COPERNICUS
MARINE ENVIRONMENT MONITORING SERVICE

European Union Programme

Copernicus
Europe's eyes on Earth

MERCATOR OCEAN
OCEAN FORECASTERS
Entrusted entity by mandate of the European Commission

Met Office

Using feature-based verification methods to explore the spatial and temporal characteristics of forecasts of the 2019 Chlorophyll-*a* bloom season in a European regional ocean model

Marion Mittermaier, Rachel North, Jan Maksymczuk and Christine Pequignet

Outline

1. Study objectives
2. Model description
3. Short introduction to MODE/MTD
4. Results
5. Conclusions

Objectives

Work package B:

To understand the skill of CMEMS products in forecasting events or features of interest in space and time; e.g. chlorophyll blooms.

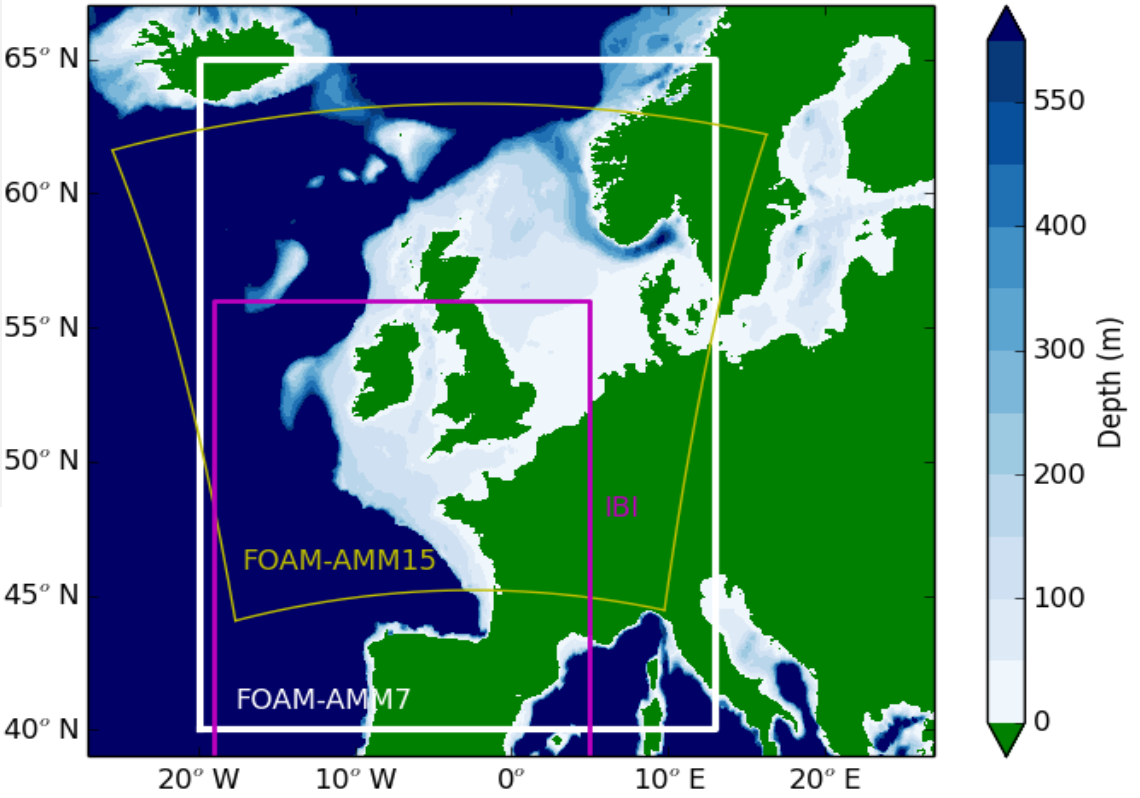
Model Domains

Assessments were made using products from the CMEMS catalogue

- **NWS – AMM7v8-ERSEM (1/10°)**
- IBI (1/36°)
- NWS - AMM15 (1.5km)

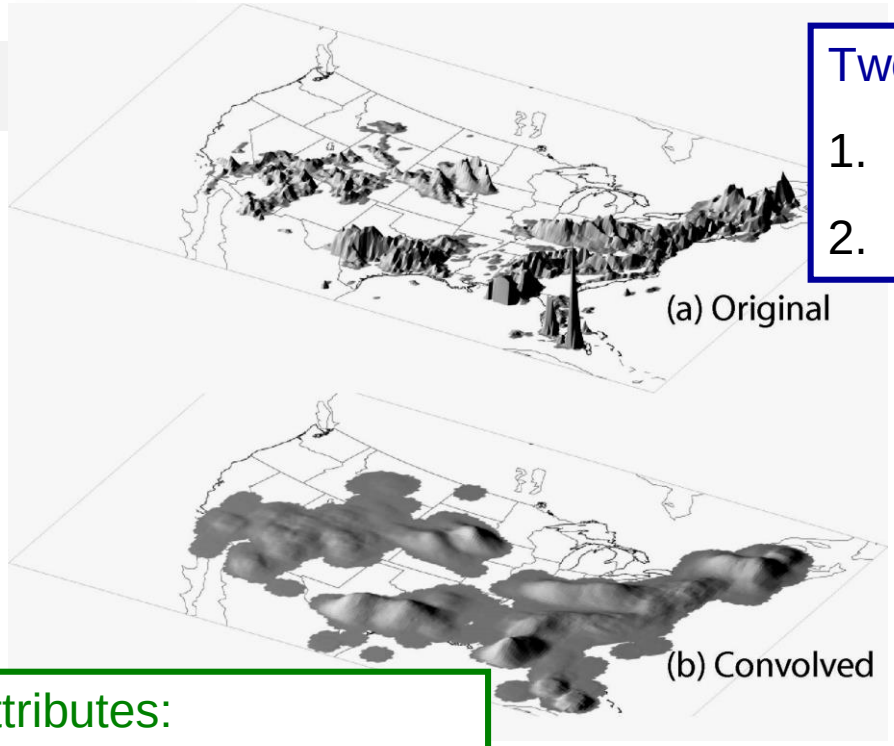
Compared to:

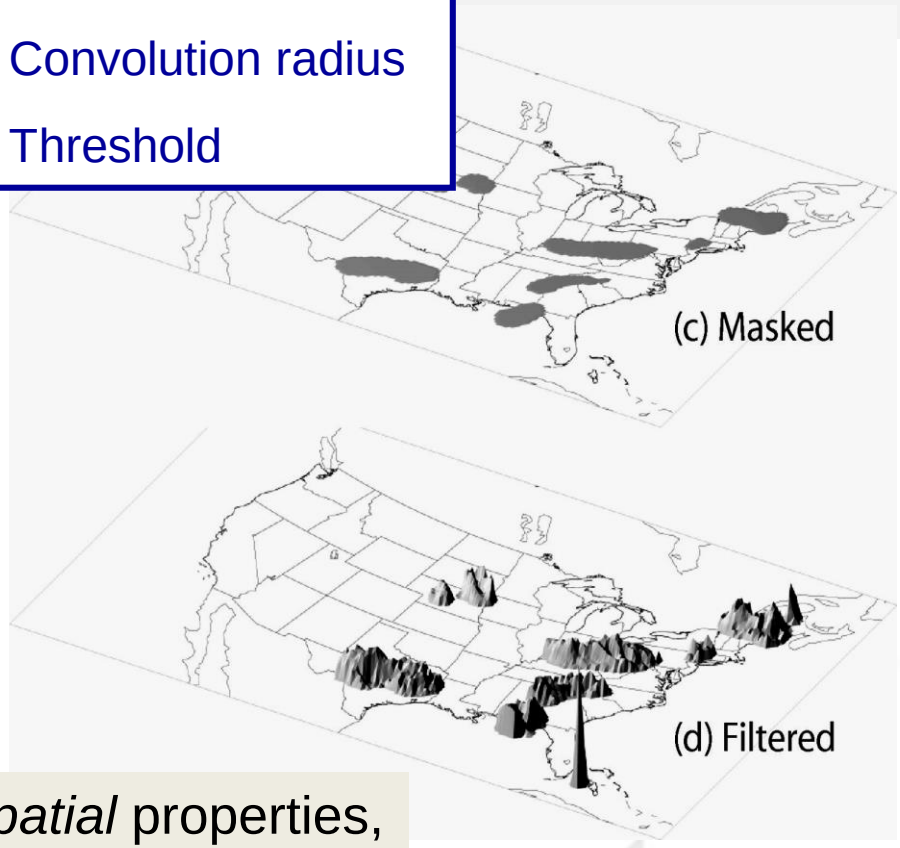
- L4 satellite product and
- AMM7v11 analysis (with BIO DA)



MODE – Method for Object-based Diagnostic Evaluation
Davis et al., *MWR*, 2006

Highly configurable





Two parameters:

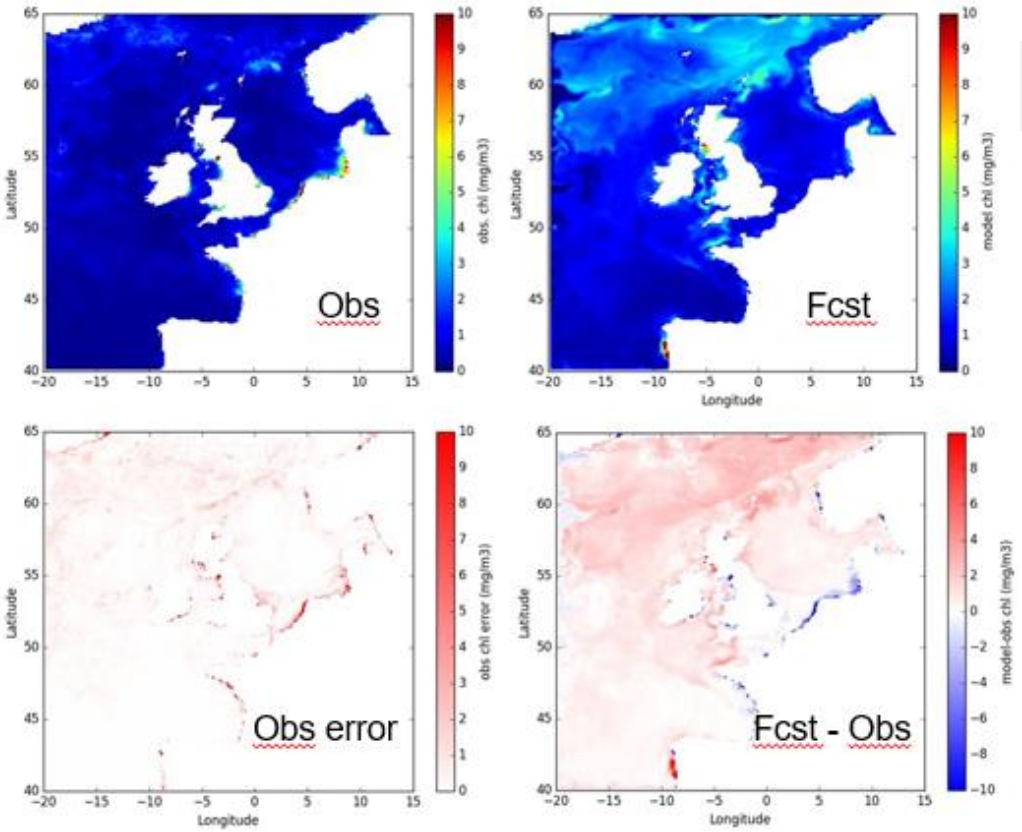
1. Convolution radius
2. Threshold

Attributes:

- Centroid difference,
- Angle difference,
- Area ratio etc

Focus is on *spatial* properties, especially the *spatial* biases

Visual inspection

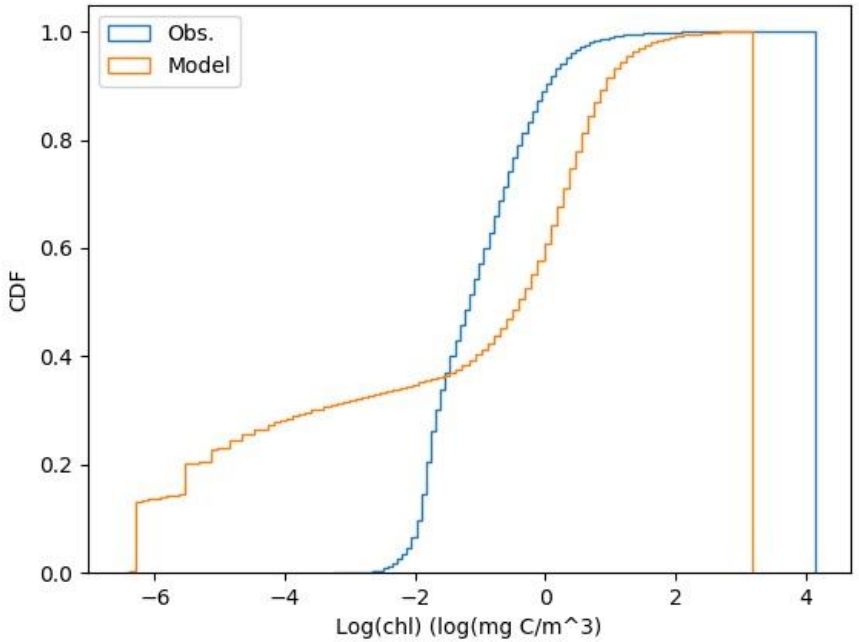


- Any threshold-based method can be sensitive to bias.
- A visual/subjective inspection of the AMM7 analyses and L4 product shows that some biases exist which must be considered during the results analysis. The biases appear to be largest near the coast.

Daily mean L4 multi-sensor regridded observations (left) and AMM7-ERSEM output (right) chlorophyll for 10 July 2018. Bottom: Error estimates on the multi-sensor chlorophyll (left) and difference between model and observations (right).

Bias

- Significant bias between the AMM7v8 forecasts and the L4 satellite product
- The model produces many very low concentrations (at the numerical noise level) which are not observed
- Whilst the shape of the upper half of the forecast CDF shows the same rate of increase, by this stage there are too many forecast concentrations compared to those observed, though with a fairly constant offset
- In the tail, forecast concentrations are underdone compared to those observed
- Due to this bias and the influence it has on identifying objects to be compared across forecasts and observations a decision was taken to minimize the effect of the bias where possible.
- Bias corrected using quantile mapping

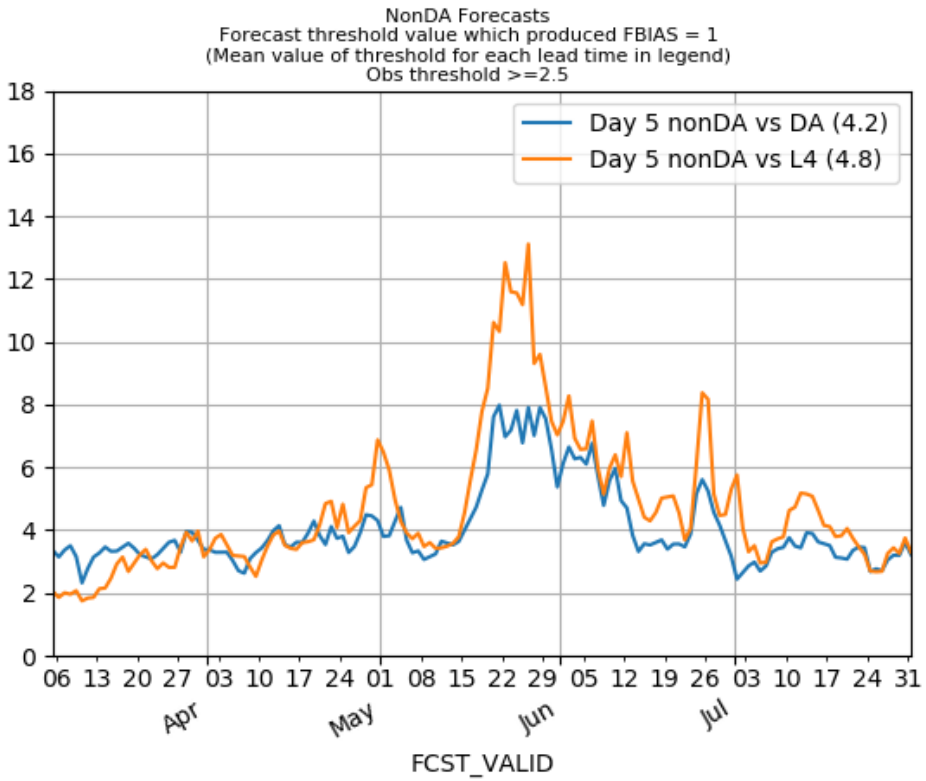


Cumulative distribution of observation and +12h hour forecast log Chlorophyll concentration for the bloom season 2019.

Threshold comparison

Threshold used with AMM7 forecast fields to ensure a frequency bias of 1 (equivalence in proportion of observed threshold exceedances)

- A quantile mapping approach means that whilst observation threshold is fixed the forecast threshold will vary in time to ensure that the frequency bias of the paired fields is equal to one
- This means that the threshold-exceedance seen in the forecasts occurs at the same proportion as that seen in the observations
- There was very little variation with lead time, so only the day 5 forecast data are shown here.
- The AMM7v11 BIO DA analysis is far less biased than the AMM7v8 forecasts and much closer to the L4 product.



Sensitivity analysis

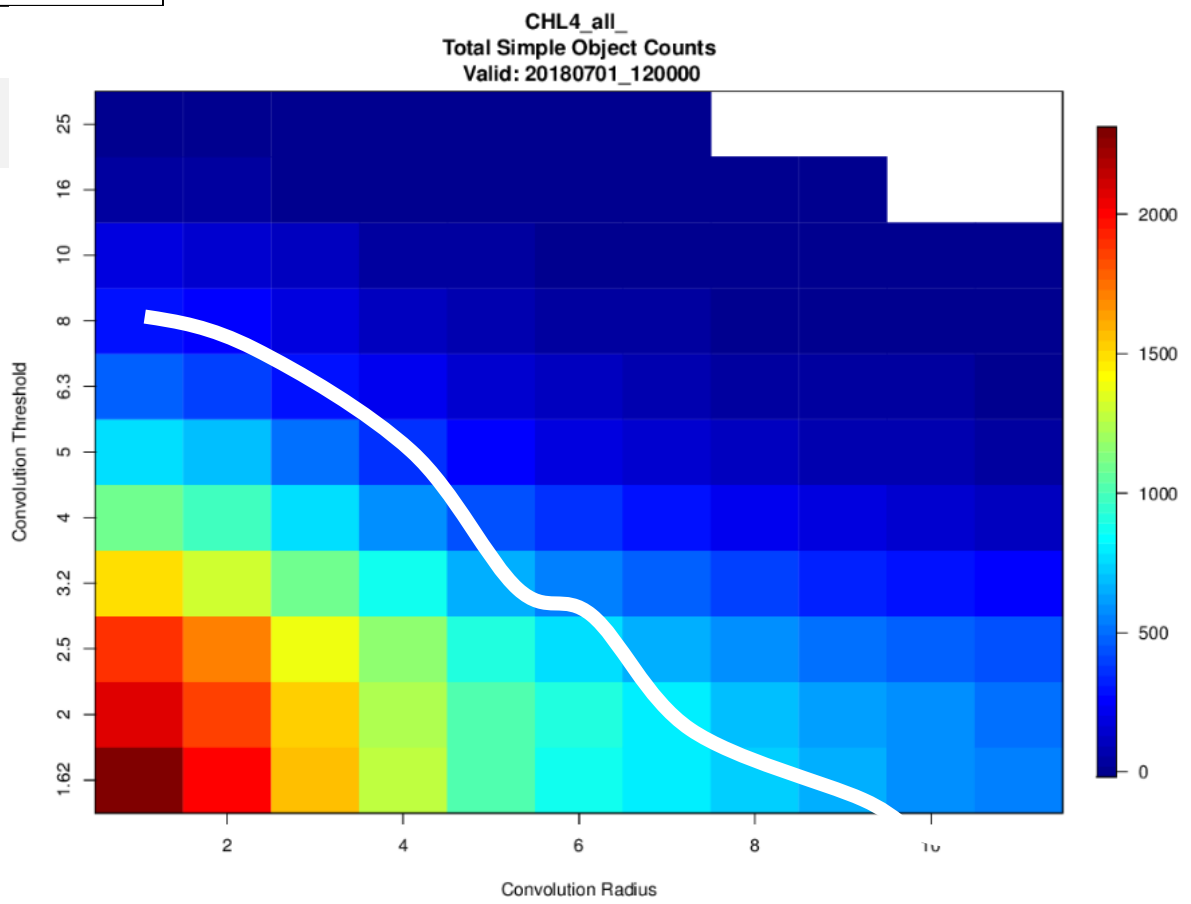
Quilt plots

Explore the relationship of threshold and smoothing radius.

This helps in selecting what the appropriate smoothing radius is for each threshold.

It suggests that for the larger thresholds there are few objects anyway without smoothing so the number of objects may be manageable without smoothing.

For the lowest thresholds you may need to use the highest convolution radius to get the number of objects under control.

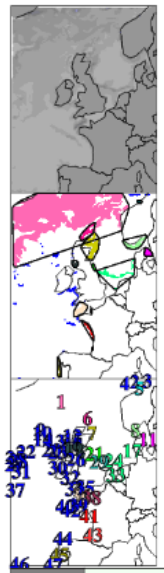


Quilt plot for sensitivity analysis: number of objects identified as a function of convolution radius R (number of grid squares) and threshold T (mg/m^3)

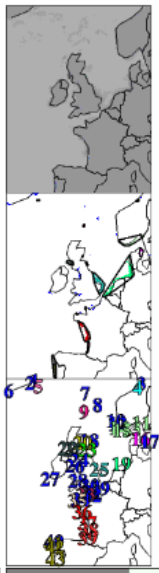
Impact of smoothing radius

MODE: CHL at 0,0,*,* vs CHL at 0,0,*,*

Forecast



Observation



Fcst	Obs	Interest
41	37	0.9562
45	43	0.9542
41	36	0.9538
19	22	0.9404
18	21	0.9120
21	23	0.9107
8	13	0.9004
5	4	0.8953
11	16	0.8884
45	42	0.8740
29	25	0.8725
1	5	0.8696
33	19	0.8615
7	20	0.8545
36	35	0.8486
8	15	0.8292
8	11	0.8267
8	14	0.8149
6	9	0.8041
45	41	0.8035
17	19	0.7961
19	21	0.7627
43	38	0.7617
24	19	0.7277
38	33	0.7154
43	39	0.7107
24	25	0.6970
44	41	0.6772
16	22	0.6759

	Forecast	Observation
Model	AMM7	
Field	CHL	CHL
Level	0,0,*,*	0,0,*,*
Units	mg C/m^3	milligram m-3
Initial	2018 07 22 12:00:00	2018 07 22 00:00:00
Valid	2018 07 22 12:00:00	2018 07 22 00:00:00
Accum	00:00:00	00:00:00

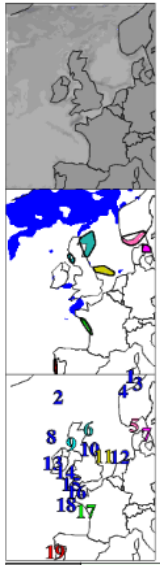
Centroid/Boundary	2.00	4.00
Convex Hull/Angle	0.00	1.00
Aspect/Area	0.00	1.00
Int Area/Curvature	2.00	0.00
Complexity/Intensity	0.00	0.00
Total Interest Thresh	0.70	

	Forecast	Observation
Mask M/G/P	on/off/off	on/off/off
Conv Radius	1	1
Conv Thresh	>=1.62	>=1.62
Obj Filters	1	1
Inten Perc	p50	
Merge Thresh	NA	NA
Merging	none	none
Matching	match/merge	
Simple/M/U	47/18/29	43/23/20
Area	18565	1767
Area M/U	17431/1134	1346/421
Cluster	15	15
MMI	0.6348	0.7154
MMI (F+O)	0.6595	

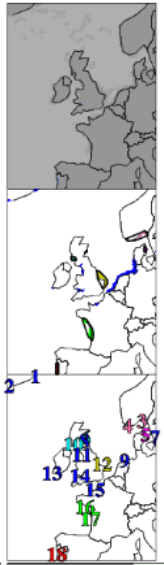
(a)

MODE: CHL at 0,0,*,* vs CHL at 0,0,*,*

Forecast



Observation



Fcst	Obs	Interest
19	18	1.0000
17	17	0.9778
17	16	0.9520
9	10	0.9486
5	3	0.9043
7	5	0.8761
6	6	0.8604
5	4	0.8428
11	12	0.7534
9	11	0.6851
15	14	0.6707
14	14	0.6416
12	9	0.6286
11	9	0.6228
7	7	0.5841
13	13	0.5715
14	13	0.5587
6	8	0.5541
10	11	0.5337
10	12	0.5337
10	8	0.5259
16	15	0.5191
7	3	0.4958
16	16	0.4911
16	14	0.4888
10	6	0.4848
5	5	0.4840
15	15	0.4648
9	8	0.4502

	Forecast	Observation
Model	AMM7	
Field	CHL	CHL
Level	0,0,*,*	0,0,*,*
Units	mg C/m^3	milligram m-3
Initial	2018 07 22 12:00:00	2018 07 22 00:00:00
Valid	2018 07 22 12:00:00	2018 07 22 00:00:00
Accum	00:00:00	00:00:00

Centroid/Boundary	2.00	4.00
Convex Hull/Angle	0.00	1.00
Aspect/Area	0.00	1.00
Int Area/Curvature	2.00	0.00
Complexity/Intensity	0.00	0.00
Total Interest Thresh	0.70	

	Forecast	Observation
Mask M/G/P	on/off/off	on/off/off
Conv Radius	6	6
Conv Thresh	>=1.62	>=1.62
Obj Filters	1	1
Inten Perc	p50	
Merge Thresh	NA	NA
Merging	none	none
Matching	match/merge	
Simple/M/U	19/7/12	18/9/9
Area	18930	1777
Area M/U	2239/16691	787/990
Cluster	7	7
MMI	0.6286	0.7192
MMI (F+O)	0.6707	

(b)

To be able to make a sensible analysis one has to find the balance between threshold and smoothing. Too many objects makes it difficult to analyse. Too much smoothing may mean events become less distinct.

(a) Shows no smoothing for the lowest threshold of 1.62 $mg.m^{-3}$, showing a large number (too many) of objects.

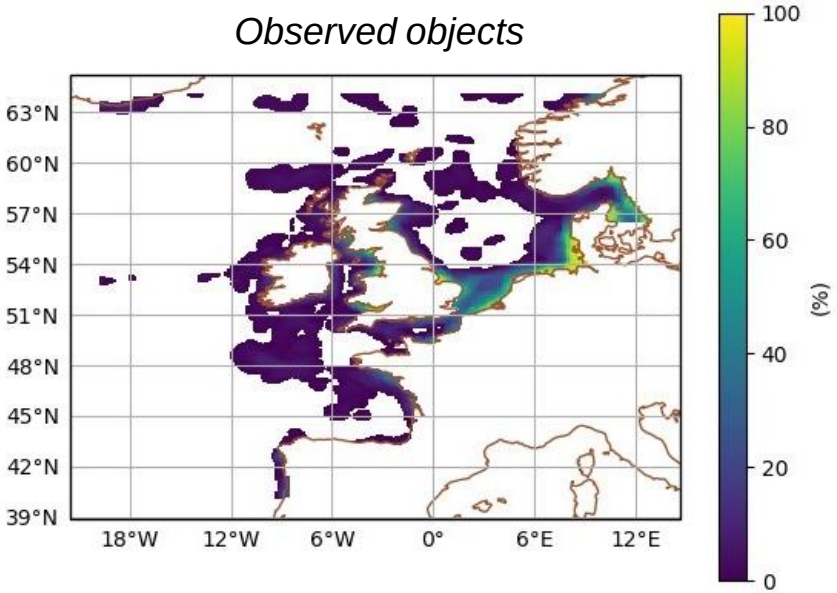
(b) A convolution radius of 6 is applied, reducing the number of objects. This is becoming more manageable.

Seasonal composite

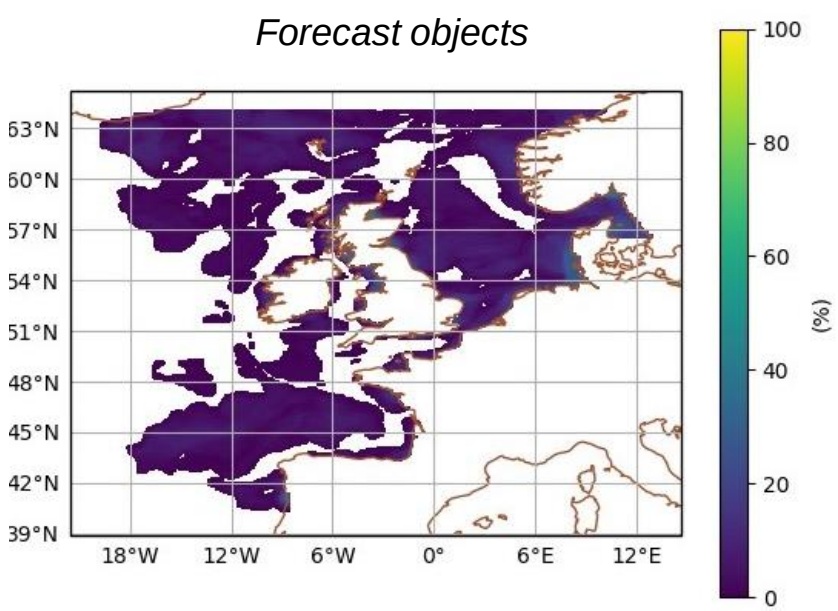
- Composite spatial coverage of objects identified through the 2019 bloom season for both observed and forecast objects.

- The maps show the proportion of time (in days) that an object occurs at that grid point.

Observed objects

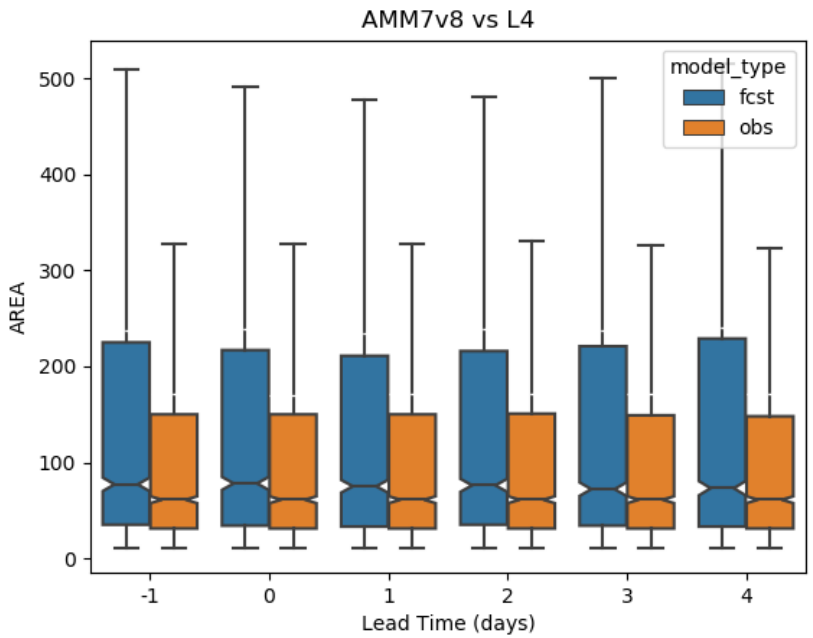


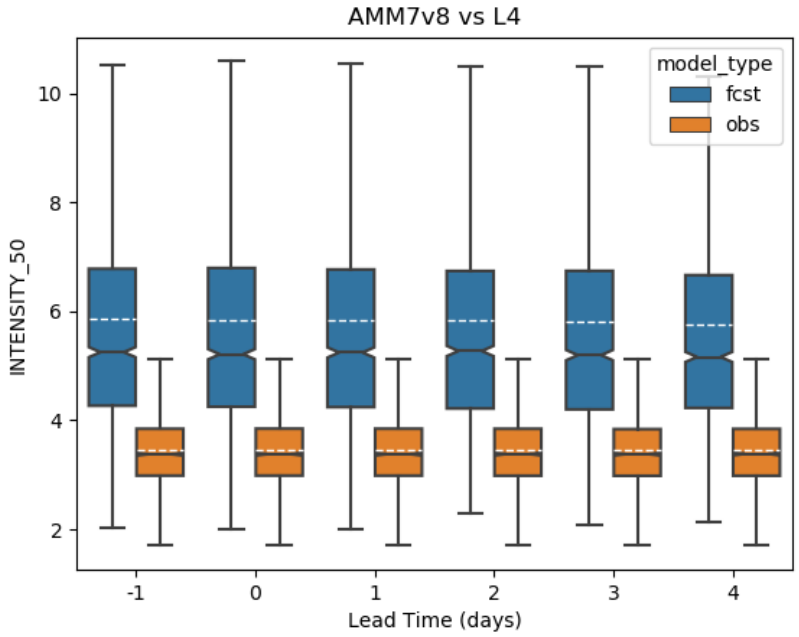
Forecast objects



Percentage of objects identified over the season

Object attributes



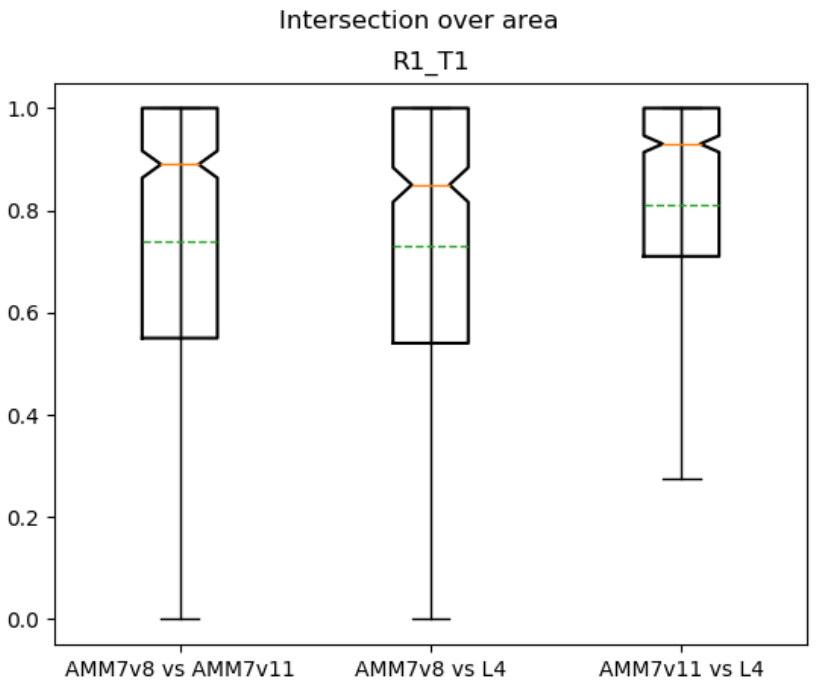


The object areas in grid squares (right) show the distribution of 50th (median) percentile values from all the identified objects in the period.

- There is very little variation with lead time
- AMM7v8 forecasts have a broader distribution in size and are bigger

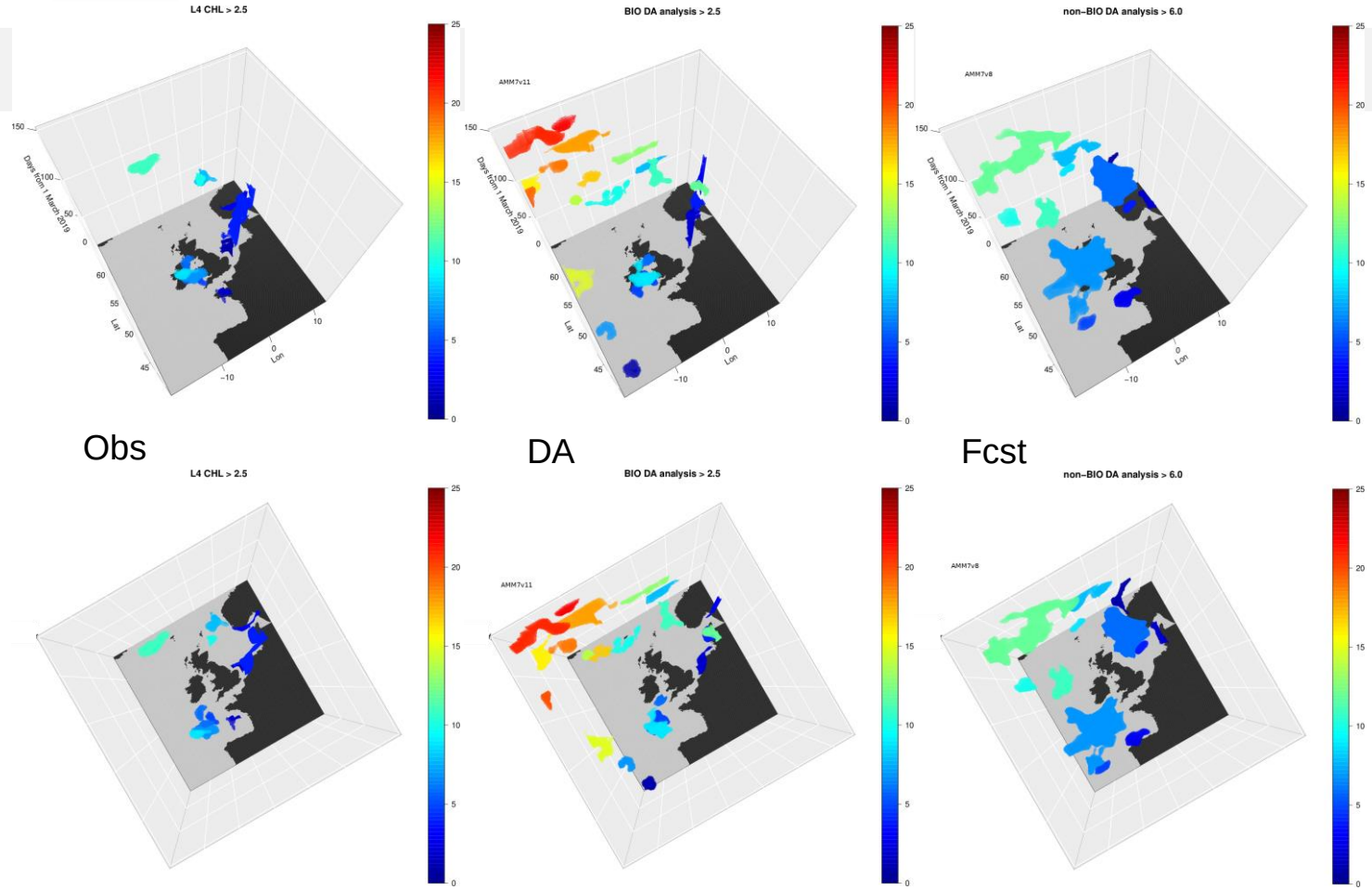
Paired (matched) object attributes

- The intersection-over-area gives a measure of how much the paired forecast-observed objects overlap in space
- If the objects do not intersect, this metric is 0
- Here many of the matched areas overlap perfectly (it is easy for smaller L4 areas to be completely enveloped by the model analyses).
- However, there is a very long whisker which shows that there are instances where this is 0.
- It is clear that the AMM7v11 BIO DA analysis is closest to the L4 product, with all pairs overlapping in some way. There is quite a difference between the median (notch) and the mean (dashed line).



Paired object attributes for the Day 0 results at the 2.5 mg.m^{-3} threshold and a smoothing radius of 5 grid squares. Ratio of the intersection area over the largest of the forecast or observed object area.

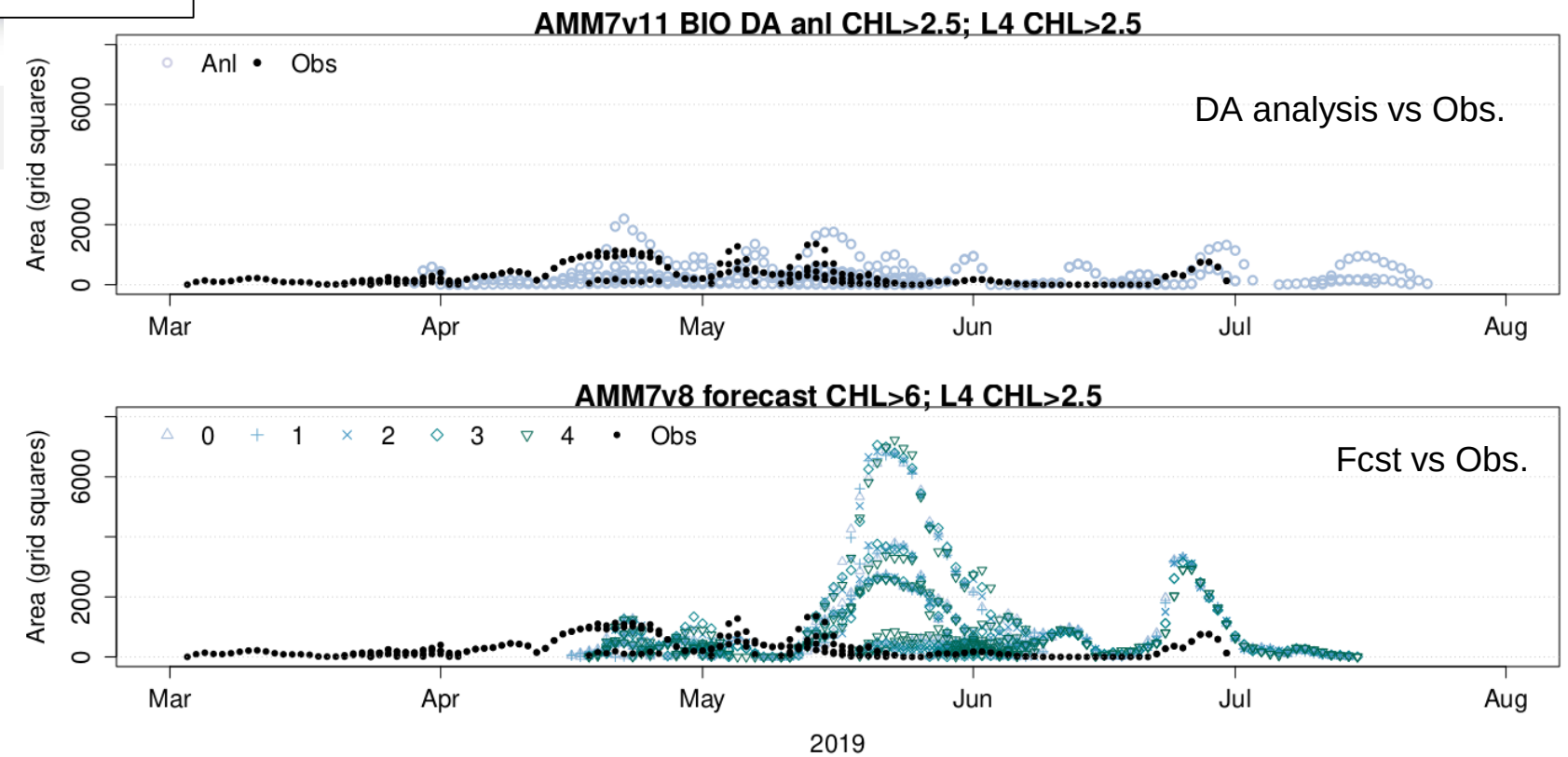
Space-time objects from MTD



Temporal evolution of identified chlorophyll space-time objects. Colours represent the object numbers, which increase with time. Thresholds in mg.m^{-3}.

Temporal evolution

Time series of all identified MTD object areas.



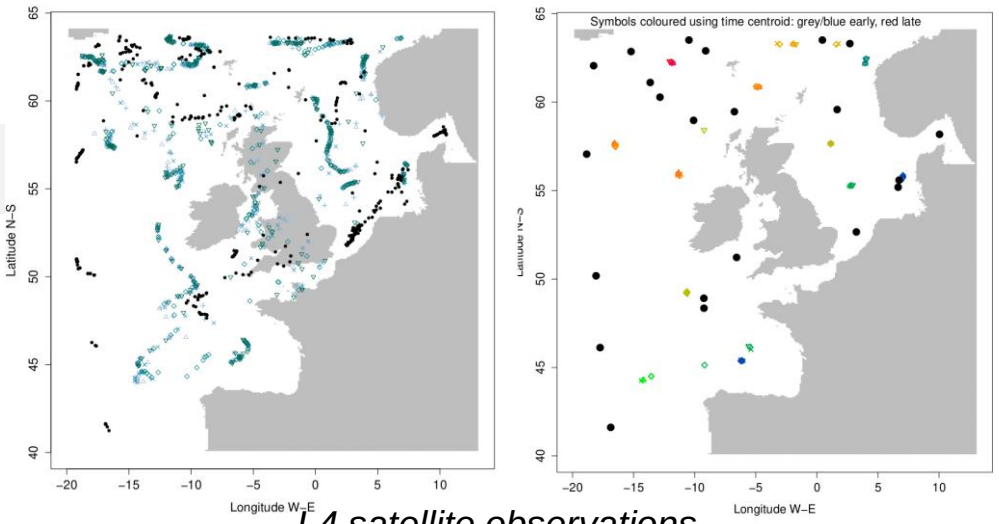
- AMM7v8 forecasts only picked up the first event of the season in April 2019, almost a month after the first identified chlorophyll object was identified in the L4 product
- The overlap between the different forecast lead times indicates that there is very little difference in the forecasts as a function of lead time
- Model also struggles to capture the end of the bloom season, stopping too soon by at least a week

MTD centroids

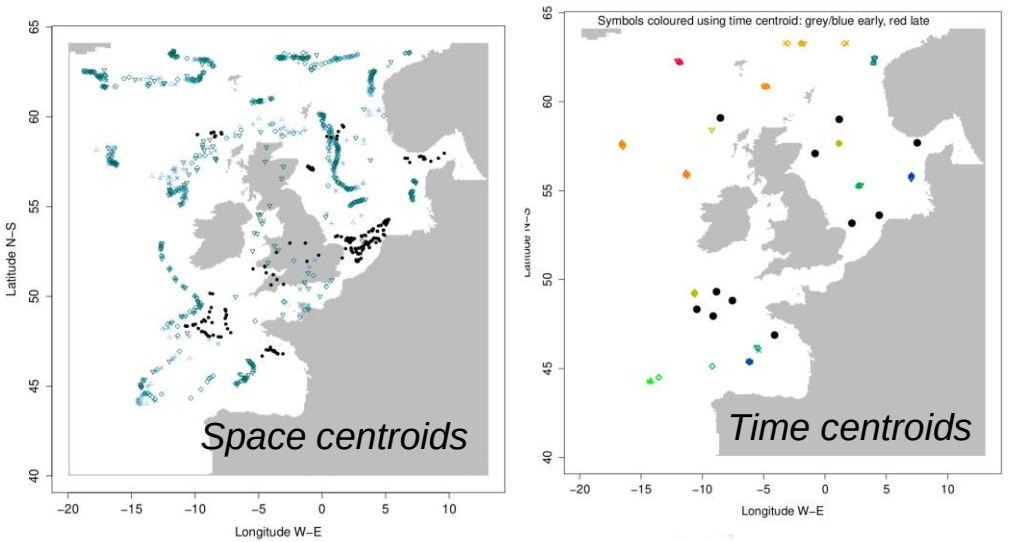
- The black dots indicate (space or time) observed centroid.
- Differences between the AMM7v11 BIO DA analysis and the L4 product, especially in the southern North Sea and also in the north and west.
- There are some areas/times where black dots and forecast centroids can be found in in reasonable spatial proximity (though this may not indicate temporal proximity).

- Time centroid is derived from spatial centroids.
- The forecast time centroids for the different lead times on top of each other showing there is no change with lead time.
- Impact of DA analysis compared to L4 product is evident in the observed centroids, with the AMM7v11 BIO DA analysis producing many more objects in deeper waters to the north and west, which are not evident in L4.

AMM7v11 BIO DA analysis



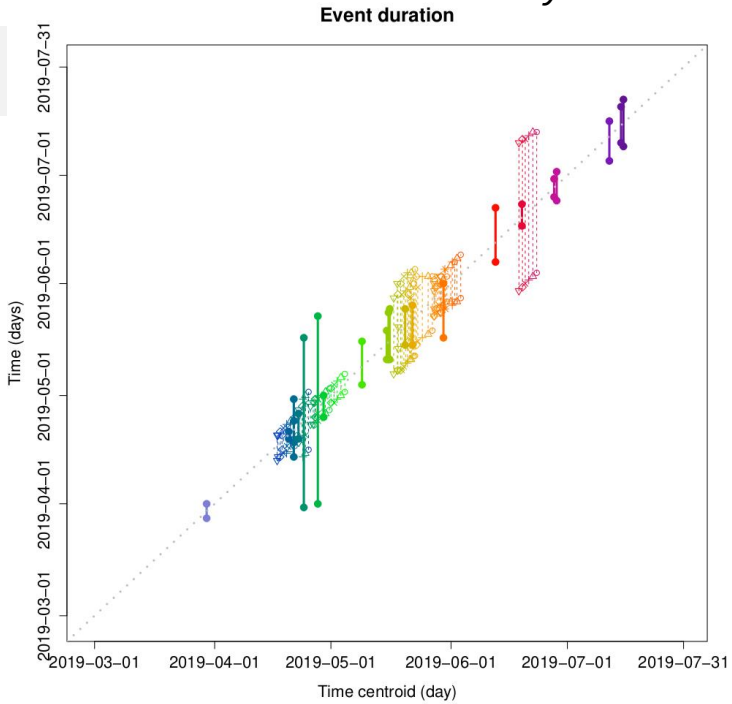
L4 satellite observations



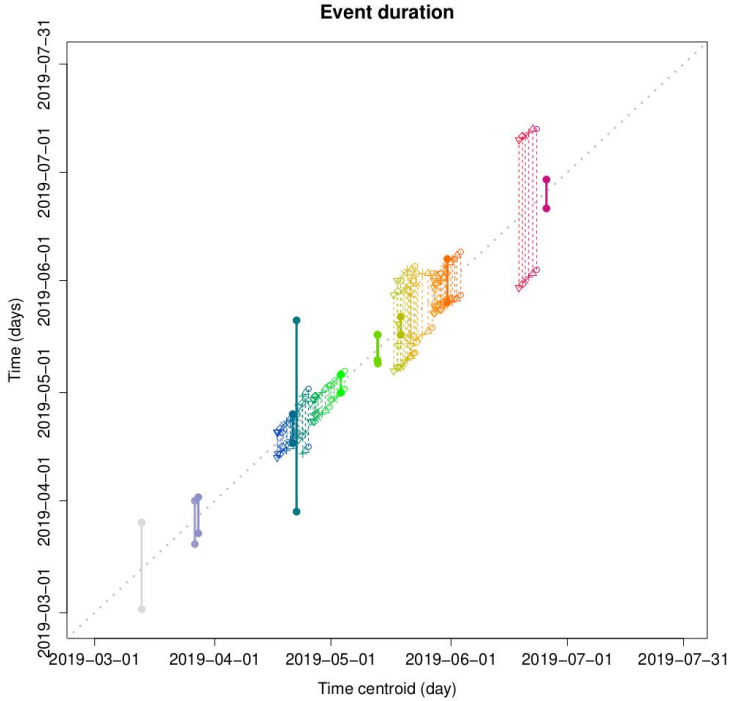
Composites of the entire 2019 season

Duration of time objects

AMM7v11 BIO DA analysis



L4 satellite observations



- The x-axis represents elapsed time, vertical lines and y-axis gives the duration of the objects centred on the time centroid.
- Solid lines represent the observed events whereas dashed lines are the forecast events
- Overall, most groups of forecast objects have some association with an observed object around about the same time (though this does not mean they are close in space)

Conclusions

Bias –forecast biased compared to the observations, which must be mitigated against before using MODE or MTD; bias is improved with DA

- *Timing issues* –initial onset of the bloom is almost a month late (25 days) in AMM7v8. Predicting the onset of the bloom seems problematic.
- *Location* – model does produce chlorophyll objects (blooms) in the right areas, but not necessarily at the right time.
- *Evolution with lead time* – there is very little change, suggesting little benefit in having longer lead time forecasts.
- *Number and size of objects* – too few blooms that are too large. Many of the coastal objects identified in the L4 product cannot be resolved by the model due to the coarseness of the coastline in the 7 km model. This situation would improve with increase in resolution.
- *Benefit of AMM7v11 BIO DA* – it is likely that forecast bias will be improved when initialised from the BIO DA.

Questions and discussion

Using feature-based verification methods to explore the spatial and temporal characteristics of forecasts of the 2019 Chlorophyll-*a* bloom season over the European North-West Shelf

Marion Mittermaier[1], Rachel North[1], Jan Maksymczuk[2], Christine Pequignet[2], David Ford[2]

In preparation for *Ocean Science.*