

Further Development of the Automated Data Quality Control Scheme with the application for Ocean Climatologies

Viktor Gouretski

Center for Earth System Research and Sustainability
University of Hamburg

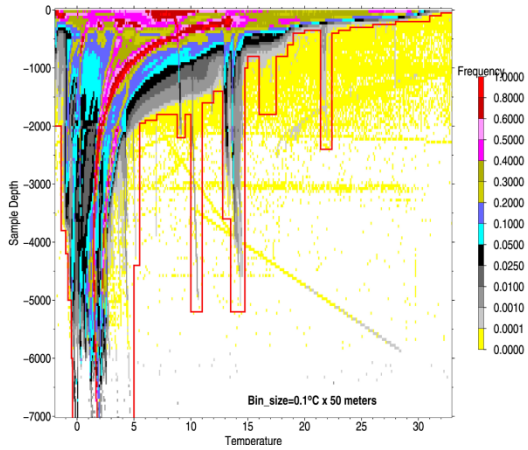


ICDC Auto QC procedure

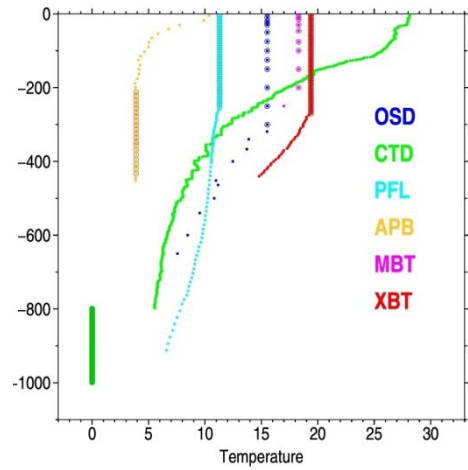
- Applied now both to temperature and salinity profiles from the WOD13
- A suite of quality checks:
 - 1) Overall checks, e.g. the threshold parameters being globally valid
 - 2) Local checks
- AQC-Result Diagnostics

Quality Checks: Temperature

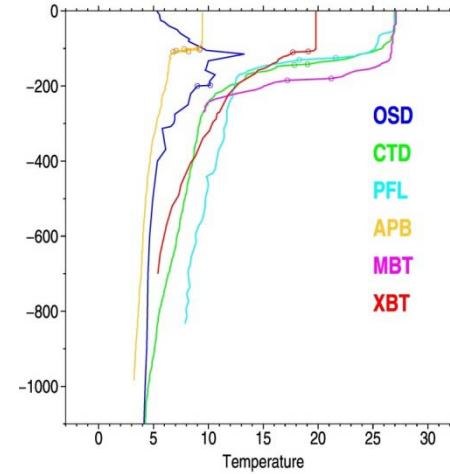
Crude_range



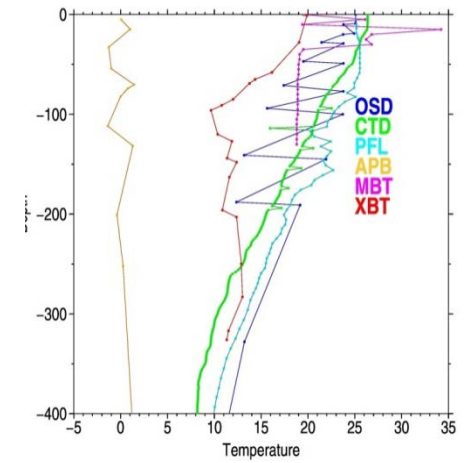
Stuck_value



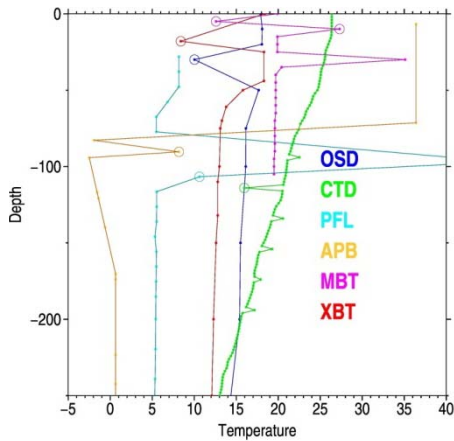
Vertical_gradient



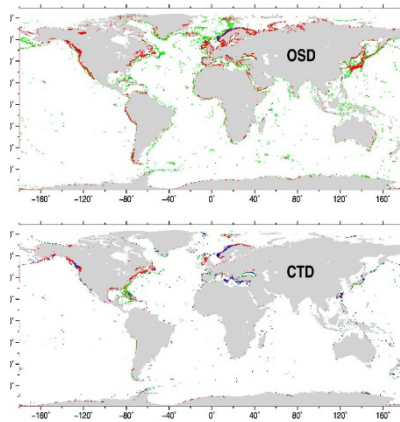
Number_of_extrema



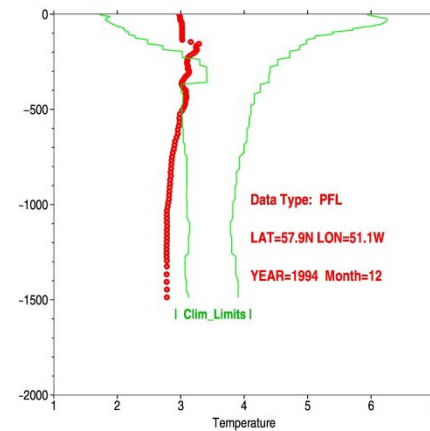
Spikes



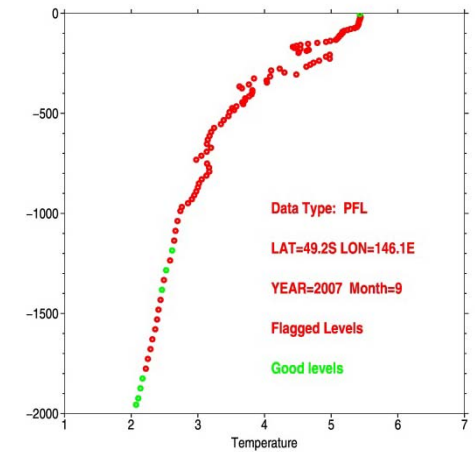
Obs_level_depth



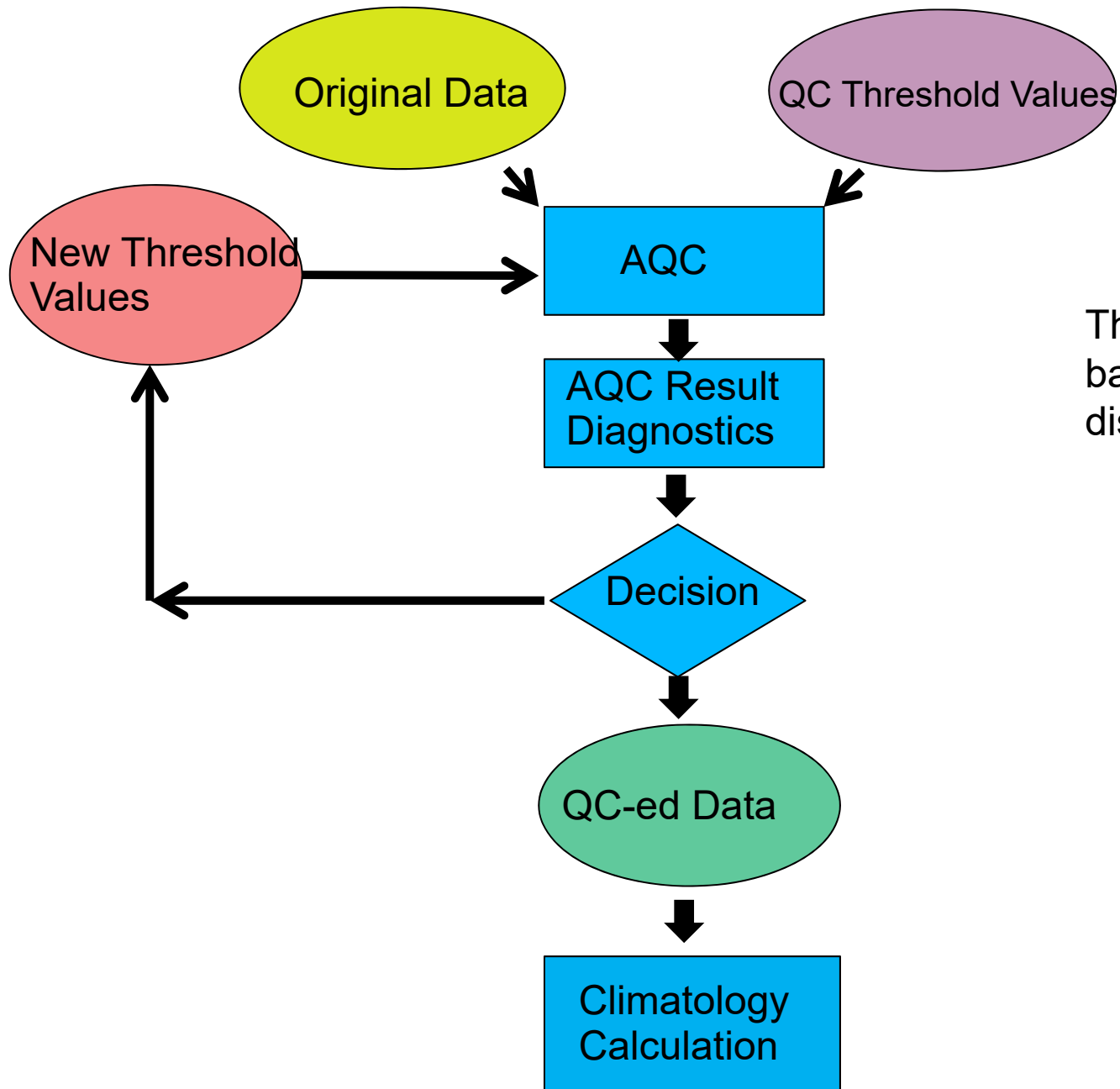
Climatological_range



Profile_gray_list



AQC Flow Chart



Threshold values are based on respective frequency distributions

Calculation of climatological temperature ranges for the Global Ocean

- Each T-observation $T(\text{Lat}, \text{Lon}, \text{Depth}, \text{Month})$ is checked against the local climatological T-range $[T_{\min}; T_{\max}]$
- Only preliminary qc-ed OSD, CTD, and PFL profiles contribute
- Grid: 0.5LAT x 0.5LON
- Variable influence radius (100 – 500 km) depending on the data abundance. Target $N = 300$
- Acceptable T-range is defined as:

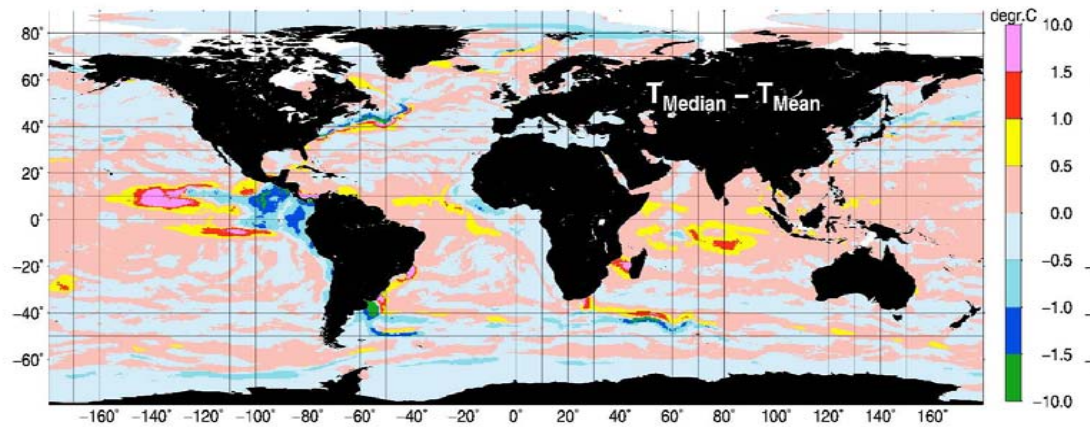
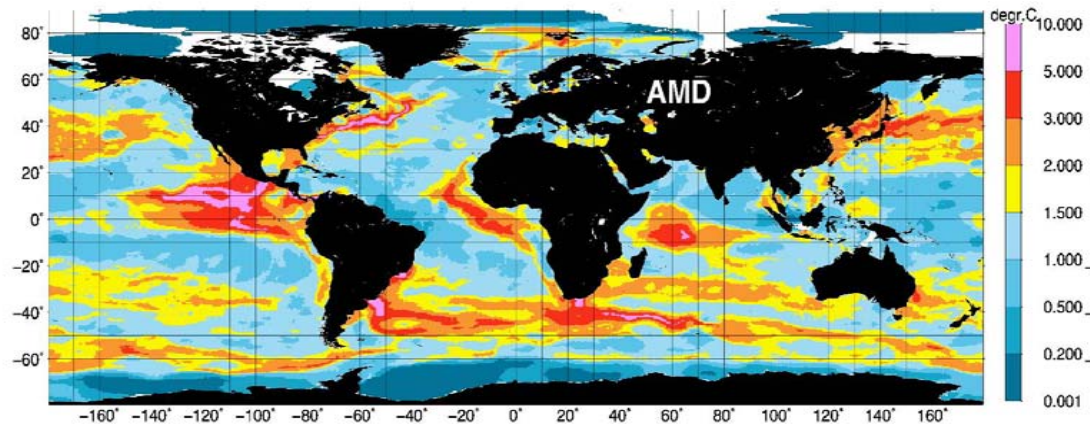
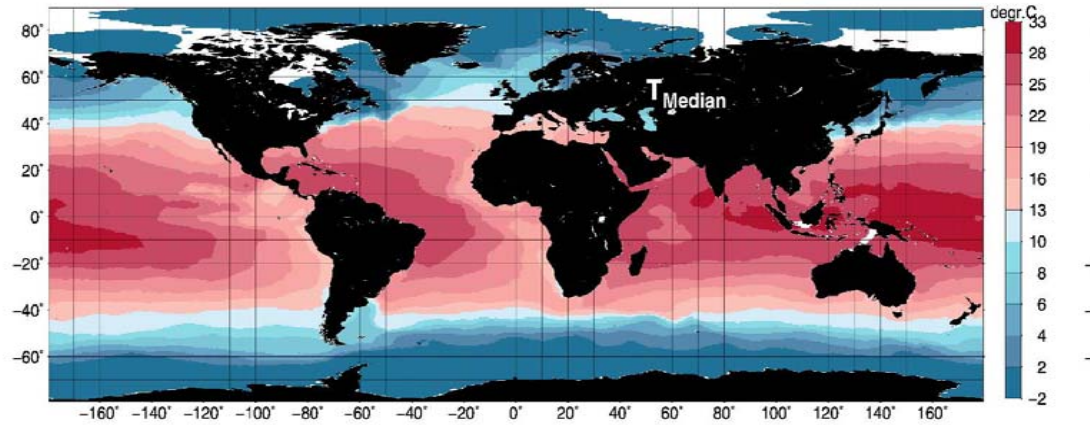
$$T_{\text{Min}} = T_{\text{MEDIAN}} - 2.5 * \text{AMD}$$

$$T_{\text{MAX}} = T_{\text{MEDIAN}} + 2.5 * \text{AMD}$$

- Monthly ranges above 500 m

Climatological Ranges

DEC+JAN+JAN Level: 55m Data_Types: OSD+CTD+PFL Grid: 0.5°LATx0.5°LON



Changes introduced since Dec 2015:

- Minimum Temperature (Crude range) check added: T is compared with T_{freezing}

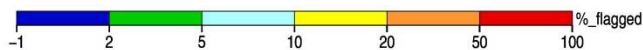
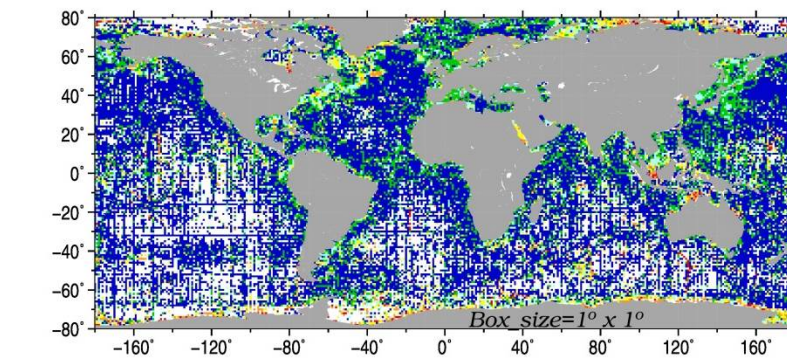
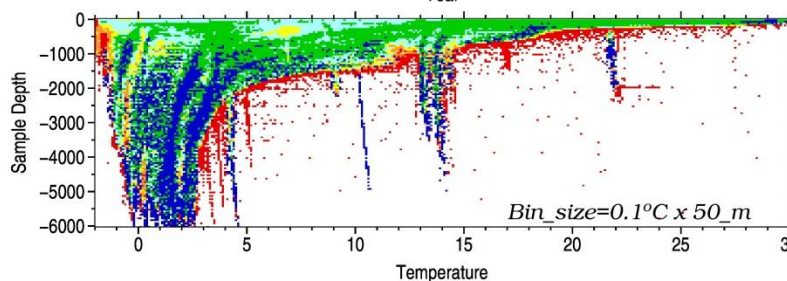
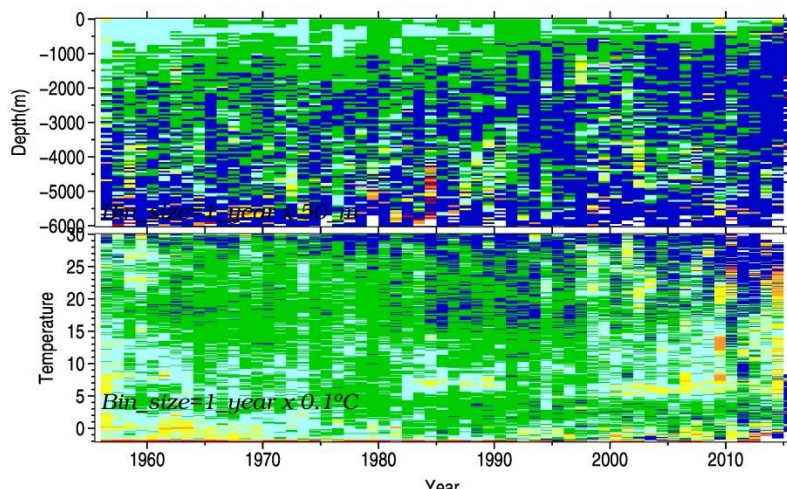
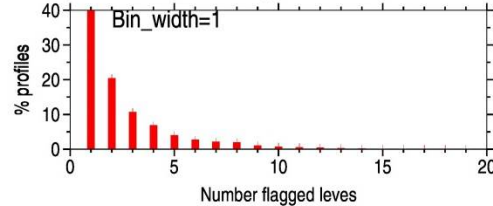
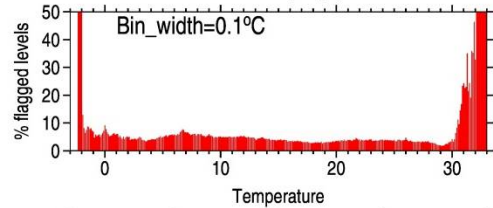
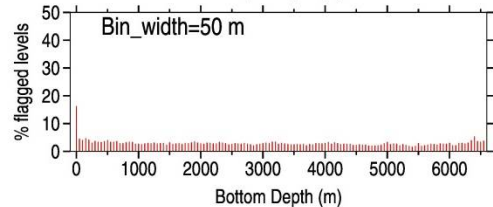
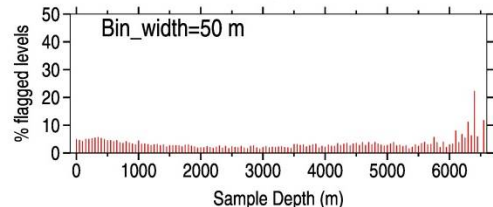
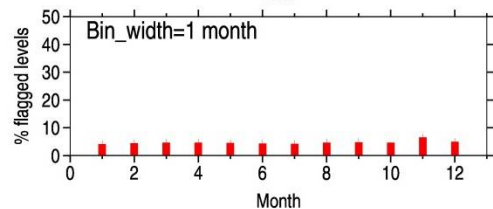
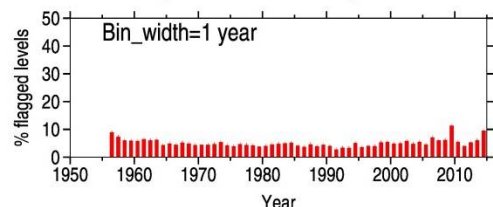
$$T_{\text{freezing}} = -0.0575 * S + 1.710523 * 10^{-3} * S^{3/2} - 2.154996 * S^2$$

(S=36PSU)

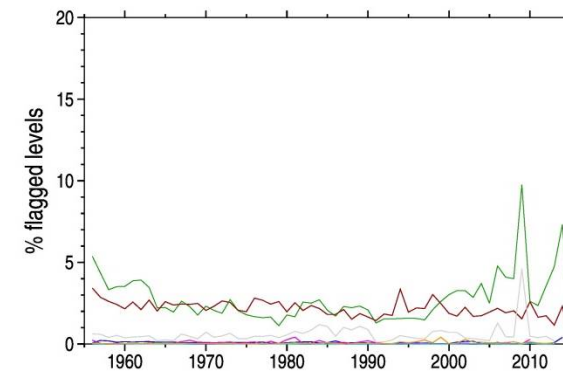
- Modified last sample depth check
- New local climatological ranges
- New set of standard Levels
- Additional datatype criteria

AQC Diagnostics: OSD Temperature Profiles

Quality Check Diagnostics for the OSD Dataset



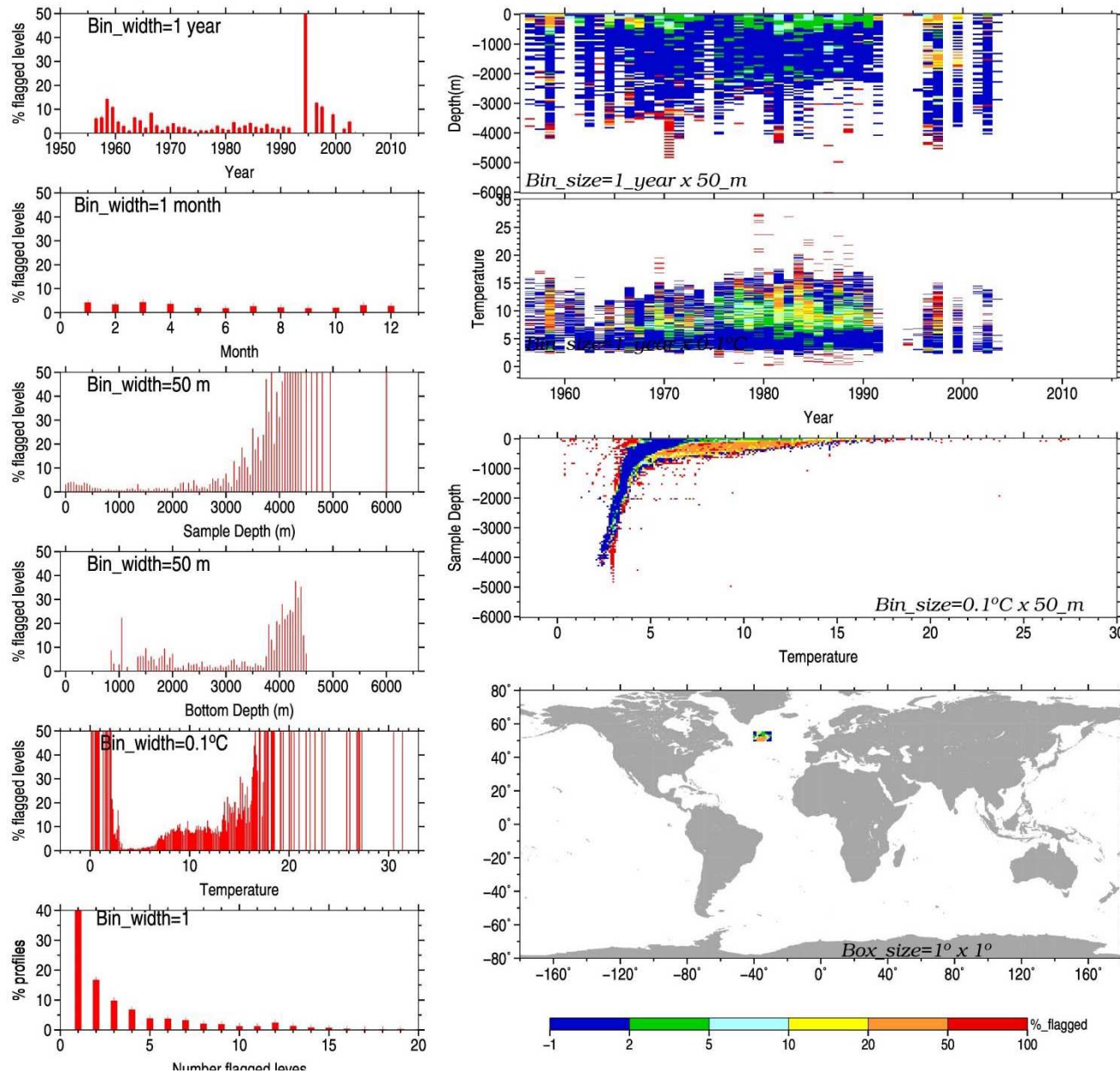
| | |
|---------------------------------------|-----------|
| Number non-dummy levels | 22372821 |
| Number / % dummy levels | 0 / 0.000 |
| Percent Flagged levels | 4.65 |
| Total Profiles | 2023214 |
| Percent profiles >= one flagged level | 17.58 |
| Overall Limits: | -2.0 33.0 |



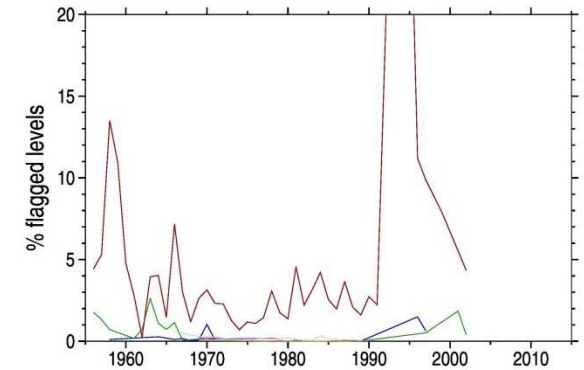
| QC-Check | % flagged levels |
|-----------------------------------|------------------|
| Sample_depth_order (1) | 0.000 |
| Crude-T_range (2) | 0.090 |
| Last_Level_Depth_vs_GEBCO (3) | 2.276 |
| Datatype_Max_Obs_Depth (4) | 0.000 |
| Stuck_Value (5) | 0.004 |
| Number_of_T-extremuma (6) | 0.097 |
| Spikes (7) | 0.009 |
| Overall_DT/DZ_Range (8) | 0.051 |
| Local_Climatological_T-range_ (9) | 2.216 |
| Flagged_Level_Percentage (10) | 0.602 |

AQC Diagnostics: OSD Temperature Profiles (selected area)

Quality Check Diagnostics for the OSD Dataset



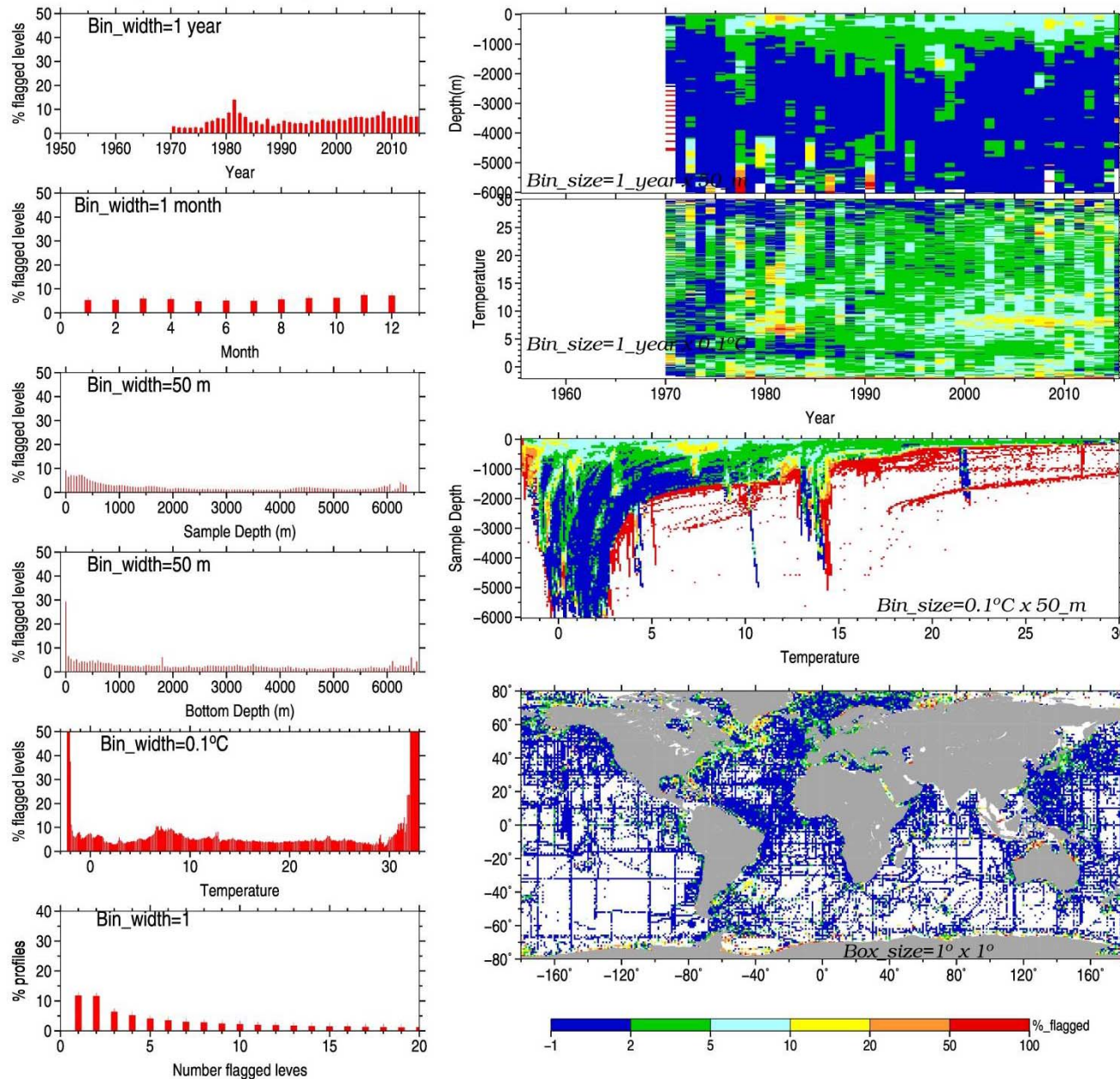
| | |
|---------------------------------------|-----------|
| Number non-dummy levels | 553367 |
| Number / % dummy levels | 0 / 0.000 |
| Percent Flagged levels | 2.67 |
| Total Profiles | 29232 |
| Percent profiles >= one flagged level | 13.65 |
| Overall Limits: | -2.0 33.0 |



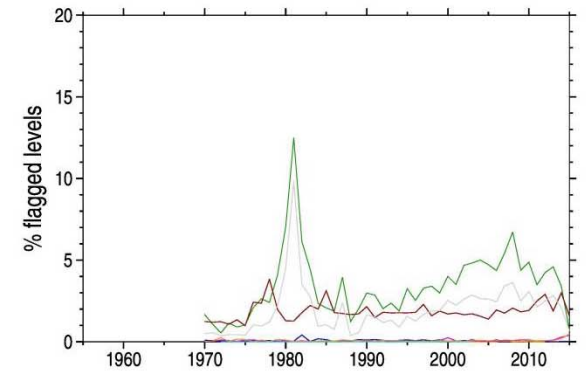
| QC-Check | % flagged levels |
|----------------------------------|------------------|
| Sample_depth_order (1) | 0.000 |
| Crude-T_range (2) | 0.031 |
| Last_Level_Depth_vs_GEBCO (3) | 0.025 |
| Datatype_Max_Obs_Depth (4) | 0.000 |
| Stuck_Value (5) | 0.000 |
| Number_of_T-extremuma (6) | 0.013 |
| Spikes (7) | 0.005 |
| Overall_DT/DZ_Range (8) | 0.012 |
| Local_Climatological_T-range (9) | 2.609 |
| Flagged_level_Percentage (10) | 0.120 |

AQC Diagnostics: CTD Temperature Profiles

Quality Check Diagnostics for the CTD Dataset



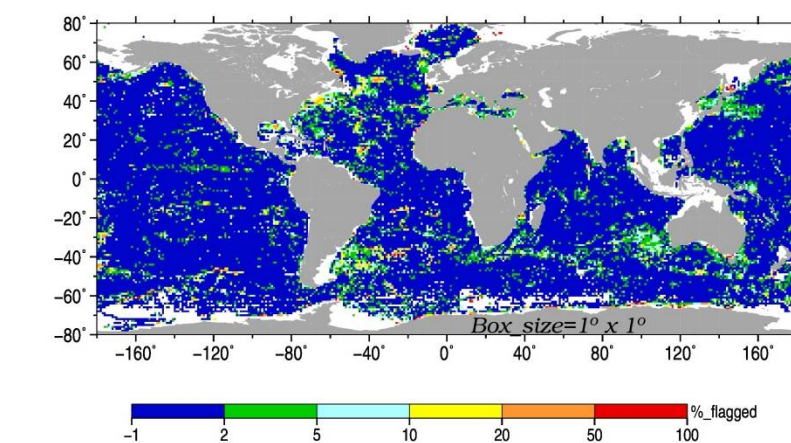
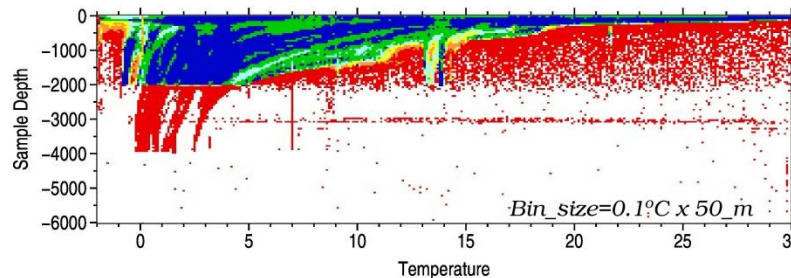
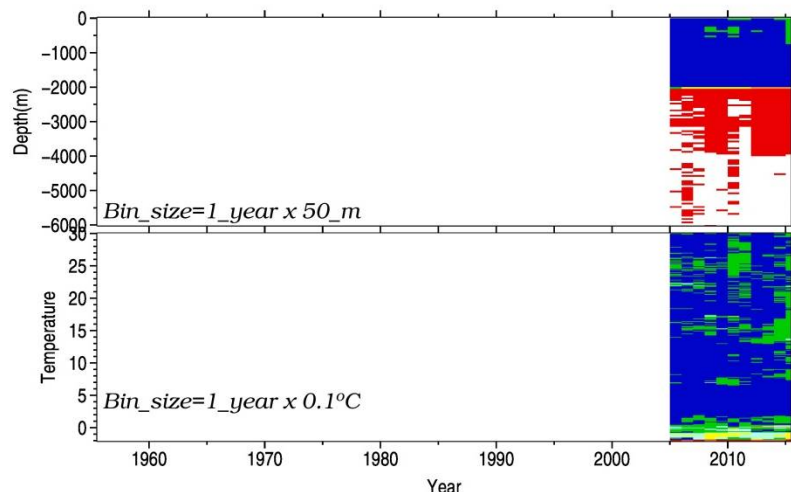
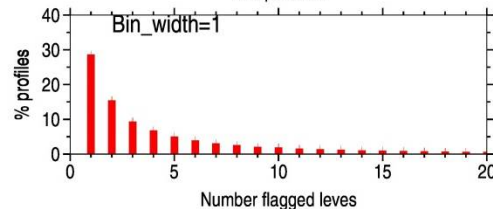
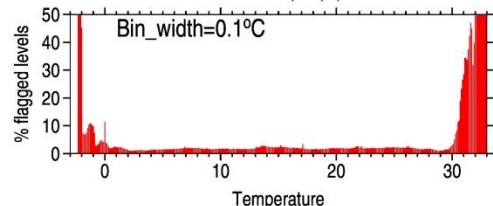
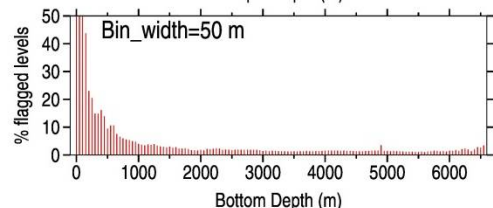
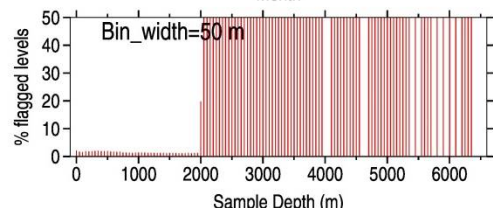
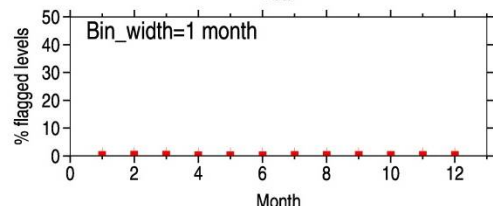
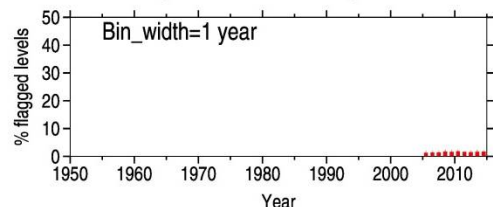
| | |
|---------------------------------------|-----------|
| Number non-dummy levels | 92272502 |
| Number / % dummy levels | 0 / 0.000 |
| Percent Flagged levels | 5.65 |
| Total Profiles | 928168 |
| Percent profiles >= one flagged level | 23.25 |
| Overall Limits: | -2.0 33.0 |



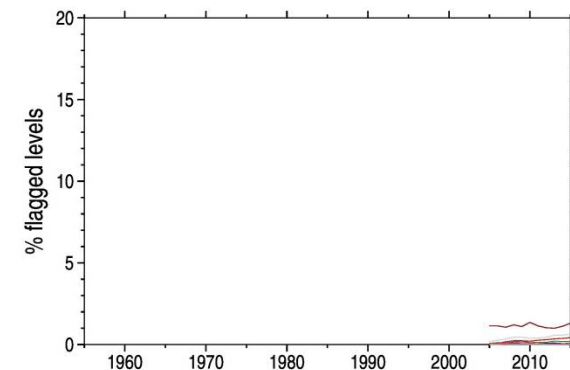
| QC-Check | % flagged levels |
|----------------------------------|------------------|
| Sample_depth_order (1) | 0.000 |
| Crude-T_range (2) | 0.074 |
| Last_Level_Depth_vs_GEBCO (3) | 3.599 |
| Datatype_Max_Obs_Depth (4) | 0.000 |
| Stuck_Value (5) | 0.007 |
| Number_of_T-extremuma (6) | 0.056 |
| Spikes (7) | 0.001 |
| Overall_DT/DZ_Range (8) | 0.069 |
| Local_Climatological_T-range (9) | 1.874 |
| Flagged_level_Percentage (10) | 2.060 |

AQC Diagnostics: PFL Temperature Profiles

Quality Check Diagnostics for the PFL Dataset



Number non-dummy levels 106446596
 Number / % dummy levels 0 / 0.000
 Percent Flagged levels 1.63
 Total Profiles 1223946
 Percent profiles >= one flagged level 14.43
 Overall Limits: -2.0 33.0



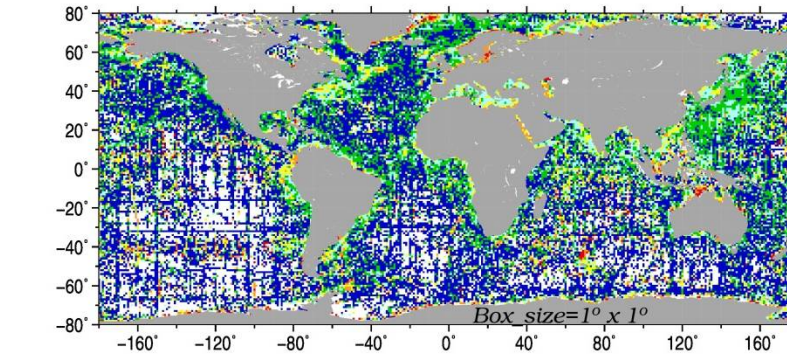
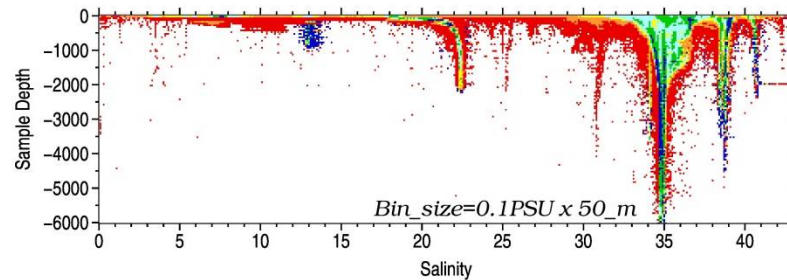
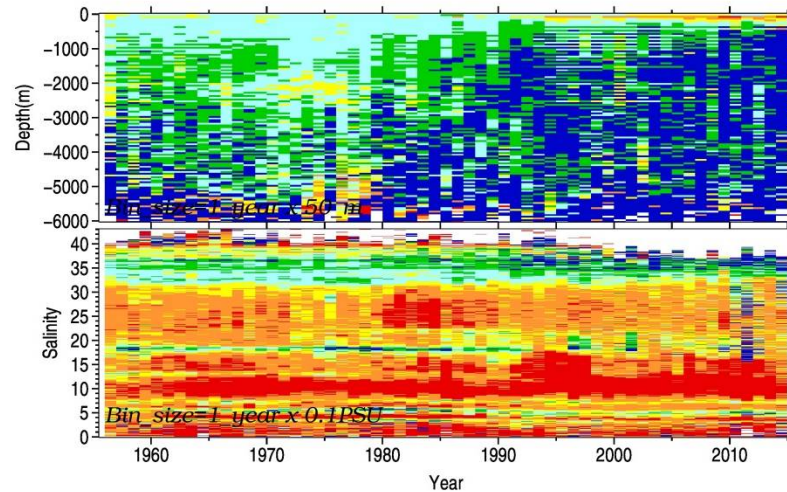
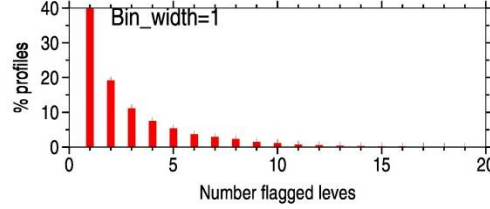
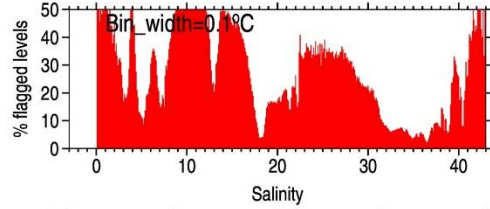
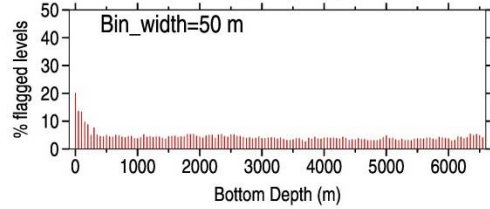
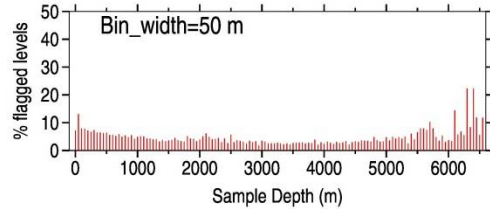
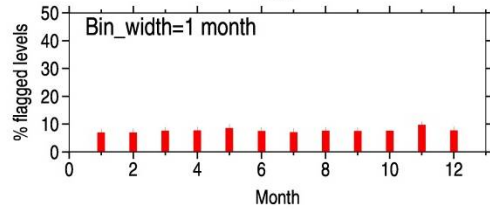
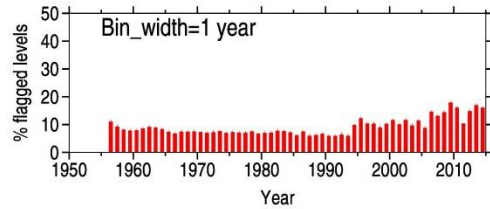
| QC-Check | % flagged levels |
|----------------------------------|------------------|
| Sample_depth_order (1) | 0.000 |
| Crude-T_range (2) | 0.066 |
| Last_Level_Depth_vs_GEBCO (3) | 0.161 |
| Datatype_Max_Obs_Depth (4) | 0.273 |
| Stuck_Value (5) | 0.005 |
| Number_of_T-extremuma (6) | 0.066 |
| Spikes (7) | 0.005 |
| Overall_DT/DZ_Range (8) | 0.025 |
| Local_Climatological_T-range (9) | 1.136 |
| Flagged_level_Percentage (10) | 0.467 |

Quality Control of Salinity Data:

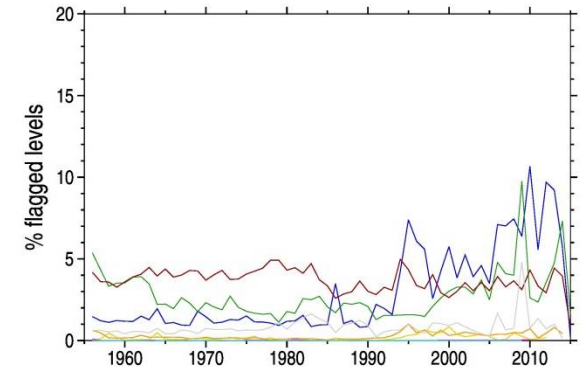
- Same sequence of quality checks
- Only OSD, CTD and PFL data used for climatology (APB; MRB; XCTD excluded)

AQC Diagnostics: OSD Salinity Profiles

Quality Check Diagnostics for the OSD Dataset



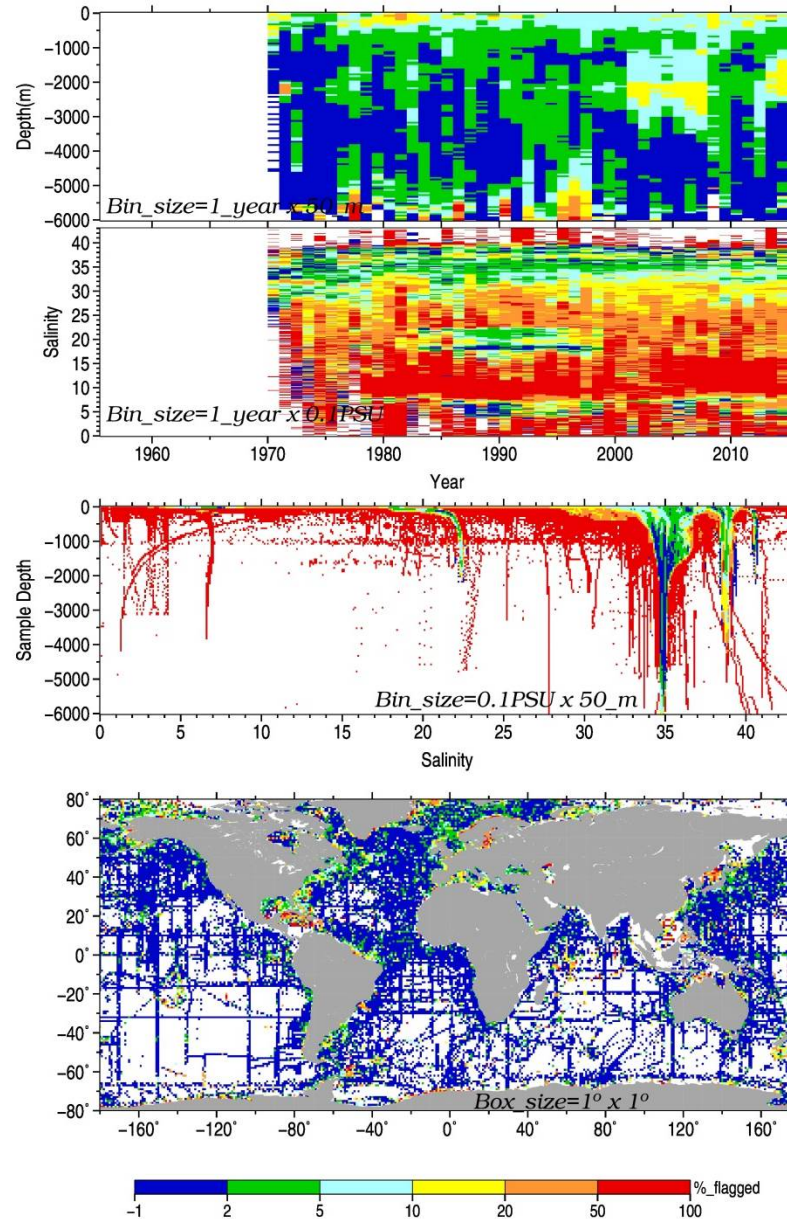
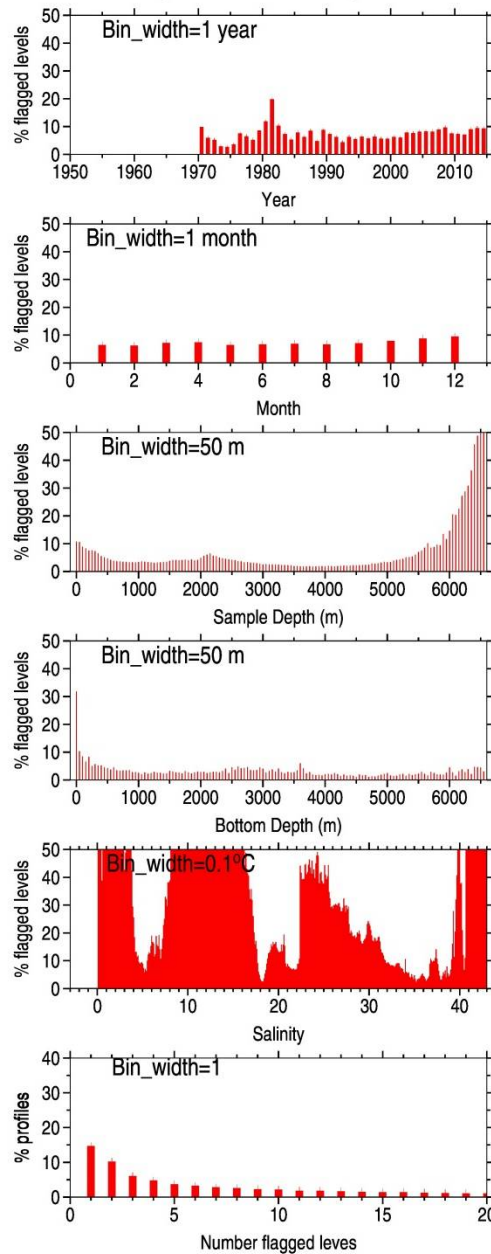
| | |
|---------------------------------------|-----------|
| Number non-dummy levels | 22372821 |
| Number / % dummy levels | 0 / 0.000 |
| Percent Flagged levels | 7.69 |
| Total Profiles | 2023214 |
| Percent profiles >= one flagged level | 25.67 |
| Overall Limits: | 0.0 |



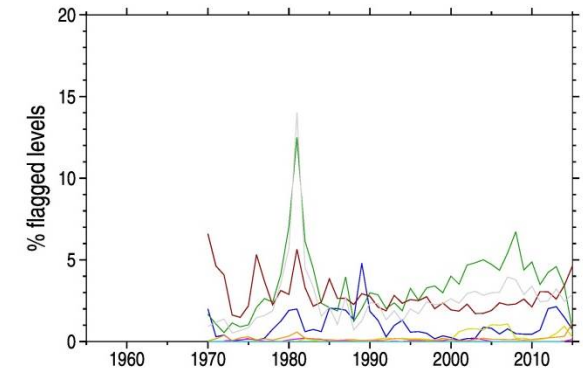
| QC-Check | % flagged levels |
|----------------------------------|------------------|
| Sample_depth_order (1) | 0.000 |
| Crude-S_range (2) | 1.815 |
| Last_Level_Depth_vs_GEBCO (3) | 2.276 |
| Datatype_Max_Obs_Depth (4) | 0.000 |
| Stuck_Value (5) | 0.123 |
| Number_of_S-extremuma (6) | 0.008 |
| Spikes (7) | 0.003 |
| Overall_DS/DZ_Range (8) | 0.192 |
| Local_Climatological_S-range (9) | 3.827 |
| Flagged_level_Percentage (10) | 0.826 |

AQC Diagnostics: CTD Salinity Profiles

Quality Check Diagnostics for the CTD Dataset



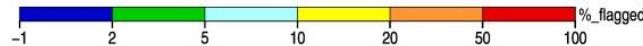
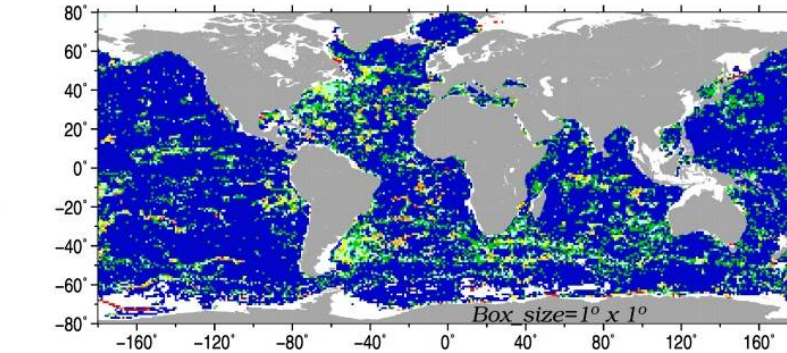
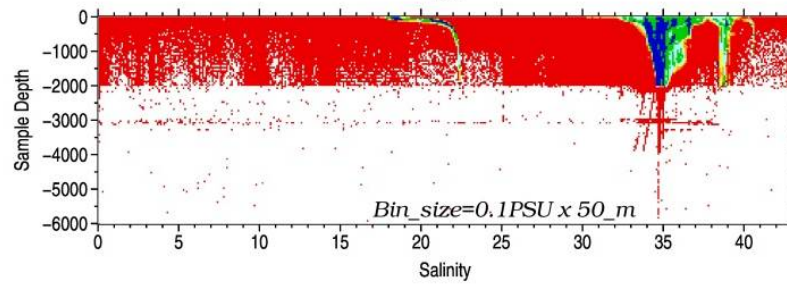
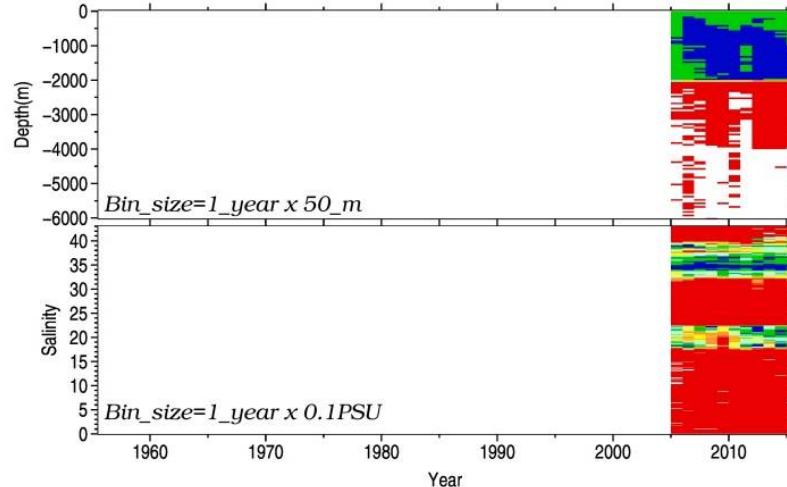
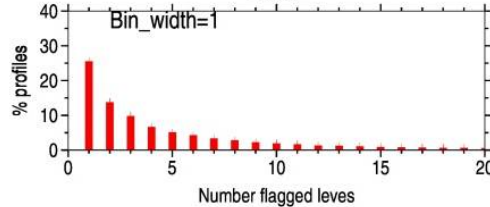
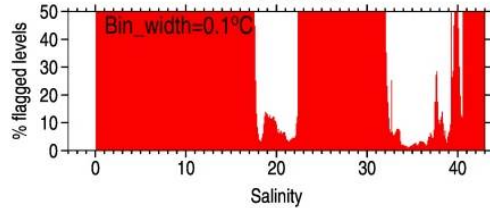
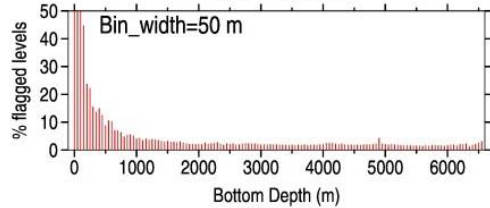
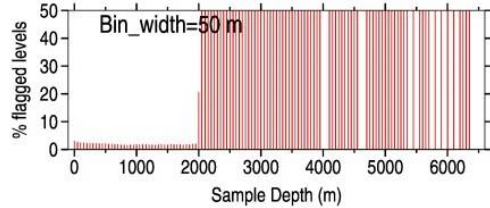
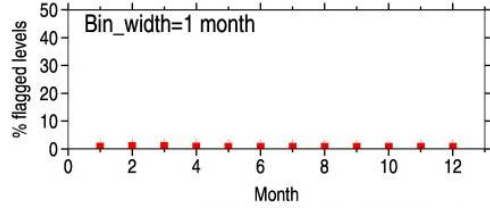
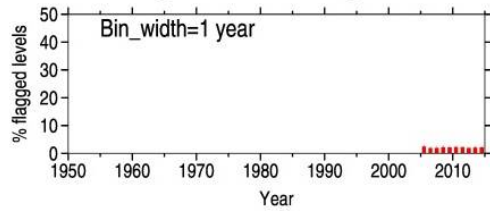
| | |
|---------------------------------------|-----------|
| Number non-dummy levels | 92272502 |
| Number / % dummy levels | 0 / 0.000 |
| Percent Flagged levels | 7.10 |
| Total Profiles | 928168 |
| Percent profiles >= one flagged level | 26.84 |
| Overall Limits: | 0.0 |



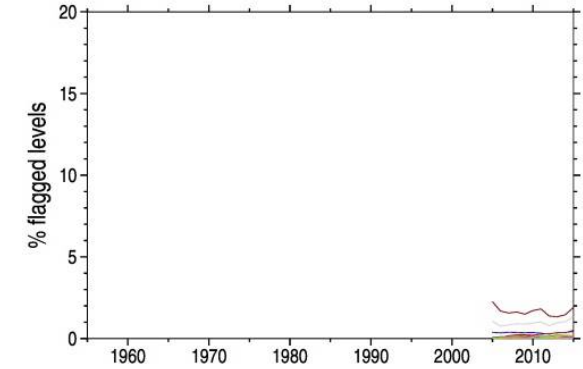
| QC-Check | % flagged levels |
|----------------------------------|------------------|
| Sample_depth_order (1) | 0.000 |
| Crude-S_range (2) | 0.900 |
| Last_Level_Depth_vs_GEBCO (3) | 3.599 |
| Datatype_Max_Obs_Depth (4) | 0.001 |
| Stuck_Value (5) | 0.288 |
| Number_of_S-extremuma (6) | 0.044 |
| Spikes (7) | 0.001 |
| Overall_DS/DZ_Range (8) | 0.143 |
| Local_Climatological_S-range (9) | 2.485 |
| Flagged_level_Percentage (10) | 2.567 |

AQC Diagnostics: PFL Salinity Profiles

Quality Check Diagnostics for the PFL Dataset

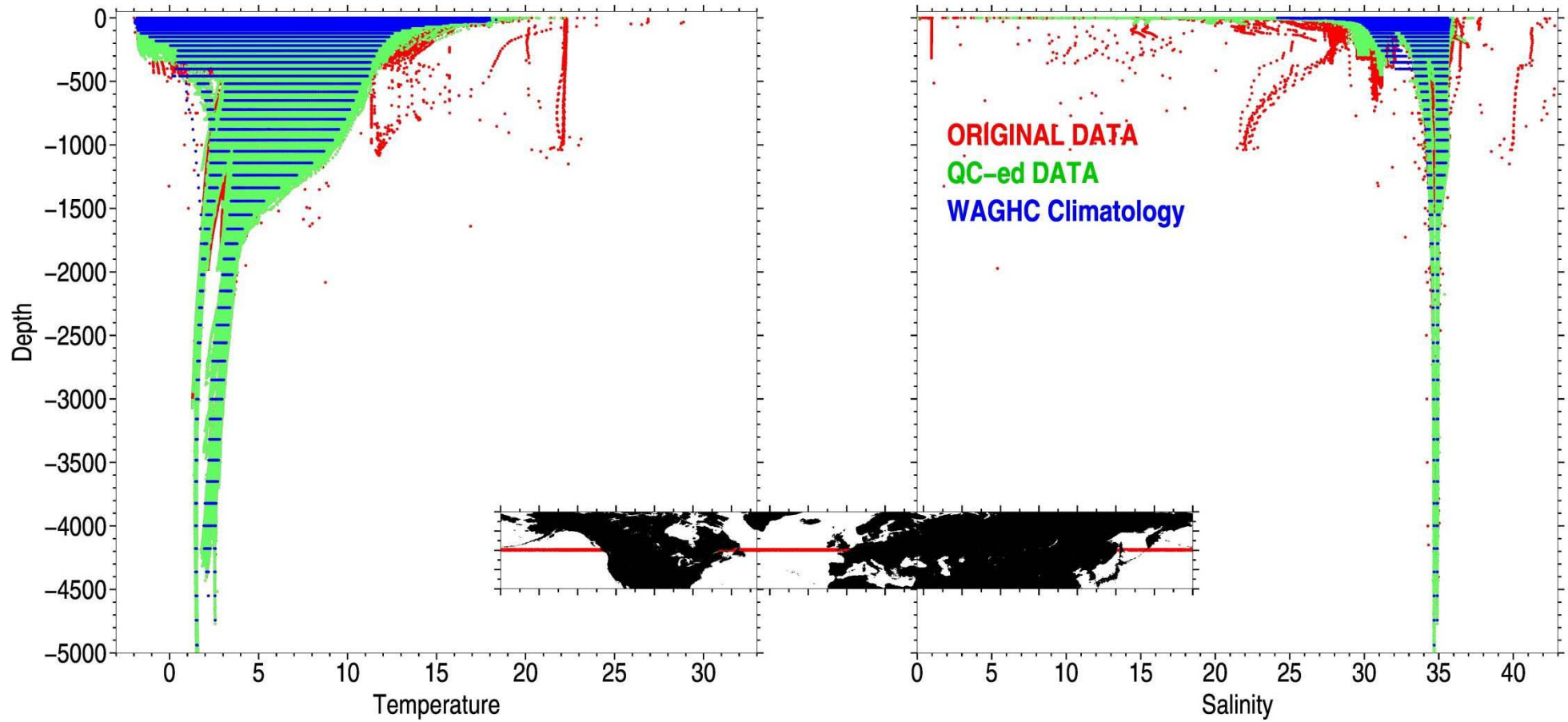


| | |
|---------------------------------------|-----------|
| Number non-dummy levels | 106443705 |
| Number / % dummy levels | 0 / 0.000 |
| Percent Flagged levels | 2.14 |
| Total Profiles | 1223901 |
| Percent profiles >= one flagged level | 14.43 |
| Overall Limits: | 0.0 |



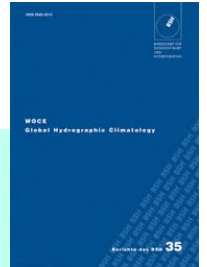
| QC-Check | % flagged levels |
|----------------------------------|------------------|
| Sample_depth_order (1) | 0.000 |
| Crude-S_range (2) | 0.362 |
| Last_Level_Depth_vs_GEBCO (3) | 0.158 |
| Datatype_Max_Obs_Depth (4) | 0.273 |
| Stuck_Value (5) | 0.109 |
| Number_of_S-extremuma (6) | 0.074 |
| Spikes (7) | 0.009 |
| Overall_DS/DZ_Range (8) | 0.045 |
| Local_Climatological_S-range (9) | 1.595 |
| Flagged_Level_Percentage (10) | 0.963 |

Example of the AQC application



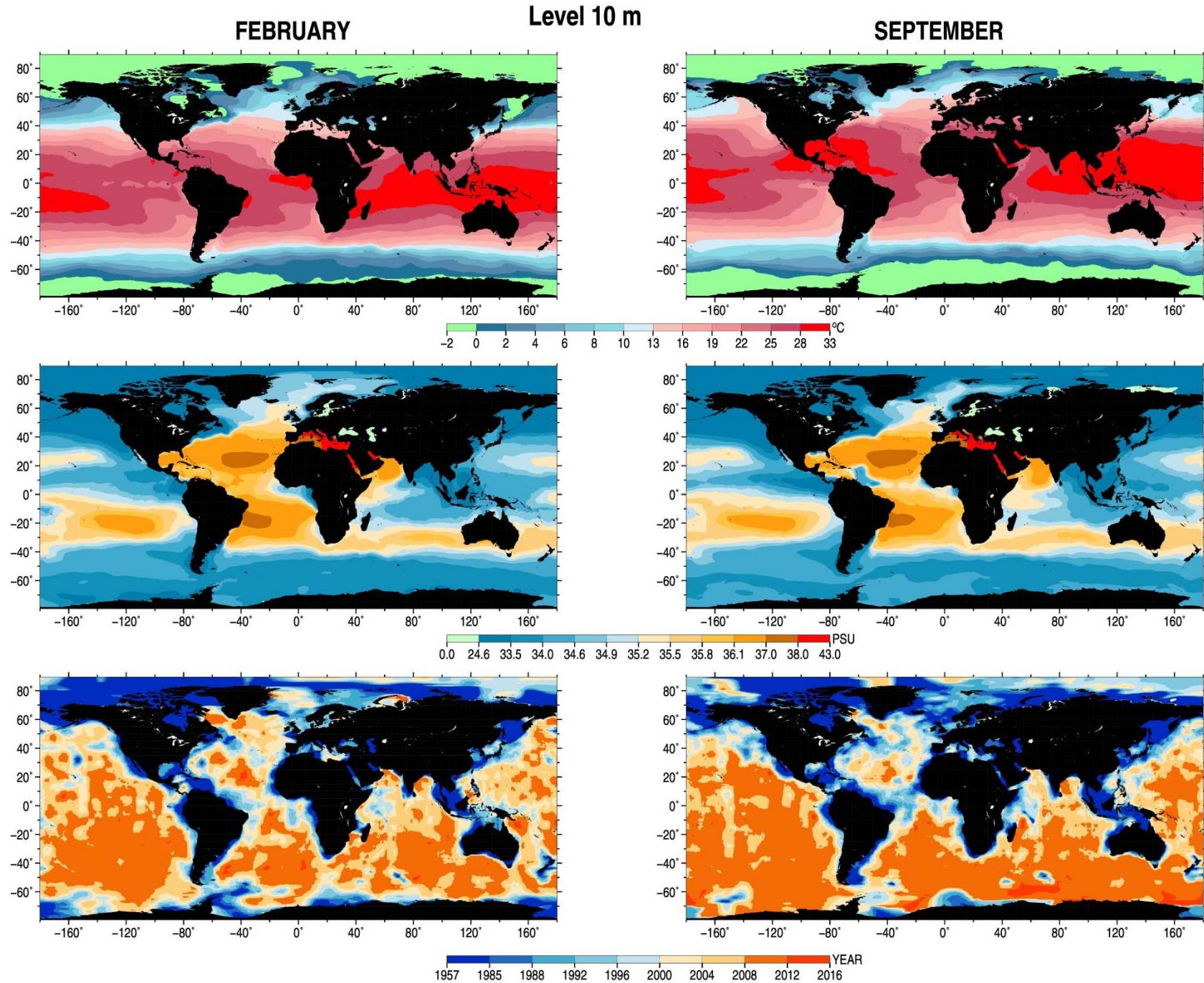
WOCE-ARGO Global Hydrography Climatology

update of the WOCE Global Hydrographic Climatology (*Gouretski&Koltermann 2004*)

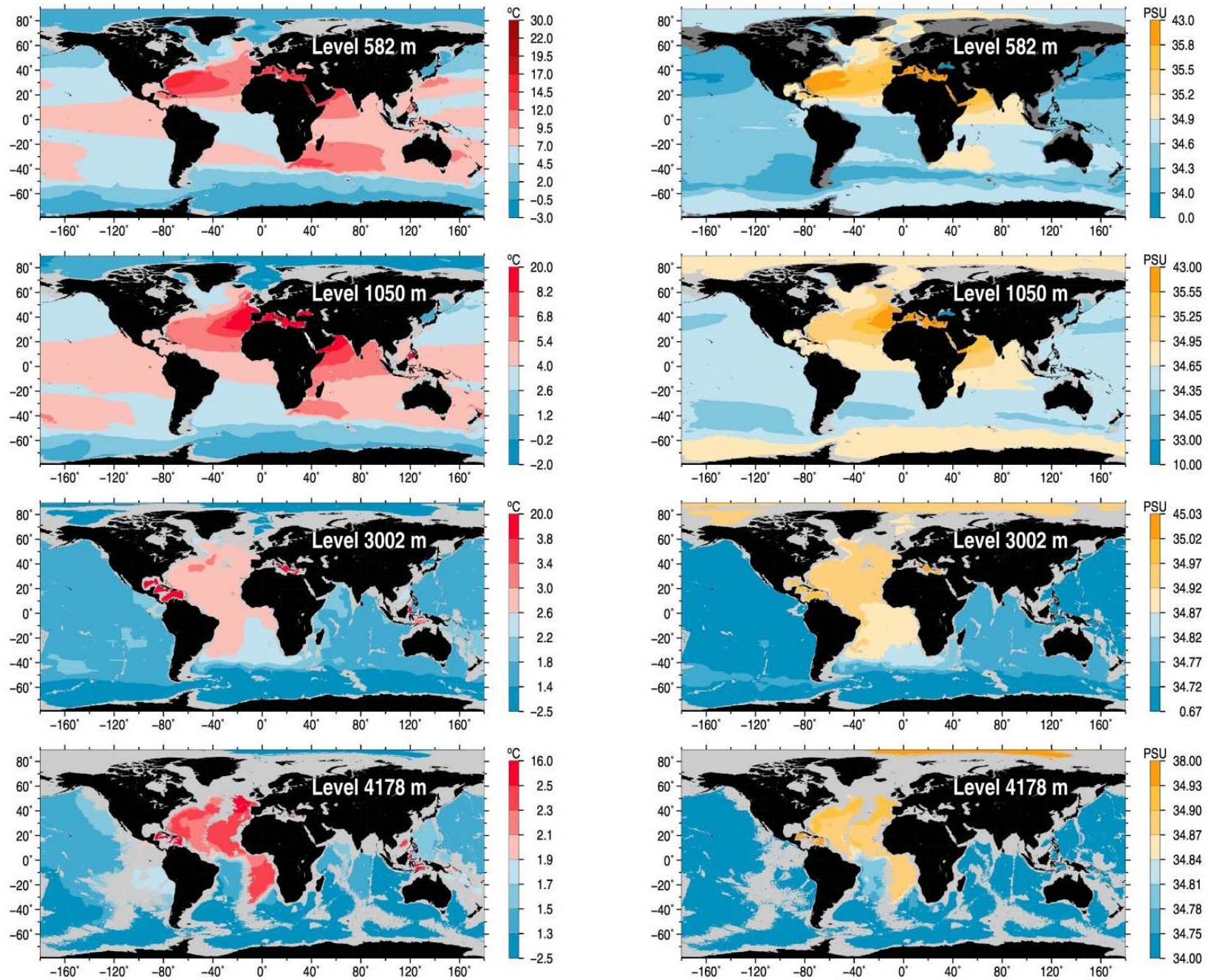


| | |
|----------------------------|---|
| Data base: | WOD13 OSD+CTD+PFL (March 2016) |
| Quality Control Procedure: | ICDC/IQuOD |
| Time-window: | 1985-2016 (extended to 1957-2016 for data-poor areas) |
| Spatial Resolution: | 0.25x0.25 LAT/LON |
| Vertical Resolution: | 65 Levels (0 to 6650 m) |
| Monthly Resolution: | Surface to 500 m |
| Interpolation Method: | Optimal Interpolation of 0.25x0.25-degree binned data on potential density surfaces |
| Digital Bottom Relief: | GEBCO 30" resolution |
| Mean year of observation | YES |
| Hydrostatically stable | YES |

WAGHC: Example of monthly property distributions



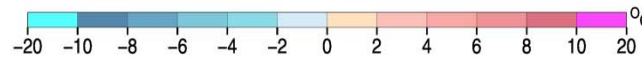
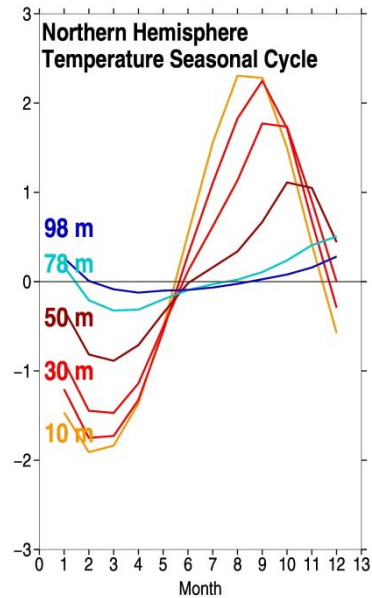
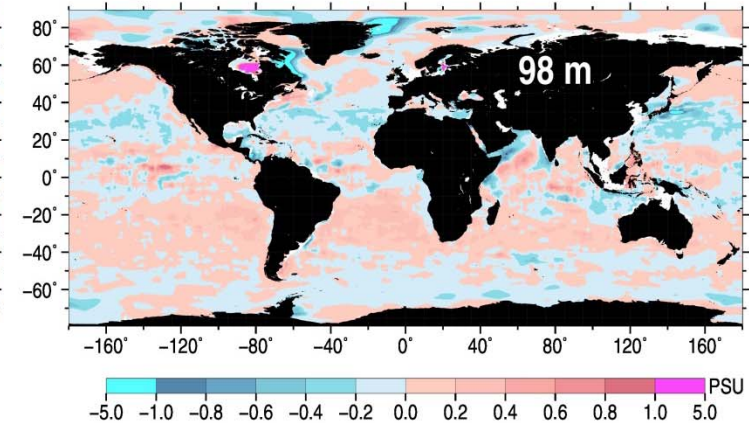
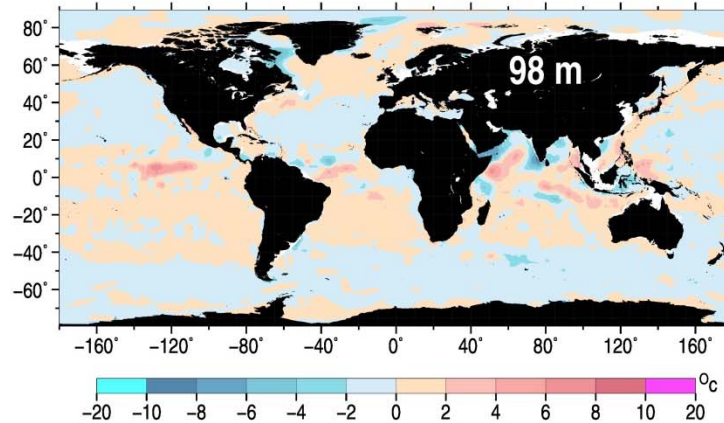
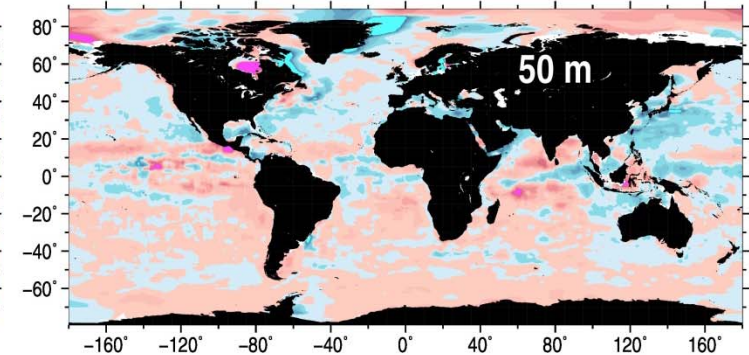
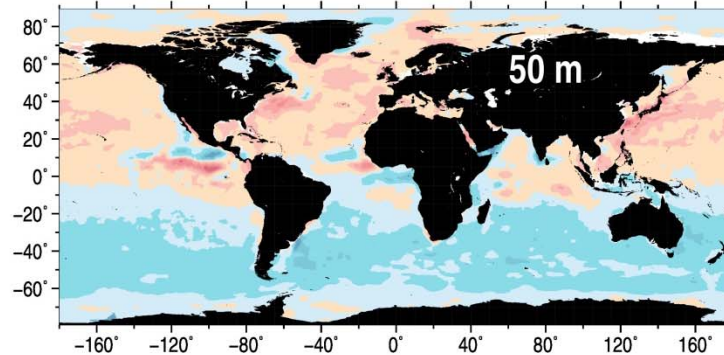
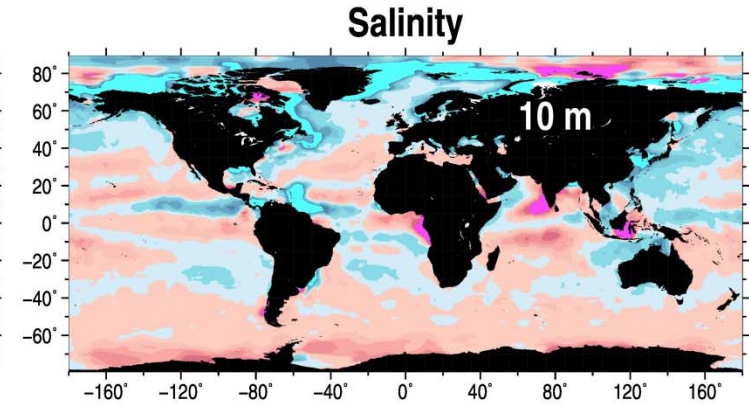
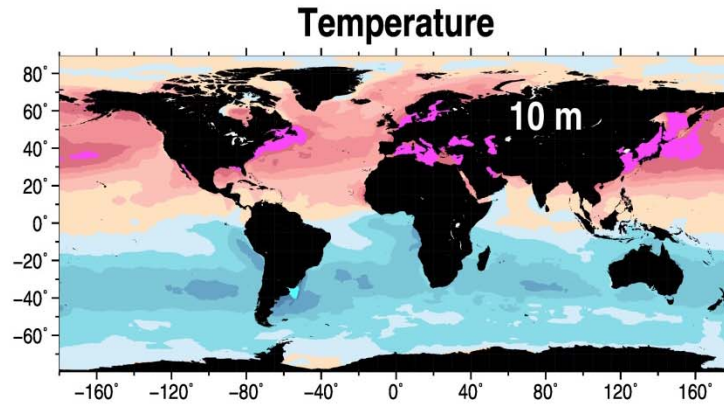
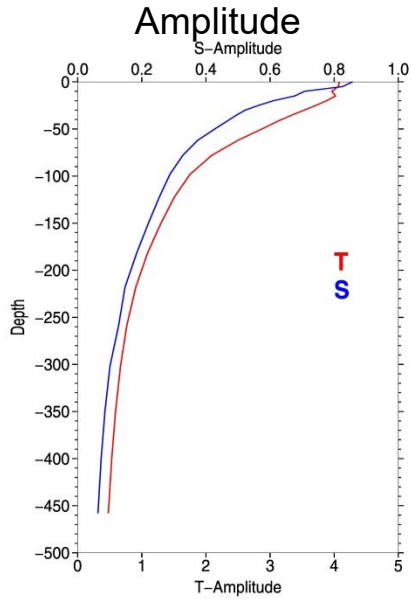
WAGHC: Example of property distributions at depths



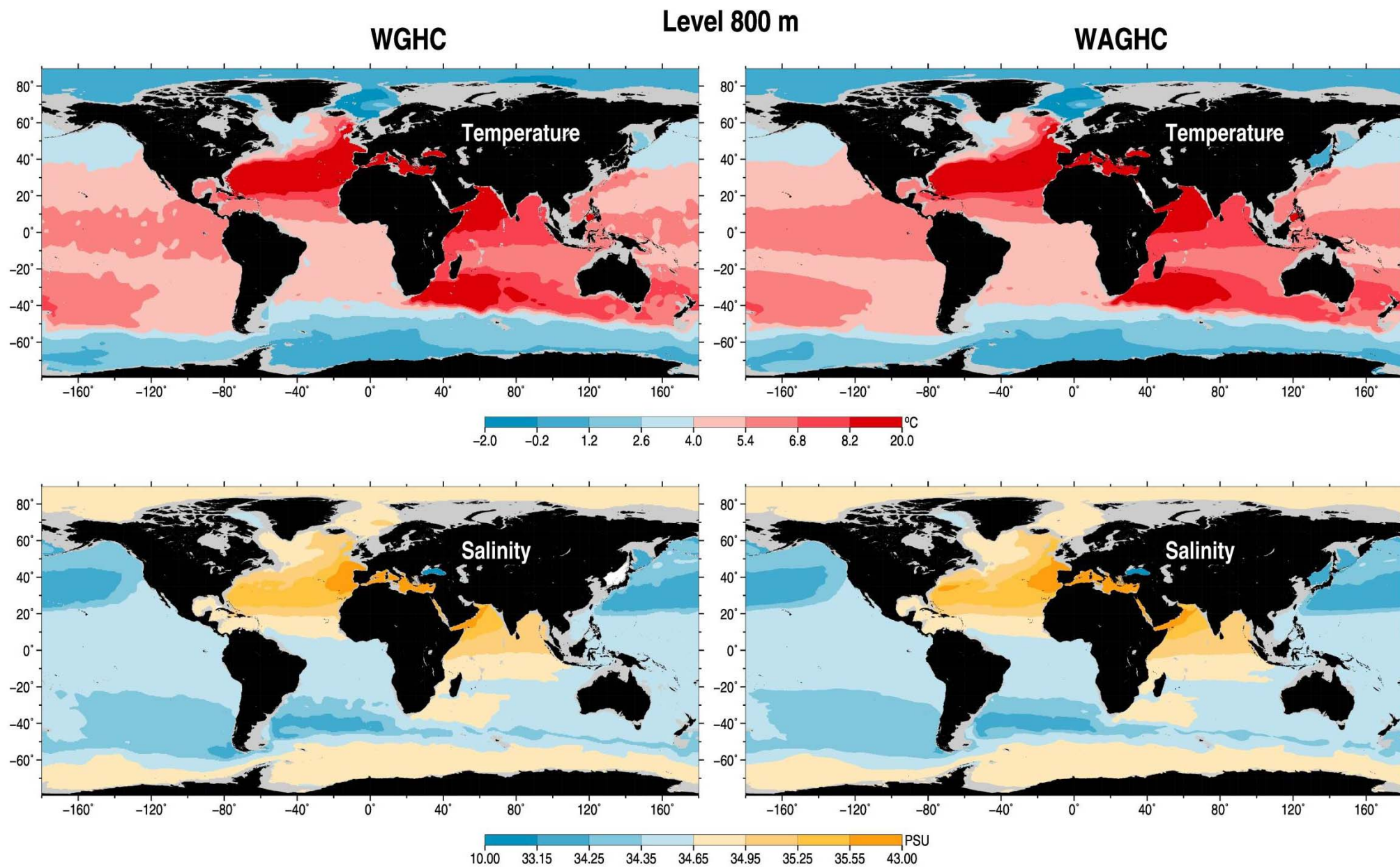
Seasonal Temperature and Salinity Variations

August – February

Global-Mean Seasonal

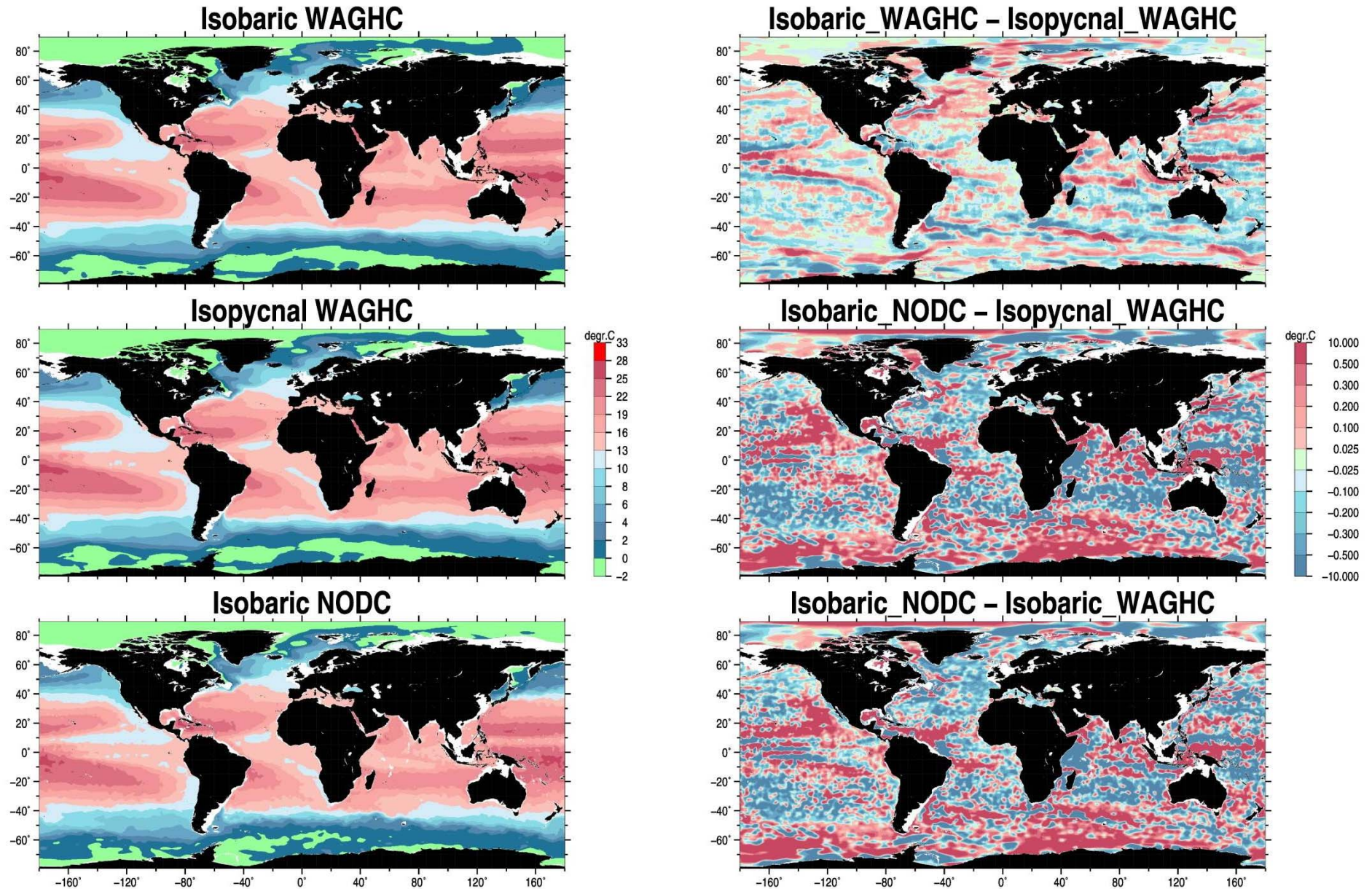


WAGHC vs WGHC (*Gouretski&Koltermann 2004*)



Isopycnal Interpolation vs Isobaric Interpolation

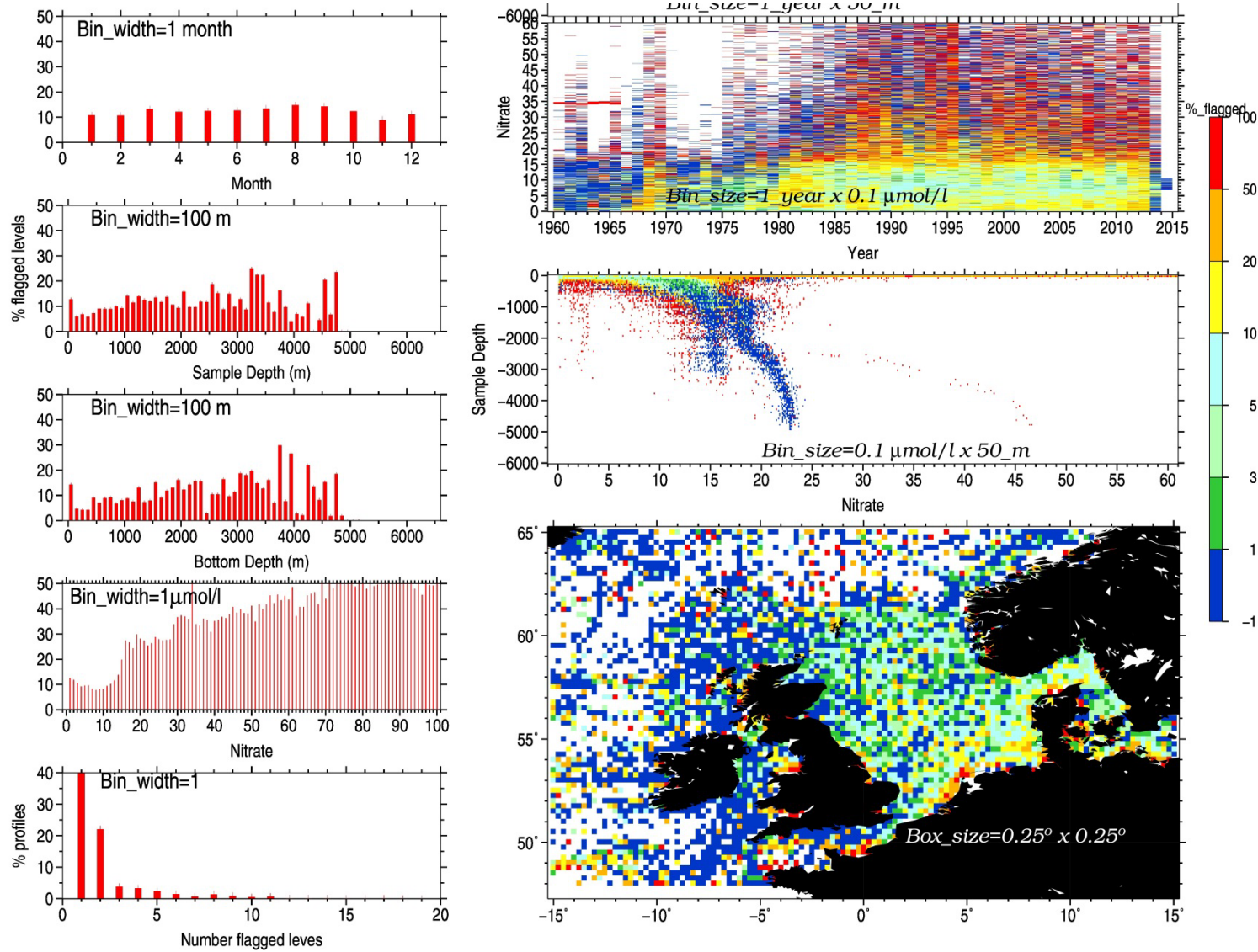
JULY Level 150m



Regional Climatologies

- The ICDC AQC procedure was implemented for the regional North Sea and Baltic Sea Climatologies
- For the North Sea the quality control procedure was tailored for the following geochemical parameters: nitrate, phosphate, oxygen, silicate, and ammonium

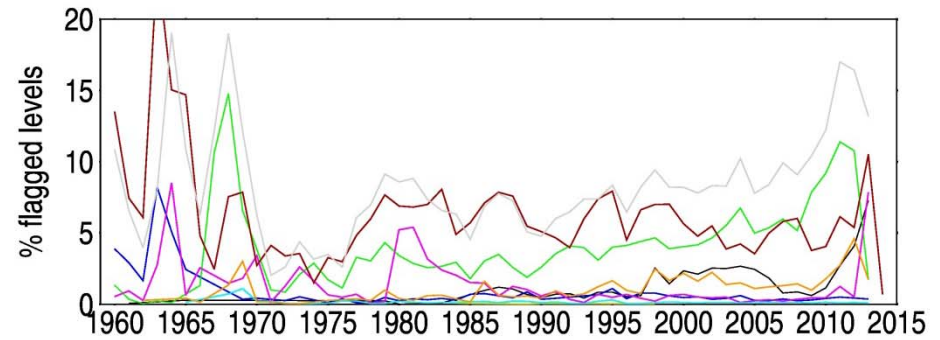
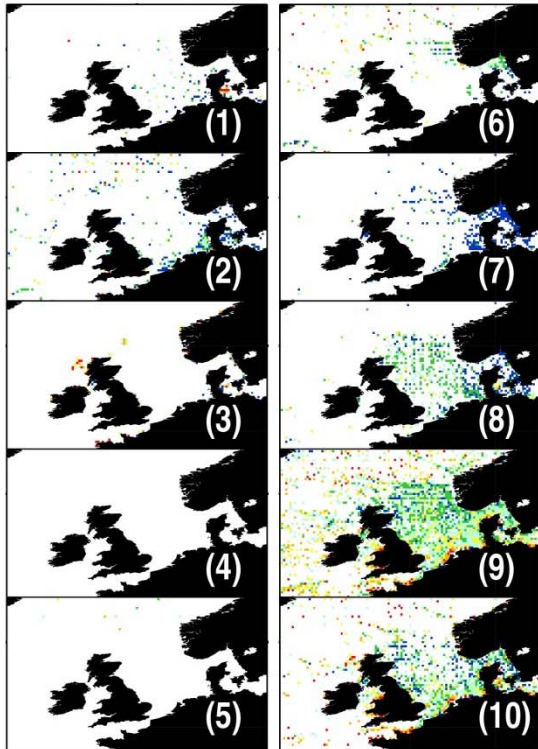
North Sea Nitrate Dataset: Nitrate Combined Quality Checks



Percent Flagged levels: 12.07

North Sea Nitrate Dataset: Distinct Quality Checks for Nitrate

% obs. flagged by QC-checks in 0.25x0.25 boxes



| QC-check | % flagged | |
|---|-----------|------|
| Sample_depth_order | 1.084 | (1) |
| Crude-NO ₃ range | .514 | (2) |
| Last_Level_Depth_vs_GEBCO | 4.169 | (3) |
| Datatype_Max_Obs_Depth | .000 | (4) |
| Stuck_Value | .008 | (5) |
| Number_of_NO3_extremuma | .924 | (6) |
| Spikes | .133 | (7) |
| Overall_NO ₃ /DZ_Range | 1.103 | (8) |
| Local_Climatological_NO ₃ -range | 5.678 | (9) |
| Flagged_level_Percentage | 7.737 | (10) |
| Overall_NO ₃ _Limits: | .0 330.0 | |

$N_{AMD} = 3.0$

Summary

- Update of the AQC procedure for temperature
- Update of the AQC diagnostic tool
- AQC-Procedure tailored for salinity and nutrients
- AQC procedure applied for the calculation of the global T,S-climatology and the nutrient climatology of the North Sea

Future Plans

- Perform AQC procedure in parameter-density space
- Test asymmetric local climatological ranges
- Use more realistic spatial correlation radii based on observations
- Instead of ad hoc values
- Repeat climatology calculations using the recommended IQuOD AQC-procedure, estimate the impact