# HYDROGNOMON – OPEN SOURCE SOFTWARE FOR THE ANALYSIS OF HYDROLOGICAL DATA

**European Geosciences Union (EGU) General Assembly, Vienna, Austria, 2 - 7 May 2010**
**Session HS5.5/NP6.10: *Stochastics in hydrometeorological processes: from point to global spatial scales and from minute to climatic time scales***

**Stefanos Kozanis, Antonis Christofides, Nikos Mamassis, Andreas Efstratiadis, and Demetris Koutsoyiannis**
**Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Greece**

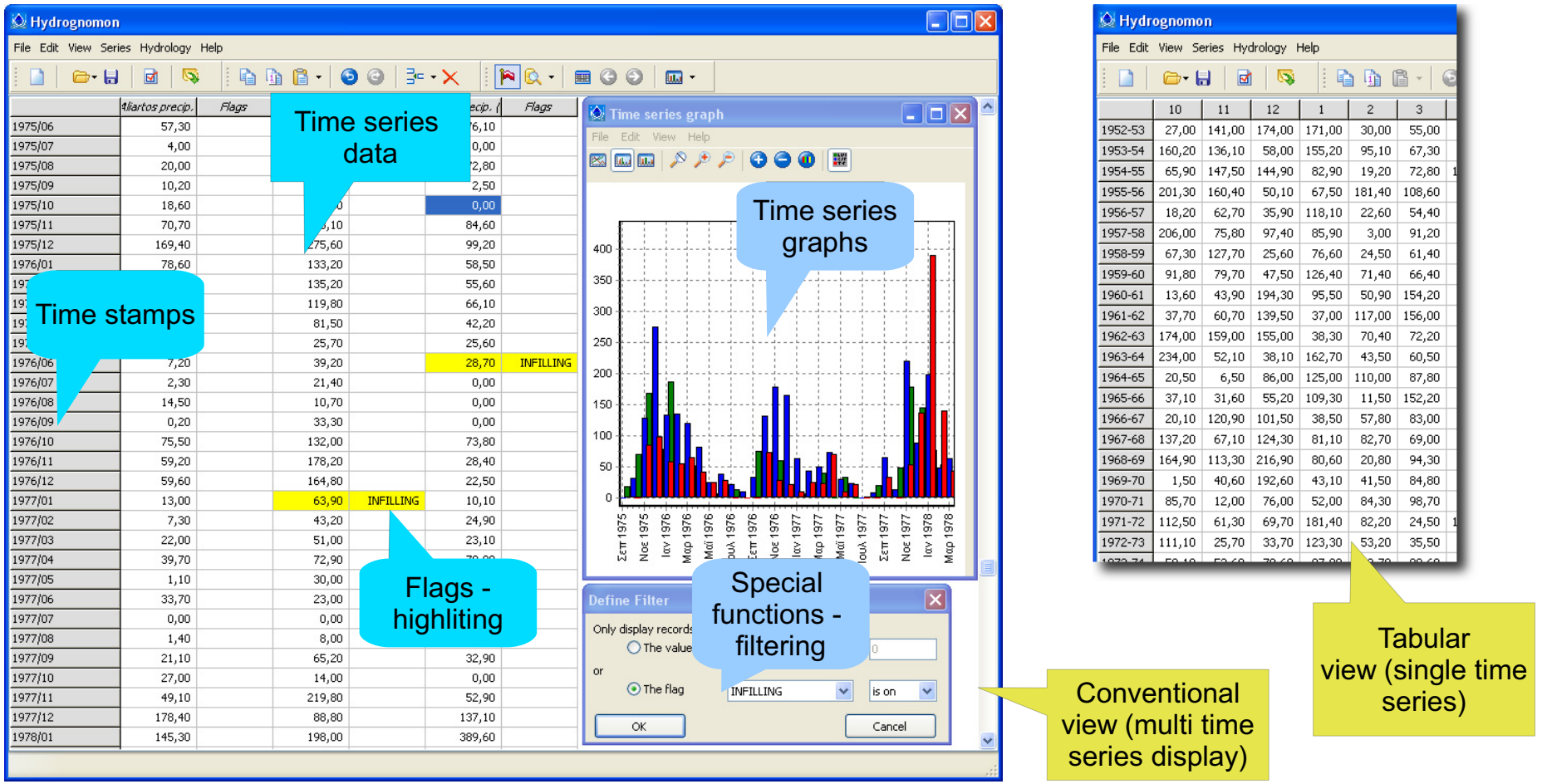
# 1. Hydrognomon — Introduction

Hydrognomon is a software tool for the processing of hydrological data. Data are imported through standard text files, spreadsheets or by typing. Available processing techniques include time step aggregation and regularization, interpolation, regression analysis and infilling of missing values, consistency tests, data filtering, graphical and tabular visualisation of time series, and more. Hydrognomon is free software licensed under the GPLv3.

It supports several time steps, from the finest minute scales up to decades; specific cases of irregular time steps and offsets are also supported. The program also includes common hydrological applications, such as evapotranspiration modelling, stage-discharge analysis, homogeneity tests, areal integration of point data series, processing of hydrometric data, as well as lumped hydrological modelling with automatic calibration facilities.

Here the emphasis is given on the statistical module of Hydrognomon, which provides tools for data exploration, fitting of distribution functions, statistical prediction, Monte-Carlo simulation, determination of confidence limits, analysis of extremes, and construction of ombrian (IDF) curves.

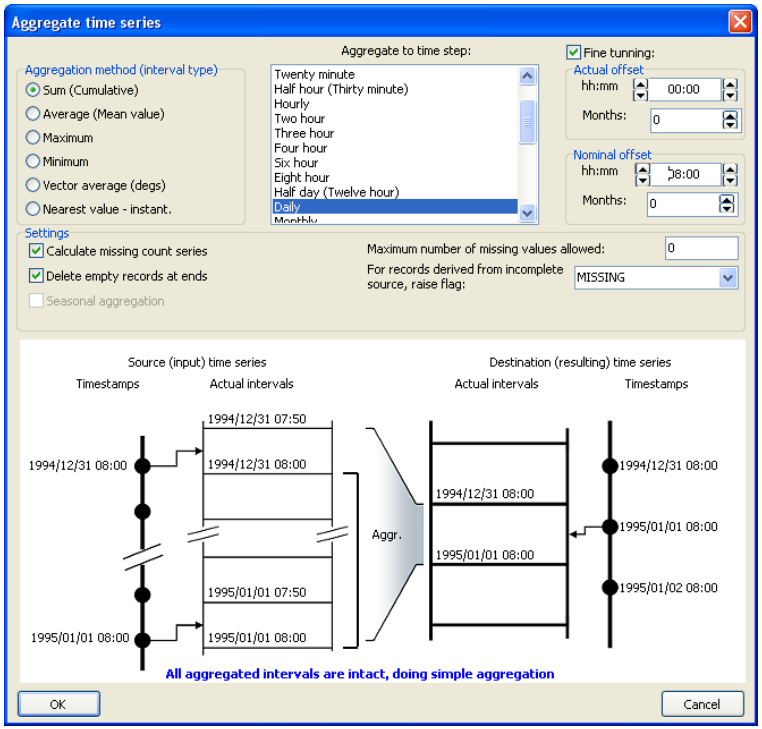
# 2. User interface



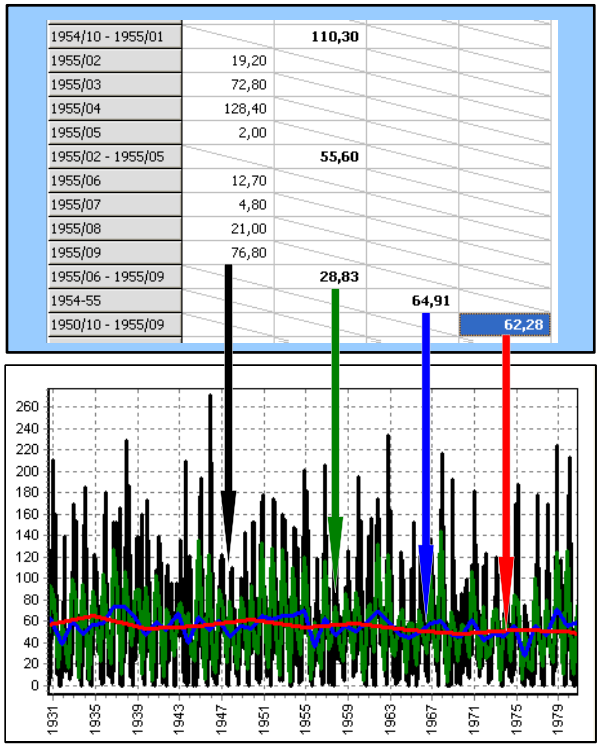
# 3. Time integration of series

Hydrognomon's time integration algorithms can handle from minute time steps of automatic gauging stations, up to daily / monthly / seasonal / annual or over-year time steps.

Any transformation is allowed between different time steps, with features such as time offsets, interpolated values for shifted time intervals, hydrological year scales with adjusted origin, seasonal aggregation, and more.



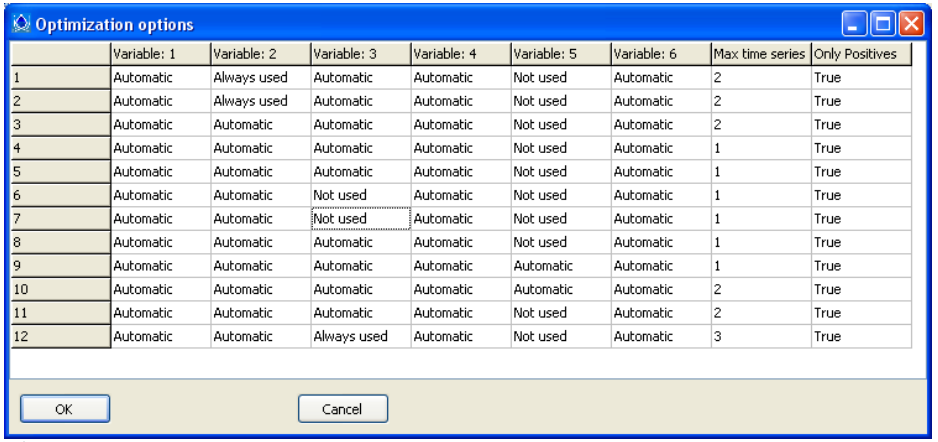

*Interface – Aggregation parameters*




*Average values from 1 month to a term period (4 months), annual, 5-year*

# 4. Time series regression and infilling

The time series regression and infilling tool can determine the correlation between different time series. A linear combination relating several independent variables to a dependent variable can be established using optimization routines. This can work seasonally, if needed, and several constraints can be defined.

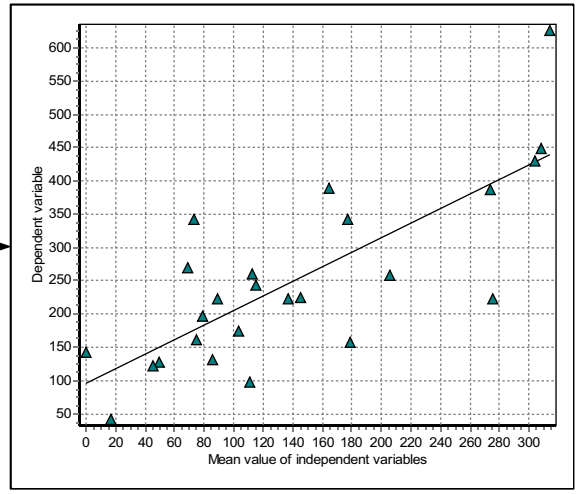



| | Variable: 1 | Variable: 2 | Variable: 3 | Variable: 4 | Variable: 5 | Variable: 6 | Max time series | Only Positives |
|---|---|---|---|---|---|---|---|---|
| 1 | Automatic | Always used | Automatic | Automatic | Not used | Automatic | 2 | True |
| 2 | Automatic | Always used | Automatic | Automatic | Not used | Automatic | 2 | True |
| 3 | Automatic | Automatic | Automatic | Automatic | Not used | Automatic | 2 | True |
| 4 | Automatic | Automatic | Automatic | Automatic | Not used | Automatic | 1 | True |
| 5 | Automatic | Automatic | Automatic | Automatic | Not used | Automatic | 1 | True |
| 6 | Automatic | Automatic | Not used | Automatic | Not used | Automatic | 1 | True |
| 7 | Automatic | Automatic | Not used | Automatic | Not used | Automatic | 1 | True |
| 8 | Automatic | Automatic | Automatic | Automatic | Not used | Automatic | 1 | True |
| 9 | Automatic | Automatic | Automatic | Automatic | Automatic | Automatic | 1 | True |
| 10 | Automatic | Automatic | Automatic | Automatic | Automatic | Automatic | 2 | True |
| 11 | Automatic | Automatic | Automatic | Automatic | Not used | Automatic | 2 | True |
| 12 | Automatic | Automatic | Always used | Automatic | Not used | Automatic | 3 | True |

*Model constraints, individually for each month*


| | N | D=R^2 | R | Dcr. | a0 | a1 | a2 | a3 | a4 | a5 | a6 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 25 | 0,508 | 0,713 | 0,160 | 36,279 | | 0,403 | 0,892 | | | |
| 11 | 25 | 0,528 | 0,726 | 0,160 | 66,687 | | | 1,358 | | | |
| 12 | 25 | 0,590 | 0,768 | 0,160 | 96,702 | | | 1,091 | | | |
| 1 | 24 | 0,056 | 0,236 | 0,167 | 139,670 | | 0,563 | | | | |
| 2 | 24 | 0,038 | 0,194 | 0,167 | 135,364 | | 0,141 | | 0,131 | | |
| 3 | 24 | 0,362 | 0,601 | 0,167 | 40,135 | | | 0,767 | 0,071 | | |
| 4 | 25 | 0,643 | 0,802 | 0,160 | 43,064 | | | | 1,032 | | |
| 5 | 25 | 0,168 | 0,410 | 0,160 | 74,001 | | | | 0,287 | | |
| 6 | 22 | 0,313 | 0,559 | 0,182 | 37,233 | 1,045 | | | | | |
| 7 | 17 | 0,609 | 0,781 | 0,235 | 26,290 | 1,662 | | | | | |
| 8 | 24 | 0,350 | 0,592 | 0,167 | 20,083 | | | | | | 0,562 |
| 9 | 22 | 0,438 | 0,662 | 0,182 | 17,797 | | | 1,095 | | | |

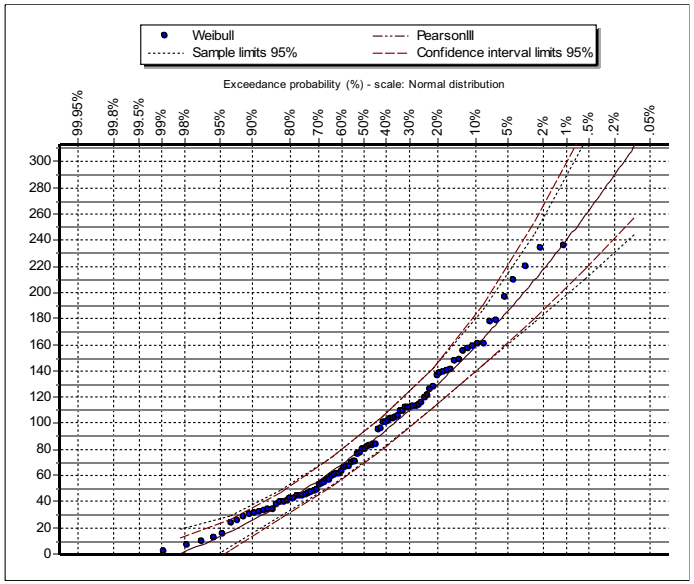
*Model ouput – regression parametres for each month*



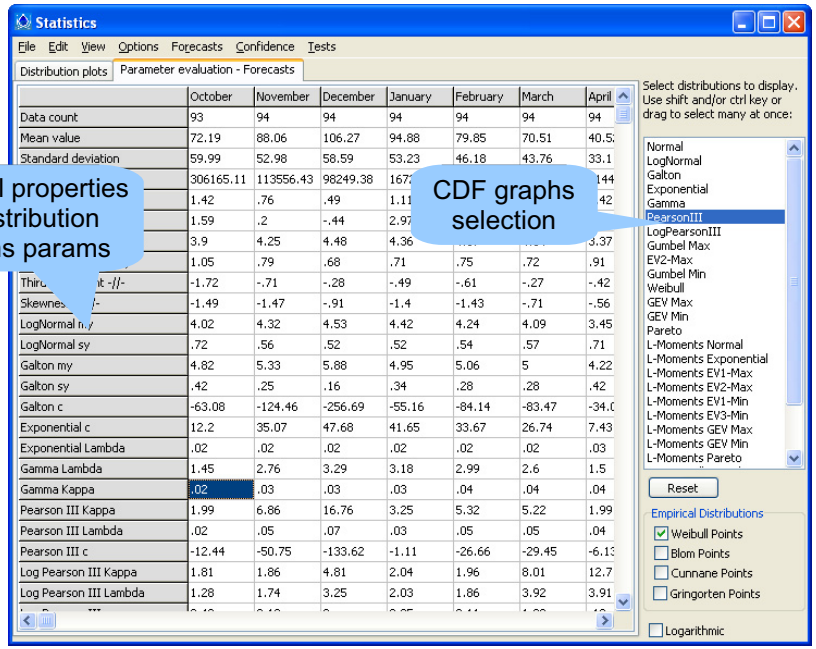

*December variogram – Dependent variable vs independent vars*

# 5. Statistical analysis of time series — “Pythia”

Hydrognomon's module for statistical time series analysis, called “Pythia”, can estimate properties of samples and parameters of distribution functions; includes 27 statistical distributions for data fitting; provides statistical forecasts and tests ($x^2$ and Kolmogorov – Smirnov); and can determine confidence intervals of any distribution function using an original Monte-Carlo algorithm.




*Confidence zones for December precipitation on Pearson III CDF*

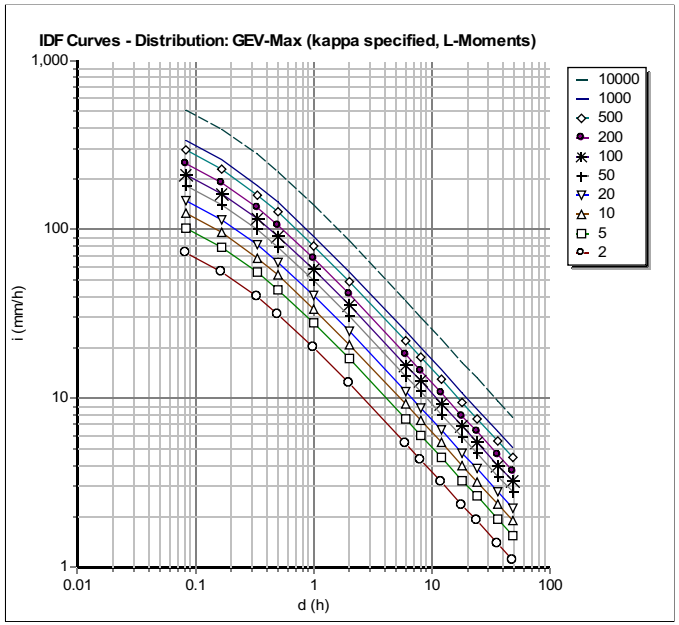


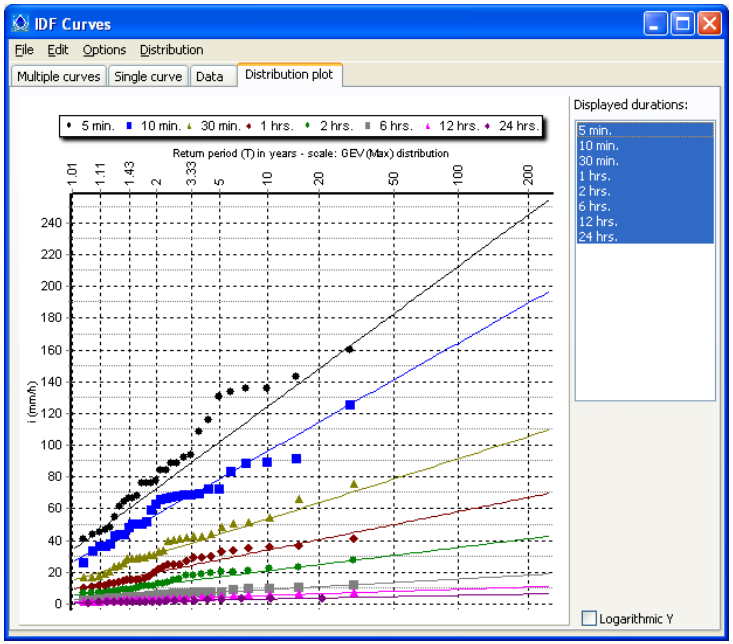

| | October | November | December | January | February | March | April |
|---|---|---|---|---|---|---|---|
| Data count | 93 | 94 | 94 | 94 | 94 | 94 | 94 |
| Mean value | 72.19 | 88.06 | 106.27 | 94.88 | 79.85 | 70.51 | 40.5 |
| Standard deviation | 59.99 | 52.98 | 58.59 | 53.23 | 46.18 | 43.76 | 33.1 |
| | 306165.11 | 113556.43 | 98249.38 | 167 | | | 144 |
| | 1.42 | .76 | .49 | 1.1 | | | 42 |
| | 1.59 | .2 | -.44 | 2.97 | | | |
| | 3.9 | 4.25 | 4.48 | 4.36 | | | 3.37 |
| | 1.05 | .79 | .68 | .71 | .75 | .72 | .91 |
| Third | -1.72 | -.71 | -.28 | -.49 | -.61 | -.27 | -.42 |
| Skewnes | -1.49 | -1.47 | -.91 | -1.4 | -1.43 | -.71 | -.56 |
| LogNormal m | 4.02 | 4.32 | 4.53 | 4.42 | 4.24 | 4.09 | 3.45 |
| LogNormal sy | .72 | .56 | .52 | .52 | .54 | .57 | .71 |
| Galton my | 4.82 | 5.33 | 5.88 | 4.95 | 5.06 | 5 | 4.22 |
| Galton sy | .42 | .25 | .16 | .34 | .28 | .28 | .42 |
| Galton c | -63.08 | -124.46 | -256.69 | -55.16 | -84.14 | -83.47 | -34.0 |
| Exponential c | 12.2 | 35.07 | 47.68 | 41.65 | 33.67 | 26.74 | 7.43 |
| Exponential Lambda | .02 | .02 | .02 | .02 | .02 | .02 | .03 |
| Gamma Lambda | 1.45 | 2.76 | 3.29 | 3.18 | 2.99 | 2.6 | 1.5 |
| Gamma Kappa | .02 | .03 | .03 | .03 | .04 | .04 | .04 |
| Pearson III Kappa | 1.99 | 6.86 | 16.76 | 3.25 | 5.32 | 5.22 | 1.99 |
| Pearson III Lambda | .02 | .05 | .07 | .03 | .05 | .05 | .04 |
| Pearson III c | -12.44 | -50.75 | -133.62 | -1.11 | -26.66 | -29.45 | -6.13 |
| Log Pearson III Kappa | 1.81 | 1.86 | 4.81 | 2.04 | 1.96 | 8.01 | 12.7 |
| Log Pearson III Lambda | 1.28 | 1.74 | 3.25 | 2.03 | 1.86 | 3.92 | 3.91 |

# 6. Ombrian curves — “Ombros”

Ombrian curves (also known as IDF curves) can be constructed from annual maxima of precipitation events, which can either be provided or estimated from measurements (with a time step of 1 minute to 24 hours). Ombrian curves, that is, intensity functions of episode duration and return period, are calculated using an advanced algorithm.

“Ombros” takes advantage of “Pythia”, allowing to choose from a list of distribution functions and calculating confidence intervals with the Monte-Carlo method.



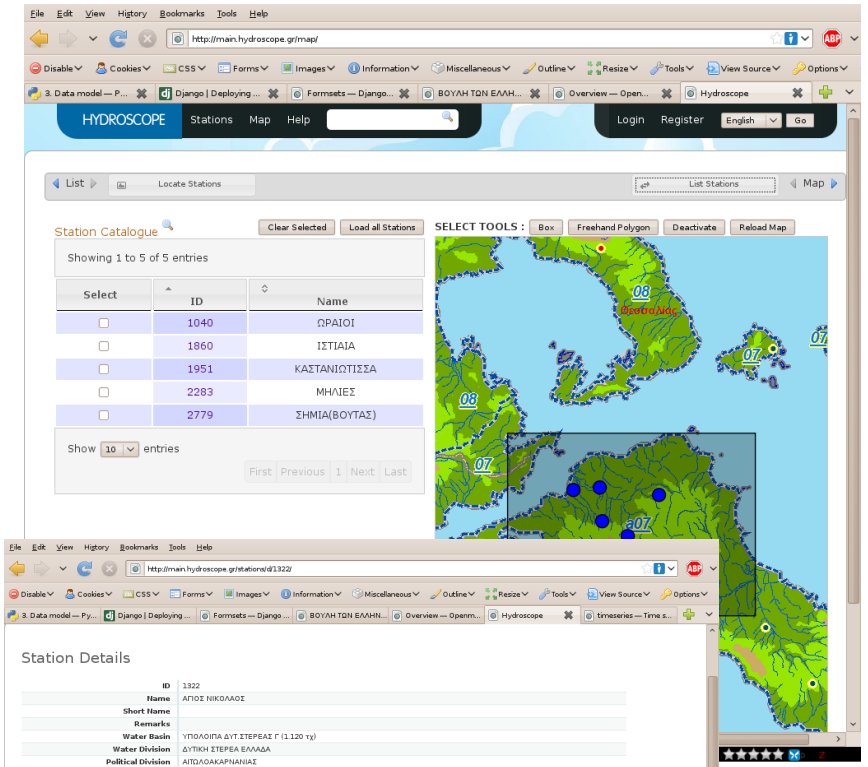
# 7. Create a time series database using Enhydris

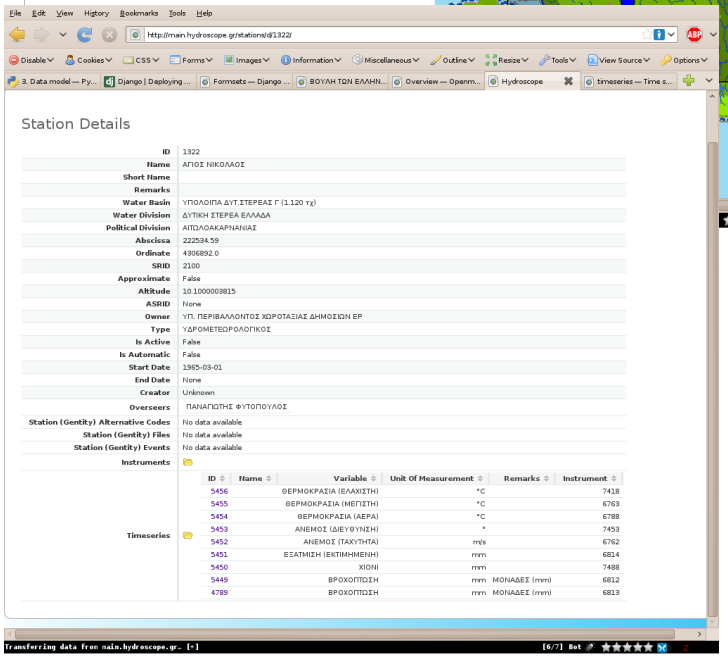
Hydrognomon is a standalone program, which stores its data in files—as simple as a word processor. But if you need to store your data centrally, we have Enhydris: a database system with a web interface and several features:

- It can optionally work in a distributed way. Many organisations can install one instance each, but an additional instance, common to all organisations, can be setup as a common portal. This additional instance can be configured to replicate data from the other databases, but without the space-consuming time series, which it retrieves from the other databases on demand. A user can transparently use this portal to access the data of all participating organisations collectively.

- It opens downloaded timeseries directly in Hydrognomon.

- It is extensible and has an API for communication with other systems.




Written in Python/Django and using PostgreSQL, Enhydris can be installed on every operating system on which Python runs, including GNU/Linux and Windows. It is free software, available under the GNU General Public License version 3 or any later version.

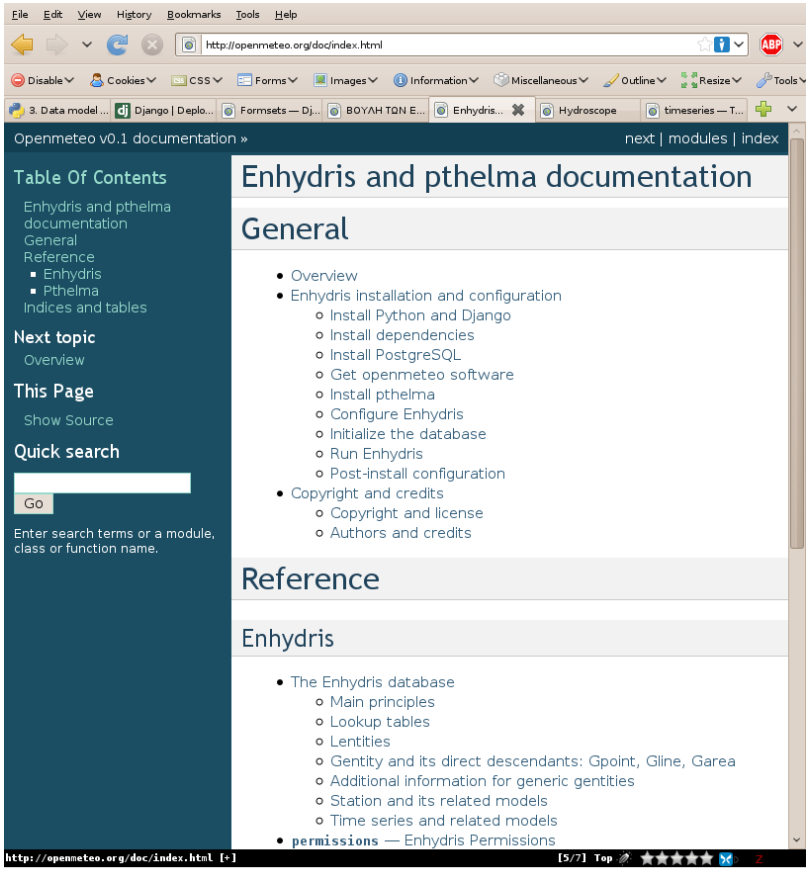
Enhydris is still under development and needs polishing, but it is already being used operationally by the Hydroscope project in Greece.

# 8. Support possibilities

Not only is Enhydris free software, it also has detailed documentation for administrators and developers. Therefore, your computer guy can probably install and maintain it, especially if they have a background in Python.

If you don't have such a computer guy,  there are many Python specialists and companies on the market who would love to offer you commercial support. For example, indifex.com, the Greek company who created Enhydris under supervision by the NTUA, will be more than happy to take your money.

Free software also allows you to have new features developed without the need to ask us for permission; but it is still better to first consult with us.



## 9. Data is like software: it's better when it's free

Maybe you don't need to maintain your own database, because we are creating openmeteo.org, an international, public meteorological database.

Enhydris has a security system that allows it to be used either in an organisational setting or in a public setting. In an organisational setting, there are priviliged users who have write access to all the data. In a public setting, users can subscribe, create stations, and add data for them, but they are not allowed to touch stations of other users.

openmeteo.org, currently under construction, will host such a public database, where you will be able to store and manage your data, provided you make it available under a free license.

Hydrognomon on the web: **`http://www.hydrognomon.org/`**

Get poster: http://www.itia.ntua.gr/en/docinfo/962/

More information on the openmeteo.org project, as well as source code for Hydrognomon and Enhydris: **`http://openmeteo.org/`**