Predicting Future Ocean Acidification

Importing Data and Packages

In [2]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
pd.plotting.register_matplotlib_converters()
import matplotlib.dates as mdates
from sklearn.feature_selection import chi2
from sklearn.model_selection import train_test_split
import numpy as np
from scipy import stats
from statsmodels.tsa.vector_ar.vecm import coint_johansen
from statsmodels.tsa.vector_ar.var_model import VAR
from statsmodels.tsa.stattools import adfuller
from statsmodels.tools.eval_measures import rmse, aic
import matplotlib.pyplot as plt
%matplotlib inline
from statsmodels.stats.stattools import durbin_watson
```

In [3]:
```
# reading CSV
df1 = pd.read_csv(r"C:\Users\datre\OneDrive\Documents\Graduate School\Spring '21\Project\2013_2014.csv")
```

In [4]:
```
# Looking at data types and numbers
print(df1.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4123 entries, 0 to 4122
Data columns (total 26 columns):
 #   Column                     Non-Null Count  Dtype  
---  ------                     --------------  -----  
 0   Mooring Name               4123 non-null   object 
 1   Latitude                   4123 non-null   float64
 2   Longitude                  4123 non-null   float64
 3   Date                       4123 non-null   object 
 4   Time                       4123 non-null   object 
 5   xCO2 SW (wet) (umol/mol)   4123 non-null   float64
 6   CO2 SW QF                  4123 non-null   int64  
 7   H2O SW (mmol/mol)          4123 non-null   float64
 8   xCO2 Air (wet) (umol/mol)  4123 non-null   float64
 9   CO2 Air QF                 4123 non-null   int64  
 10  H2O Air (mmol/mol)         4123 non-null   float64
 11  Licor Atm Pressure (hPa)   4123 non-null   float64
 12  Licor Temp (C)             4123 non-null   float64
 13  MAPCO2 %O2                 4123 non-null   float64
 14  SST (C)                    4123 non-null   float64
 15  Salinity                   4123 non-null   float64
 16  xCO2 SW (dry) (umol/mol)   4123 non-null   float64
 17  xCO2 Air (dry) (umol/mol)  4123 non-null   float64
 18  fCO2 SW (sat) uatm         4123 non-null   float64
 19  fCO2 Air (sat) uatm        4123 non-null   float64
 20  dfCO2                      4123 non-null   float64
 21  pCO2 SW (sat) uatm         4123 non-null   float64
 22  pCO2 Air (sat) uatm        4123 non-null   float64
 23  dpCO2                      4123 non-null   float64
 24  pH (Total Scale)           4123 non-null   float64
 25  pH QF                      4123 non-null   int64  
dtypes: float64(20), int64(3), object(3)
memory usage: 837.6+ KB
None
```

In [4]:
```
# Removing features where every instance has the same thing
# Data was taken from the same mooring, with the same latitude and longitude
del df1["Mooring Name"]
del df1["Latitude"]
del df1["Longitude"]
```

In [5]:
```
# creating one datetime column with seperate date and time columns
df1["Datetime"] = pd.to_datetime(df1["Date"] + " " + df1["Time"])
```

In [6]:
```
# setting index as datetime column
df1 = df1.set_index("Datetime")
```

In [7]:
```python
# grouping data by day versus by hour
df1 = df1.groupby(pd.Grouper(freq='1D')).mean()
df1.head(3)
```

Out[7]:

	Latitude	Longitude	xCO2 SW (wet) (umol/mol)	CO2 SW QF	H2O SW (mmol/mol)	xCO2 Air (wet) (umol/mol)	CO2 Air QF	H2O Air (mmol/mol)	Licor Atm Pressure (hPa)	Licor Temp (C)	...	xCO2 SW (dry) (umol/mol)	x((um
Datetime													
2013-04-24	7.464	151.898	397.9250	2.0	9.56000	394.375	2.0	8.70500	1006.9750	30.6750	...	401.7750	
2013-04-25	7.464	151.898	396.5750	2.0	10.11625	393.775	2.0	9.10875	1006.8125	30.9125	...	400.6250	
2013-04-26	7.464	151.898	398.0875	2.0	10.34625	393.600	2.0	9.18375	1006.7750	31.2625	...	402.2625	

3 rows × 23 columns

In [8]:
```python
# Checking for Null values
df1.isnull().sum()
```

Out[8]:
```
Latitude                   0
Longitude                  0
xCO2 SW (wet) (umol/mol)   0
CO2 SW QF                  0
H2O SW (mmol/mol)          0
xCO2 Air (wet) (umol/mol)  0
CO2 Air QF                 0
H2O Air (mmol/mol)         0
Licor Atm Pressure (hPa)   0
Licor Temp (C)             0
MAPCO2 %O2                 0
SST (C)                    0
Salinity                   0
xCO2 SW (dry) (umol/mol)   0
xCO2 Air (dry) (umol/mol)  0
fCO2 SW (sat) uatm         0
fCO2 Air (sat) uatm        0
dfCO2                      0
pCO2 SW (sat) uatm         0
pCO2 Air (sat) uatm        0
dpCO2                      0
pH (Total Scale)           0
pH QF                      0
dtype: int64
```

In [9]:
```python
# Checking for recording errors, either extreme highs or extreme lows
df1.describe()
```

Out[9]:

	Latitude	Longitude	xCO2 SW (wet) (umol/mol)	CO2 SW QF	H2O SW (mmol/mol)	xCO2 Air (wet) (umol/mol)	CO2 Air QF	H2O Air (mmol/mol)	Licor Atm Pressure (hPa)	Licor Temp (C)
count	5.160000e+02	5.160000e+02	516.000000	516.000000	516.000000	516.000000	516.000000	516.000000	516.000000	516.000000
mean	7.464000e+00	1.518980e+02	406.902962	2.014535	8.971283	381.955655	2.014535	8.165870	1006.099062	30.692577
std	3.467257e-14	7.396815e-13	46.034660	0.061589	3.128491	42.350616	0.061589	2.975783	1.378662	1.067312
min	7.464000e+00	1.518980e+02	63.300000	2.000000	3.516250	43.437500	2.000000	3.041429	999.387500	27.537500
25%	7.464000e+00	1.518980e+02	398.065625	2.000000	6.370625	389.371875	2.000000	5.702813	1005.262500	29.971875
50%	7.464000e+00	1.518980e+02	410.993750	2.000000	9.059375	391.293750	2.000000	8.245000	1006.125000	30.706250
75%	7.464000e+00	1.518980e+02	430.593750	2.000000	11.713750	392.862500	2.000000	10.827188	1007.090625	31.450000
max	7.464000e+00	1.518980e+02	463.575000	2.500000	14.833750	398.862500	2.500000	13.052500	1008.925000	33.162500

8 rows × 23 columns

In [10]:
```
chuuk = df1
```

In [11]:
```
# Removing all extreme value instances
chuuk = chuuk[chuuk["xCO2 SW (wet) (umol/mol)"] != -999]
chuuk = chuuk[chuuk["xCO2 Air (wet) (umol/mol)"] != -999]
chuuk = chuuk[chuuk["SST (C)"] != -999]
chuuk = chuuk[(chuuk["SST (C)"]>0) & (chuuk["pH (Total Scale)"]>0) & (chuuk["dfCO2"]>0) & (chuuk["dpCO2"]>0)]
chuuk = chuuk[chuuk["xCO2 SW (dry) (umol/mol)"] != -999]
chuuk = chuuk[chuuk["xCO2 Air (dry) (umol/mol)"] != -999]
chuuk = chuuk[chuuk["fCO2 SW (sat) uatm"] != -999]
chuuk = chuuk[chuuk["fCO2 Air (sat) uatm"] != -999]
chuuk = chuuk[chuuk["dfCO2"] != -999]
chuuk = chuuk[chuuk["dpCO2"] != -999]
chuuk = chuuk[chuuk["pH (Total Scale)"] != -999]
# Dropping QF values
chuuk = chuuk.drop(["CO2 SW QF", "CO2 Air QF"], axis = 1)
```

In [12]:
```
# checking that all extreme values were removed
chuuk.describe()
```

Out[12]:

	Latitude	Longitude	xCO2 SW (wet) (umol/mol)	H2O SW (mmol/mol)	xCO2 Air (wet) (umol/mol)	H2O Air (mmol/mol)	Licor Atm Pressure (hPa)	Licor Temp (C)	MAPCO2 %O2	SST (C)
count	3.790000e+02	3.790000e+02	379.000000	379.000000	379.000000	379.000000	379.000000	379.000000	379.000000	379.000000
mean	7.464000e+00	1.518980e+02	416.996735	9.524825	391.358212	8.681078	1006.205541	30.793701	99.151405	29.439762
std	2.668057e-14	6.545634e-13	17.664458	3.127114	2.358320	2.942537	1.282512	1.037521	0.428422	0.489045
min	7.464000e+00	1.518980e+02	390.600000	4.002500	386.837500	3.295000	999.387500	27.937500	97.632500	28.331125
25%	7.464000e+00	1.518980e+02	400.931250	6.601875	389.400000	5.907500	1005.393750	30.131250	98.908125	29.089437
50%	7.464000e+00	1.518980e+02	413.337500	9.965000	391.575000	9.195000	1006.212500	30.800000	99.158750	29.429125
75%	7.464000e+00	1.518980e+02	430.687500	12.386875	393.006250	11.499375	1007.125000	31.493750	99.441250	29.799625
max	7.464000e+00	1.518980e+02	463.575000	14.833750	398.862500	13.052500	1008.925000	33.162500	100.767500	30.775875

8 rows × 21 columns

In [13]:
```
# Copying dataset
chuuk_corr = chuuk.copy()
```

In [14]:
```
chuuk_corr.describe()
```

Out[14]:

	Latitude	Longitude	xCO2 SW (wet) (umol/mol)	H2O SW (mmol/mol)	xCO2 Air (wet) (umol/mol)	H2O Air (mmol/mol)	Licor Atm Pressure (hPa)	Licor Temp (C)	MAPCO2 %O2	SST (C)
count	3.790000e+02	3.790000e+02	379.000000	379.000000	379.000000	379.000000	379.000000	379.000000	379.000000	379.000000
mean	7.464000e+00	1.518980e+02	416.996735	9.524825	391.358212	8.681078	1006.205541	30.793701	99.151405	29.439762
std	2.668057e-14	6.545634e-13	17.664458	3.127114	2.358320	2.942537	1.282512	1.037521	0.428422	0.489045
min	7.464000e+00	1.518980e+02	390.600000	4.002500	386.837500	3.295000	999.387500	27.937500	97.632500	28.331125
25%	7.464000e+00	1.518980e+02	400.931250	6.601875	389.400000	5.907500	1005.393750	30.131250	98.908125	29.089437
50%	7.464000e+00	1.518980e+02	413.337500	9.965000	391.575000	9.195000	1006.212500	30.800000	99.158750	29.429125
75%	7.464000e+00	1.518980e+02	430.687500	12.386875	393.006250	11.499375	1007.125000	31.493750	99.441250	29.799625
max	7.464000e+00	1.518980e+02	463.575000	14.833750	398.862500	13.052500	1008.925000	33.162500	100.767500	30.775875

8 rows × 21 columns

In [15]:
```
# Creating correlation matrix
corr_matrix = chuuk_corr.corr()
corr_matrix
```

Out[15]:

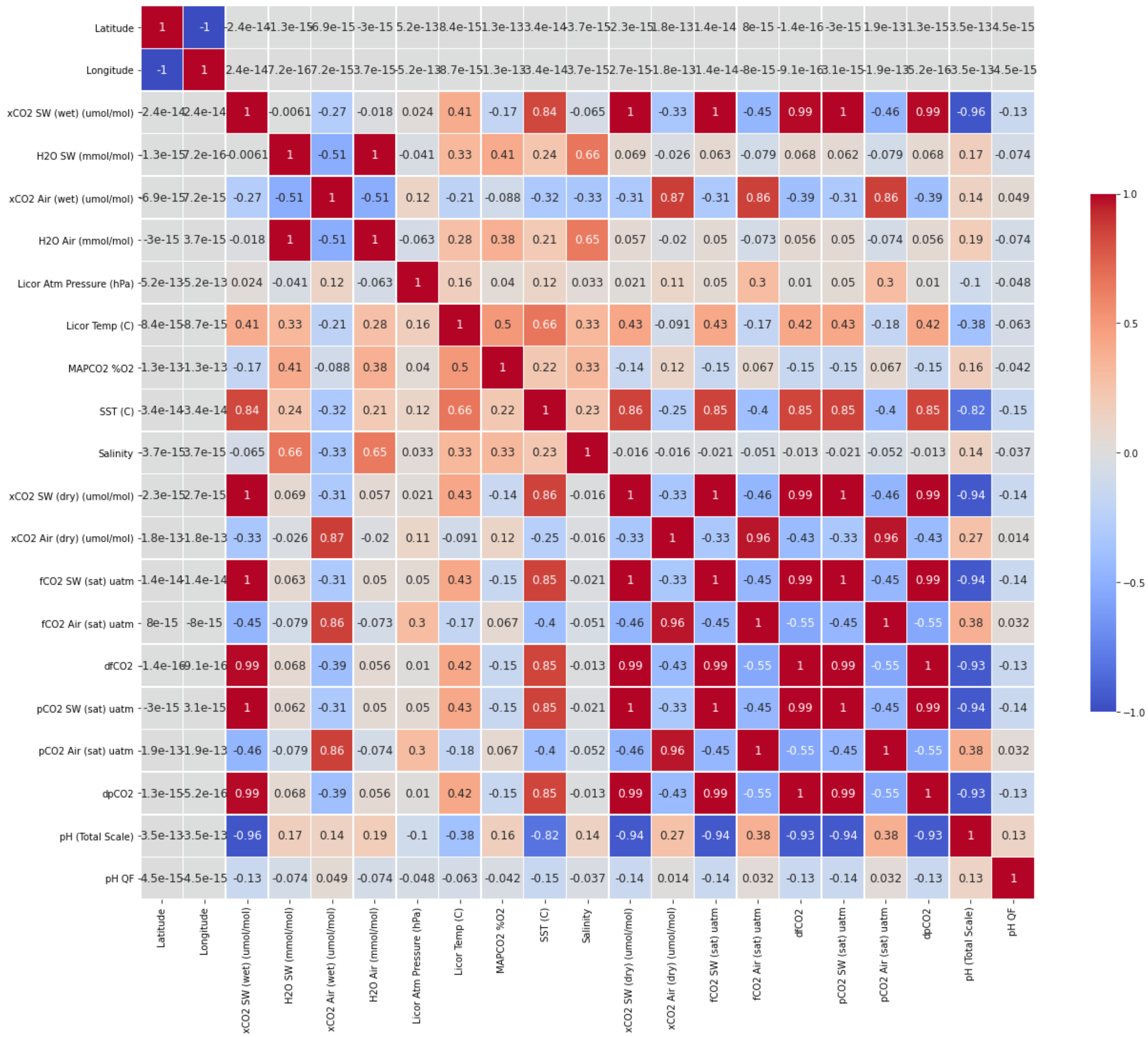
	Latitude	Longitude	xCO2 SW (wet) (umol/mol)	H2O SW (mmol/mol)	xCO2 Air (wet) (umol/mol)	H2O Air (mmol/mol)	Licor Atm Pressure (hPa)	Licor Temp (C)	I
Latitude	1.000000e+00	-1.000000e+00	-2.429664e-14	-1.339909e-15	-6.852107e-15	-2.982314e-15	5.213248e-13	8.414180e-15	1.25
Longitude	-1.000000e+00	1.000000e+00	2.442846e-14	7.177709e-16	7.151686e-15	3.742963e-15	-5.213147e-13	-8.666888e-15	-1.2
xCO2 SW (wet) (umol/mol)	-2.429664e-14	2.442846e-14	1.000000e+00	-6.065765e-03	-2.730888e-01	-1.793784e-02	2.442229e-02	4.062362e-01	-1.7
H2O SW (mmol/mol)	-1.339909e-15	7.177709e-16	-6.065765e-03	1.000000e+00	-5.132088e-01	9.960826e-01	-4.135096e-02	3.332468e-01	4.114
xCO2 Air (wet) (umol/mol)	-6.852107e-15	7.151686e-15	-2.730888e-01	-5.132088e-01	1.000000e+00	-5.100476e-01	1.221732e-01	-2.138960e-01	-8.8
H2O Air (mmol/mol)	-2.982314e-15	3.742963e-15	-1.793784e-02	9.960826e-01	-5.100476e-01	1.000000e+00	-6.266097e-02	2.766740e-01	3.82
Licor Atm Pressure (hPa)	5.213248e-13	-5.213147e-13	2.442229e-02	-4.135096e-02	1.221732e-01	-6.266097e-02	1.000000e+00	1.644344e-01	3.98
Licor Temp (C)	8.414180e-15	-8.666888e-15	4.062362e-01	3.332468e-01	-2.138960e-01	2.766740e-01	1.644344e-01	1.000000e+00	5.03
MAPCO2 %O2	1.253603e-13	-1.252280e-13	-1.748042e-01	4.114359e-01	-8.814678e-02	3.821659e-01	3.983958e-02	5.033865e-01	1.000
SST (C)	3.363916e-14	-3.361552e-14	8.417755e-01	2.369253e-01	-3.200275e-01	2.104537e-01	1.166726e-01	6.572118e-01	2.19
Salinity	-3.747761e-15	3.710776e-15	-6.453285e-02	6.636009e-01	-3.326869e-01	6.461309e-01	3.318696e-02	3.317037e-01	3.25
xCO2 SW (dry) (umol/mol)	-2.282160e-15	2.678035e-15	9.971867e-01	6.885311e-02	-3.114377e-01	5.670892e-02	2.112733e-02	4.301949e-01	-1.4
xCO2 Air (dry) (umol/mol)	1.811631e-13	-1.812928e-13	-3.284913e-01	-2.579151e-02	8.700751e-01	-1.984712e-02	1.059009e-01	-9.062602e-02	1.16
fCO2 SW (sat) uatm	1.370838e-14	-1.360384e-14	9.971339e-01	6.254226e-02	-3.059022e-01	5.018098e-02	5.026538e-02	4.268245e-01	-1.5
fCO2 Air (sat) uatm	8.024809e-15	-7.979930e-15	-4.537848e-01	-7.866870e-02	8.630614e-01	-7.285439e-02	2.978609e-01	-1.747817e-01	6.73
dfCO2	-1.385843e-16	-9.104510e-16	9.917377e-01	6.834997e-02	-3.938823e-01	5.604111e-02	1.029266e-02	4.220163e-01	-1.5
pCO2 SW (sat) uatm	-2.994030e-15	3.123264e-15	9.971347e-01	6.242631e-02	-3.058729e-01	5.007342e-02	5.037229e-02	4.266586e-01	-1.5
pCO2 Air (sat) uatm	1.934314e-13	-1.934268e-13	-4.555326e-01	-7.934766e-02	8.630323e-01	-7.353538e-02	2.977787e-01	-1.759563e-01	6.72
dpCO2	1.312365e-15	-5.178742e-16	9.917409e-01	6.835485e-02	-3.938582e-01	5.604738e-02	1.038784e-02	4.219782e-01	-1.5
pH (Total Scale)	3.525734e-13	-3.526486e-13	-9.552045e-01	1.708242e-01	1.416996e-01	1.860399e-01	-1.006334e-01	-3.797438e-01	1.64
pH QF	4.491916e-15	-4.463443e-15	-1.330350e-01	-7.422357e-02	4.878880e-02	-7.427450e-02	-4.801234e-02	-6.347305e-02	-4.2

21 rows × 21 columns

In [16]:

```
# setting figure size
f, ax = plt.subplots(figsize=(21,25))
heatmap = sns.heatmap(corr_matrix,
                      square = True,
                      linewidths = .5,
```

```
                  cmap = "coolwarm", # choosing color type for graph
                  cbar_kws = {"shrink": .4,
                              "ticks" : [-1, -.5, 0, 0.5, 1]},
                  vmin = -1,
                  vmax = 1,
                  annot = True,
                  annot_kws = {"size": 12})

#add the column names as labels
ax.set_yticklabels(corr_matrix.columns, rotation = 0)
ax.set_xticklabels(corr_matrix.columns)
sns.set_style({'xtick.bottom': True}, {'ytick.left': True})
```




```
In [116...
chuuk_df = chuuk.copy()
chuuk_df
```

Out[116...

	Latitude	Longitude	xCO2 SW (wet) (umol/mol)	H2O SW (mmol/mol)	xCO2 Air (wet) (umol/mol)	H2O Air (mmol/mol)	Licor Atm Pressure (hPa)	Licor Temp (C)	MAPCO2 %O2	SST (C)	...	xCO2 ((umol/
Datetime												
2013-04-24	7.464	151.898	397.9250	9.56000	394.3750	8.70500	1006.9750	30.6750	99.67750	28.855500	...	401.7
2013-04-25	7.464	151.898	396.5750	10.11625	393.7750	9.10875	1006.8125	30.9125	99.57625	28.933125	...	400.6
2013-04-26	7.464	151.898	398.0875	10.34625	393.6000	9.18375	1006.7750	31.2625	99.49250	28.989875	...	402.2

	Latitude	Longitude	xCO2 SW (wet) (umol/mol)	H2O SW (mmol/mol)	xCO2 Air (wet) (umol/mol)	H2O Air (mmol/mol)	Licor Atm Pressure (hPa)	Licor Temp (C)	MAPCO2 %O2	SST (C)	...	xCO2 ((umol/
Datetime												
2013-04-27	7.464	151.898	399.6750	10.49250	393.6625	9.42500	1006.7125	31.1000	99.44250	28.981000	...	403.
2013-04-28	7.464	151.898	395.5625	10.24000	393.0375	9.39875	1007.1500	30.6375	99.57750	28.957750	...	399.
...	...	...	...	...	...	...	...	...	...	...	...	
2014-08-04	7.464	151.898	437.6375	4.93750	389.7375	3.80625	1008.1000	33.1125	99.15875	30.236125	...	439.
2014-08-05	7.464	151.898	435.6250	5.13125	391.2875	3.98875	1007.7000	32.1500	99.17625	30.186875	...	437.
2014-08-08	7.464	151.898	415.6625	4.00250	390.9625	3.29500	1008.0250	30.1875	99.24625	29.501750	...	417.
2014-08-09	7.464	151.898	430.1250	4.17125	392.2500	3.63750	1006.9750	28.7625	98.73875	29.435375	...	431.
2014-08-10	7.464	151.898	437.9625	4.35625	393.3250	3.76375	1006.3750	29.7625	98.43250	29.442875	...	439.

379 rows × 21 columns

In [18]:
```python
# Dropping all features that did not have significant correlation with target, pH
chuuk_df = chuuk_df.drop(["H2O SW (mmol/mol)", "xCO2 Air (wet) (umol/mol)", "H2O Air (mmol/mol)",
                    "Licor Temp (C)", "MAPCO2 %O2", "Salinity", "xCO2 Air (dry) (umol/mol)", "fCO2 Air (sat) uatm",
                    "pCO2 Air (sat) uatm", "pH QF"],
                   axis = 1)
chuuk_df.head(3)
```

Out[18]:

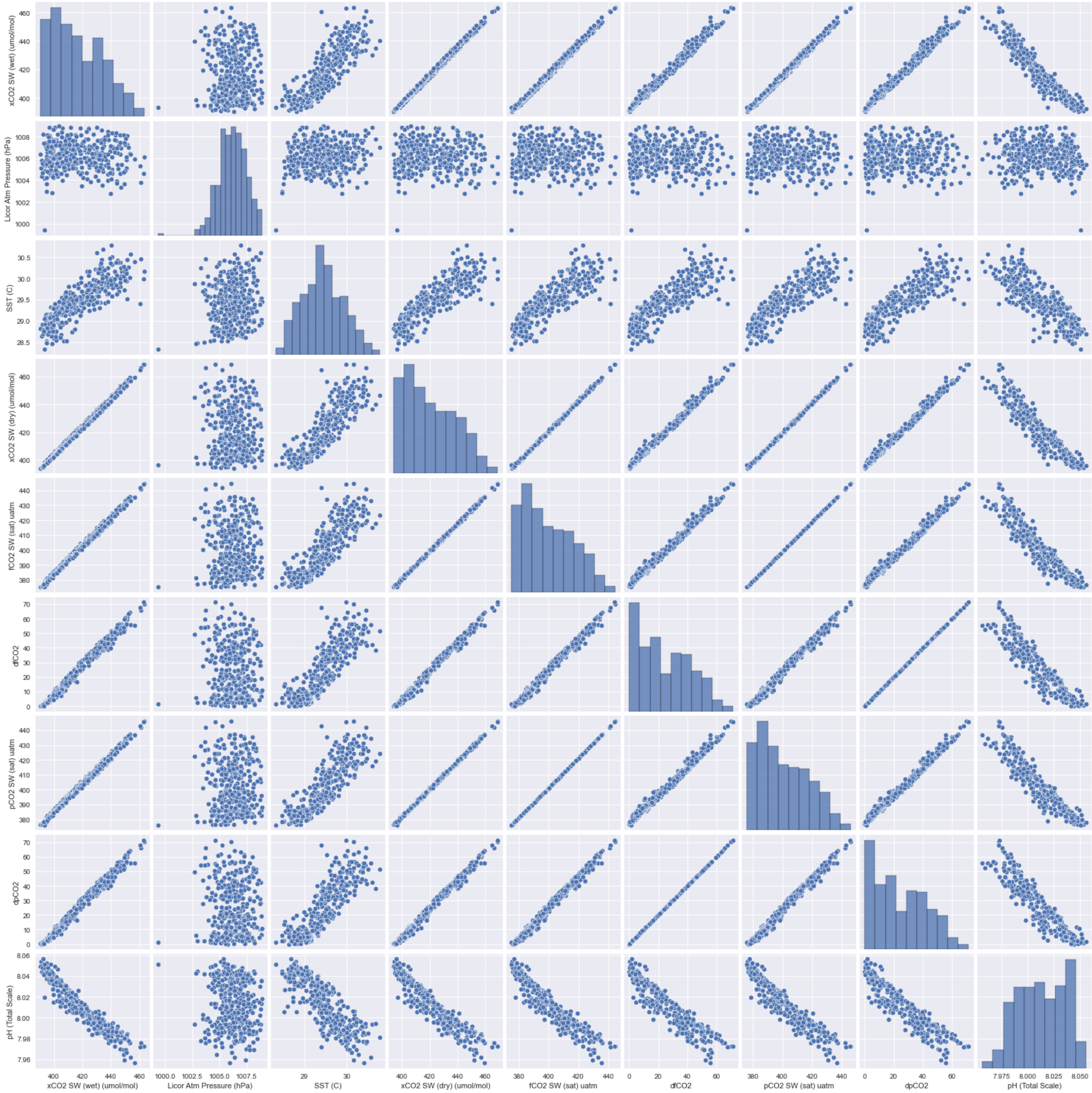
	Latitude	Longitude	xCO2 SW (wet) (umol/mol)	Licor Atm Pressure (hPa)	SST (C)	xCO2 SW (dry) (umol/mol)	fCO2 SW (sat) uatm	dfCO2	pCO2 SW (sat) uatm	dpCO2	pH (Total Scale)
Datetime											
2013-04-24	7.464	151.898	397.9250	1006.9750	28.855500	401.7750	382.6750	3.7000	383.8250	3.7250	8.03300
2013-04-25	7.464	151.898	396.5750	1006.8125	28.933125	400.6250	381.4250	3.0500	382.5875	3.0625	8.04000
2013-04-26	7.464	151.898	398.0875	1006.7750	28.989875	402.2625	382.9375	4.7875	384.0875	4.8000	8.03475

In [24]:
```python
chuuk_df = chuuk_df.drop(["Latitude", "Longitude"], axis = 1)
```

In [19]:
```python
# creating easier reference for features
xCO2w = chuuk_df["xCO2 SW (wet) (umol/mol)"]
AtmP = chuuk_df["Licor Atm Pressure (hPa)"]
SST = chuuk_df["SST (C)"]
xCO2d = chuuk_df["xCO2 SW (dry) (umol/mol)"]
fCO2 = chuuk_df["fCO2 SW (sat) uatm"]
dfCO2 = chuuk_df["dfCO2"]
pCO2 = chuuk_df["pCO2 SW (sat) uatm"]
dpCO2 = chuuk_df["dpCO2"]
pH = chuuk_df["pH (Total Scale)"]
```

In [25]:
```python
# Pairplot to see correaltions and histograms
sns.pairplot(chuuk_df)
```

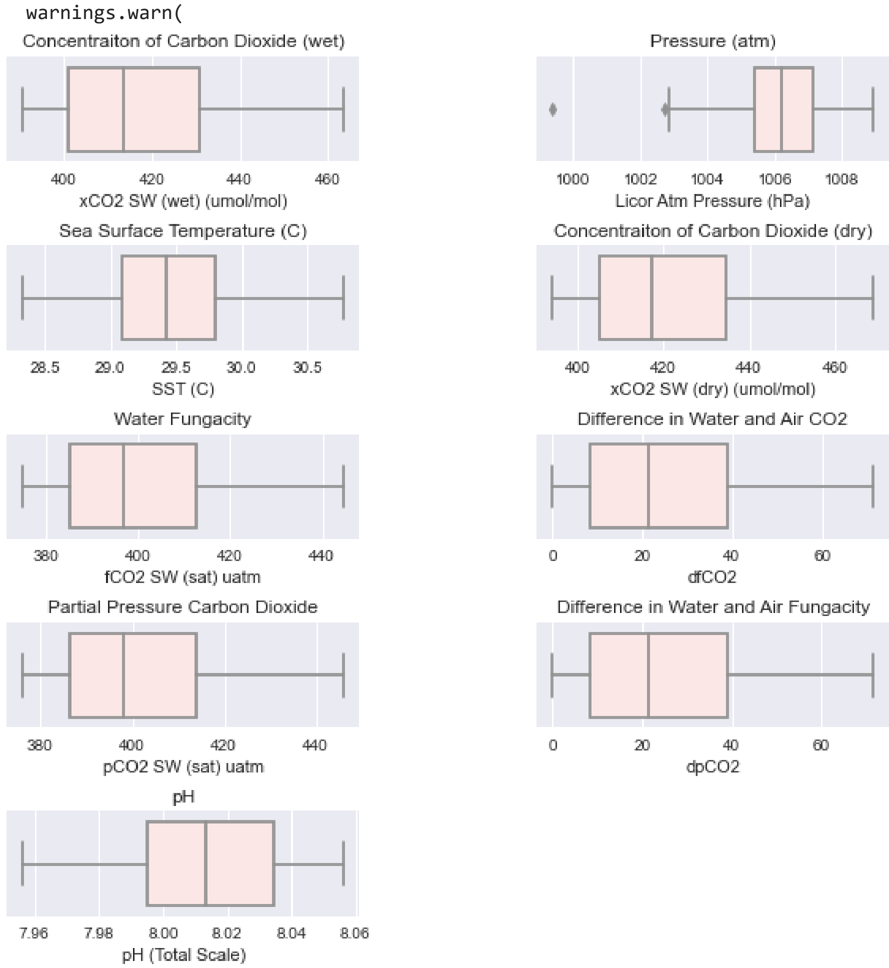
Out[25]: <seaborn.axisgrid.PairGrid at 0x1f24d065d90>



In [26]:
```python
# Subplot for boxplots of each feature
plt.style.use('seaborn')
fig, axis = plt.subplots(nrows = 5, ncols = 2)
fig.set_size_inches(10,10)
fig.subplots_adjust(wspace = 0.5, hspace = 0.8)

# plot 1
sns.boxplot(xCO2w, color = 'mistyrose', ax = axis[0,0]).set_title("Concentraiton of Carbon Dioxide (wet)")
# plot 2
sns.boxplot(AtmP, color = 'mistyrose', ax = axis[0,1]).set_title("Pressure (atm)")
# plot 3
sns.boxplot(SST, color = 'mistyrose', ax = axis[1,0]).set_title("Sea Surface Temperature (C)")
# plot 4
sns.boxplot(xCO2d, color = 'mistyrose', ax = axis[1,1]).set_title("Concentraiton of Carbon Dioxide (dry)")
# plot 5
sns.boxplot(fCO2, color = 'mistyrose', ax = axis[2,0]).set_title("Water Fungacity")
# plot 6
sns.boxplot(dfCO2, color = 'mistyrose', ax = axis[2,1]).set_title("Difference in Water and Air CO2")
# plot 7
sns.boxplot(pCO2, color = 'mistyrose', ax = axis[3,0]).set_title("Partial Pressure Carbon Dioxide")
# plot 8
sns.boxplot(dpCO2, color = 'mistyrose', ax = axis[3,1]).set_title("Difference in Water and Air Fungacity")
# plot 9
sns.boxplot(pH, color = 'mistyrose', ax = axis[4,0]).set_title("pH")
```

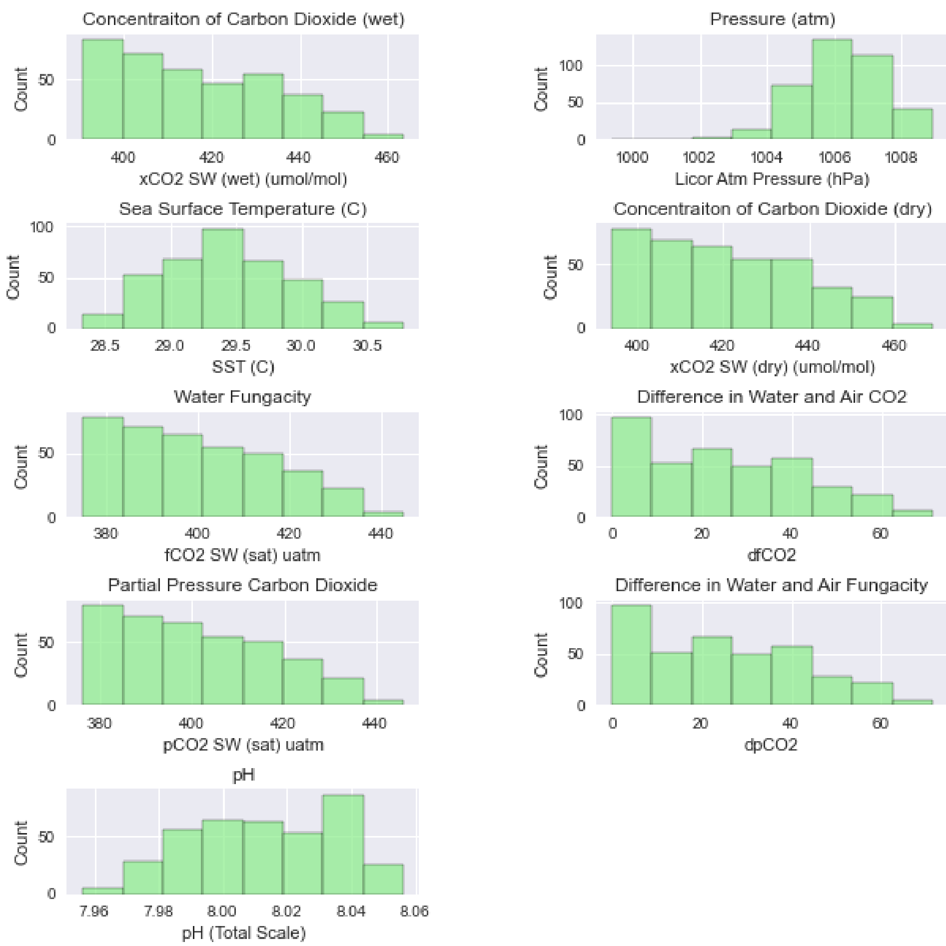
```
ax = axis[4,1].set_visible(False)
```

```
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
```



In [27]:
```
# Subplot for boxplots of each feature
```

```
plt.style.use('seaborn')
fig, axis = plt.subplots(nrows = 5, ncols = 2)
fig.set_size_inches(10,10)
fig.subplots_adjust(wspace = 0.5, hspace = 0.8)

# plot 1
sns.histplot(xCO2w, color = 'lightgreen', bins = 8, ax = axis[0,0]).set_title("Concentraiton of Carbon Dioxide (wet
# plot 2
sns.histplot(AtmP, color = 'lightgreen', bins = 8, ax = axis[0,1]).set_title("Pressure (atm)")
# plot 3
sns.histplot(SST, color = 'lightgreen', bins = 8, ax = axis[1,0]).set_title("Sea Surface Temperature (C)")
# plot 4
sns.histplot(xCO2d, color = 'lightgreen', bins = 8, ax = axis[1,1]).set_title("Concentraiton of Carbon Dioxide (dry
# plot 5
sns.histplot(fCO2, color = 'lightgreen', bins = 8, ax = axis[2,0]).set_title("Water Fungacity")
# plot 6
sns.histplot(dfCO2, color = 'lightgreen', bins = 8, ax = axis[2,1]).set_title("Difference in Water and Air CO2")
# plot 7
sns.histplot(pCO2, color = 'lightgreen', bins = 8, ax = axis[3,0]).set_title("Partial Pressure Carbon Dioxide")
# plot 8
sns.histplot(dpCO2, color = 'lightgreen', bins = 8, ax = axis[3,1]).set_title("Difference in Water and Air Fungacit
# plot 9
sns.histplot(pH, color = 'lightgreen', bins = 8, ax = axis[4,0]).set_title("pH")

ax = axis[4,1].set_visible(False)
```

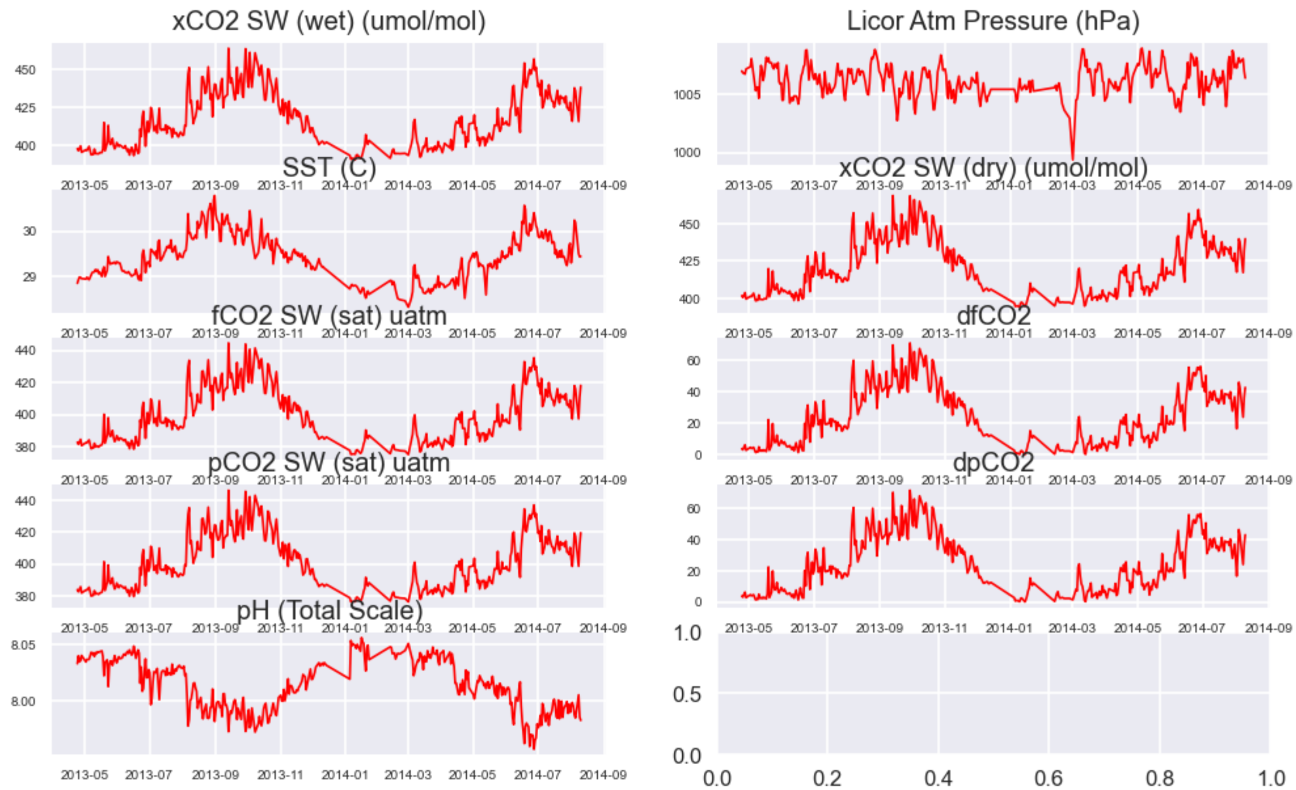


In [28]:

```
# looking at trends of all data
fig, axes = plt.subplots(nrows=5, ncols=2, dpi=120, figsize=(10,6), )
for i, ax in enumerate(axes.flatten()):
    data = chuuk_df[chuuk_df.columns[i]]
    ax.plot(data, color='red', linewidth=1)
    # Decorations
    ax.set_title(chuuk_df.columns[i])
    ax.xaxis.set_ticks_position('none')
    ax.yaxis.set_ticks_position('none')
    ax.spines["top"].set_alpha(0)
    ax.tick_params(labelsize=6)
```

```
subplots.adjust(wspace=4, hspace=20)
plt.show();
```

```
---------------------------------------------------------------------------
IndexError                                Traceback (most recent call last)
<ipython-input-28-5383668cfdfb> in <module>
      2 fig, axes = plt.subplots(nrows=5, ncols=2, dpi=120, figsize=(10,6), )
      3 for i, ax in enumerate(axes.flatten()):
----> 4     data = chuuk_df[chuuk_df.columns[i]]
      5     ax.plot(data, color='red', linewidth=1)
      6     # Decorations

~\anaconda3\lib\site-packages\pandas\core\indexes\base.py in __getitem__(self, key)
   4099         if is_scalar(key):
   4100             key = com.cast_scalar_indexer(key, warn_float=True)
-> 4101             return getitem(key)
   4102 
   4103         if isinstance(key, slice):

IndexError: index 9 is out of bounds for axis 0 with size 9
```



In [29]:

```
# setting subplots to look at different ways pH can be grouped
fig = plt.figure(figsize=(18,16))
fig.subplots_adjust(hspace=.4)
# Danily average
ax1 = fig.add_subplot(5,1,1)
ax1.plot(chuuk_df['pH (Total Scale)'].resample('D').mean(),linewidth=1)
ax1.set_title('Mean pH (total scale) resampled over day')
ax1.tick_params(axis='both', which='major')
# Weekly Average
ax2 = fig.add_subplot(5,1,2, sharex=ax1)
ax2.plot(chuuk_df['pH (Total Scale)'].resample('W').mean(),linewidth=1)
ax2.set_title('Mean pH (total scale) resampled over week')
ax2.tick_params(axis='both', which='major')
# Monthly Average
ax3 = fig.add_subplot(5,1,3, sharex=ax1)
ax3.plot(chuuk_df['pH (Total Scale)'].resample('M').mean(),linewidth=1)
ax3.set_title('Mean pH (total scale) resampled over month')
ax3.tick_params(axis='both', which='major')
# Quarterly Average
ax4  = fig.add_subplot(5,1,4, sharex=ax1)
ax4.plot(chuuk_df['pH (Total Scale)'].resample('Q').mean(),linewidth=1)
ax4.set_title('Mean pH (total scale) resampled over quarter')
ax4.tick_params(axis='both', which='major')
# Yearly Average
ax5  = fig.add_subplot(5,1,5, sharex=ax1)
```

```
ax5.plot(chuuk_df['pH (Total Scale)'].resample('A').mean(),linewidth=1)
ax5.set_title('Mean pH (total scale) resampled over year')
ax5.tick_params(axis='both', which='major');
```

In [26]:
```
chuuk_df.index = pd.DatetimeIndex(chuuk_df.index).to_period("D")
chuuk_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
PeriodIndex: 379 entries, 2013-04-24 to 2014-08-10
Freq: D
Data columns (total 9 columns):
 #   Column                     Non-Null Count  Dtype  
---  ------                     --------------  -----  
 0   xCO2 SW (wet) (umol/mol)   379 non-null    float64
 1   Licor Atm Pressure (hPa)   379 non-null    float64
 2   SST (C)                    379 non-null    float64
 3   xCO2 SW (dry) (umol/mol)   379 non-null    float64
 4   fCO2 SW (sat) uatm         379 non-null    float64
 5   dfCO2                      379 non-null    float64
 6   pCO2 SW (sat) uatm         379 non-null    float64
 7   dpCO2                      379 non-null    float64
 8   pH (Total Scale)           379 non-null    float64
dtypes: float64(9)
memory usage: 29.6 KB
```

In [30]:
```
# Splitting test and train set
train = chuuk_df[:int(0.8*(len(chuuk_df)))]
valid = chuuk_df[int(0.8*(len(chuuk_df))):]
print(train.shape)
print(valid.shape)
```

```
(303, 9)
```

```
(76, 9)
```

In [31]:

```
def cointegration_test(df, alpha=0.05):
    """Perform Johanson's Cointegration Test and Report Summary"""
    out = coint_johansen(df,-1,5)
    d = {'0.90':0, '0.95':1, '0.99':2}
    traces = out.lr1
    cvts = out.cvt[:, d[str(1-alpha)]]
    def adjust(val, length= 6): return str(val).ljust(length)

    # Summary
    print('Name   ::  Test Stat > C(95%)    =>   Signif  \n', '--'*20)
    for col, trace, cvt in zip(df.columns, traces, cvts):
        print(adjust(col), ':: ', adjust(round(trace,2), 9), ">", adjust(cvt, 8), ' =>  ' , trace > cvt)

cointegration_test(chuuk_df)
```

```
Name   ::  Test Stat > C(95%)    =>   Signif
 ----------------------------------------
xCO2 SW (wet) (umol/mol) ::  285.21    > 179.5199  =>   True
Licor Atm Pressure (hPa) ::  202.63    > 143.6691  =>   True
SST (C) ::  131.67    > 111.7797  =>   True
xCO2 SW (dry) (umol/mol) ::  81.44     > 83.9383   =>   False
fCO2 SW (sat) uatm ::  47.12     > 60.0627   =>   False
dfCO2  ::  23.24     > 40.1749   =>   False
pCO2 SW (sat) uatm ::  9.84      > 24.2761   =>   False
dpCO2  ::  3.21      > 12.3212   =>   False
pH (Total Scale) ::  0.16      > 4.1296    =>   False
```

In [32]:

```
def adfuller_test(series, signif=0.05, name='', verbose=False):
    """Perform ADFuller to test for Stationarity of given series and print report"""
    r = adfuller(series, autolag='AIC')
    output = {'test_statistic':round(r[0], 4), 'pvalue':round(r[1], 4), 'n_lags':round(r[2], 4), 'n_obs':r[3]}
    p_value = output['pvalue']
    def adjust(val, length= 6): return str(val).ljust(length)

    # Print Summary
    print(f'    Augmented Dickey-Fuller Test on "{name}"', "\n   ", '-'*47)
    print(f' Null Hypothesis: Data has unit root. Non-Stationary.')
    print(f' Significance Level    = {signif}')
    print(f' Test Statistic        = {output["test_statistic"]}')
    print(f' No. Lags Chosen       = {output["n_lags"]}')

    for key,val in r[4].items():
        print(f' Critical value {adjust(key)} = {round(val, 3)}')

    if p_value <= signif:
        print(f" => P-Value = {p_value}. Rejecting Null Hypothesis.")
        print(f" => Series is Stationary.")
    else:
        print(f" => P-Value = {p_value}. Weak evidence to reject the Null Hypothesis.")
        print(f" => Series is Non-Stationary.")
```

In [33]:

```
# ADF Test on each column
for name, column in train.iteritems():
    adfuller_test(column, name=column.name)
    print('\n')
```

```
    Augmented Dickey-Fuller Test on "xCO2 SW (wet) (umol/mol)"
    -----------------------------------------------
 Null Hypothesis: Data has unit root. Non-Stationary.
 Significance Level    = 0.05
 Test Statistic        = -1.4415
 No. Lags Chosen       = 10
 Critical value 1%     = -3.453
 Critical value 5%     = -2.871
 Critical value 10%    = -2.572
 => P-Value = 0.5622. Weak evidence to reject the Null Hypothesis.
 => Series is Non-Stationary.


    Augmented Dickey-Fuller Test on "Licor Atm Pressure (hPa)"
    -----------------------------------------------
 Null Hypothesis: Data has unit root. Non-Stationary.
 Significance Level    = 0.05
 Test Statistic        = -6.1375
 No. Lags Chosen       = 3
 Critical value 1%     = -3.452
 Critical value 5%     = -2.871
```

```
Critical value 10%    = -2.572
=> P-Value = 0.0. Rejecting Null Hypothesis.
=> Series is Stationary.


   Augmented Dickey-Fuller Test on "SST (C)"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -1.6993
No. Lags Chosen       = 5
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.4315. Weak evidence to reject the Null Hypothesis.
=> Series is Non-Stationary.


   Augmented Dickey-Fuller Test on "xCO2 SW (dry) (umol/mol)"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -1.398
No. Lags Chosen       = 10
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.5832. Weak evidence to reject the Null Hypothesis.
=> Series is Non-Stationary.


   Augmented Dickey-Fuller Test on "fCO2 SW (sat) uatm"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -1.4348
No. Lags Chosen       = 10
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.5655. Weak evidence to reject the Null Hypothesis.
=> Series is Non-Stationary.


   Augmented Dickey-Fuller Test on "dfCO2"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -1.3926
No. Lags Chosen       = 10
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.5858. Weak evidence to reject the Null Hypothesis.
=> Series is Non-Stationary.


   Augmented Dickey-Fuller Test on "pCO2 SW (sat) uatm"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -1.4353
No. Lags Chosen       = 10
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.5653. Weak evidence to reject the Null Hypothesis.
=> Series is Non-Stationary.


   Augmented Dickey-Fuller Test on "dpCO2"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -1.3925
No. Lags Chosen       = 10
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.5858. Weak evidence to reject the Null Hypothesis.
=> Series is Non-Stationary.
```

```
   Augmented Dickey-Fuller Test on "pH (Total Scale)"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -1.5639
No. Lags Chosen       = 6
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.5017. Weak evidence to reject the Null Hypothesis.
=> Series is Non-Stationary.
```

In [34]:
```python
# 1st difference
train_diff = train.diff().dropna()
```

In [35]:
```python
# ADF Test on each column
for name, column in train_diff.iteritems():
    adfuller_test(column, name=column.name)
    print('\n')
```

```
   Augmented Dickey-Fuller Test on "xCO2 SW (wet) (umol/mol)"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -8.012
No. Lags Chosen       = 9
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.0. Rejecting Null Hypothesis.
=> Series is Stationary.


   Augmented Dickey-Fuller Test on "Licor Atm Pressure (hPa)"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -7.3988
No. Lags Chosen       = 13
Critical value 1%     = -3.453
Critical value 5%     = -2.872
Critical value 10%    = -2.572
=> P-Value = 0.0. Rejecting Null Hypothesis.
=> Series is Stationary.


   Augmented Dickey-Fuller Test on "SST (C)"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -8.4613
No. Lags Chosen       = 8
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.0. Rejecting Null Hypothesis.
=> Series is Stationary.


   Augmented Dickey-Fuller Test on "xCO2 SW (dry) (umol/mol)"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -7.9435
No. Lags Chosen       = 9
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.0. Rejecting Null Hypothesis.
=> Series is Stationary.


   Augmented Dickey-Fuller Test on "fCO2 SW (sat) uatm"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -7.8566
No. Lags Chosen       = 9
```

```
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.0. Rejecting Null Hypothesis.
=> Series is Stationary.


   Augmented Dickey-Fuller Test on "dfCO2"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -7.7811
No. Lags Chosen       = 9
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.0. Rejecting Null Hypothesis.
=> Series is Stationary.


   Augmented Dickey-Fuller Test on "pCO2 SW (sat) uatm"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -7.8536
No. Lags Chosen       = 9
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.0. Rejecting Null Hypothesis.
=> Series is Stationary.


   Augmented Dickey-Fuller Test on "dpCO2"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -7.7804
No. Lags Chosen       = 9
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.0. Rejecting Null Hypothesis.
=> Series is Stationary.


   Augmented Dickey-Fuller Test on "pH (Total Scale)"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -10.6787
No. Lags Chosen       = 5
Critical value 1%     = -3.453
Critical value 5%     = -2.871
Critical value 10%    = -2.572
=> P-Value = 0.0. Rejecting Null Hypothesis.
=> Series is Stationary.
```

In [36]:

```
model = VAR(train_diff)
x = model.select_order(maxlags=10)
x.summary()
```

```
C:\Users\datre\anaconda3\lib\site-packages\statsmodels\tsa\base\tsa_model.py:581: ValueWarning: A date index has be
en provided, but it has no associated frequency information and so will be ignored when e.g. forecasting.
  warnings.warn('A date index has been provided, but it has no'
```

Out[36]:

VAR Order Selection (* highlights the minimums)

	AIC	BIC	FPE	HQIC
0	-34.86	-34.75*	7.230e-16	-34.82
1	-35.81	-34.67	2.814e-16	-35.35
2	-36.28	-34.13	1.753e-16	-35.42*
3	-36.64*	-33.46	1.234e-16*	-35.37
4	-36.62	-32.43	1.263e-16	-34.94
5	-36.55	-31.34	1.370e-16	-34.46

6	-36.59	-30.36	1.333e-16	-34.10
7	-36.52	-29.26	1.475e-16	-33.61
8	-36.44	-28.17	1.646e-16	-33.13
9	-36.45	-27.16	1.698e-16	-32.73
10	-36.30	-25.99	2.088e-16	-32.17

In [37]:

```
# Using fit 3 from results above
model_fitted = model.fit(3)
model_fitted.summary()
```

Out[37]:

```
  Summary of Regression Results   
==================================
Model:                         VAR
Method:                        OLS
Date:           Thu, 03, Jun, 2021
Time:                     16:51:17
--------------------------------------------------------------------
No. of Equations:         9.00000    BIC:                   -33.6672
Nobs:                     299.000    HQIC:                  -35.5377
Log likelihood:           1933.14    FPE:                1.06221e-16
AIC:                     -36.7860    Det(Omega_mle):     4.74593e-17
--------------------------------------------------------------------
Results for equation xCO2 SW (wet) (umol/mol)
===============================================================================================
                                 coefficient       std. error           t-stat            prob
-----------------------------------------------------------------------------------------------
const                               0.092659         0.370431            0.250           0.802
L1.xCO2 SW (wet) (umol/mol)        -1.697108         1.572509           -1.079           0.280
L1.Licor Atm Pressure (hPa)         4.253268         5.490276            0.775           0.439
L1.SST (C)                         -7.749009        13.626843           -0.569           0.570
L1.xCO2 SW (dry) (umol/mol)        12.187276        12.809705            0.951           0.341
L1.fCO2 SW (sat) uatm              11.733445        22.291591            0.526           0.599
L1.dfCO2                           26.521244        23.037665            1.151           0.250
L1.pCO2 SW (sat) uatm             -21.782485        23.028528           -0.946           0.344
L1.dpCO2                          -27.726482        22.966879           -1.207           0.227
L1.pH (Total Scale)               -47.685990        92.412470           -0.516           0.606
L2.xCO2 SW (wet) (umol/mol)        -1.783413         1.663797           -1.072           0.284
L2.Licor Atm Pressure (hPa)        -8.321198         5.367704           -1.550           0.121
L2.SST (C)                         18.195817        13.228480            1.376           0.169
L2.xCO2 SW (dry) (umol/mol)       -17.627894        12.379860           -1.424           0.154
L2.fCO2 SW (sat) uatm              -8.647903        24.211936           -0.357           0.721
L2.dfCO2                           10.246204        26.618213            0.385           0.700
L2.pCO2 SW (sat) uatm              30.189244        24.505240            1.232           0.218
L2.dpCO2                          -11.701242        26.539818           -0.441           0.659
L2.pH (Total Scale)               -12.199291       102.448626           -0.119           0.905
L3.xCO2 SW (wet) (umol/mol)        -0.575889         1.507157           -0.382           0.702
L3.Licor Atm Pressure (hPa)        -4.663817         5.459729           -0.854           0.393
L3.SST (C)                         12.061142        13.441163            0.897           0.370
L3.xCO2 SW (dry) (umol/mol)        -8.038954        12.651793           -0.635           0.525
L3.fCO2 SW (sat) uatm              26.234863        22.114256            1.186           0.235
L3.dfCO2                          -16.486678        22.938355           -0.719           0.472
L3.pCO2 SW (sat) uatm             -16.372867        22.520665           -0.727           0.467
L3.dpCO2                           15.549094        22.872544            0.680           0.497
L3.pH (Total Scale)               -91.223112        89.219420           -1.022           0.307
===============================================================================================

Results for equation Licor Atm Pressure (hPa)
===============================================================================================
                                 coefficient       std. error           t-stat            prob
-----------------------------------------------------------------------------------------------
const                               0.003464         0.046752            0.074           0.941
L1.xCO2 SW (wet) (umol/mol)        -0.298861         0.198467           -1.506           0.132
L1.Licor Atm Pressure (hPa)        -1.743880         0.692928           -2.517           0.012
L1.SST (C)                          4.267175         1.719846            2.481           0.013
L1.xCO2 SW (dry) (umol/mol)        -4.092935         1.616714           -2.532           0.011
L1.fCO2 SW (sat) uatm              -0.143403         2.813424           -0.051           0.959
L1.dfCO2                           -3.431241         2.907587           -1.180           0.238
L1.pCO2 SW (sat) uatm               4.758612         2.906433            1.637           0.102
L1.dpCO2                            3.428710         2.898653            1.183           0.237
L1.pH (Total Scale)                 8.785099        11.663389            0.753           0.451
L2.xCO2 SW (wet) (umol/mol)        -0.086561         0.209988           -0.412           0.680
L2.Licor Atm Pressure (hPa)        -1.529642         0.677459           -2.258           0.024
L2.SST (C)                          2.828131         1.669568            1.694           0.090
L2.xCO2 SW (dry) (umol/mol)        -2.707458         1.562464           -1.733           0.083
L2.fCO2 SW (sat) uatm               0.535578         3.055791            0.175           0.861
L2.dfCO2                           -3.612027         3.359488           -1.075           0.282
L2.pCO2 SW (sat) uatm               2.464582         3.092809            0.797           0.426
```

```
L2.dpCO2                          3.561548         3.349594            1.063           0.288
L2.pH (Total Scale)              16.531170        12.930054            1.279           0.201
L3.xCO2 SW (wet) (umol/mol)      -0.329881         0.190218           -1.734           0.083
L3.Licor Atm Pressure (hPa)      -0.732929         0.689073           -1.064           0.287
L3.SST (C)                        2.133611         1.696411            1.258           0.208
L3.xCO2 SW (dry) (umol/mol)      -1.386550         1.596784           -0.868           0.385
L3.fCO2 SW (sat) uatm            -0.388050         2.791043           -0.139           0.889
L3.dfCO2                         -3.455190         2.895053           -1.193           0.233
L3.pCO2 SW (sat) uatm             2.256208         2.842336            0.794           0.427
L3.dpCO2                          3.393023         2.886747            1.175           0.240
L3.pH (Total Scale)              14.029620        11.260394            1.246           0.213
============================================================================================

Results for equation SST (C)
============================================================================================
                               coefficient       std. error           t-stat            prob
--------------------------------------------------------------------------------------------
const                             0.002209         0.009134            0.242           0.809
L1.xCO2 SW (wet) (umol/mol)      -0.056959         0.038773           -1.469           0.142
L1.Licor Atm Pressure (hPa)      -0.258584         0.135373           -1.910           0.056
L1.SST (C)                        0.484565         0.335995            1.442           0.149
L1.xCO2 SW (dry) (umol/mol)      -0.538131         0.315847           -1.704           0.088
L1.fCO2 SW (sat) uatm             0.137743         0.549640            0.251           0.802
L1.dfCO2                         -0.033927         0.568036           -0.060           0.952
L1.pCO2 SW (sat) uatm             0.527215         0.567810            0.929           0.353
L1.dpCO2                         -0.004108         0.566290           -0.007           0.994
L1.pH (Total Scale)               0.143478         2.278598            0.063           0.950
L2.xCO2 SW (wet) (umol/mol)      -0.027693         0.041024           -0.675           0.500
L2.Licor Atm Pressure (hPa)      -0.150573         0.132351           -1.138           0.255
L2.SST (C)                        0.134937         0.326172            0.414           0.679
L2.xCO2 SW (dry) (umol/mol)      -0.311925         0.305248           -1.022           0.307
L2.fCO2 SW (sat) uatm            -0.001185         0.596989           -0.002           0.998
L2.dfCO2                          0.074056         0.656321            0.113           0.910
L2.pCO2 SW (sat) uatm             0.382840         0.604221            0.634           0.526
L2.dpCO2                         -0.098967         0.654388           -0.151           0.880
L2.pH (Total Scale)               0.531486         2.526058            0.210           0.833
L3.xCO2 SW (wet) (umol/mol)      -0.003007         0.037162           -0.081           0.936
L3.Licor Atm Pressure (hPa)      -0.391636         0.134620           -2.909           0.004
L3.SST (C)                        0.916084         0.331416            2.764           0.006
L3.xCO2 SW (dry) (umol/mol)      -0.977053         0.311953           -3.132           0.002
L3.fCO2 SW (sat) uatm             0.576808         0.545267            1.058           0.290
L3.dfCO2                         -0.258749         0.565587           -0.457           0.647
L3.pCO2 SW (sat) uatm             0.466794         0.555288            0.841           0.401
L3.dpCO2                          0.250450         0.563964            0.444           0.657
L3.pH (Total Scale)               1.721753         2.199867            0.783           0.434
============================================================================================

Results for equation xCO2 SW (dry) (umol/mol)
============================================================================================
                               coefficient       std. error           t-stat            prob
--------------------------------------------------------------------------------------------
const                             0.083434         0.375026            0.222           0.824
L1.xCO2 SW (wet) (umol/mol)      -1.249596         1.592017           -0.785           0.433
L1.Licor Atm Pressure (hPa)       4.392606         5.558385            0.790           0.429
L1.SST (C)                       -7.713459        13.795889           -0.559           0.576
L1.xCO2 SW (dry) (umol/mol)      12.107051        12.968614            0.934           0.351
L1.fCO2 SW (sat) uatm            11.731047        22.568127            0.520           0.603
L1.dfCO2                         27.344869        23.323456            1.172           0.241
L1.pCO2 SW (sat) uatm           -22.163864        23.314206           -0.951           0.342
L1.dpCO2                        -28.543549        23.251792           -1.228           0.220
L1.pH (Total Scale)             -44.702176        93.558882           -0.478           0.633
L2.xCO2 SW (wet) (umol/mol)      -1.504923         1.684437           -0.893           0.372
L2.Licor Atm Pressure (hPa)      -8.298165         5.434293           -1.527           0.127
L2.SST (C)                       17.940149        13.392584            1.340           0.180
L2.xCO2 SW (dry) (umol/mol)     -17.857475        12.533437           -1.425           0.154
L2.fCO2 SW (sat) uatm            -8.518789        24.512295           -0.348           0.728
L2.dfCO2                         10.173068        26.948422            0.378           0.706
L2.pCO2 SW (sat) uatm            30.014711        24.809237            1.210           0.226
L2.dpCO2                        -11.621240        26.869054           -0.433           0.665
L2.pH (Total Scale)              -3.494614       103.719541           -0.034           0.973
L3.xCO2 SW (wet) (umol/mol)      -0.335608         1.525854           -0.220           0.826
L3.Licor Atm Pressure (hPa)      -5.530004         5.527459           -1.000           0.317
L3.SST (C)                       13.978028        13.607905            1.027           0.304
L3.xCO2 SW (dry) (umol/mol)     -10.324755        12.808743           -0.806           0.420
L3.fCO2 SW (sat) uatm            28.965959        22.388592            1.294           0.196
L3.dfCO2                        -16.622964        23.222914           -0.716           0.474
L3.pCO2 SW (sat) uatm           -16.931223        22.800043           -0.743           0.458
L3.dpCO2                         15.687631        23.156287            0.677           0.498
L3.pH (Total Scale)             -82.156275        90.326221           -0.910           0.363
============================================================================================

Results for equation fCO2 SW (sat) uatm
```

	coefficient	std. error	t-stat	prob
const	0.078580	0.352747	0.223	0.824
L1.xCO2 SW (wet) (umol/mol)	-1.248528	1.497441	-0.834	0.404
L1.Licor Atm Pressure (hPa)	3.788152	5.228180	0.725	0.469
L1.SST (C)	-6.232155	12.976321	-0.480	0.631
L1.xCO2 SW (dry) (umol/mol)	10.515089	12.198192	0.862	0.389
L1.fCO2 SW (sat) uatm	10.332736	21.227430	0.487	0.626
L1.dfCO2	24.456185	21.937888	1.115	0.265
L1.pCO2 SW (sat) uatm	-19.182407	21.929187	-0.875	0.382
L1.dpCO2	-25.559329	21.870480	-1.169	0.243
L1.pH (Total Scale)	-38.770750	88.000861	-0.441	0.660
L2.xCO2 SW (wet) (umol/mol)	-1.435642	1.584370	-0.906	0.365
L2.Licor Atm Pressure (hPa)	-8.255483	5.111460	-1.615	0.106
L2.SST (C)	17.825060	12.596975	1.415	0.157
L2.xCO2 SW (dry) (umol/mol)	-17.517951	11.788867	-1.486	0.137
L2.fCO2 SW (sat) uatm	-8.427194	23.056101	-0.366	0.715
L2.dfCO2	7.953231	25.347506	0.314	0.754
L2.pCO2 SW (sat) uatm	29.437847	23.335403	1.262	0.207
L2.dpCO2	-9.326528	25.272853	-0.369	0.712
L2.pH (Total Scale)	2.477995	97.557909	0.025	0.980
L3.xCO2 SW (wet) (umol/mol)	-0.445770	1.435208	-0.311	0.756
L3.Licor Atm Pressure (hPa)	-5.051109	5.199091	-0.972	0.331
L3.SST (C)	12.949352	12.799505	1.012	0.312
L3.xCO2 SW (dry) (umol/mol)	-9.141870	12.047818	-0.759	0.448
L3.fCO2 SW (sat) uatm	26.392153	21.058560	1.253	0.210
L3.dfCO2	-16.817885	21.843318	-0.770	0.441
L3.pCO2 SW (sat) uatm	-15.514153	21.445568	-0.723	0.469
L3.dpCO2	15.912611	21.780649	0.731	0.465
L3.pH (Total Scale)	-73.972096	84.960242	-0.871	0.384

Results for equation dfCO2

	coefficient	std. error	t-stat	prob
const	0.085502	0.358798	0.238	0.812
L1.xCO2 SW (wet) (umol/mol)	-1.558477	1.523126	-1.023	0.306
L1.Licor Atm Pressure (hPa)	3.519844	5.317857	0.662	0.508
L1.SST (C)	-4.803112	13.198900	-0.364	0.716
L1.xCO2 SW (dry) (umol/mol)	10.764447	12.407424	0.868	0.386
L1.fCO2 SW (sat) uatm	11.371767	21.591538	0.527	0.598
L1.dfCO2	26.477545	22.314181	1.187	0.235
L1.pCO2 SW (sat) uatm	-19.472306	22.305331	-0.873	0.383
L1.dpCO2	-28.246858	22.245618	-1.270	0.204
L1.pH (Total Scale)	-41.538403	89.510313	-0.464	0.643
L2.xCO2 SW (wet) (umol/mol)	-1.336675	1.611546	-0.829	0.407
L2.Licor Atm Pressure (hPa)	-7.995554	5.199135	-1.538	0.124
L2.SST (C)	17.217440	12.813047	1.344	0.179
L2.xCO2 SW (dry) (umol/mol)	-16.817832	11.991078	-1.403	0.161
L2.fCO2 SW (sat) uatm	-12.888049	23.451575	-0.550	0.583
L2.dfCO2	13.767077	25.782284	0.534	0.593
L2.pCO2 SW (sat) uatm	33.442265	23.735668	1.409	0.159
L2.dpCO2	-15.503424	25.706351	-0.603	0.546
L2.pH (Total Scale)	3.010460	99.231290	0.030	0.976
L3.xCO2 SW (wet) (umol/mol)	0.164513	1.459826	0.113	0.910
L3.Licor Atm Pressure (hPa)	-5.485900	5.288269	-1.037	0.300
L3.SST (C)	14.873767	13.019051	1.142	0.253
L3.xCO2 SW (dry) (umol/mol)	-10.874368	12.254471	-0.887	0.375
L3.fCO2 SW (sat) uatm	25.532102	21.419771	1.192	0.233
L3.dfCO2	-12.792057	22.217990	-0.576	0.565
L3.pCO2 SW (sat) uatm	-13.286067	21.813417	-0.609	0.542
L3.dpCO2	11.755703	22.154246	0.531	0.596
L3.pH (Total Scale)	-69.829534	86.417539	-0.808	0.419

Results for equation pCO2 SW (sat) uatm

	coefficient	std. error	t-stat	prob
const	0.079086	0.353763	0.224	0.823
L1.xCO2 SW (wet) (umol/mol)	-1.262789	1.501752	-0.841	0.400
L1.Licor Atm Pressure (hPa)	3.791559	5.243233	0.723	0.470
L1.SST (C)	-6.255130	13.013684	-0.481	0.631
L1.xCO2 SW (dry) (umol/mol)	10.532871	12.233314	0.861	0.389
L1.fCO2 SW (sat) uatm	11.041147	21.288550	0.519	0.604
L1.dfCO2	24.509455	22.001053	1.114	0.265
L1.pCO2 SW (sat) uatm	-19.889548	21.992327	-0.904	0.366
L1.dpCO2	-25.616366	21.933451	-1.168	0.243
L1.pH (Total Scale)	-38.570252	88.254239	-0.437	0.662
L2.xCO2 SW (wet) (umol/mol)	-1.451897	1.588932	-0.914	0.361

```
L2.Licor Atm Pressure (hPa)       -8.279183        5.126177          -1.615           0.106
L2.SST (C)                        17.845996       12.633245           1.413           0.158
L2.xCO2 SW (dry) (umol/mol)      -17.555099       11.822810          -1.485           0.138
L2.fCO2 SW (sat) uatm             -7.937543       23.122486          -0.343           0.731
L2.dfCO2                           7.966565       25.420488           0.313           0.754
L2.pCO2 SW (sat) uatm             29.011724       23.402592           1.240           0.215
L2.dpCO2                          -9.346441       25.345621          -0.369           0.712
L2.pH (Total Scale)                2.614778       97.838804           0.027           0.979
L3.xCO2 SW (wet) (umol/mol)       -0.455358        1.439340          -0.316           0.752
L3.Licor Atm Pressure (hPa)       -5.065401        5.214060          -0.971           0.331
L3.SST (C)                        12.989717       12.836358           1.012           0.312
L3.xCO2 SW (dry) (umol/mol)       -9.159099       12.082507          -0.758           0.448
L3.fCO2 SW (sat) uatm             26.680817       21.119194           1.263           0.206
L3.dfCO2                         -16.974872       21.906211          -0.775           0.438
L3.pCO2 SW (sat) uatm            -15.768379       21.507316          -0.733           0.463
L3.dpCO2                          16.062845       21.843362           0.735           0.462
L3.pH (Total Scale)              -74.292738       85.204865          -0.872           0.383
==============================================================================================

Results for equation dpCO2
==============================================================================================
                                coefficient       std. error           t-stat              prob
----------------------------------------------------------------------------------------------
const                              0.085789        0.359846           0.238           0.812
L1.xCO2 SW (wet) (umol/mol)       -1.565772        1.527576          -1.025           0.305
L1.Licor Atm Pressure (hPa)        3.518926        5.333394           0.660           0.509
L1.SST (C)                        -4.795990       13.237463          -0.362           0.717
L1.xCO2 SW (dry) (umol/mol)       10.773844       12.443675           0.866           0.387
L1.fCO2 SW (sat) uatm             11.400152       21.654621           0.526           0.599
L1.dfCO2                          27.310832       22.379376           1.220           0.222
L1.pCO2 SW (sat) uatm            -19.497779       22.370500          -0.872           0.383
L1.dpCO2                         -29.083445       22.310613          -1.304           0.192
L1.pH (Total Scale)              -41.271882       89.771834          -0.460           0.646
L2.xCO2 SW (wet) (umol/mol)       -1.341758        1.616254          -0.830           0.406
L2.Licor Atm Pressure (hPa)       -8.036132        5.214325          -1.541           0.123
L2.SST (C)                        17.312291       12.850483           1.347           0.178
L2.xCO2 SW (dry) (umol/mol)      -16.910472       12.026112          -1.406           0.160
L2.fCO2 SW (sat) uatm            -12.892249       23.520093          -0.548           0.584
L2.dfCO2                          14.324418       25.857611           0.554           0.580
L2.pCO2 SW (sat) uatm             33.554133       23.805016           1.410           0.159
L2.dpCO2                         -16.064717       25.781457          -0.623           0.533
L2.pH (Total Scale)                3.098095       99.521212           0.031           0.975
L3.xCO2 SW (wet) (umol/mol)        0.169428        1.464091           0.116           0.908
L3.Licor Atm Pressure (hPa)       -5.488298        5.303720          -1.035           0.301
L3.SST (C)                        14.886692       13.057089           1.140           0.254
L3.xCO2 SW (dry) (umol/mol)      -10.877461       12.290275          -0.885           0.376
L3.fCO2 SW (sat) uatm             25.540628       21.482353           1.189           0.234
L3.dfCO2                         -12.531884       22.282904          -0.562           0.574
L3.pCO2 SW (sat) uatm            -13.293905       21.877149          -0.608           0.543
L3.dpCO2                          11.493638       22.218974           0.517           0.605
L3.pH (Total Scale)              -69.867510       86.670023          -0.806           0.420
==============================================================================================

Results for equation pH (Total Scale)
==============================================================================================
                                coefficient       std. error           t-stat              prob
----------------------------------------------------------------------------------------------
const                             -0.000193        0.000432          -0.446           0.655
L1.xCO2 SW (wet) (umol/mol)        0.000427        0.001833           0.233           0.816
L1.Licor Atm Pressure (hPa)       -0.003322        0.006400          -0.519           0.604
L1.SST (C)                        -0.000410        0.015885          -0.026           0.979
L1.xCO2 SW (dry) (umol/mol)       -0.007417        0.014932          -0.497           0.619
L1.fCO2 SW (sat) uatm             -0.025563        0.025986          -0.984           0.325
L1.dfCO2                          -0.033246        0.026855          -1.238           0.216
L1.pCO2 SW (sat) uatm              0.030793        0.026845           1.147           0.251
L1.dpCO2                           0.034841        0.026773           1.301           0.193
L1.pH (Total Scale)               -0.525504        0.107727          -4.878           0.000
L2.xCO2 SW (wet) (umol/mol)        0.001799        0.001940           0.927           0.354
L2.Licor Atm Pressure (hPa)        0.002632        0.006257           0.421           0.674
L2.SST (C)                        -0.005165        0.015421          -0.335           0.738
L2.xCO2 SW (dry) (umol/mol)        0.002826        0.014431           0.196           0.845
L2.fCO2 SW (sat) uatm              0.003436        0.028224           0.122           0.903
L2.dfCO2                          -0.003430        0.031029          -0.111           0.912
L2.pCO2 SW (sat) uatm             -0.009973        0.028566          -0.349           0.727
L2.dpCO2                           0.005079        0.030938           0.164           0.870
L2.pH (Total Scale)               -0.324060        0.119426          -2.713           0.007
L3.xCO2 SW (wet) (umol/mol)        0.000850        0.001757           0.484           0.629
L3.Licor Atm Pressure (hPa)        0.004782        0.006364           0.751           0.452
L3.SST (C)                        -0.013854        0.015669          -0.884           0.377
L3.xCO2 SW (dry) (umol/mol)        0.009078        0.014748           0.616           0.538
L3.fCO2 SW (sat) uatm             -0.041563        0.025779          -1.612           0.107
L3.dfCO2                           0.038496        0.026740           1.440           0.150
```

```
L3.pCO2 SW (sat) uatm              0.030067         0.026253            1.145           0.252
L3.dpCO2                          -0.037602         0.026663           -1.410           0.158
L3.pH (Total Scale)               -0.154808         0.104004           -1.488           0.137
===============================================================================================

Correlation matrix of residuals
                            xCO2 SW (wet) (umol/mol)  Licor Atm Pressure (hPa)   SST (C)  xCO2 SW (dry) (umol/mol)
fCO2 SW (sat) uatm     dfCO2  pCO2 SW (sat) uatm     dpCO2  pH (Total Scale)
xCO2 SW (wet) (umol/mol)                    1.000000                 -0.016383  0.374658                  0.999053
0.998142  0.973060          0.998138  0.973085        -0.832916
Licor Atm Pressure (hPa)                   -0.016383                  1.000000  0.020791                 -0.021967
0.031772 -0.010622          0.031790 -0.010468        -0.041978
SST (C)                                     0.374658                  0.020791  1.000000                  0.395850
0.375797  0.414584          0.375321  0.414352        -0.267610
xCO2 SW (dry) (umol/mol)                    0.999053                 -0.021967  0.395850                  1.000000
0.998284  0.976087          0.998272  0.976102        -0.830490
fCO2 SW (sat) uatm                          0.998142                  0.031772  0.375797                  0.998284
1.000000  0.974394          0.999996  0.974423        -0.833739
dfCO2                                       0.973060                 -0.010622  0.414584                  0.976087
0.974394  1.000000          0.974385  0.999997        -0.814025
pCO2 SW (sat) uatm                          0.998138                  0.031790  0.375321                  0.998272
0.999996  0.974385          1.000000  0.974414        -0.834033
dpCO2                                       0.973085                 -0.010468  0.414352                  0.976102
0.974423  0.999997          0.974414  1.000000        -0.814034
pH (Total Scale)                           -0.832916                 -0.041978 -0.267610                 -0.830490
-0.833739 -0.814025          -0.834033 -0.814034         1.000000
```

In [38]:

```python
out = durbin_watson(model_fitted.resid)

for col, val in zip(chuuk_df.columns, out):
    print(col, ':', round(val, 2))
```

```
xCO2 SW (wet) (umol/mol) : 1.97
Licor Atm Pressure (hPa) : 2.0
SST (C) : 2.05
xCO2 SW (dry) (umol/mol) : 1.97
fCO2 SW (sat) uatm : 1.98
dfCO2 : 1.99
pCO2 SW (sat) uatm : 1.98
dpCO2 : 1.99
pH (Total Scale) : 1.98
```

In [39]:

```python
# Get the lag order
lag_order = model_fitted.k_ar
print(lag_order)

# Input data for forecasting
forecast_input = train_diff.values[-lag_order:]
forecast_input
```

```
3
```

Out[39]:

```
array([[ 5.11250e+00,  8.75000e-02,  4.53750e-02,  5.13750e+00,
         4.87500e+00,  5.27500e+00,  4.88750e+00,  5.31250e+00,
        -6.00000e-03],
       [ 4.52500e+00,  3.12500e-01,  1.25125e-01,  4.73750e+00,
         4.52500e+00,  5.32500e+00,  4.51250e+00,  5.33750e+00,
        -8.87500e-03],
       [-7.50000e-02, -6.25000e-02, -1.50000e-02, -1.87500e-01,
        -2.00000e-01, -3.50000e-01, -2.00000e-01, -3.25000e-01,
        -2.25000e-03]])
```

In [40]:

```python
# Forecast
fc = model_fitted.forecast(y=forecast_input, steps = len(train_diff))
df_forecast = pd.DataFrame(fc, index=train_diff.index, columns = chuuk_df.columns + "_1d")
df_forecast
```

Out[40]:

	xCO2 SW (wet) (umol/mol)_1d	Licor Atm Pressure (hPa)_1d	SST (C)_1d	xCO2 SW (dry) (umol/mol)_1d	fCO2 SW (sat) uatm_1d	dfCO2_1d	pCO2 SW (sat) uatm_1d	dpCO2_1d	pH (Total Scale)_1d
Datetime									
2013-04-25	-3.011511	-0.113039	-0.032469	-3.052355	-2.919747	-3.185919	-2.920522	-3.220280	0.005145
2013-04-26	1.503981	-0.063022	0.016578	1.531379	1.411335	1.539328	1.414204	1.548054	-0.002005
2013-04-27	1.754687	-0.029876	0.007649	1.746079	1.636270	1.805490	1.640750	1.812185	-0.002062

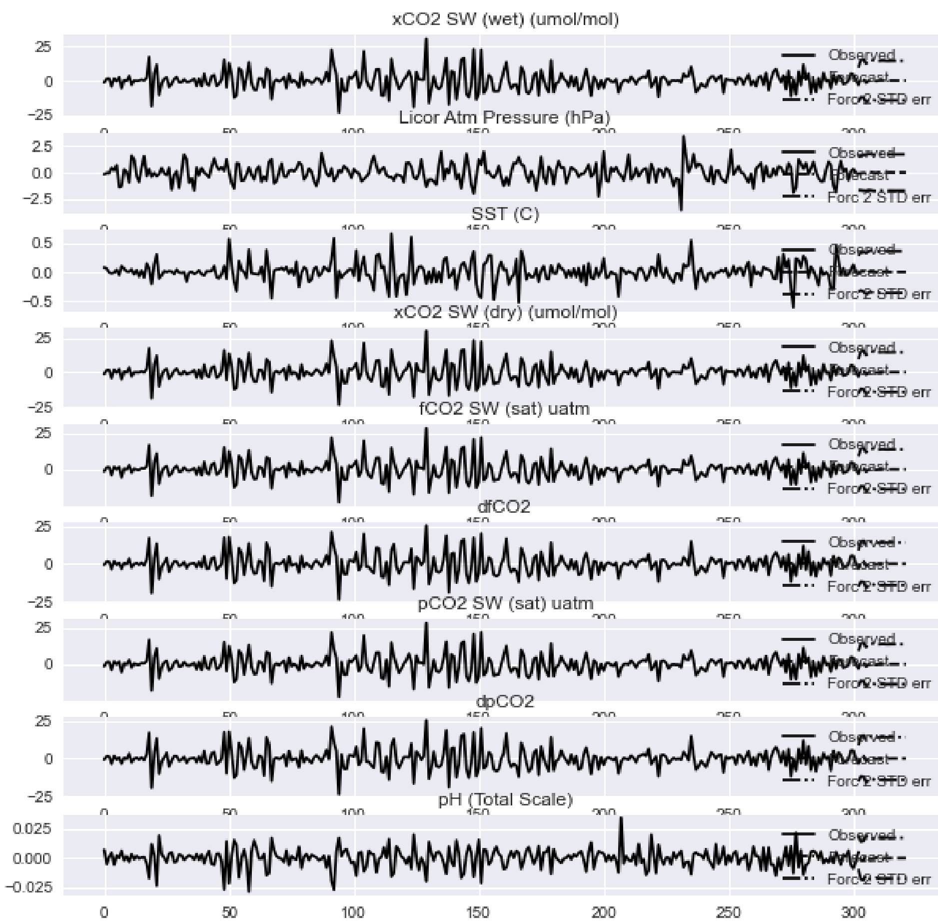
Datetime	xCO2 SW (wet) (umol/mol)_1d	Licor Atm Pressure (hPa)_1d	SST (C)_1d	xCO2 SW (dry) (umol/mol)_1d	fCO2 SW (sat) uatm_1d	dfCO2_1d	pCO2 SW (sat) uatm_1d	dpCO2_1d	pH (Total Scale)_1d
2013-04-28	-1.666134	-0.044382	-0.027470	-1.692304	-1.589906	-1.692350	-1.597855	-1.692765	0.002181
2013-05-02	-0.337657	0.052205	0.000268	-0.351231	-0.312635	-0.333211	-0.310113	-0.339033	0.000169
...	...	...	...	...	...	...	...	...	...
2014-05-18	0.043320	-0.001188	0.001062	0.038493	0.035133	0.041571	0.035227	0.041692	-0.000092
2014-05-19	0.043320	-0.001188	0.001062	0.038493	0.035133	0.041571	0.035227	0.041692	-0.000092
2014-05-20	0.043320	-0.001188	0.001062	0.038493	0.035133	0.041571	0.035227	0.041692	-0.000092
2014-05-21	0.043320	-0.001188	0.001062	0.038493	0.035133	0.041571	0.035227	0.041692	-0.000092
2014-05-22	0.043320	-0.001188	0.001062	0.038493	0.035133	0.041571	0.035227	0.041692	-0.000092

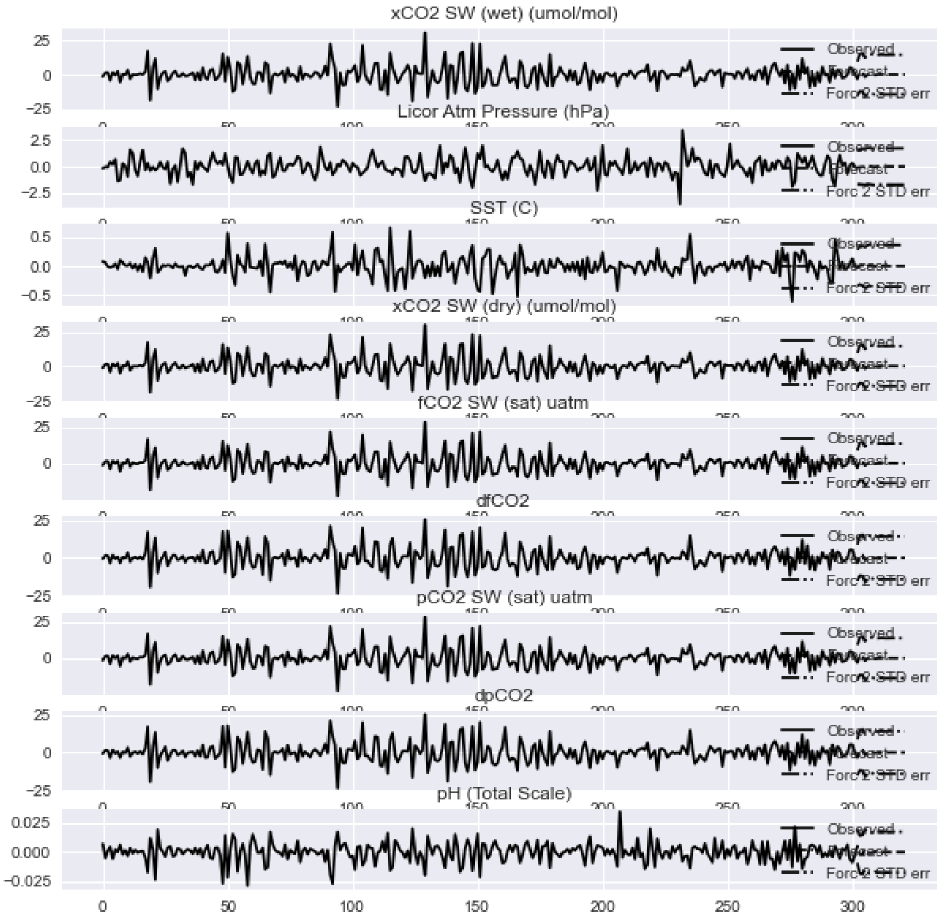
302 rows × 9 columns

In [41]:

```
model_fitted.plot_forecast(20)
```

Out[41]:





In [42]:

```
fevd = model_fitted.fevd(5)
fevd.summary()
```

```
FEVD for xCO2 SW (wet) (umol/mol)
     xCO2 SW (wet) (umol/mol)  Licor Atm Pressure (hPa)   SST (C)  xCO2 SW (dry) (umol/mol)  fCO2 SW (sat) uatm
dfCO2  pCO2 SW (sat) uatm     dpCO2  pH (Total Scale)
0                  1.000000                  0.000000  0.000000                  0.000000            0.000000  0.
000000            0.000000  0.000000          0.000000
1                  0.927590                  0.000072  0.002589                  0.000009            0.001887  0.
060801            0.001849  0.004411          0.000791
2                  0.902973                  0.001953  0.011743                  0.000025            0.005128  0.
060891            0.011055  0.005438          0.000794
3                  0.872782                  0.005558  0.013027                  0.000025            0.007405  0.
069973            0.017816  0.012222          0.001192
4                  0.865162                  0.005593  0.012950                  0.000845            0.007709  0.
073648            0.017696  0.014317          0.002080

FEVD for Licor Atm Pressure (hPa)
     xCO2 SW (wet) (umol/mol)  Licor Atm Pressure (hPa)   SST (C)  xCO2 SW (dry) (umol/mol)  fCO2 SW (sat) uatm
dfCO2  pCO2 SW (sat) uatm     dpCO2  pH (Total Scale)
0                  0.000268                  0.999732  0.000000                  0.000000            0.000000  0.
000000            0.000000  0.000000          0.000000
1                  0.000686                  0.958052  0.000033                  0.011521            0.016710  0.
000176            0.006489  0.004539          0.001794
2                  0.001939                  0.941586  0.000590                  0.010772            0.021175  0.
003848            0.006369  0.005578          0.008142
3                  0.003851                  0.929918  0.002824                  0.012698            0.021477  0.
008760            0.006271  0.005763          0.008439
4                  0.004272                  0.917413  0.003469                  0.014499            0.024451  0.
009209            0.006491  0.009469          0.010727

FEVD for SST (C)
     xCO2 SW (wet) (umol/mol)  Licor Atm Pressure (hPa)   SST (C)  xCO2 SW (dry) (umol/mol)  fCO2 SW (sat) uatm
dfCO2  pCO2 SW (sat) uatm     dpCO2  pH (Total Scale)
0                  0.140368                  0.000725  0.858906                  0.000000            0.000000  0.
000000            0.000000  0.000000          0.000000
1                  0.126070                  0.002368  0.776007                  0.001661            0.005496  0.
085941            0.002446  0.000000          0.000012
2                  0.140640                  0.002733  0.765963                  0.001673            0.005396  0.
080648            0.002473  0.000001          0.000472
3                  0.136769                  0.025534  0.718610                  0.002169            0.021389  0.
087900            0.002316  0.003809          0.001503
```

```
4               0.137000            0.026785  0.713234                0.002713            0.024966  0.
087077          0.002292  0.004114          0.001820

FEVD for xCO2 SW (dry) (umol/mol)
    xCO2 SW (wet) (umol/mol)  Licor Atm Pressure (hPa)   SST (C)  xCO2 SW (dry) (umol/mol)  fCO2 SW (sat) uatm
dfCO2  pCO2 SW (sat) uatm     dpCO2  pH (Total Scale)
0               0.998107            0.000031  0.000548                0.001313            0.000000  0.
000000          0.000000  0.000000          0.000000
1               0.927157            0.000096  0.002971                0.001486            0.001947  0.
059187          0.001905  0.004571          0.000680
2               0.900559            0.002014  0.013969                0.001421            0.005225  0.
059540          0.010840  0.005696          0.000737
3               0.870568            0.005171  0.014784                0.001380            0.008391  0.
068835          0.017560  0.012220          0.001092
4               0.862769            0.005186  0.014790                0.002416            0.008890  0.
072498          0.017537  0.014191          0.001722

FEVD for fCO2 SW (sat) uatm
    xCO2 SW (wet) (umol/mol)  Licor Atm Pressure (hPa)   SST (C)  xCO2 SW (dry) (umol/mol)  fCO2 SW (sat) uatm
dfCO2  pCO2 SW (sat) uatm     dpCO2  pH (Total Scale)
0               0.996287            0.002317  0.000000                0.001374            0.000023  0.
000000          0.000000  0.000000          0.000000
1               0.929031            0.002057  0.002876                0.001409            0.001691  0.
056594          0.001608  0.004152          0.000580
2               0.902075            0.002836  0.013318                0.001339            0.005411  0.
057863          0.010915  0.005512          0.000731
3               0.872220            0.006983  0.014427                0.001295            0.008062  0.
066079          0.017885  0.011969          0.001080
4               0.864168            0.006918  0.014381                0.002169            0.008630  0.
070076          0.017825  0.014220          0.001613

FEVD for dfCO2
    xCO2 SW (wet) (umol/mol)  Licor Atm Pressure (hPa)   SST (C)  xCO2 SW (dry) (umol/mol)  fCO2 SW (sat) uatm
dfCO2  pCO2 SW (sat) uatm     dpCO2  pH (Total Scale)
0               0.946846            0.000028  0.002896                0.005627            0.000250  0.
044353          0.000000  0.000000          0.000000
1               0.816702            0.000048  0.005469                0.005337            0.002256  0.
163522          0.001466  0.004596          0.000604
2               0.794952            0.001191  0.019797                0.005329            0.004607  0.
156728          0.011471  0.005171          0.000755
3               0.767663            0.002758  0.020441                0.005235            0.008182  0.
163452          0.019003  0.012222          0.001043
4               0.760630            0.002826  0.020395                0.007680            0.008402  0.
166014          0.018948  0.013705          0.001399

FEVD for pCO2 SW (sat) uatm
    xCO2 SW (wet) (umol/mol)  Licor Atm Pressure (hPa)   SST (C)  xCO2 SW (dry) (umol/mol)  fCO2 SW (sat) uatm
dfCO2  pCO2 SW (sat) uatm     dpCO2  pH (Total Scale)
0               0.996279            0.002318  0.000000                0.001382            0.000014  0.
000000          0.000007  0.000000          0.000000
1               0.928908            0.002057  0.002865                0.001408            0.001626  0.
056673          0.001748  0.004145          0.000570
2               0.901973            0.002825  0.013337                0.001338            0.005342  0.
057947          0.011014  0.005505          0.000719
3               0.872152            0.006973  0.014481                0.001295            0.007970  0.
066134          0.017927  0.011997          0.001072
4               0.864006            0.006907  0.014428                0.002170            0.008536  0.
070187          0.017869  0.014282          0.001612

FEVD for dpCO2
    xCO2 SW (wet) (umol/mol)  Licor Atm Pressure (hPa)   SST (C)  xCO2 SW (dry) (umol/mol)  fCO2 SW (sat) uatm
dfCO2  pCO2 SW (sat) uatm     dpCO2  pH (Total Scale)
0               0.946894            0.000030  0.002868                0.005613            0.000250  0.
044338          0.000000  0.000007          0.000000
1               0.816539            0.000052  0.005461                0.005315            0.002246  0.
163492          0.001456  0.004847          0.000592
2               0.794773            0.001213  0.019796                0.005311            0.004605  0.
156693          0.011459  0.005412          0.000738
3               0.767505            0.002791  0.020439                0.005221            0.008144  0.
163422          0.018982  0.012474          0.001022
4               0.760508            0.002861  0.020393                0.007683            0.008359  0.
165967          0.018926  0.013929          0.001374

FEVD for pH (Total Scale)
    xCO2 SW (wet) (umol/mol)  Licor Atm Pressure (hPa)   SST (C)  xCO2 SW (dry) (umol/mol)  fCO2 SW (sat) uatm
dfCO2  pCO2 SW (sat) uatm     dpCO2  pH (Total Scale)
0               0.693749            0.003095  0.002458                0.000021            0.000529  0.
000990          0.008792  0.000069          0.290298
1               0.570901            0.006868  0.015183                0.000801            0.000462  0.
082193          0.017320  0.004213          0.302061
2               0.579775            0.009198  0.020795                0.001561            0.000459  0.
078065          0.021261  0.008165          0.280721
```

```
3                0.560164              0.009211  0.020912           0.001541          0.008083  0.
089315        0.025524  0.015224     0.270025
4                0.552060              0.009076  0.020606           0.002063          0.008645  0.
089891        0.026865  0.022284     0.268510
```

In [43]:

```python
def invert_transformation(df_train, df_forecast, second_diff=False):
    """Revert back the differencing to get the forecast to original scale."""
    df_fc = df_forecast.copy()
    columns = df_train.columns
    for col in columns:
        # Roll back 2nd Diff
        if second_diff:
            df_fc[str(col)+'_1d'] = (df_train[col].iloc[-1]-df_train[col].iloc[-2]) + df_fc[str(col)+'_2d'].cumsum(
        # Roll back 1st Diff
        df_fc[col] = df_train[col].iloc[-1] + df_fc[str(col)+'_1d'].cumsum()
    return df_fc
```

In [45]:

```python
# Creating result dataframe
df_results = invert_transformation(train, df_forecast, second_diff=False)
df_results["Type"] = "Forecast"
df_results.head(2)
```

Out[45]:

	xCO2 SW (wet) (umol/mol)_1d	Licor Atm Pressure (hPa)_1d	SST (C)_1d	xCO2 SW (dry) (umol/mol)_1d	fCO2 SW (sat) uatm_1d	dfCO2_1d	pCO2 SW (sat) uatm_1d	dpCO2_1d	pH (Total Scale)_1d	xCO2 SW (wet) (umol/mol)	Lic P
Datetime											
2013-04-25	-3.011511	-0.113039	-0.032469	-3.052355	-2.919747	-3.185919	-2.920522	-3.220280	0.005145	410.188489	1006
2013-04-26	1.503981	-0.063022	0.016578	1.531379	1.411335	1.539328	1.414204	1.548054	-0.002005	411.692470	1006

In [46]:

```python
# Dropping all features that did not have significant correlation with target, pH
df_results = df_results.drop(["xCO2 SW (wet) (umol/mol)_1d", "Licor Atm Pressure (hPa)_1d", "SST (C)_1d", "xCO2 SW
                   axis = 1)
df_results["Type"] = "Forecast"
df_results.head(3)
```

Out[46]:

	xCO2 SW (wet) (umol/mol)	Licor Atm Pressure (hPa)	SST (C)	xCO2 SW (dry) (umol/mol)	fCO2 SW (sat) uatm	dfCO2	pCO2 SW (sat) uatm	dpCO2	pH (Total Scale)	Type
Datetime										
2013-04-25	410.188489	1006.661961	29.310906	412.885145	392.705253	16.126581	393.879478	16.167220	8.009645	Forecast
2013-04-26	411.692470	1006.598939	29.327484	414.416524	394.116588	17.665909	395.293681	17.715274	8.007640	Forecast
2013-04-27	413.447156	1006.569063	29.335133	416.162603	395.752858	19.471398	396.934431	19.527458	8.005578	Forecast

In [47]:

```python
df_results.index
```

Out[47]:

```
DatetimeIndex(['2013-04-25', '2013-04-26', '2013-04-27', '2013-04-28',
               '2013-05-02', '2013-05-03', '2013-05-05', '2013-05-07',
               '2013-05-08', '2013-05-09',
               ...
               '2014-05-13', '2014-05-14', '2014-05-15', '2014-05-16',
               '2014-05-17', '2014-05-18', '2014-05-19', '2014-05-20',
               '2014-05-21', '2014-05-22'],
              dtype='datetime64[ns]', name='Datetime', length=302, freq=None)
```

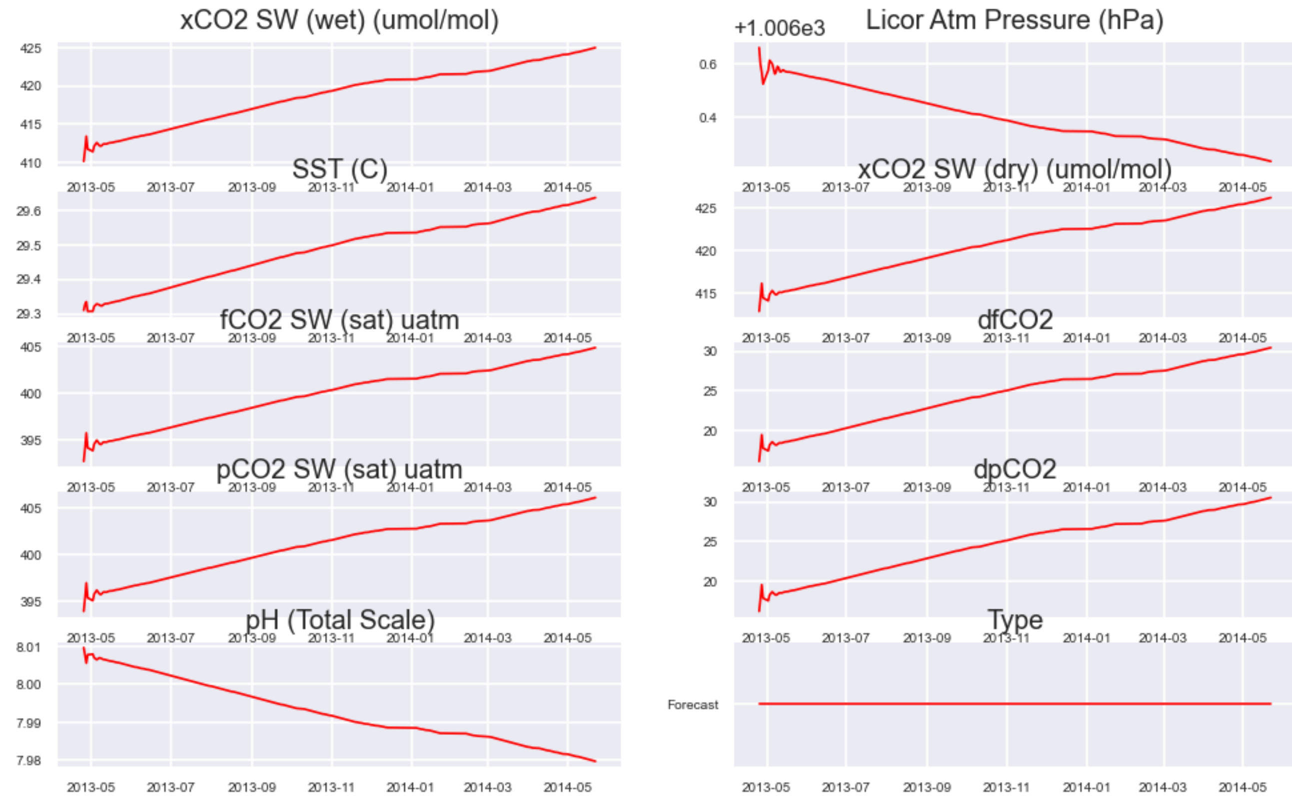
In [44]:

```python
df_results.index = df_results.index.to_timestamp()
```

In [48]:

```python
fig, axes = plt.subplots(nrows=5, ncols=2, dpi=120, figsize=(10,6), )
for i, ax in enumerate(axes.flatten()):
    data = df_results[df_results.columns[i]]
    ax.plot(data, color='red', linewidth=1)
```

```
    # Decorations
    ax.set_title(df_results.columns[i])
    ax.xaxis.set_ticks_position('none')
    ax.yaxis.set_ticks_position('none')
    ax.spines["top"].set_alpha(0)
    ax.tick_params(labelsize=6)
plt.show();
```



In [49]:

```
df_actual = pd.DataFrame(chuuk_df, index = df_results.index, columns = chuuk_df.columns)
df_actual["Type"] = "Actual"
df_actual
```

Out[49]:

Datetime	xCO2 SW (wet) (umol/mol)	Licor Atm Pressure (hPa)	SST (C)	xCO2 SW (dry) (umol/mol)	fCO2 SW (sat) uatm	dfCO2	pCO2 SW (sat) uatm	dpCO2	pH (Total Scale)	Type
2013-04-25	396.5750	1006.8125	28.933125	400.6250	381.4250	3.0500	382.5875	3.0625	8.040000	Actual
2013-04-26	398.0875	1006.7750	28.989875	402.2625	382.9375	4.7875	384.0875	4.8000	8.034750	Actual
2013-04-27	399.6750	1006.7125	28.981000	403.9000	384.5125	6.2000	385.6500	6.2125	8.035625	Actual
2013-04-28	395.5625	1007.1500	28.957750	399.6500	380.6250	2.7375	381.7625	2.7625	8.040250	Actual
2013-05-02	397.1625	1007.3375	28.946375	401.4875	382.4500	4.6000	383.6250	4.6125	8.034500	Actual
...	...	...	...	...	...	...	...	...	...	...
2014-05-18	403.8875	1007.1750	29.228875	406.6750	387.0750	9.7500	388.2625	9.7875	8.015125	Actual
2014-05-19	403.6375	1006.4375	29.187875	406.2500	386.4250	9.0625	387.6000	9.0625	8.021625	Actual
2014-05-20	408.7500	1006.5250	29.233250	411.3875	391.3000	14.3375	392.4875	14.3750	8.015625	Actual
2014-05-21	413.2750	1006.8375	29.358375	416.1250	395.8250	19.6625	397.0000	19.7125	8.006750	Actual

Datetime	xCO2 SW (wet) (umol/mol)	Licor Atm Pressure (hPa)	SST (C)	xCO2 SW (dry) (umol/mol)	fCO2 SW (sat) uatm	dfCO2	pCO2 SW (sat) uatm	dpCO2	pH (Total Scale)	Type
2014-05-22	413.2000	1006.7750	29.343375	415.9375	395.6250	19.3125	396.8000	19.3875	8.004500	Actual

302 rows × 10 columns

In [50]:
```python
# combining dataframes for better visual
dfs = [df_actual, df_results]
df_vis = df_actual.append(df_results)
df_vis
```

Out[50]:

Datetime	xCO2 SW (wet) (umol/mol)	Licor Atm Pressure (hPa)	SST (C)	xCO2 SW (dry) (umol/mol)	fCO2 SW (sat) uatm	dfCO2	pCO2 SW (sat) uatm	dpCO2	pH (Total Scale)	Type
2013-04-25	396.575000	1006.812500	28.933125	400.625000	381.425000	3.050000	382.587500	3.062500	8.040000	Actual
2013-04-26	398.087500	1006.775000	28.989875	402.262500	382.937500	4.787500	384.087500	4.800000	8.034750	Actual
2013-04-27	399.675000	1006.712500	28.981000	403.900000	384.512500	6.200000	385.650000	6.212500	8.035625	Actual
2013-04-28	395.562500	1007.150000	28.957750	399.650000	380.625000	2.737500	381.762500	2.762500	8.040250	Actual
2013-05-02	397.162500	1007.337500	28.946375	401.487500	382.450000	4.600000	383.625000	4.612500	8.034500	Actual
...	...	...	...	...	...	...	...	...	...	...
2014-05-18	424.805184	1006.237609	29.631702	426.087817	404.788438	30.292703	405.993113	30.381310	7.980094	Forecast
2014-05-19	424.848504	1006.236421	29.632764	426.126310	404.823572	30.334274	406.028341	30.423002	7.980002	Forecast
2014-05-20	424.891824	1006.235234	29.633827	426.164803	404.858705	30.375845	406.063568	30.464694	7.979910	Forecast
2014-05-21	424.935143	1006.234046	29.634889	426.203297	404.893838	30.417416	406.098795	30.506387	7.979817	Forecast
2014-05-22	424.978463	1006.232859	29.635951	426.241790	404.928971	30.458987	406.134023	30.548079	7.979725	Forecast

604 rows × 10 columns

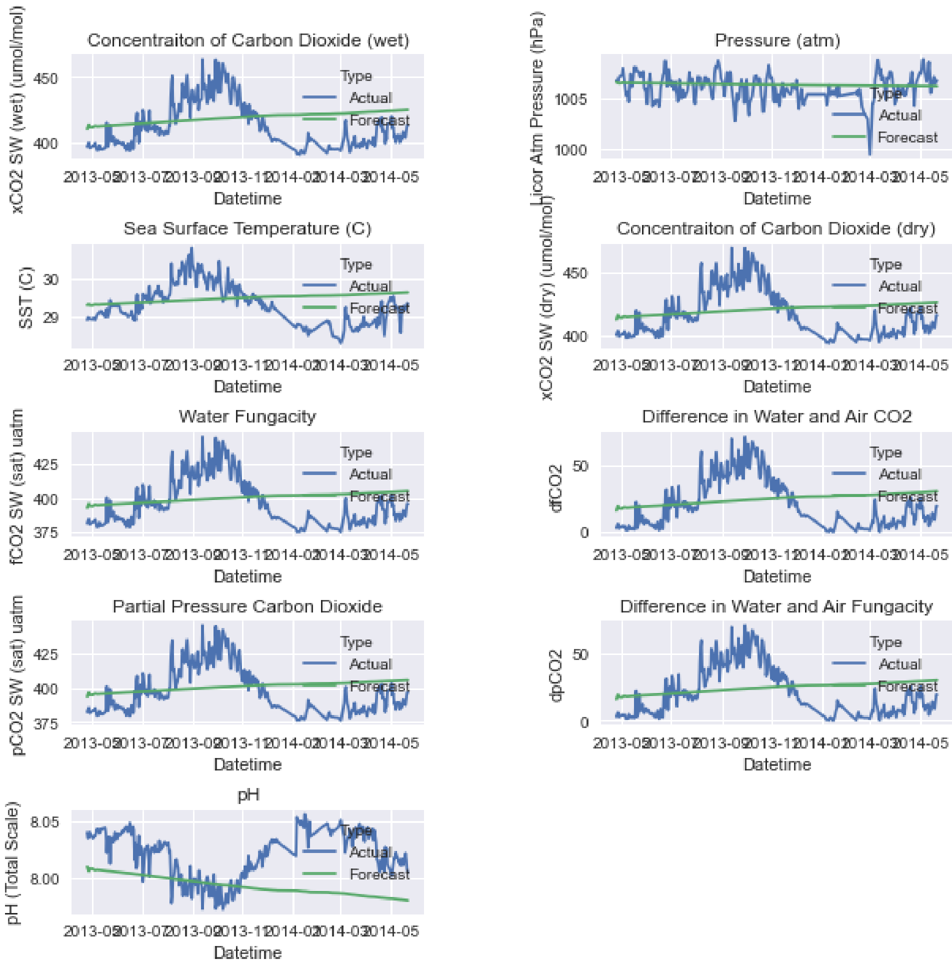
In [51]:
```python
type(df_vis.index)
```

Out[51]:
```
pandas.core.indexes.datetimes.DatetimeIndex
```

In [52]:
```python
plt.style.use('seaborn')
fig, axis = plt.subplots(nrows = 5, ncols = 2)
fig.set_size_inches(10,10)
fig.subplots_adjust(wspace = 0.5, hspace = 0.8)

# plot 1
sns.lineplot(df_vis.index, df_vis["xCO2 SW (wet) (umol/mol)"], hue = df_vis["Type"], ax = axis[0,0]).set_title("Cor
# plot 2
sns.lineplot(df_vis.index, df_vis["Licor Atm Pressure (hPa)"], hue = df_vis["Type"], ax = axis[0,1]).set_title("Pre
# plot 3
sns.lineplot(df_vis.index, df_vis["SST (C)"], hue = df_vis["Type"], ax = axis[1,0]).set_title("Sea Surface Temperat
# plot 4
sns.lineplot(df_vis.index, df_vis["xCO2 SW (dry) (umol/mol)"], hue = df_vis["Type"], ax = axis[1,1]).set_title("Cor
# plot 5
sns.lineplot(df_vis.index, df_vis["fCO2 SW (sat) uatm"], hue = df_vis["Type"], ax = axis[2,0]).set_title("Water Fur
# plot 6
sns.lineplot(df_vis.index, df_vis["dfCO2"], hue = df_vis["Type"], ax = axis[2,1]).set_title("Difference in Water ar
```

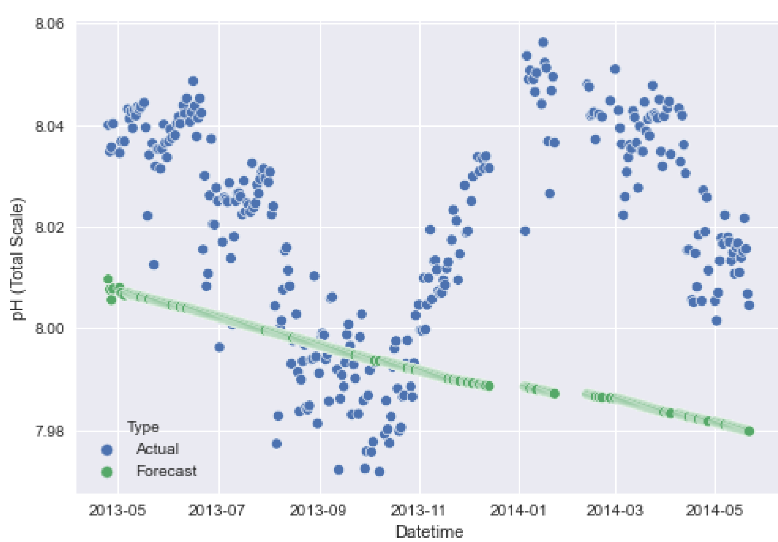
```
# plot 7
sns.lineplot(df_vis.index, df_vis["pCO2 SW (sat) uatm"], hue = df_vis["Type"], ax = axis[3,0]).set_title("Partial P
# plot 8
sns.lineplot(df_vis.index, df_vis["dpCO2"], hue = df_vis["Type"], ax = axis[3,1]).set_title("Difference in Water ar
# plot 9
sns.lineplot(df_vis.index, df_vis["pH (Total Scale)"], hue = df_vis["Type"], ax = axis[4,0]).set_title("pH")

ax = axis[4,1].set_visible(False)
```

```
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables a
s keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other argum
ents without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables a
s keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other argum
ents without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables a
s keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other argum
ents without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables a
s keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other argum
ents without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables a
s keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other argum
ents without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables a
s keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other argum
ents without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables a
s keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other argum
ents without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables a
s keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other argum
ents without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
C:\Users\datre\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables a
s keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other argum
ents without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
```



In [53]:
```
# Isolating pH results
sns.scatterplot(x = df_vis.index, y = df_vis["pH (Total Scale)"], hue = df_vis.Type)
```

Out[53]: <AxesSubplot:xlabel='Datetime', ylabel='pH (Total Scale)'>



In [ ]:

In [54]:
```
from keras.preprocessing.sequence import TimeseriesGenerator
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import LSTM
from tensorflow.keras.layers import Dense
from tensorflow.keras.layers import Flatten
from sklearn.preprocessing import MinMaxScaler
```

```
from sklearn.metrics import mean_squared_error
import math
from keras.models import load_model
from sklearn.metrics import r2_score
import plotly.graph_objects as go
```

In [55]:
```
def create_dataset(dataset, lookback=1):
    dataX = []
    dataY = []
    for i in range(len(dataset) - lookback - 1):
        a = dataset[i: (i+lookback), 0]
        dataX.append(a)
        dataY.append(dataset[i+lookback,0])
    return np.array(dataX), np.array(dataY)
```

In [56]:
```
# input data
data = chuuk_df["pH (Total Scale)"].values
data = data.astype("float32")
```

In [57]:
```
# correcting shape of data
scaler = MinMaxScaler(feature_range=(0,1))
data = data.reshape(-1,1)
data = scaler.fit_transform(data)
```

In [58]:
```
# Splitting into test and train sets
train = data[:int(len(data)*0.8), :]
test = data[int(len(data)*0.8):, :]
```

In [59]:
```
lookback = 1
trainX, trainY = create_dataset(train, lookback)
testX, testY = create_dataset(test, lookback)
```

In [60]:
```
trainX = np.reshape(trainX, (trainX.shape[0], 1, trainX.shape[1]))
testX = np.reshape(testX, (testX.shape[0], 1, testX.shape[1]))
```

In [61]:
```
model = Sequential()
# Check LSTM
model.add(LSTM(4, input_shape=(1, lookback)))
#return_sequences=True, model.add(LSTM(4))
model.add(Dense(1))
model.compile(loss='mean_squared_error', optimizer='adam', metrics=['mae'])
model.fit(trainX, trainY, epochs=100, batch_size=1, verbose=2)
```

```
Epoch 1/100
301/301 - 5s - loss: 0.1952 - mae: 0.3815
Epoch 2/100
301/301 - 1s - loss: 0.0227 - mae: 0.1272
Epoch 3/100
301/301 - 1s - loss: 0.0180 - mae: 0.1118
Epoch 4/100
301/301 - 1s - loss: 0.0156 - mae: 0.1039
Epoch 5/100
301/301 - 1s - loss: 0.0135 - mae: 0.0952
Epoch 6/100
301/301 - 1s - loss: 0.0116 - mae: 0.0867
Epoch 7/100
301/301 - 1s - loss: 0.0099 - mae: 0.0780
Epoch 8/100
301/301 - 1s - loss: 0.0087 - mae: 0.0723
Epoch 9/100
301/301 - 1s - loss: 0.0079 - mae: 0.0674
Epoch 10/100
301/301 - 1s - loss: 0.0073 - mae: 0.0625
Epoch 11/100
301/301 - 1s - loss: 0.0071 - mae: 0.0621
Epoch 12/100
301/301 - 1s - loss: 0.0071 - mae: 0.0617
Epoch 13/100
301/301 - 1s - loss: 0.0070 - mae: 0.0610
Epoch 14/100
301/301 - 1s - loss: 0.0070 - mae: 0.0612
Epoch 15/100
301/301 - 1s - loss: 0.0069 - mae: 0.0608
Epoch 16/100
301/301 - 1s - loss: 0.0069 - mae: 0.0608
Epoch 17/100
301/301 - 1s - loss: 0.0069 - mae: 0.0606
```

```
Epoch 18/100
301/301 - 1s - loss: 0.0069 - mae: 0.0608
Epoch 19/100
301/301 - 1s - loss: 0.0070 - mae: 0.0608
Epoch 20/100
301/301 - 1s - loss: 0.0069 - mae: 0.0605
Epoch 21/100
301/301 - 1s - loss: 0.0069 - mae: 0.0610
Epoch 22/100
301/301 - 0s - loss: 0.0069 - mae: 0.0606
Epoch 23/100
301/301 - 1s - loss: 0.0069 - mae: 0.0613
Epoch 24/100
301/301 - 0s - loss: 0.0069 - mae: 0.0609
Epoch 25/100
301/301 - 1s - loss: 0.0070 - mae: 0.0608
Epoch 26/100
301/301 - 1s - loss: 0.0069 - mae: 0.0608
Epoch 27/100
301/301 - 1s - loss: 0.0069 - mae: 0.0608
Epoch 28/100
301/301 - 1s - loss: 0.0069 - mae: 0.0600
Epoch 29/100
301/301 - 1s - loss: 0.0070 - mae: 0.0616
Epoch 30/100
301/301 - 1s - loss: 0.0069 - mae: 0.0601
Epoch 31/100
301/301 - 1s - loss: 0.0070 - mae: 0.0616
Epoch 32/100
301/301 - 1s - loss: 0.0070 - mae: 0.0609
Epoch 33/100
301/301 - 1s - loss: 0.0069 - mae: 0.0603
Epoch 34/100
301/301 - 1s - loss: 0.0069 - mae: 0.0606
Epoch 35/100
301/301 - 1s - loss: 0.0069 - mae: 0.0607
Epoch 36/100
301/301 - 1s - loss: 0.0070 - mae: 0.0607
Epoch 37/100
301/301 - 1s - loss: 0.0069 - mae: 0.0608
Epoch 38/100
301/301 - 1s - loss: 0.0070 - mae: 0.0608
Epoch 39/100
301/301 - 1s - loss: 0.0069 - mae: 0.0606
Epoch 40/100
301/301 - 1s - loss: 0.0069 - mae: 0.0621
Epoch 41/100
301/301 - 1s - loss: 0.0068 - mae: 0.0600
Epoch 42/100
301/301 - 1s - loss: 0.0069 - mae: 0.0617
Epoch 43/100
301/301 - 1s - loss: 0.0069 - mae: 0.0605
Epoch 44/100
301/301 - 1s - loss: 0.0070 - mae: 0.0612
Epoch 45/100
301/301 - 1s - loss: 0.0068 - mae: 0.0605
Epoch 46/100
301/301 - 1s - loss: 0.0069 - mae: 0.0608
Epoch 47/100
301/301 - 1s - loss: 0.0070 - mae: 0.0610
Epoch 48/100
301/301 - 1s - loss: 0.0069 - mae: 0.0608
Epoch 49/100
301/301 - 0s - loss: 0.0067 - mae: 0.0596
Epoch 50/100
301/301 - 0s - loss: 0.0069 - mae: 0.0601
Epoch 51/100
301/301 - 0s - loss: 0.0069 - mae: 0.0610
Epoch 52/100
301/301 - 1s - loss: 0.0069 - mae: 0.0604
Epoch 53/100
301/301 - 1s - loss: 0.0069 - mae: 0.0609
Epoch 54/100
301/301 - 1s - loss: 0.0069 - mae: 0.0608
Epoch 55/100
301/301 - 1s - loss: 0.0070 - mae: 0.0605
Epoch 56/100
301/301 - 1s - loss: 0.0069 - mae: 0.0612
Epoch 57/100
301/301 - 1s - loss: 0.0070 - mae: 0.0609
Epoch 58/100
301/301 - 1s - loss: 0.0069 - mae: 0.0604
```

```
Epoch 59/100
301/301 - 1s - loss: 0.0069 - mae: 0.0602
Epoch 60/100
301/301 - 1s - loss: 0.0069 - mae: 0.0614
Epoch 61/100
301/301 - 1s - loss: 0.0069 - mae: 0.0603
Epoch 62/100
301/301 - 1s - loss: 0.0070 - mae: 0.0605
Epoch 63/100
301/301 - 1s - loss: 0.0069 - mae: 0.0608
Epoch 64/100
301/301 - 1s - loss: 0.0069 - mae: 0.0607
Epoch 65/100
301/301 - 1s - loss: 0.0070 - mae: 0.0604
Epoch 66/100
301/301 - 1s - loss: 0.0069 - mae: 0.0608
Epoch 67/100
301/301 - 1s - loss: 0.0069 - mae: 0.0605
Epoch 68/100
301/301 - 1s - loss: 0.0069 - mae: 0.0610
Epoch 69/100
301/301 - 1s - loss: 0.0067 - mae: 0.0612
Epoch 70/100
301/301 - 1s - loss: 0.0069 - mae: 0.0609
Epoch 71/100
301/301 - 1s - loss: 0.0069 - mae: 0.0610
Epoch 72/100
301/301 - 1s - loss: 0.0070 - mae: 0.0605
Epoch 73/100
301/301 - 1s - loss: 0.0069 - mae: 0.0603
Epoch 74/100
301/301 - 1s - loss: 0.0069 - mae: 0.0606
Epoch 75/100
301/301 - 1s - loss: 0.0068 - mae: 0.0603
Epoch 76/100
301/301 - 0s - loss: 0.0070 - mae: 0.0617
Epoch 77/100
301/301 - 0s - loss: 0.0069 - mae: 0.0607
Epoch 78/100
301/301 - 0s - loss: 0.0070 - mae: 0.0607
Epoch 79/100
301/301 - 1s - loss: 0.0069 - mae: 0.0596
Epoch 80/100
301/301 - 1s - loss: 0.0067 - mae: 0.0605
Epoch 81/100
301/301 - 1s - loss: 0.0069 - mae: 0.0607
Epoch 82/100
301/301 - 1s - loss: 0.0068 - mae: 0.0602
Epoch 83/100
301/301 - 1s - loss: 0.0070 - mae: 0.0601
Epoch 84/100
301/301 - 1s - loss: 0.0069 - mae: 0.0606
Epoch 85/100
301/301 - 1s - loss: 0.0070 - mae: 0.0608
Epoch 86/100
301/301 - 1s - loss: 0.0070 - mae: 0.0616
Epoch 87/100
301/301 - 1s - loss: 0.0068 - mae: 0.0604
Epoch 88/100
301/301 - 1s - loss: 0.0070 - mae: 0.0607
Epoch 89/100
301/301 - 1s - loss: 0.0069 - mae: 0.0598
Epoch 90/100
301/301 - 1s - loss: 0.0069 - mae: 0.0603
Epoch 91/100
301/301 - 1s - loss: 0.0069 - mae: 0.0612
Epoch 92/100
301/301 - 1s - loss: 0.0069 - mae: 0.0611
Epoch 93/100
301/301 - 1s - loss: 0.0069 - mae: 0.0606
Epoch 94/100
301/301 - 1s - loss: 0.0068 - mae: 0.0606
Epoch 95/100
301/301 - 1s - loss: 0.0069 - mae: 0.0607
Epoch 96/100
301/301 - 1s - loss: 0.0070 - mae: 0.0620
Epoch 97/100
301/301 - 1s - loss: 0.0069 - mae: 0.0607
Epoch 98/100
301/301 - 1s - loss: 0.0069 - mae: 0.0612
Epoch 99/100
301/301 - 1s - loss: 0.0070 - mae: 0.0609
```

```
Epoch 100/100
301/301 - 1s - loss: 0.0068 - mae: 0.0597
```

Out[61]: `<tensorflow.python.keras.callbacks.History at 0x1f25c186970>`

In [62]:
```
# make predictions
trainPredict = model.predict(trainX)
testPredict = model.predict(testX)
```

In [63]:
```
#nsamples, nx, ny = trainY.shape
#trainY = trainY.reshape((nsamples,nx*ny))
```

In [64]:
```
# invert predictions
trainPredict = scaler.inverse_transform(trainPredict)
trainY = scaler.inverse_transform([trainY])
testPredict = scaler.inverse_transform(testPredict)
testY = scaler.inverse_transform([testY])
# calculate root mean squared error
trainScore = math.sqrt(mean_squared_error(trainY[0], trainPredict[:,0]))
print('Train Score: %.2f RMSE' % (trainScore))
testScore = math.sqrt(mean_squared_error(testY[0], testPredict[:,0]))
print('Test Score: %.2f RMSE' % (testScore))
```

```
Train Score: 0.01 RMSE
Test Score: 0.01 RMSE
```

In [65]:
```
trainPredictPlot = np.empty_like(data)
trainPredictPlot[:, :] = np.nan
trainPredictPlot[lookback:len(trainPredict)+lookback, :] = trainPredict
# shift test predictions for plotting
testPredictPlot = np.empty_like(data)
testPredictPlot[:, :] = np.nan
testPredictPlot[len(trainPredict)+(lookback*2)+1:len(data)-1, :] = testPredict
# plot baseline and predictions
plt.plot(scaler.inverse_transform(data), label="pH Data")
plt.plot(trainPredictPlot, label="Train Prediction")
plt.plot(testPredictPlot, label="Test Prediction")
plt.xlabel("Observations")
plt.ylabel("pH (total scale)")
plt.legend()
plt.show()
```

In [66]:
```
model.save("pH_model.h5")
```

In [67]:
```
model = load_model("pH_model.h5")
```

In [68]:
```
trainScore = math.sqrt(r2_score(trainY[0], trainPredict[:,0]))
print('Train Score: %.2f R2' % (trainScore))
testScore = math.sqrt(r2_score(testY[0], testPredict[:,0]))
print('Test Score: %.2f R2' % (testScore))
```

```
Train Score: 0.92 R2
Test Score: 0.77 R2
```

```
In [ ]:
```