27-29 May 2024

International conference on Marine Data and Information Systems

SeaDataNet AISBL · Havforskningsinstituttet · Ifremer · MARIS · MEDIN · National Oceanography Centre · INGV · iMagine · IODE · eosc Blue-Cloud2026

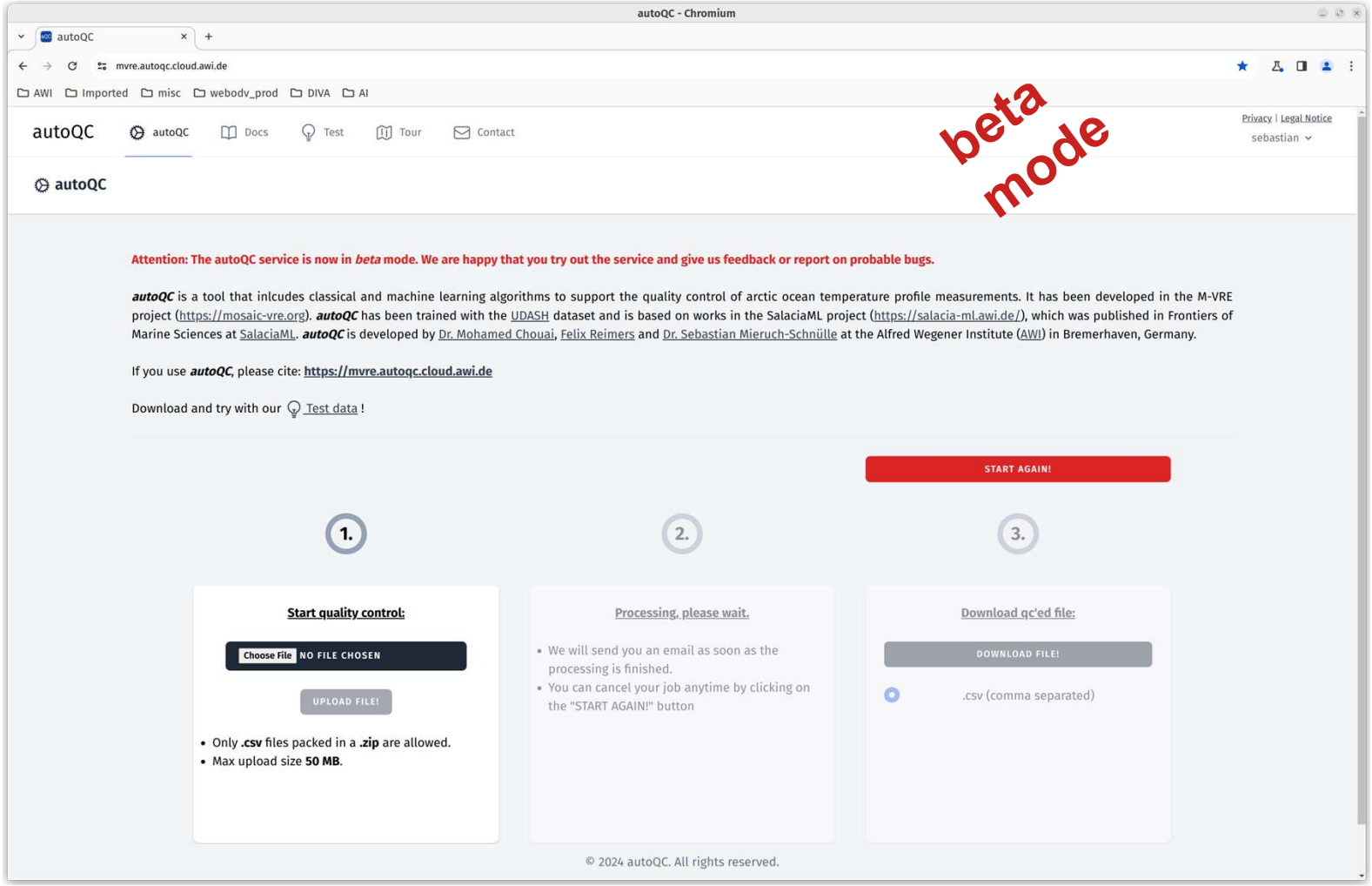
A new online app for ocean temperature quality control - An artificial intelligence approach

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

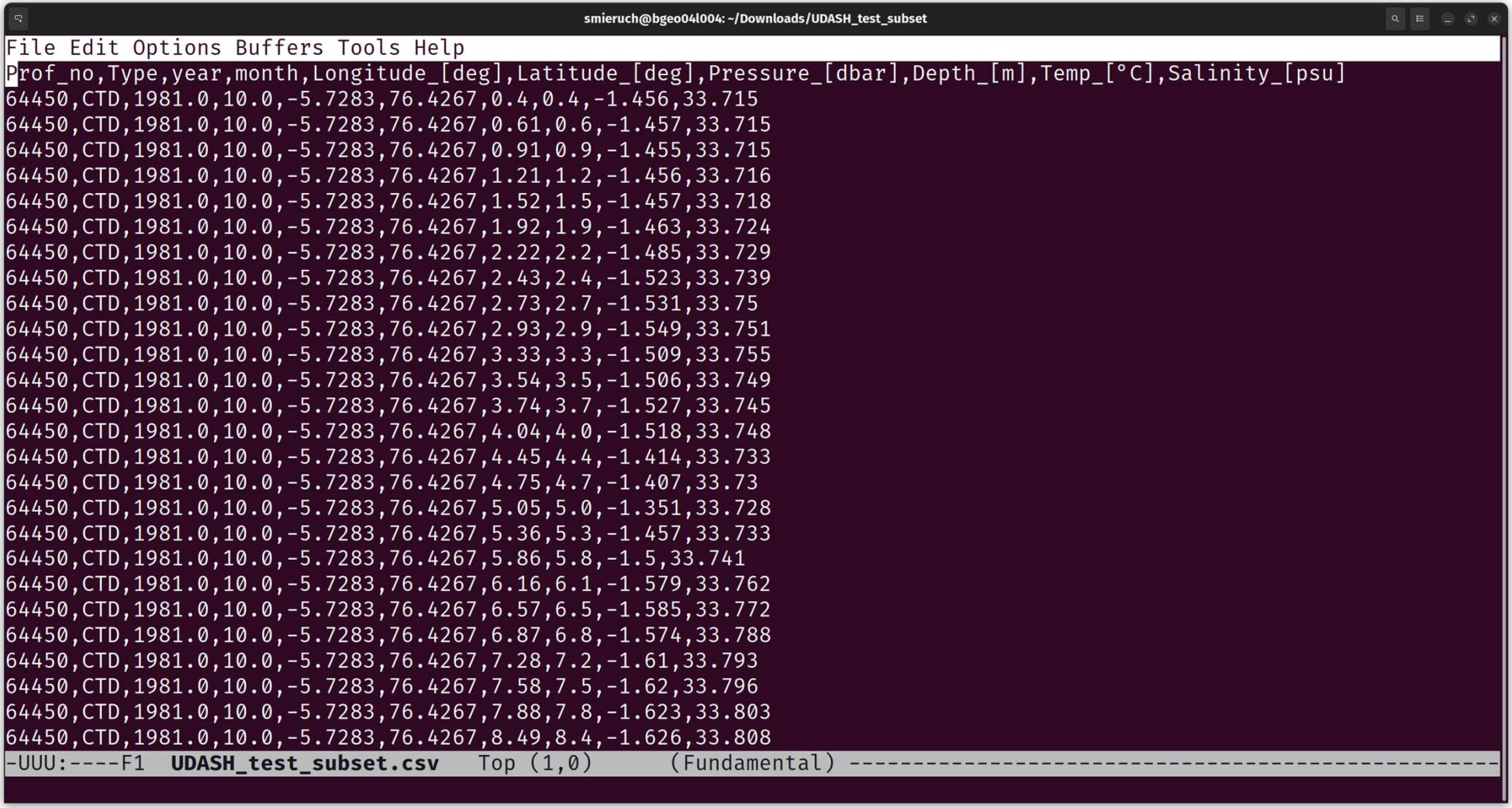
Bundesministerium
für Bildung
und Forschung

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

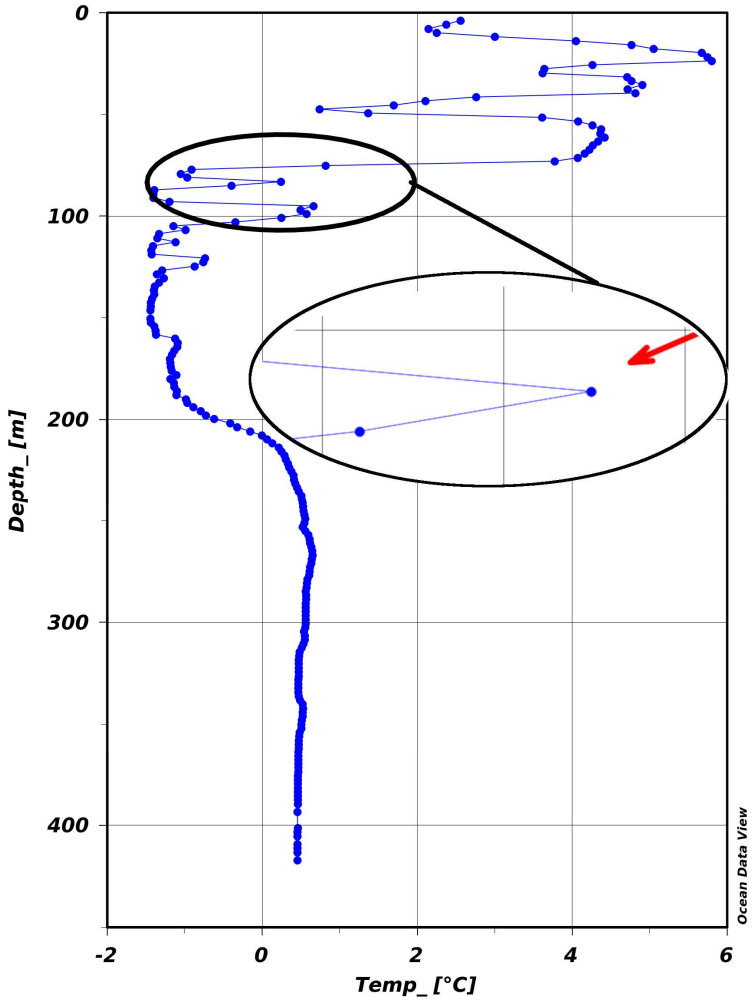
The app: https://mvre.autoqc.cloud.awi.de/

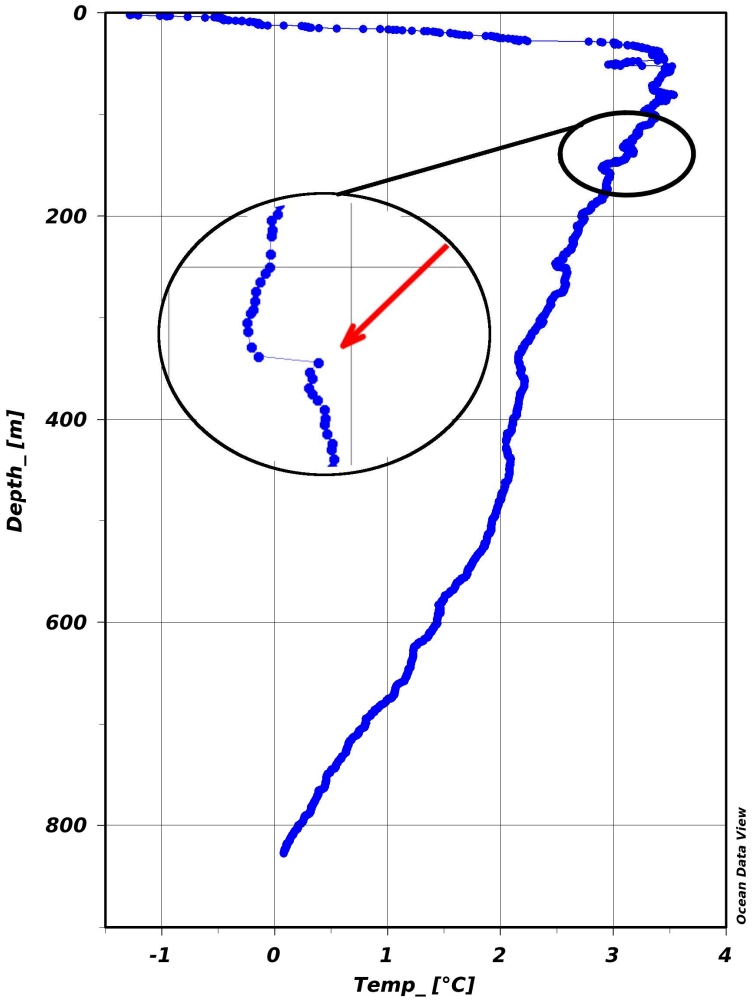


The input data

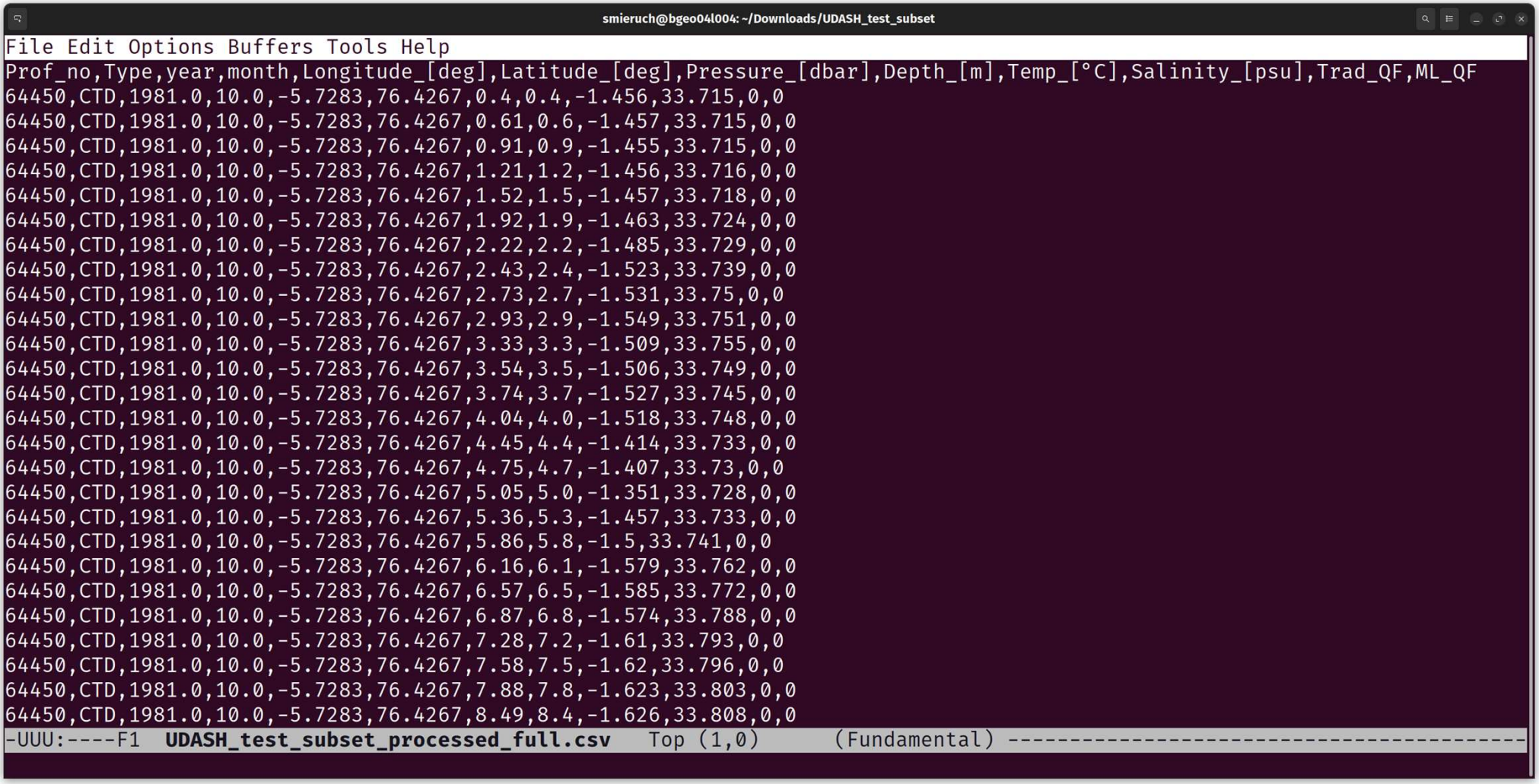


The flags: spike & suspect gradient

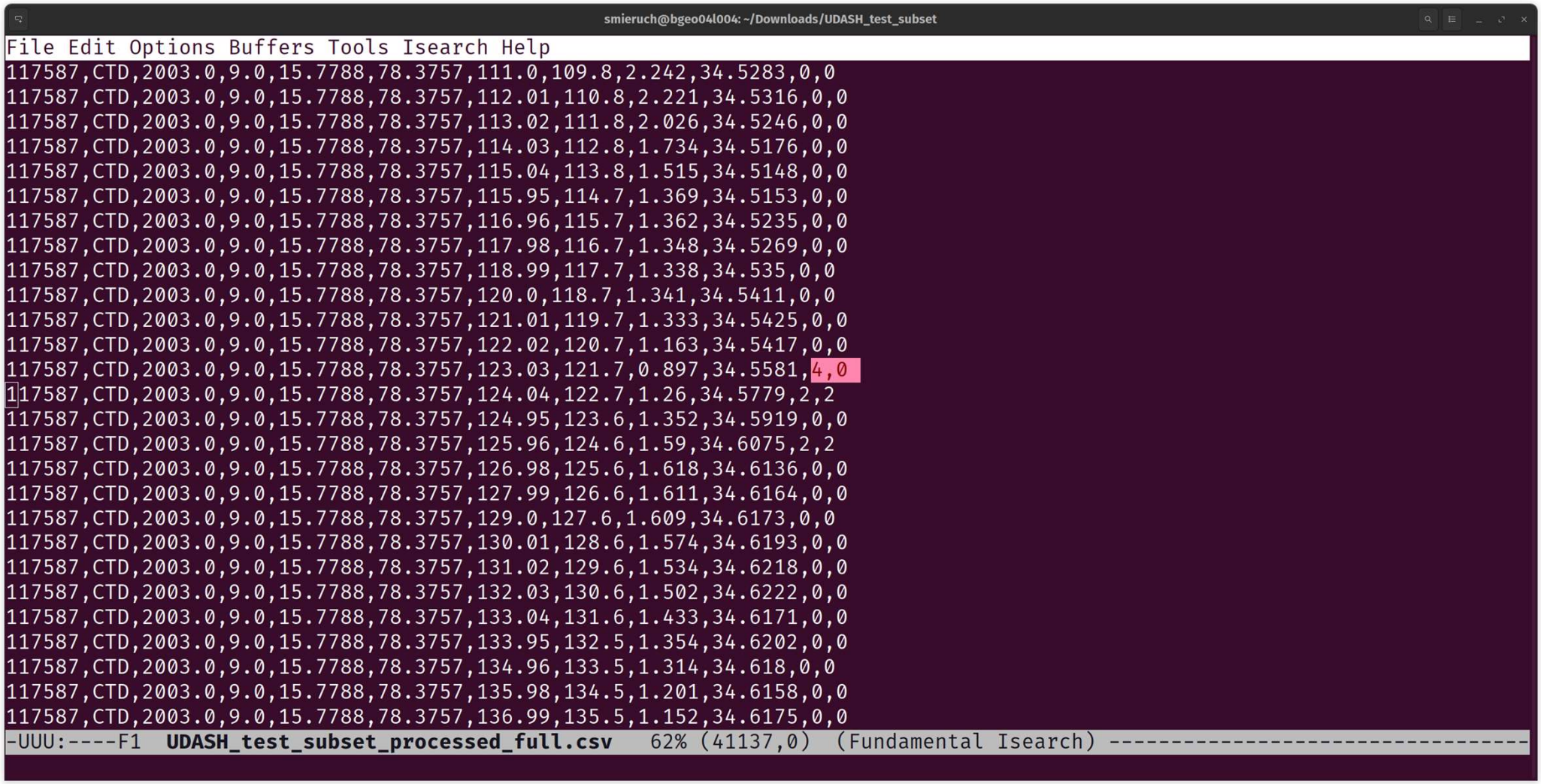




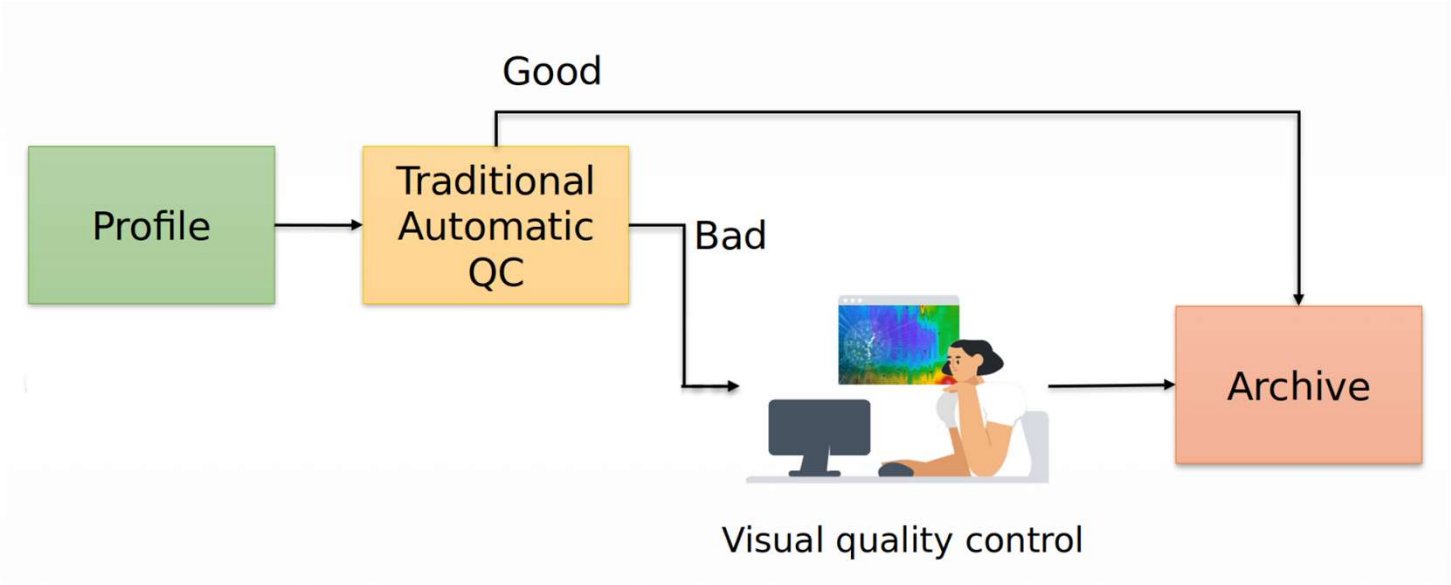
The output data



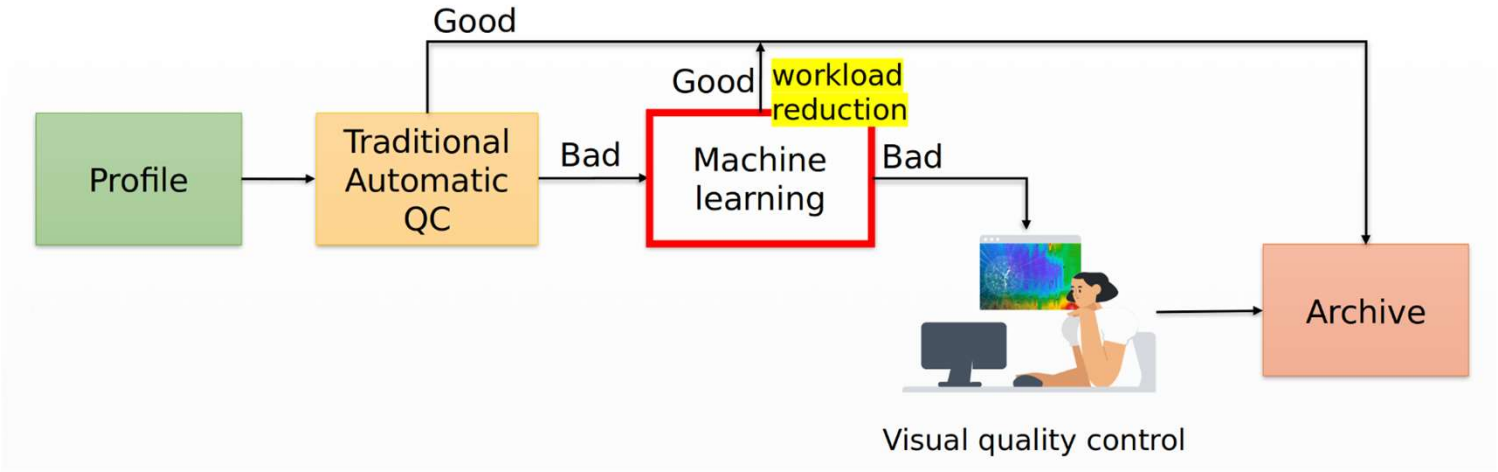
The AI


smieruch@bgeo04l004: ~/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

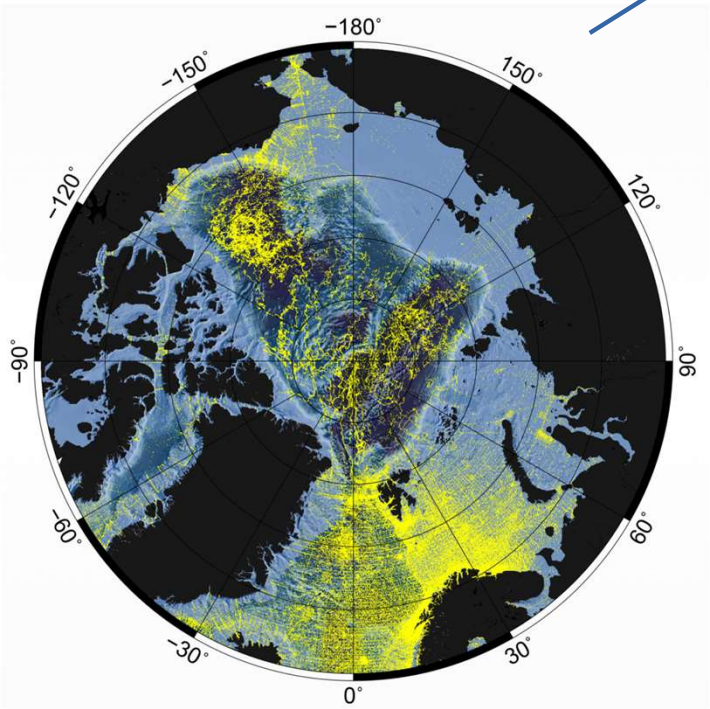
preliminar
v

How we did it

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

280,000 temperature and salinity profiles

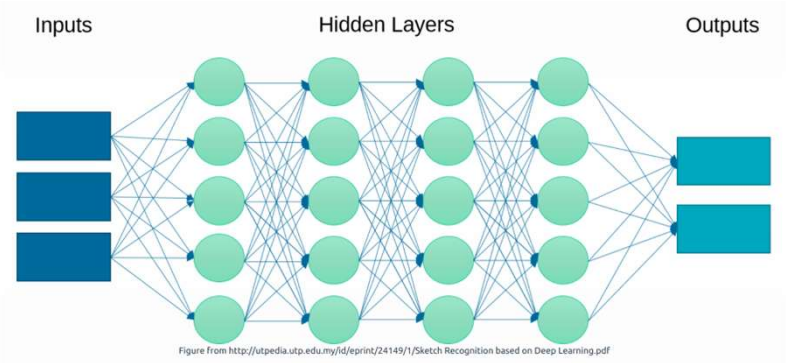
Visually controlled by human expert



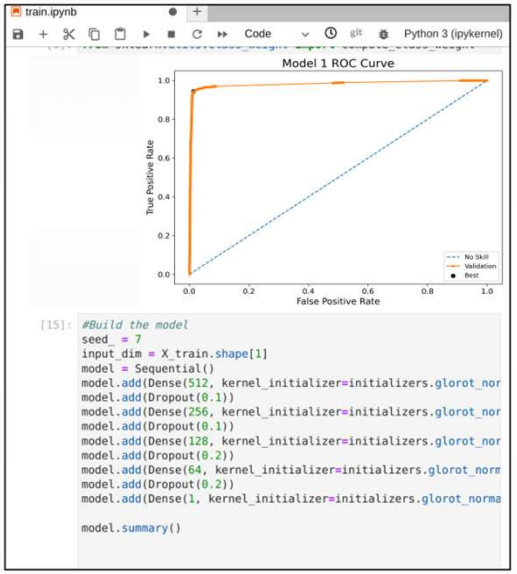
Classical QC checks

```
all_in_one.ipynb
Code
#                   Outliers in mixed layer
#
# ================================================================
def Traditional_outlier_detection(Data):
    gradient = T_Suspect_gradient_T_D(Data)
    Data['gradient2']=np.array(gradient)
    data1=Data.loc[(Data['gradient2']>=0.5) & (Data['Depth_[m]']<=100)]
    data1.loc[:,'QF_trad'] = 4
    return data1
def T_Suspect_gradient_T_D(Data):
    unique_profil=Data['Prof_no'].unique()
    d_grad=[]
    data2=[]
    for j in range(len(unique_profil)):
        profil=Data[Data["Prof_no"].isin([unique_profil[j]])].reset_index(drop=True)
        #data2=pd.concat([data2,profil])
        data2.append(profil.values)
        t=profil['Temp_[°C]'].values
        d=profil['Depth_[m]'].values
        i=0
        grad=[-999]
        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+=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)
# ================================================================
#
```

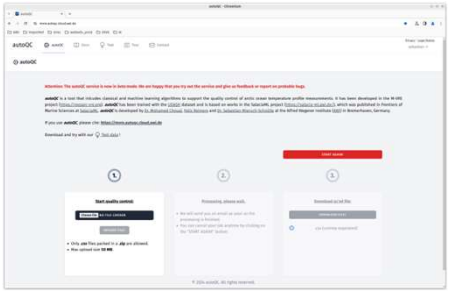
MLP (Multi-Layer-Perceptron)



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

SeaDataNet AISBL · Havforskningsinstituttet · Ifremer · MARIS · MEDIN · National Oceanography Centre · INGV · iMagine · IODE · eosc Blue-Cloud2026