

27-29 May 2024 



International conference on Marine Data and Information Systems



A new online app for ocean temperature quality control - An artificial intelligence approach

Sebastian Mieruch and Mohamed Chouai
Alfred-Wegener-Institute (AWI), Bremerhaven, Germany

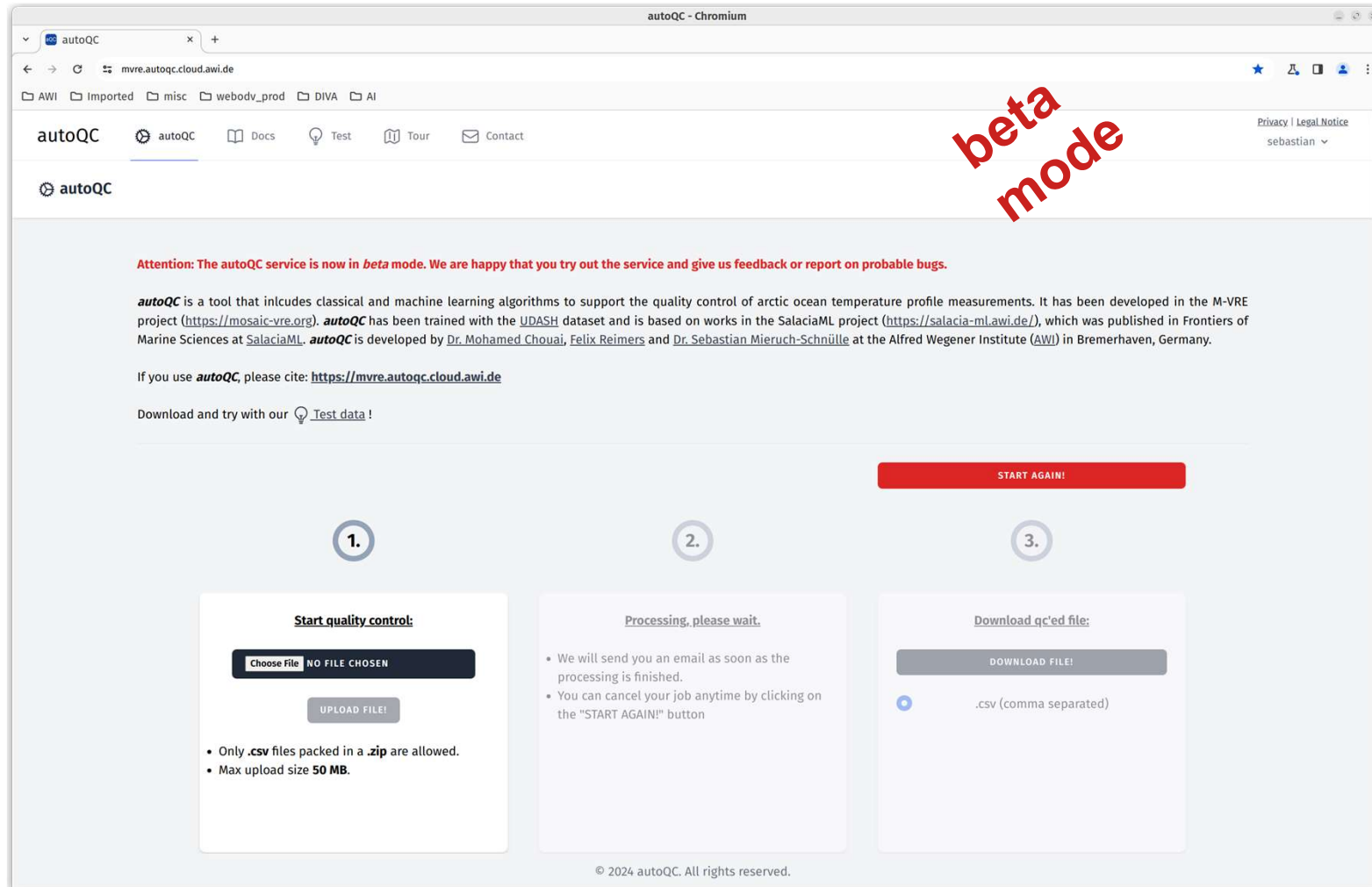


Bundesministerium
für Bildung
und Forschung



ALFRED-WEGENER-INSTITUT
HELMHOLTZ-ZENTRUM FÜR POLAR-
UND MEERESFORSCHUNG

The app: <https://mvre.autoqc.cloud.awi.de/>



autoQC - Chromium

autoQC mvre.autoqc.cloud.awi.de

AWI Imported misc webodv_prod DIVA AI

autoQC autoQC Docs Test Tour Contact

Privacy | Legal Notice
sebastian

autoQC

Attention: The autoQC service is now in beta mode. We are happy that you try out the service and give us feedback or report on probable bugs.

autoQC is a tool that includes classical and machine learning algorithms to support the quality control of arctic ocean temperature profile measurements. It has been developed in the M-VRE project (<https://mosaic-vre.org>). **autoQC** has been trained with the UDASH dataset and is based on works in the SalaciaML project (<https://salacia-ml.awi.de/>), which was published in Frontiers of Marine Sciences at SalaciaML. **autoQC** is developed by [Dr. Mohamed Chouai](#), [Felix Reimers](#) and [Dr. Sebastian Mieruch-Schnülle](#) at the Alfred Wegener Institute (AWI) in Bremerhaven, Germany.

If you use **autoQC**, please cite: <https://mvre.autoqc.cloud.awi.de>

Download and try with our [Test data](#) !

1. **Start quality control:**

Choose File NO FILE CHOSEN

UPLOAD FILE!

- Only **.csv** files packed in a **.zip** are allowed.
- Max upload size **50 MB**.

2. **Processing, please wait.**

START AGAIN!

- We will send you an email as soon as the processing is finished.
- You can cancel your job anytime by clicking on the "START AGAIN!" button

3. **Download qc'ed file:**

DOWNLOAD FILE!

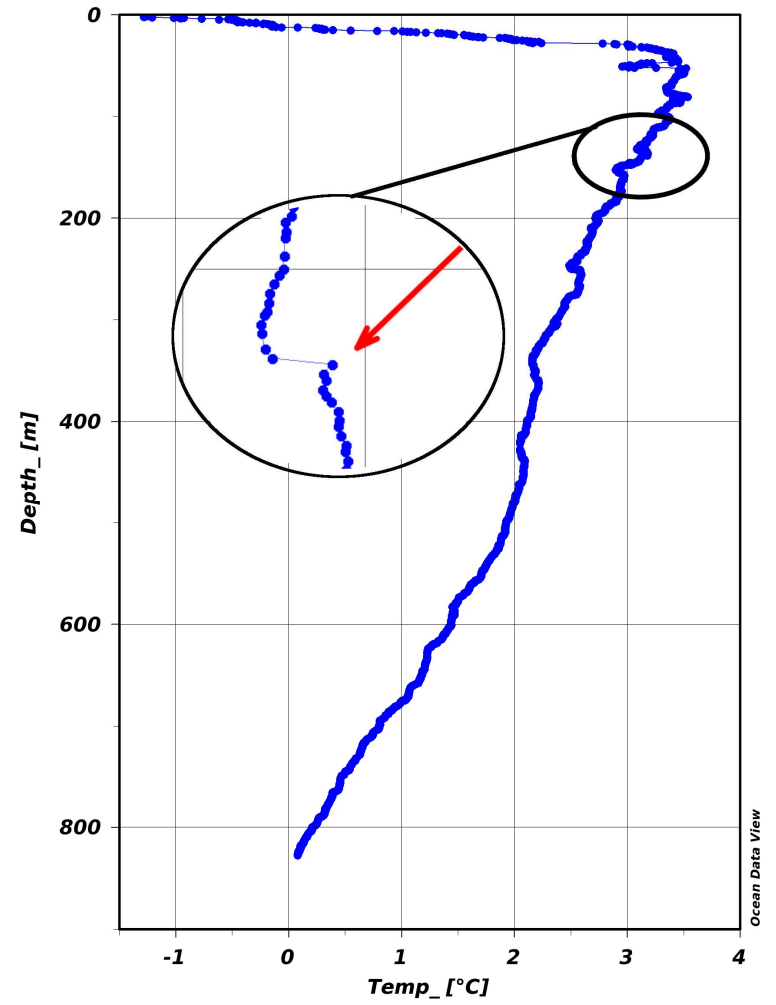
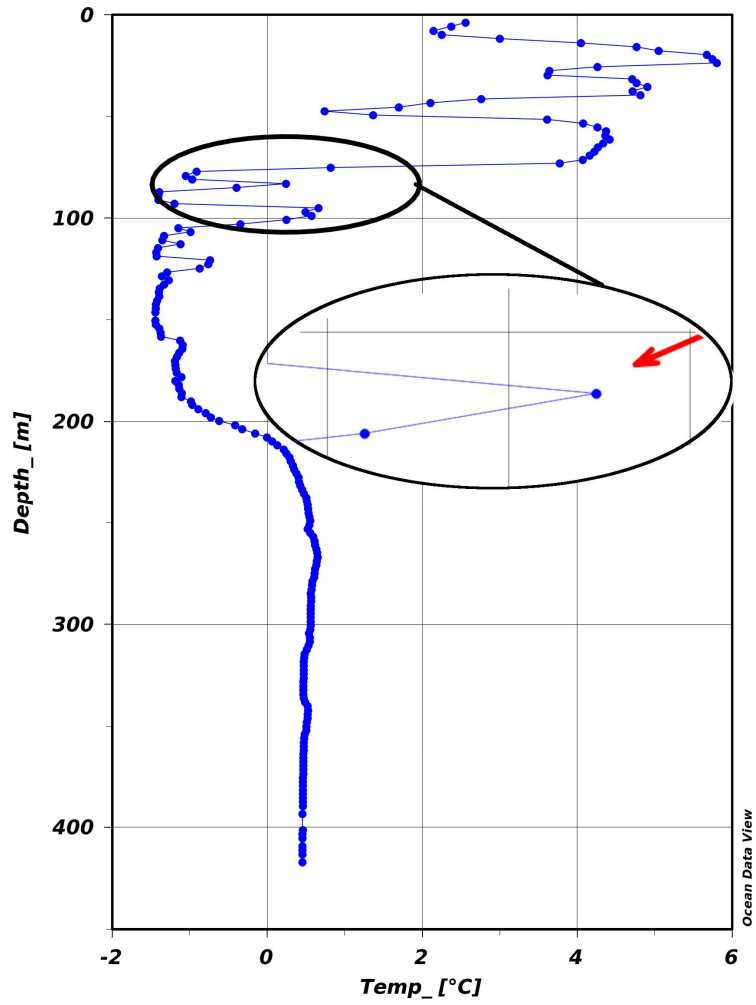
.csv (comma separated)

© 2024 autoQC. All rights reserved.

The input data

```
smieruch@bgeo04l004: ~/Downloads/UDASH_test_subset
File Edit Options Buffers Tools Help
Prof_no,Type,year,month,Longitude_[deg],Latitude_[deg],Pressure_[dbar],Depth_[m],Temp_[°C],Salinity_[psu]
64450,CTD,1981.0,10.0,-5.7283,76.4267,0.4,0.4,-1.456,33.715
64450,CTD,1981.0,10.0,-5.7283,76.4267,0.61,0.6,-1.457,33.715
64450,CTD,1981.0,10.0,-5.7283,76.4267,0.91,0.9,-1.455,33.715
64450,CTD,1981.0,10.0,-5.7283,76.4267,1.21,1.2,-1.456,33.716
64450,CTD,1981.0,10.0,-5.7283,76.4267,1.52,1.5,-1.457,33.718
64450,CTD,1981.0,10.0,-5.7283,76.4267,1.92,1.9,-1.463,33.724
64450,CTD,1981.0,10.0,-5.7283,76.4267,2.22,2.2,-1.485,33.729
64450,CTD,1981.0,10.0,-5.7283,76.4267,2.43,2.4,-1.523,33.739
64450,CTD,1981.0,10.0,-5.7283,76.4267,2.73,2.7,-1.531,33.75
64450,CTD,1981.0,10.0,-5.7283,76.4267,2.93,2.9,-1.549,33.751
64450,CTD,1981.0,10.0,-5.7283,76.4267,3.33,3.3,-1.509,33.755
64450,CTD,1981.0,10.0,-5.7283,76.4267,3.54,3.5,-1.506,33.749
64450,CTD,1981.0,10.0,-5.7283,76.4267,3.74,3.7,-1.527,33.745
64450,CTD,1981.0,10.0,-5.7283,76.4267,4.04,4.0,-1.518,33.748
64450,CTD,1981.0,10.0,-5.7283,76.4267,4.45,4.4,-1.414,33.733
64450,CTD,1981.0,10.0,-5.7283,76.4267,4.75,4.7,-1.407,33.73
64450,CTD,1981.0,10.0,-5.7283,76.4267,5.05,5.0,-1.351,33.728
64450,CTD,1981.0,10.0,-5.7283,76.4267,5.36,5.3,-1.457,33.733
64450,CTD,1981.0,10.0,-5.7283,76.4267,5.86,5.8,-1.5,33.741
64450,CTD,1981.0,10.0,-5.7283,76.4267,6.16,6.1,-1.579,33.762
64450,CTD,1981.0,10.0,-5.7283,76.4267,6.57,6.5,-1.585,33.772
64450,CTD,1981.0,10.0,-5.7283,76.4267,6.87,6.8,-1.574,33.788
64450,CTD,1981.0,10.0,-5.7283,76.4267,7.28,7.2,-1.61,33.793
64450,CTD,1981.0,10.0,-5.7283,76.4267,7.58,7.5,-1.62,33.796
64450,CTD,1981.0,10.0,-5.7283,76.4267,7.88,7.8,-1.623,33.803
64450,CTD,1981.0,10.0,-5.7283,76.4267,8.49,8.4,-1.626,33.808
-UUU:----F1 UDASH_test_subset.csv Top (1,0) (Fundamental) -----
```

The flags: spike & suspect gradient



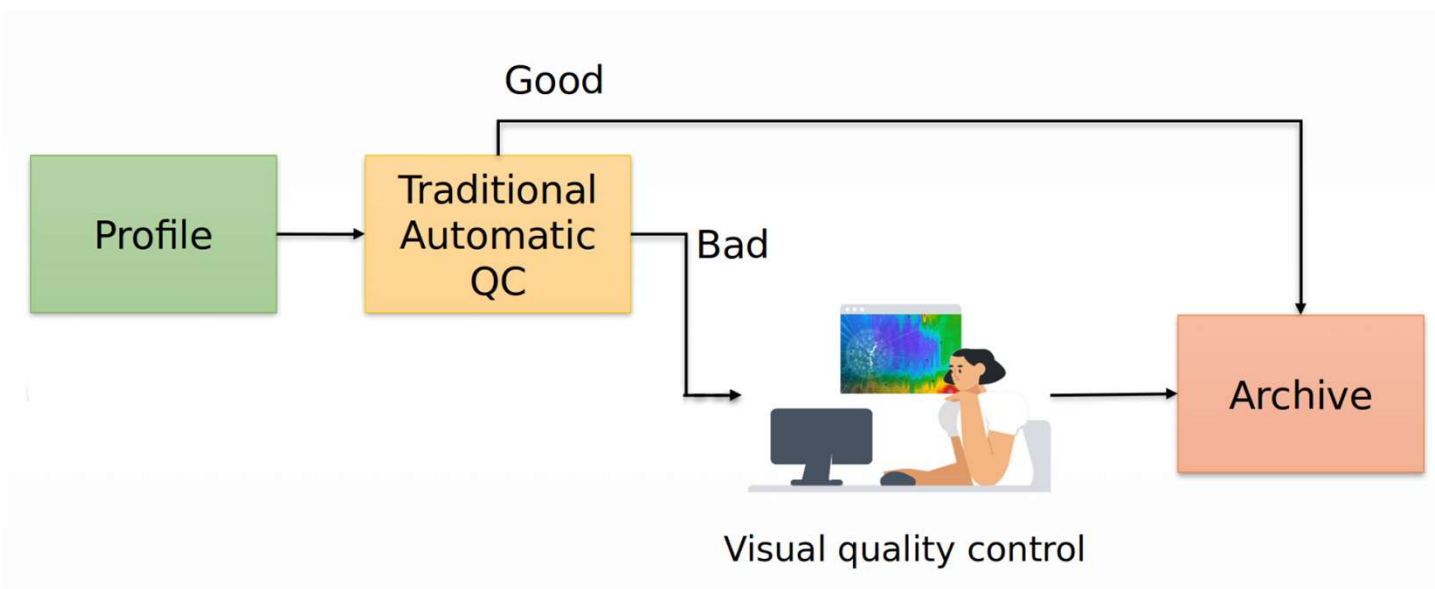
The output data

```
smieruch@bgeo041004: ~/Downloads/UDASH_test_subset
File Edit Options Buffers Tools Help
Prof_no,Type,year,month,Longitude_[deg],Latitude_[deg],Pressure_[dbar],Depth_[m],Temp_[°C],Salinity_[psu],Trad_QF,ML_QF
64450,CTD,1981.0,10.0,-5.7283,76.4267,0.4,0.4,-1.456,33.715,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,0.61,0.6,-1.457,33.715,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,0.91,0.9,-1.455,33.715,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,1.21,1.2,-1.456,33.716,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,1.52,1.5,-1.457,33.718,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,1.92,1.9,-1.463,33.724,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,2.22,2.2,-1.485,33.729,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,2.43,2.4,-1.523,33.739,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,2.73,2.7,-1.531,33.75,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,2.93,2.9,-1.549,33.751,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,3.33,3.3,-1.509,33.755,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,3.54,3.5,-1.506,33.749,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,3.74,3.7,-1.527,33.745,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,4.04,4.0,-1.518,33.748,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,4.45,4.4,-1.414,33.733,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,4.75,4.7,-1.407,33.73,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,5.05,5.0,-1.351,33.728,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,5.36,5.3,-1.457,33.733,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,5.86,5.8,-1.5,33.741,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,6.16,6.1,-1.579,33.762,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,6.57,6.5,-1.585,33.772,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,6.87,6.8,-1.574,33.788,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,7.28,7.2,-1.61,33.793,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,7.58,7.5,-1.62,33.796,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,7.88,7.8,-1.623,33.803,0,0
64450,CTD,1981.0,10.0,-5.7283,76.4267,8.49,8.4,-1.626,33.808,0,0
-UUU:----F1 UDASH_test_subset_processed_full.csv Top (1,0) (Fundamental) -----
```

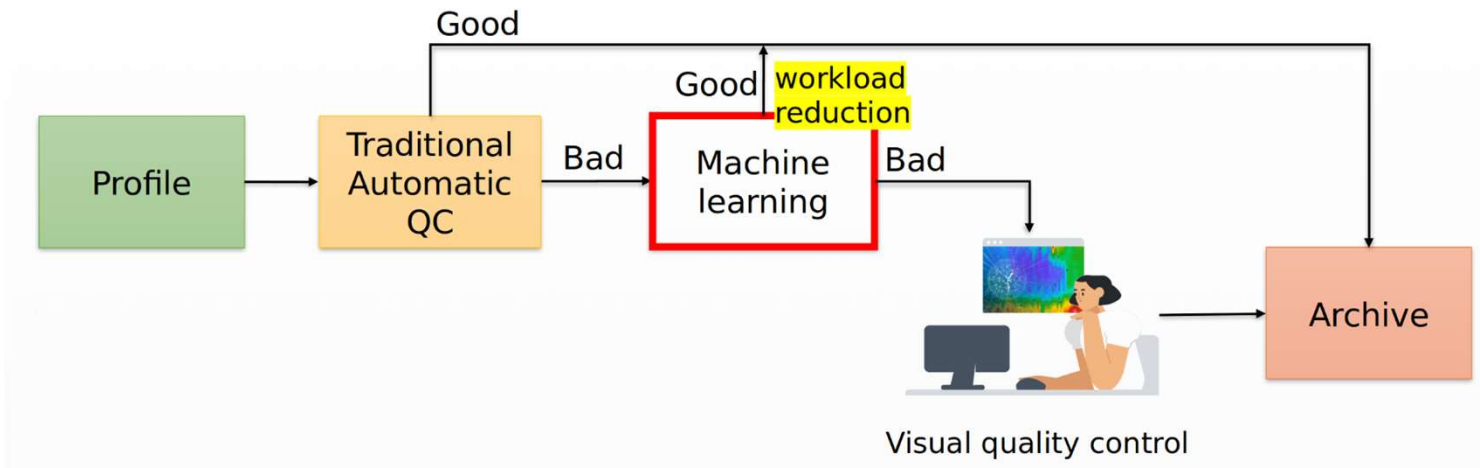
The AI

```
smieruch@bgeo041004: ~/Downloads/UDASH_test_subset
File Edit Options Buffers Tools Isearch Help
117587,CTD,2003.0,9.0,15.7788,78.3757,111.0,109.8,2.242,34.5283,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,112.01,110.8,2.221,34.5316,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,113.02,111.8,2.026,34.5246,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,114.03,112.8,1.734,34.5176,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,115.04,113.8,1.515,34.5148,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,115.95,114.7,1.369,34.5153,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,116.96,115.7,1.362,34.5235,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,117.98,116.7,1.348,34.5269,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,118.99,117.7,1.338,34.535,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,120.0,118.7,1.341,34.5411,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,121.01,119.7,1.333,34.5425,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,122.02,120.7,1.163,34.5417,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,123.03,121.7,0.897,34.5581,4,0
117587,CTD,2003.0,9.0,15.7788,78.3757,124.04,122.7,1.26,34.5779,2,2
117587,CTD,2003.0,9.0,15.7788,78.3757,124.95,123.6,1.352,34.5919,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,125.96,124.6,1.59,34.6075,2,2
117587,CTD,2003.0,9.0,15.7788,78.3757,126.98,125.6,1.618,34.6136,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,127.99,126.6,1.611,34.6164,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,129.0,127.6,1.609,34.6173,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,130.01,128.6,1.574,34.6193,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,131.02,129.6,1.534,34.6218,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,132.03,130.6,1.502,34.6222,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,133.04,131.6,1.433,34.6171,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,133.95,132.5,1.354,34.6202,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,134.96,133.5,1.314,34.618,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,135.98,134.5,1.201,34.6158,0,0
117587,CTD,2003.0,9.0,15.7788,78.3757,136.99,135.5,1.152,34.6175,0,0
-UUU:----F1 UDASH_test_subset_processed_full.csv 62% (41137,0) (Fundamental Isearch) -----
```

The classical approach



The classical + AI approach



The evaluation

Classic checks:

TP=1,229,357	FP=335
FN=2,443	TN=2,595

Classic + AI checks:

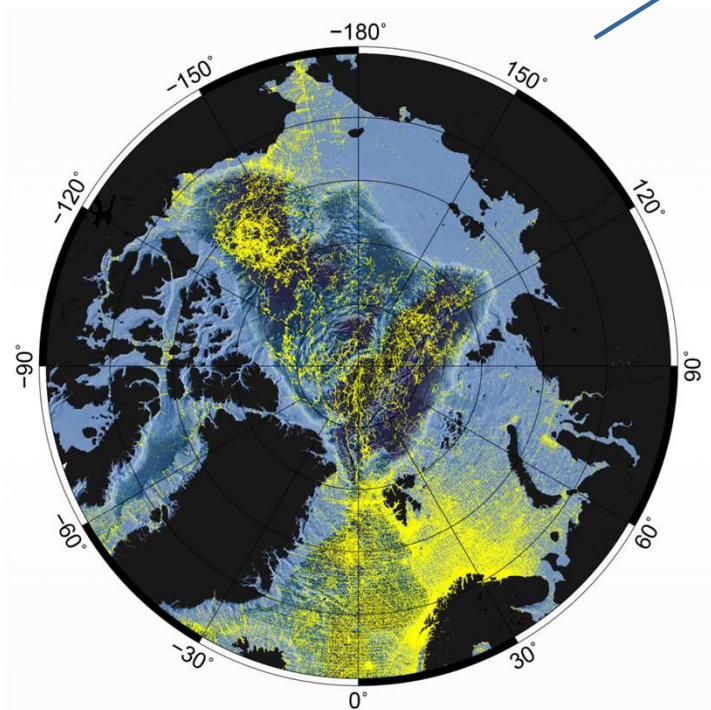
TP=1,231,439	FP=339
FN=361	TN=2,591

preliminar
v

UDASH (Unified Database for Arctic and Subarctic Hydrography)
<https://doi.pangaea.de/10.1594/PANGAEA.872931>

280,000 temperature and salinity profiles

Visually controlled by human expert



How we did it

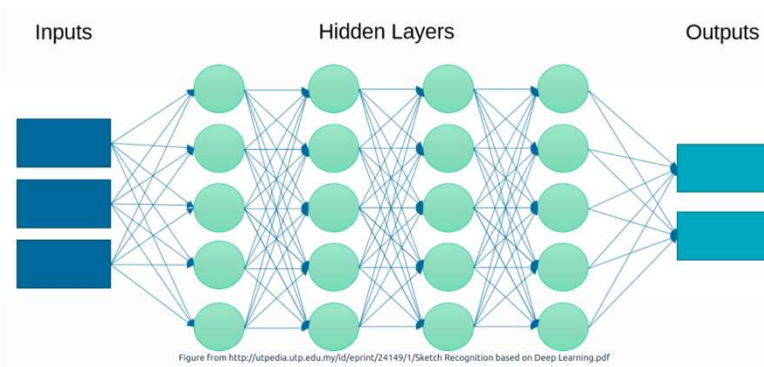
Classical QC checks

```

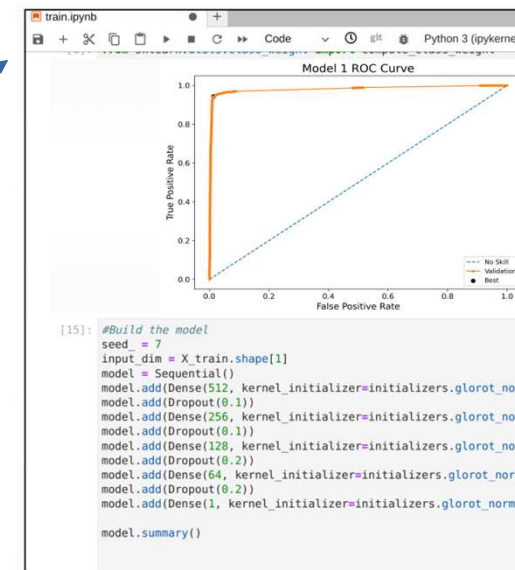
# all_in_one.ipynb
#
# Outliers in mixed layer
#
# =====
def Traditional_outlier_detection(Data):
    gradient = T_Suspect_gradient_T_D(Data)
    Data['gradient2'] = np.array(gradient)
    data = Data.loc[(Data['gradient2'] >= 0.5) & (Data['Depth_m'] <= 100)]
    data['loc'] = 'OF_trad'
    return data
def T_Suspect_gradient_T_D(Data):
    unique_profils = Data['Prof_no'].unique()
    d_grad = []
    data2 = []
    for j in range(len(unique_profils)):
        profil = Data[Data['Prof_no'] == unique_profils[j]].reset_index(drop=True)
        #data2 = pd.concat([data2, profil])
        data2.append(profil.values)
        t = profil['Temp_1C'].values
        d = profil['Depth_m'].values
        i = 0
        while i < d.size - 1:
            #case of -999 in temperature
            if -999 in (t[i], t[i+1], d[i+1], d[i]):
                grad = np.append(grad, -999) #give this value 1000
            else:
                if (d[i+1] - d[i]) >= 0:
                    grad = np.append(grad, (t[i+1] - t[i]) / (d[i+1] - d[i]))
                else:
                    grad = np.append(grad, -999) #give this value 1000
            i = i + 1
        #grad = np.append(grad, grad[-1])
        d_grad.append(grad)
    d_grad_flat = [item for sublist in d_grad for item in sublist]
    return np.array(d_grad_flat)
#
# =====

```

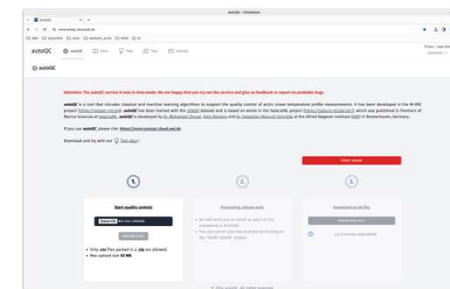
MLP (Multi-Layer-Perceptron)



AI – training, validation, evaluation



App



Future

- Mature and finalize
- GitHub: <https://github.com/mchouai27/SalaciaML-Arctic>
- Apply on salinity
- Improve algorithms
- Approach communities, e.g. SeaDataNet, IQuOD, ...
- Develop model for other error: statistical screening
- Use other datasets (from other qc-operators)
- ...

27-29 May 2024 



International conference on Marine Data and Information Systems

