Rainfall_4_meteostations_Samos_01

November 22, 2017

1 Post-processing of the meteorological data

Here we look at the rainfall data obtained from the meteorological stations of Samos, Myloi, Airport, Ydroussa, Karlovassi

1.1 Meteorological stations' coordinates

Meteorological station
I. Samos (Vathi)
II. Myloi
III. Airport
IV. Ydroussa
V. Karlovassi

```
In [1]: import pandas as pd
        %matplotlib inline
        data = pd.read_csv ('Rainfall_meteostations_5_Samos_new_02.csv',
                            header=0,
                            index_col='Hydrological Year',
                            decimal=',')
        data.head(n=10)
```

Out[1]:

Hydrological Year	City of Samos	Myloi	Airport	Ydroussa	Karlovassi	Station	Count (Number of values)
1930-31	1226.8	NaN	NaN	NaN			
1931-32	753.5	NaN	NaN	NaN			
1932-33	600.7	NaN	NaN	NaN			
1933-34	803.5	NaN	NaN	NaN			
1934-35	672.0	NaN	NaN	NaN			
1935-36	1444.9	NaN	NaN	NaN			
1936-37	525.6	NaN	NaN	NaN			
1937-38	1246.0	NaN	NaN	NaN			
1938-39	1152.4	NaN	NaN	NaN			
1939-40	1169.2	NaN	NaN	NaN			

```
Hydrological Year
1930-31                          NaN                      1.0
1931-32                          NaN                      1.0
1932-33                          NaN                      1.0
1933-34                          NaN                      1.0
1934-35                          NaN                      1.0
1935-36                          NaN                      1.0
1936-37                          NaN                      1.0
1937-38                          NaN                      1.0
1938-39                          NaN                      1.0
1939-40                          NaN                      1.0
```

In [2]:
```
# Delete final column (Count)
data_raw = data.loc[:,['City of Samos','Myloi','Airport','Ydroussa','Karlovassi Statio
data_raw.head(n=5)
```

Out[2]:

Hydrological Year	City of Samos	Myloi	Airport	Ydroussa	Karlovassi Station
1930-31	1226.8	NaN	NaN	NaN	NaN
1931-32	753.5	NaN	NaN	NaN	NaN
1932-33	600.7	NaN	NaN	NaN	NaN
1933-34	803.5	NaN	NaN	NaN	NaN
1934-35	672.0	NaN	NaN	NaN	NaN

In [3]:
```
data_raw.describe()
```

Out[3]:

	City of Samos	Myloi	Airport	Ydroussa	\
count	38.000000	43.000000	37.000000	7.000000	
mean	876.163158	868.644186	698.151351	815.685714	
std	278.236624	251.218348	207.439960	244.875978	
min	501.700000	450.100000	333.600000	571.500000	
25%	622.675000	680.350000	596.400000	618.550000	
50%	825.900000	839.900000	661.000000	832.300000	
75%	1128.500000	1030.900000	795.200000	903.950000	
max	1444.900000	1381.100000	1268.300000	1261.000000	

	Karlovassi Station
count	34.000000
mean	904.311765
std	233.374754
min	515.000000
25%	717.250000
50%	879.700000
75%	1033.125000
max	1366.500000

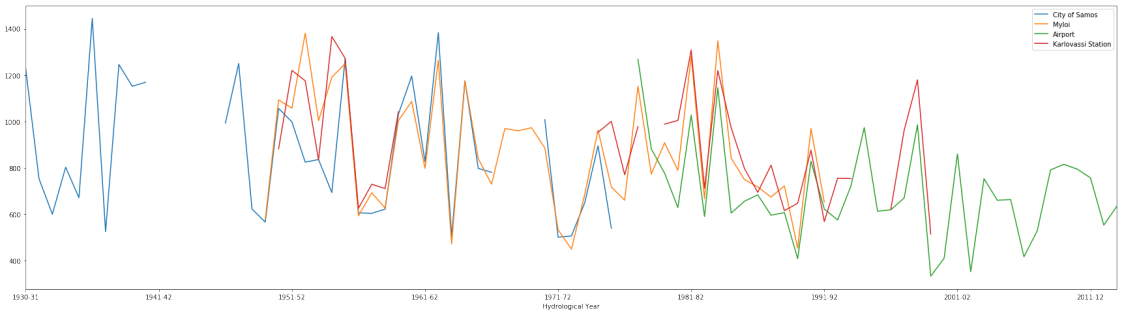
In [19]:
```
#Since Ydroussa's meteorological station data of only 7 hydrological years will be ex
data_new=data_raw.loc[:,['City of Samos','Myloi','Airport','Karlovassi Station']]
data_new.describe()
```

Out[19]:

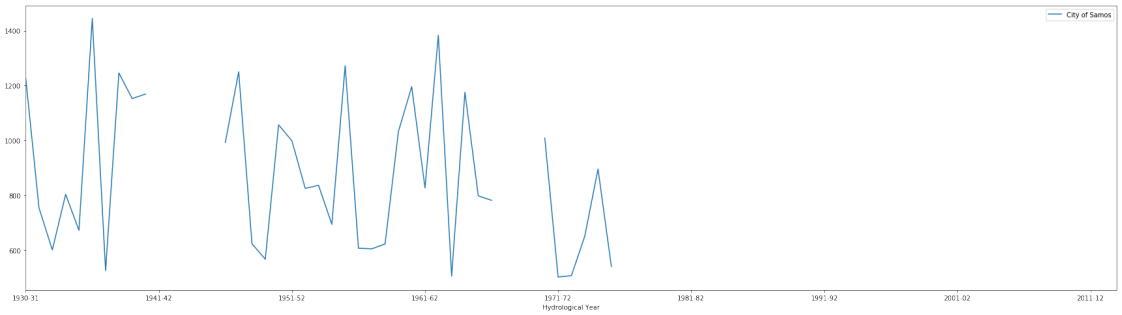
	City of Samos	Myloi	Airport	Karlovassi Station
count	38.000000	43.000000	37.000000	34.000000
mean	876.163158	868.644186	698.151351	904.311765
std	278.236624	251.218348	207.439960	233.374754
min	501.700000	450.100000	333.600000	515.000000
25%	622.675000	680.350000	596.400000	717.250000
50%	825.900000	839.900000	661.000000	879.700000
75%	1128.500000	1030.900000	795.200000	1033.125000
max	1444.900000	1381.100000	1268.300000	1366.500000

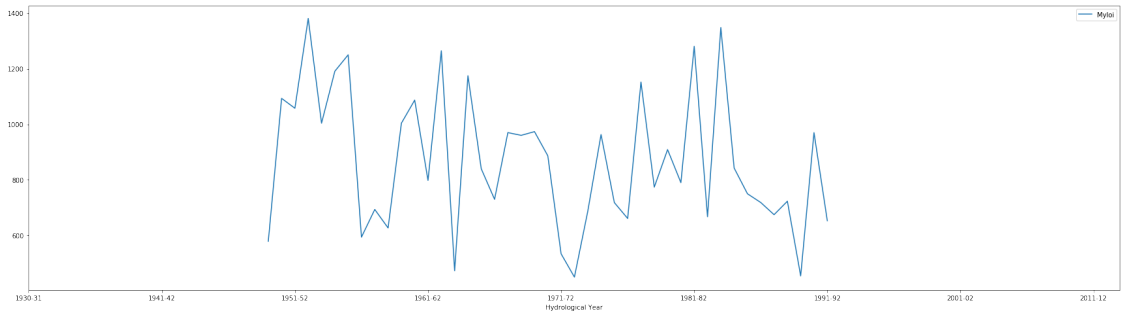
In [11]:
```
data_new.plot(figsize=(30,8))
```

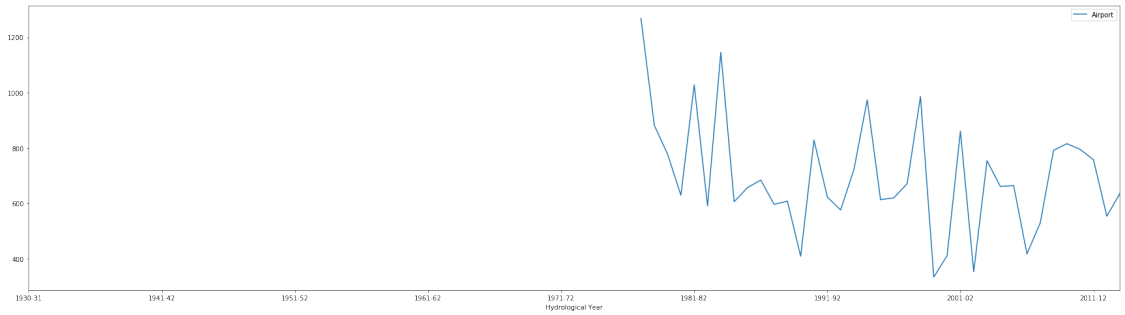
Out[11]: <matplotlib.axes._subplots.AxesSubplot at 0x7f409d8adb10>

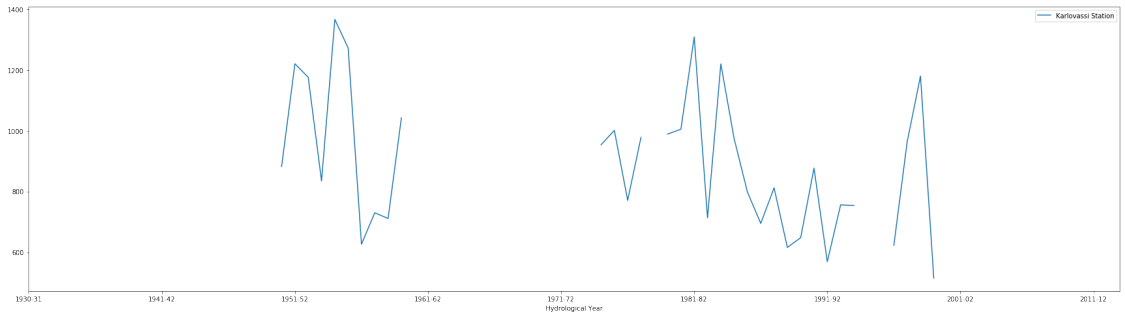


In [18]:
```
#Let's check each dataset in regard to continuity
data_samos=data_raw.loc[:,['City of Samos']].plot(figsize=(30,8))
data_myloi=data_raw.loc[:,['Myloi']].plot(figsize=(30,8))
data_airp=data_raw.loc[:,['Airport']].plot(figsize=(30,8))
data_karl=data_raw.loc[:,['Karlovassi Station']].plot(figsize=(30,8))
```





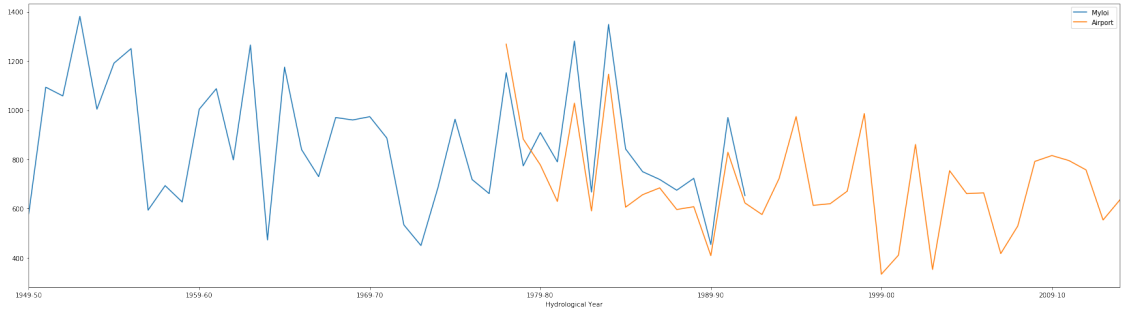




Plots indicate that Samos and Karlovassi station have major gaps and will be not be taken into account for this analysis

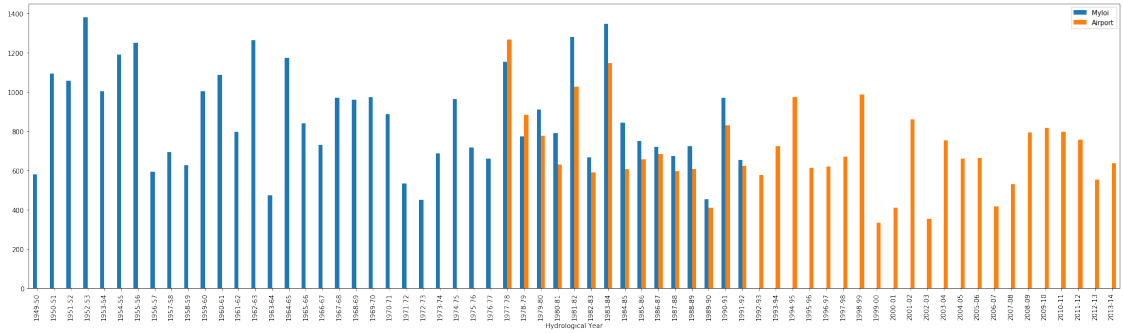
```
In [32]: #Let's concentrate at Myloi and Airport stations
         data_mylair=data_raw.loc[:,['Myloi','Airport']]

         # Creation of a new dataset that considers only the complete hydrological years where
         data_mylair_short = pd.concat([data_mylair['1949-50':'2013-14']])
         data_mylair_short.head()

         data_mylair_short.plot(figsize=(30,8))
```

Out[32]: <matplotlib.axes._subplots.AxesSubplot at 0x7f409d561e10>



In [34]:
```
# Histogramm of the total value per hydrological year for Myloi and airport station
data_mylair_short.plot.bar(yerr=data_mylair_short.std(),figsize=(30,8))
```

Out[34]: <matplotlib.axes._subplots.AxesSubplot at 0x7f409c91c8d0>



In [35]:
```
#Main statistics on these 2 datasets are as following:
data_mylair_short.describe()
```

Out[35]:

	Myloi	Airport
count	43.000000	37.000000
mean	868.644186	698.151351
std	251.218348	207.439960
min	450.100000	333.600000
25%	680.350000	596.400000
50%	839.900000	661.000000
75%	1030.900000	795.200000
max	1381.100000	1268.300000

In [36]:
```
data_mylair_overl = pd.concat([data_mylair_short['1977-78':'1991-92']])
```

In [39]:
```
data_mylair_overl.head(n=30)
```

Out[39]:

Hydrological Year	Myloi	Airport
1977-78	1152.0	1268.3
1978-79	773.9	882.5
1979-80	908.9	777.7
1980-81	790.0	629.3
1981-82	1280.7	1028.3
1982-83	667.3	590.8
1983-84	1348.3	1146.1
1984-85	842.6	605.8
1985-86	750.0	656.6
1986-87	718.5	684.3
1987-88	674.7	596.4
1988-89	723.1	607.8
1989-90	454.3	409.0
1990-91	970.0	828.9
1991-92	653.0	623.0