KC1 KC2 KC3 KC4 KC5 LH0 LW1 LW2 MY1 NM2 NM3 RB1 RB4 ST3 ST4 TH1 TH4 WM0

Strt/end Ind: 1-8784
Model (s): ADMS2012
Parameter: NO2
Scen: 2012
Extra Values: No
Season: Summer-Winter
Day hours: All 24h
Time Average: Preserve
Daily stats: preserve

- Forecast



BG2 BL0 BQ7 BT5 BT6 BX1 BX2 CD1 CR5 CT3 CT6 EA6 EA8 EI1 EN1 EN4 EN5 GB6 GN0 GN3 GN4 GR4 GR7 GR8 GR9 HG1 HI0 HK6 HR2 HV1 IS2 IS6 KC1 LB5 LB6 LW1 LW2 MY1 RB4 RI1 RI2 ST4 ST5 ST6 TH2

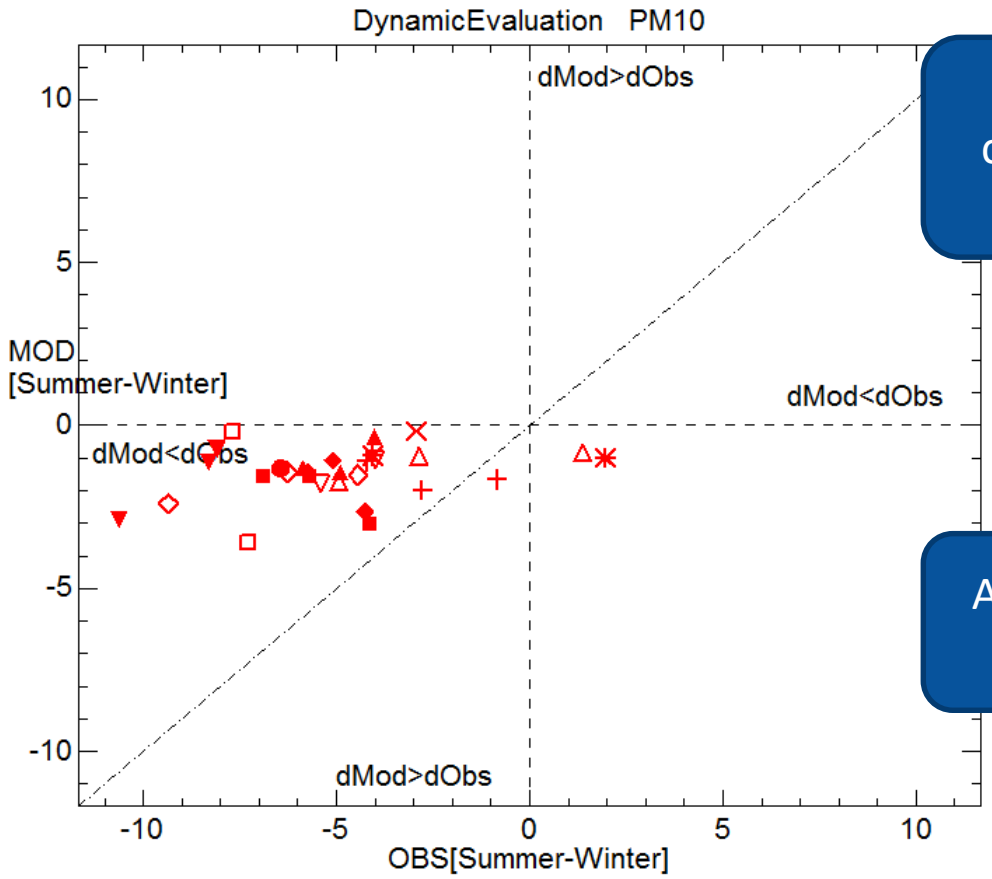
Strt/end Ind: 1-8760
Model (s): Day0
Parameter: NO2
Scen: 2018
Extra Values: No
Season: Summer-Winter
Day hours: All 24h
Time Average: Preserve
Daily stats: preserve

Aiming for points to be on the 1:1 line

…but is summer over-predicting or winter under-predicting…?

Compare performance: Summer – Winter PM10

- Hindcast



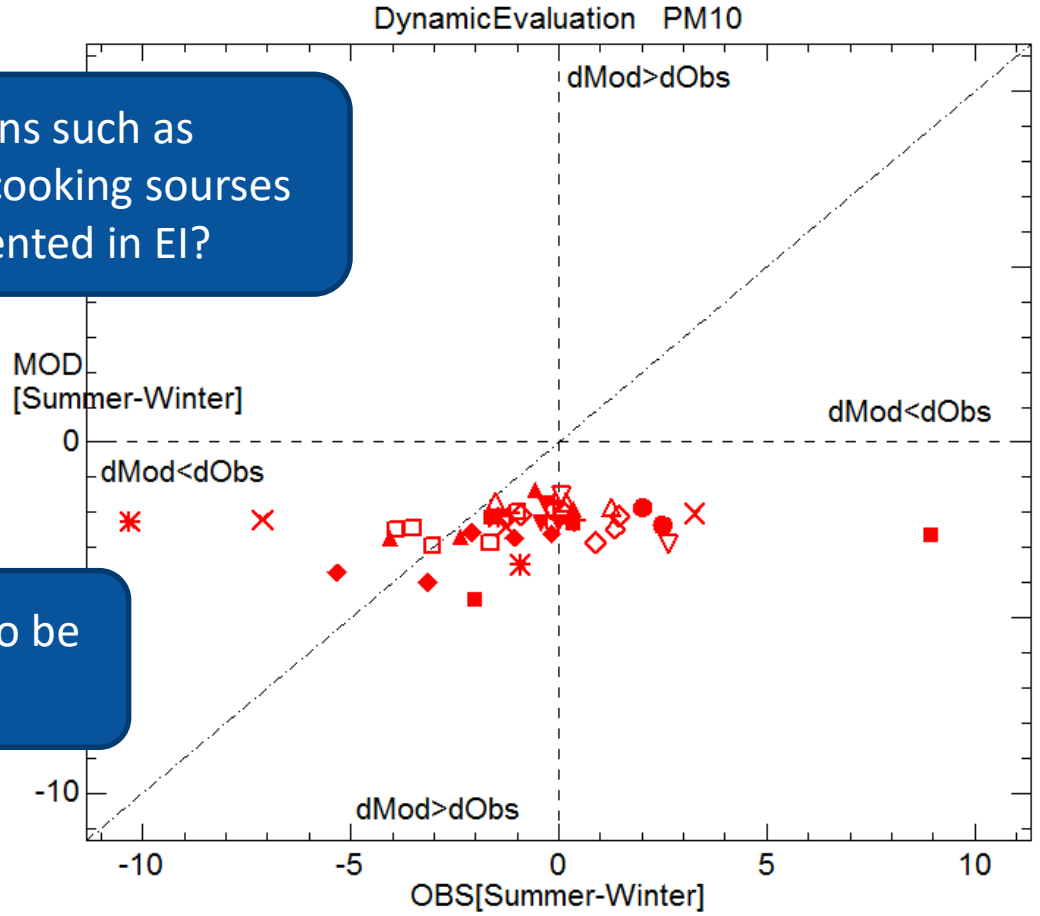
BL0 BX1 BX2 CD1 CD3 CR3 CR4 CT3 GN0 GR4 GR5 GR7 GR8 GR9 HG1 HK6 HR1 HR2 HV3 IS2 KC1 KC5 LH0 MY1 RB4 ST4 TH1 TH4 WM0

Strt/end Ind: 1-8784
Model (s): ADMS2012
Parameter: PM10
Scen: 2012
Extra Values: No
Season: Summer-Winter
Day hours: All 24h
Time Average: Preserve
Daily stats: Mean

Local emissions such as construction and cooking sourses poorly represented in EI?

Aiming for points to be on the 1:1 line

- Forecast

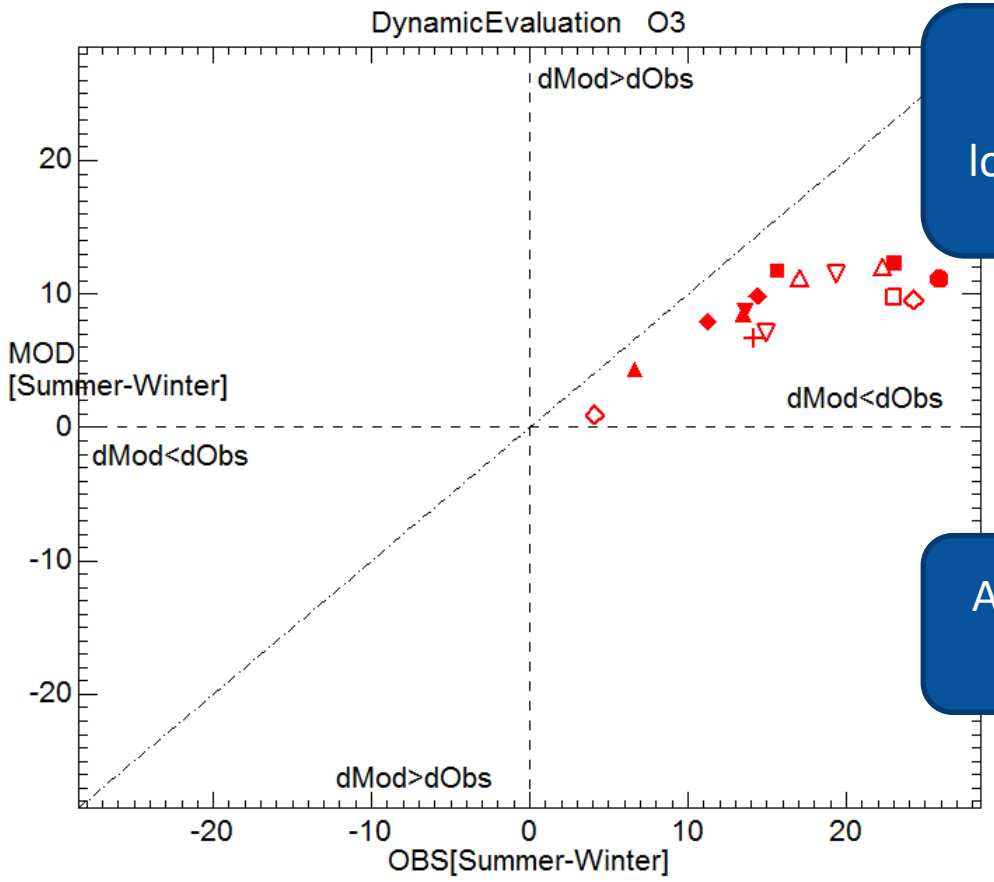


BG2 BL0 BQ7 BQ8 BT4 BT5 BT6 BX1 BX2 CD1 CT3 CT8 EA6 EA8 EI1 GB6 GN0 GN3 GN4 GR4 GR7 GR8 GR9 HR2 HV1 HV3 IS2 IS6 LB5 LB6 LW2 MY1 MY7 RB4 RI1 RI2 ST4 ST5 ST6 TH4 TK8 WA7 WA9 WM0

Strt/end Ind: 1-8760
Model (s): Day0
Parameter: PM10
Scen: 2018
Extra Values: No
Season: Summer-Winter
Day hours: All 24h
Time Average: Preserve
Daily stats: Mean

Compare performance: Summer – Winter O3

- Hindcast



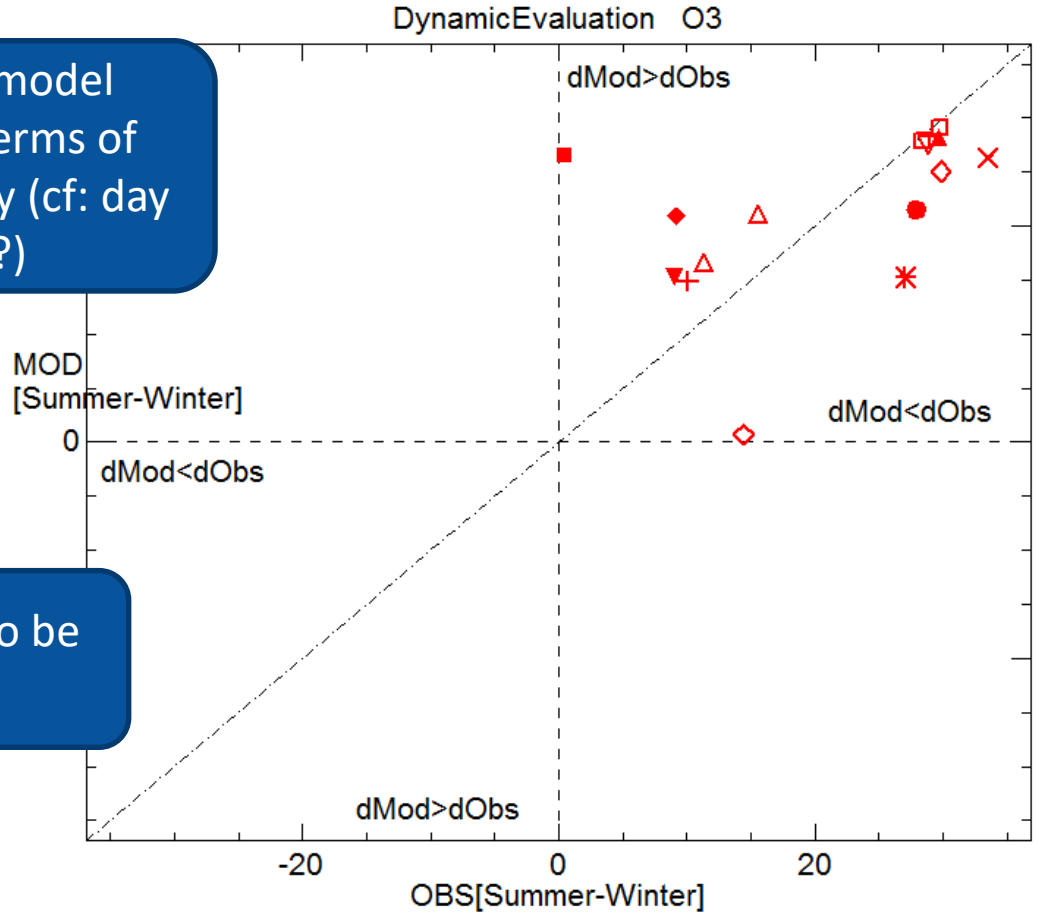
BL0 BX1 GB6 GN3 GR4 GR8 GR9 HK6 KC1 MY1 NM2 NM3 RB1 ST3 TH1 TH4

Strt/end Ind: 1-8784
Model (s): ADMS2012
Parameter: O3
Scen: 2012
Extra Values: No
Season: Summer-Winter
Day hours: All 24h
Time Average: 8h
Daily stats: Max

Generally good model performance in terms of local NOx chemistry (cf: day / night plot?)

Aiming for points to be on the 1:1 line

- Forecast

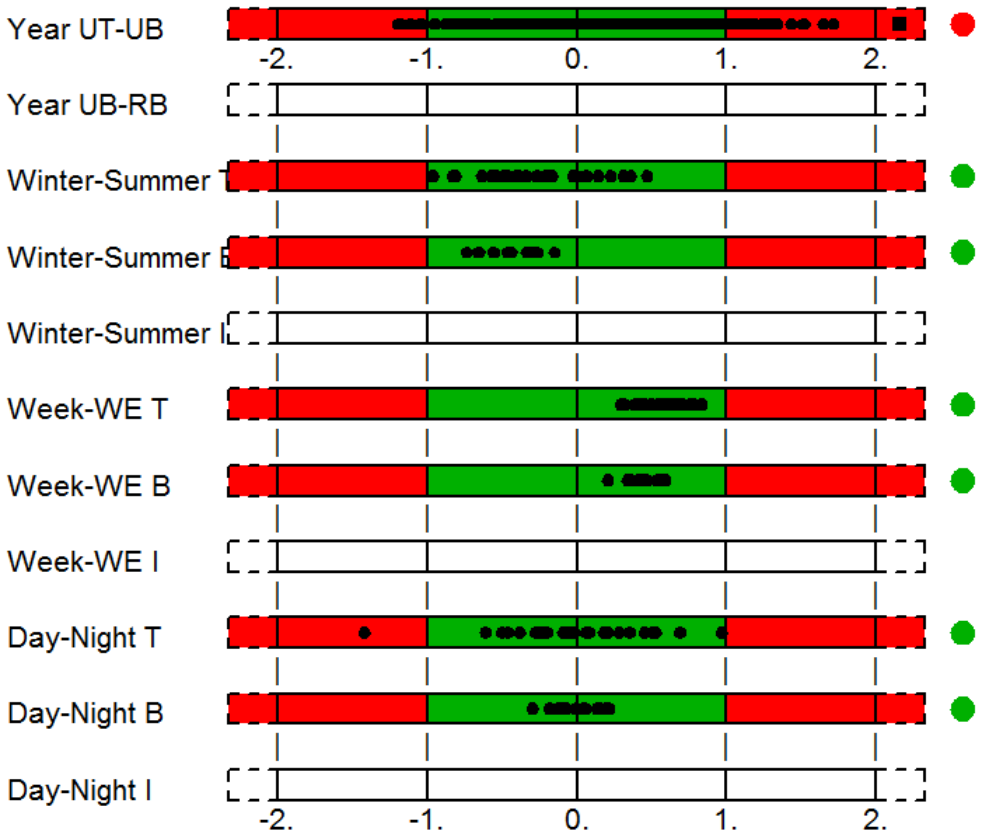


BL0 BQ7 BT4 BX1 GB6 GN3 GR4 GR8 GR9 HI0 HK6 KC1 MY1 RI2 TH4

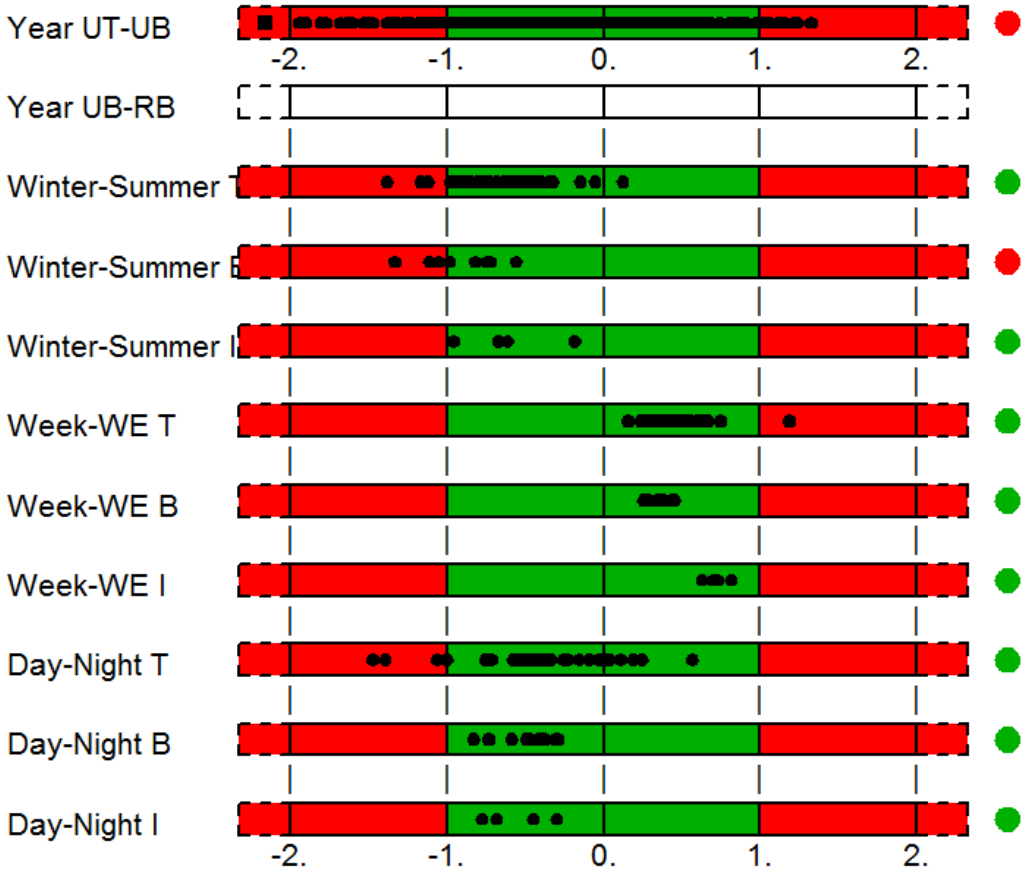
Strt/end Ind: 1-8760
Model (s): Day0
Parameter: O3
Scen: 2018
Extra Values: No
Season: Summer-Winter
Day hours: All 24h
Time Average: 8h
Daily stats: Max

Compare performance: Dynamic evaluation NO2

- Hindcast



- Forecast



Weekend / weekday looks fine according to these metrics

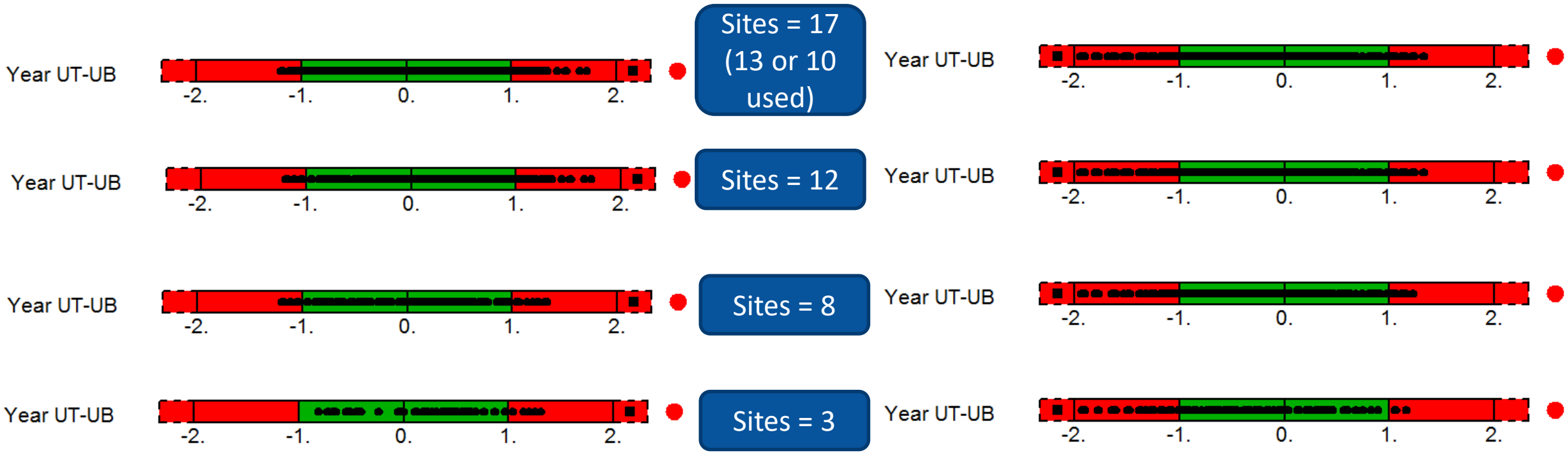
Strt/end Ind: 1-8784
Model (s): ADMS2012
Parameter: NO2
Scen: 2012
Extra Values: 12
Season: Year
Day hours: All 24h
Time Average: Preserve
Daily stats: preserve

Sites considered in spatial metrics = 12

Strt/end Ind: 1-8760
Model (s): Day0
Parameter: NO2
Scen: 2018
Extra Values: 12
Season: Year
Day hours: All 24h
Time Average: Preserve
Daily stats: preserve

Compare performance: Dynamic evaluation NO2

- Hindcast
- Forecast



- It would be useful to understand more about how many sites should be used in the evaluation
- QA/QC document says 'if a choice of 3 is made, three urban background (UB) stations will be used to calculate three gradients with each available urban traffic (UT) station' – but Delta appears to use 1 site less than the threshold

Summary

- Tested the proposed QA/QC protocols for 2 applications, one that we would expect to pass the majority of criteria set, one where we would understand if model performance was not so good due to forecasting inputs
- Target plots behaved as expected: more difficult to satisfy criteria at traffic sites for pollutants strongly influenced by traffic emissions (NO2)
- With the day/night, summer/winter, weekend/weekday plots:
  - O3 day/night values seem very odd
  - Could supplementary results be output to enable users to compare absolute values in addition to differences?
  - It is good practice for model users to be able to explain the results in terms of model formulation and model inputs (e.g. we should look into the weekend / weekday NO2 results)
  - Repeated symbols not helpful
- Dynamic evaluation: the spatial metric may not be robust for large city evaluations where an urban background site in the city centre records a higher concentration than a roadside site in a suburban location
- Evaluation binning according to wind speed or atmospheric stability would also be helpful, for example because some models perform less well in some meteorological conditions

(PM2.5 results not presented due to formatting issues)

Thank you for listening

Any Questions?