

How to use WISKI for CCRN data

Kevin Shook

Centre for Hydrology,
University of Saskatchewan



UNIVERSITY OF
SASKATCHEWAN

Global Institute for
Water Security

What is WISKI?

- Water Information System Kisters

<http://www.kisters.net>

- Commercial environmental time series database
- Widely used in Europe
- Used in Canada by Alberta Environment, BC Hydro, Ontario Ministry of the Environment

Where WISKI is strong

- Formalized disciplined processing, retaining multiple processing and quality levels
- Many steps can be easily automated including simple QA/QC, infilling missing values, computation of daily or annual statistics, graphing, report generation, output to standard flat CSV files
- Lots of utilities to process streamflow
- Visualization tools are pretty good
- Prevents the need to reinvent the wheel

Where WISKI is weak

- Steep learning curve
- Often difficult to adapt to special needs, need to find creative workarounds
- Not well designed for infilling time series data with strong daily cycle as needed for most process studies
- None of these apply to users who are downloading data from the WISKI database

Advantages for us

- Robust
 - Sits on industrial database software (SQL, Oracle)
- Flexible
 - Works for wide variety of data types
- Rigorous
 - Original data protected from change
 - Provides complete audit trail

What's in it for you?

- Develop good data management attitudes and habits early in your career: help develop a scientific culture that is committed to data management for sharing and legacy
- CCRN data policy protects needs of students, intended to foster a culture of trust
- After steep learning curves, WISKI and its CCRN implementation can streamline DM tasks

Ways to access data

- GUI
 - Requires U of S ID
- Web interface
 - Easier, but can't add/change data
- KiWIS
 - Allows access through http queries
 - Can be done through R/MATLAB

Web interface

- <http://giws.usask.ca/KistersWeb/main.php>
 - Username: public
 - Password: public
- Blog
 - <http://words.usask.ca/wiski/>

Web interface

KISTERS Web public - Mozilla Firefox

File Edit View History Bookmarks Tools Help

KISTERS Web ... x +

giws.usask.ca/KistersWeb/main.php

200

- Sites
- Station list
- Groups
- Favorites
- History



WISKI Data Hierarchy: Four Levels

Site is a geographic region

Station is a grouped set of measurements within a site, may refer to a particular location and/or a functional grouping

Parameter is a single measurement (variable)

Time series is one processing level within a parameter

Example

Site:

BERMS

Station:

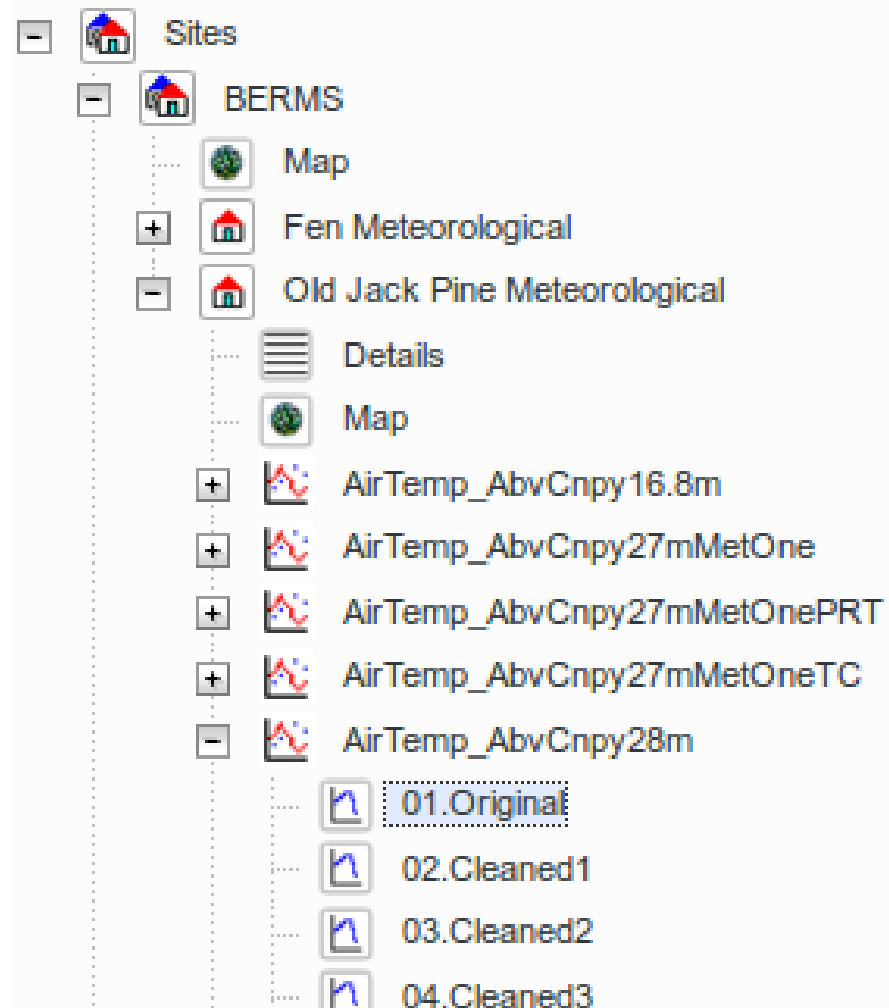
*Old Jack Pine
Meteorological*

Parameter:

AirTemp_AbvCnpy28m

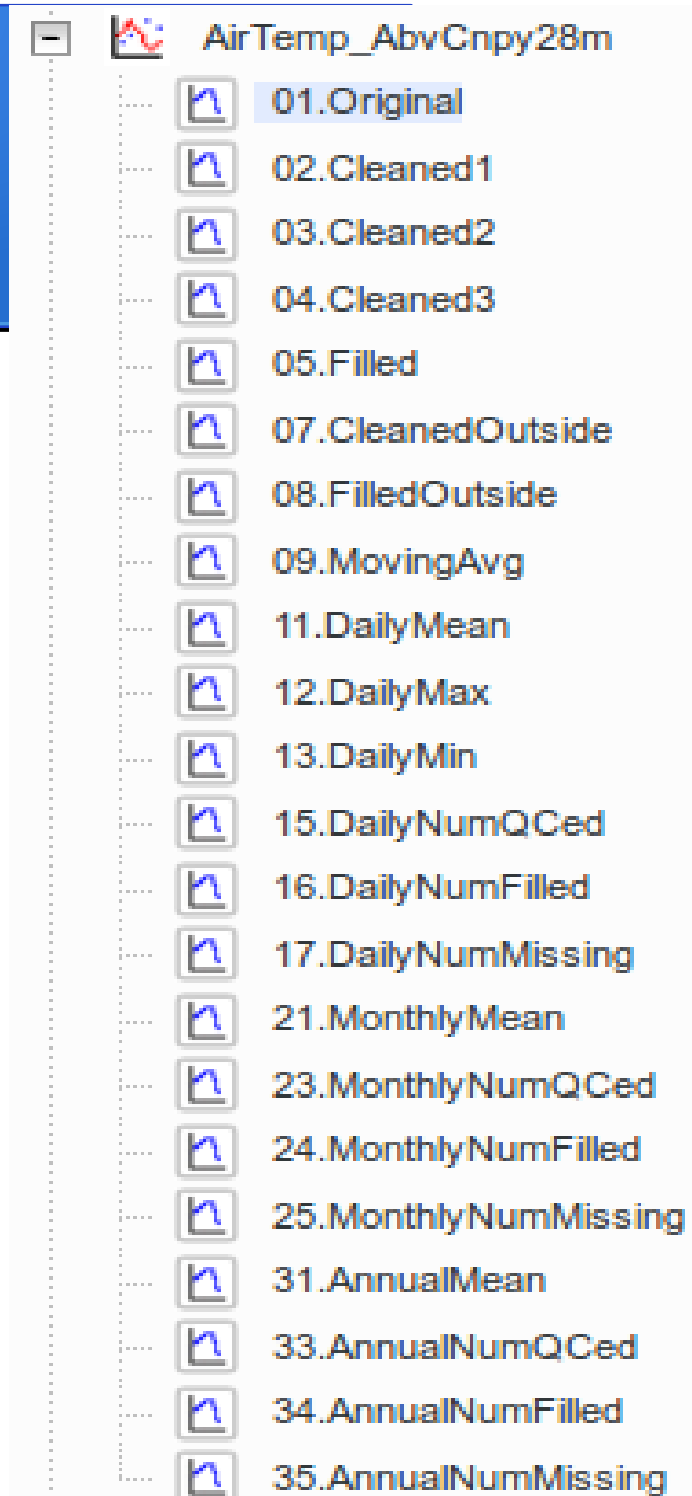
Time series:

01.Original



Time series

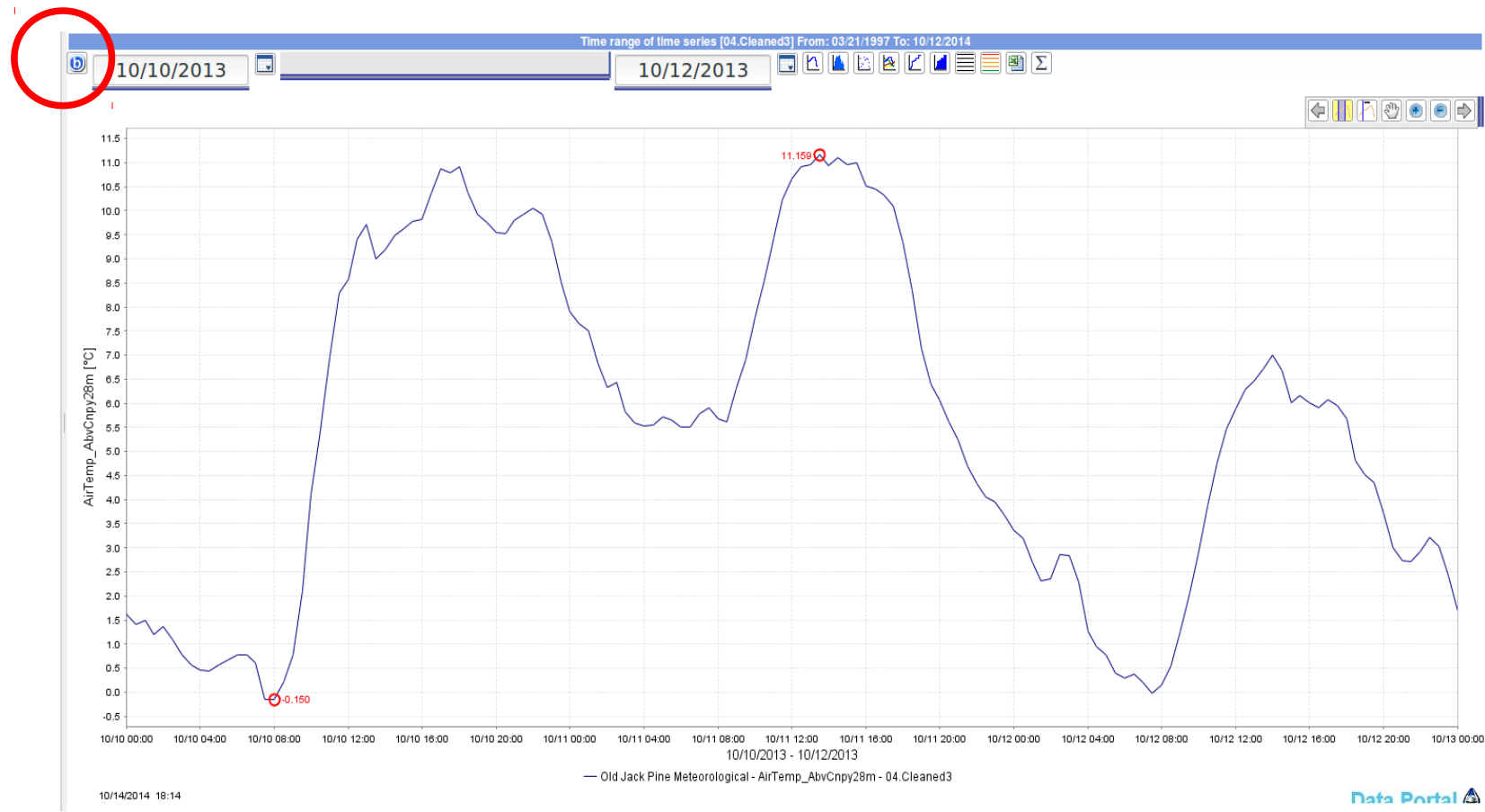
- 01.Original
 - Raw data
- 04.Cleaned3
 - Best unfilled
- 05.Filled
 - Missing values infilled
- 06.CleanedOutside
 - Cleaned outside of WISKI



	AirTemp_AbvCnpy28m
...	 01.Original
...	 02.Cleaned1
...	 03.Cleaned2
...	 04.Cleaned3
...	 05.Filled
...	 07.CleanedOutside
...	 08.FilledOutside
...	 09.MovingAvg
...	 11.DailyMean
...	 12.DailyMax
...	 13.DailyMin
...	 15.DailyNumQCed
...	 16.DailyNumFilled
...	 17.DailyNumMissing
...	 21.MonthlyMean
...	 23.MonthlyNumQCed
...	 24.MonthlyNumFilled
...	 25.MonthlyNumMissing
...	 31.AnnualMean
...	 33.AnnualNumQCed
...	 34.AnnualNumFilled
...	 35.AnnualNumMissing

Plotting

- Automatic, as soon as you select a time series
- Mostly self-explanatory



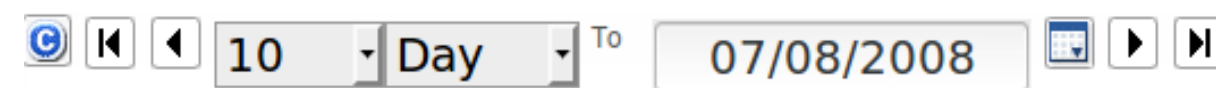
WISKI time browsing



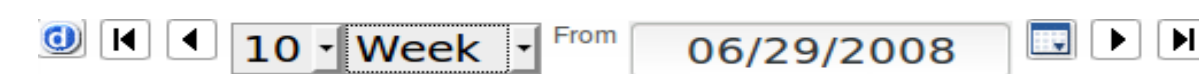
a – Start and end dates



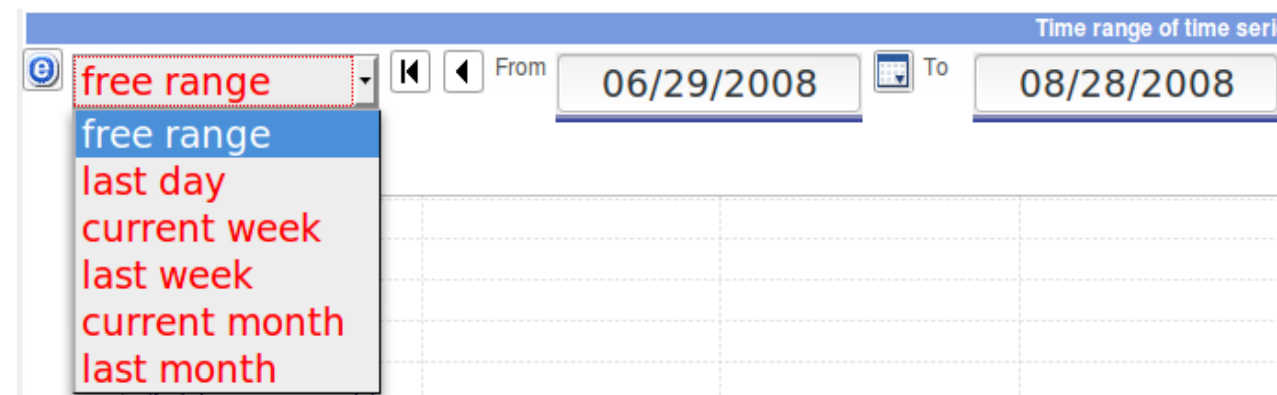
b – Movable window



c – Specified interval to end date



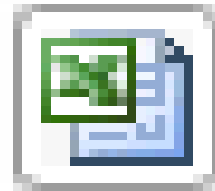
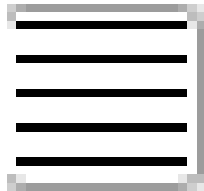
d – Specified interval from start date



e – Other intervals

Exporting data

Show as text file



Show as Excel
file

Only works for up to 1 year of data

Adding data to WISKI

- Directly adding data requires U of S ID
- Others must submit data files to Branko Zdradovic

branko.zdravkovic@usask.ca

- Files need to be in WISKI's native ZRXP format
 - Simple ASCII file

ZRXP files

- Can hold one or several time series
- Single time series example:

```
#REXCHANGE11042 | * | RINVAL -777 | * | LAYOUT(timestamp,value,remark) | * |
19020101 0
19020102 0
19020103 0
19020104 0
19020105 0
19020106 0
19020107 0
19020108 0
19020109 0
19020110 0
```

The diagram illustrates the structure of a ZRXP file. The first line is the header line, which contains the prefix, time series name, and layout information. The subsequent lines represent the time series data. Annotations point to specific parts of the header line: 'Prefix' points to '#REXCHANGE', 'Invalid data' points to '-777', 'Time series name' points to '11042', and 'Header line' points to the entire first line.

How to create ZRXP files

- Single time series
ZRXP files can easily be created in Excel
- Good instructions from Branko on the blog:
<http://words.usask.ca/wiski/>

1. Creating the ZRXP output from the Excel file

Posted on **January 10, 2013**

Here is one example how to create the ZRXP file directly from Excel. The format of the sample Excel file is given as:

	A	B	C	D	E
1	year	day	hour	T	RH
2	1966	1	0	-777	-777
3	1966	1	1	-777	-777
4	1966	1	2	-777	-777
5	1966	1	3	-777	-777
6	1966	1	4	-777	-777
7	1966	1	5	-777	-777
8	1966	1	6	-777	-777
9	1966	1	7	-777	-777
10	1966	1	8	-777	-777

Parameter names

- Need to follow the conventions laid out in *WISKI CCRN-GIWS Data Management Plan*
- When in doubt, follow the examples in BERMS

Parameter naming conventions

Type	Name	Optional Locations	Optional Adjectives	Example
TA	AirTemp	[Descriptors][#m] or [#cm] Descriptors: AbvCnpy, Cnpy, BlwCnpy, UndrStry	[Instrument]	AirTemp_AbvCnpy37mHMP45
TS	SoilTemp	[#cm][ProfileXX]		SoilTemp_5cmNW

Multiple time series ZRXP files

- Excel is usually not the right tool for the job
- Particularly true with files produced by dataloggers
- Branko has created a VB program called ZRXPCustom.exe to create ZRZP files
- Not a GUI program – it requires a parameter file, and is run from the command line, e.g.

```
C:\>ZRXPCustom.exe parameters.txt
```

Parameters

Parameter	Example
Path to the folder where the data file(s) are located	C:\zrxp\ECsample\DAT\
Path to the folder where the log file will be stored	C:\zrxp\ECsample\LOG\
Path to the folder where the FSL file is located	C:\zrxp\ECsample\FSL\
Path to the folder where zrxp files will be stored	C:\zrxp\ECsample\OUT\
Columns for the time stamp.	4,5,6
First row with the data.	3
Line delimiter	,
Number of characters to compare and match in the FSL and data file name.	6
Time zone information	TZUTC+0
Delete dat files after completion	DELETEDAT=NO

FSL file

- Contains the Final Storage Labels

Example:

```
Final Storage Label File for: FEN Calculated Flux  
Date:7/15/2013  
Time:12:00:00
```

```
HeaderSection:  
1,DataType,DoNotFile  
2,Site,DoNotFile  
3,SubSite,DoNotFile  
4,Year,DoNotFile  
5,Day,DoNotFile  
6,End_Time,DoNotFile  
7,NetEcosystemExchange,Import#=Test1  
8,FourWay_NetRad_AbvCnpy,Import#=Test2  
9,LatentHeatFlux,Import#=Test3  
10,SensibleHeatFlux,Import#=Test4  
11,FrictionVelocity_AbvCnpy_4m,Import#=Test5  
12,C02Flux_AbvCnpy_4m,Import#=Test6
```

Bulk data exports

- You can get large quantities of data from the ftp site <ftp://giws.usask.ca/ccrn/scheduledexports/>
- Contact Branko for user name and password

Index of <ftp://giws.usask.ca/ccrn/scheduledexports/>

[↑ Up to higher level directory](#)

Name	Size	Last Modified	
2002_FEN_Summary.csv	1973 KB	14-10-08	02:10:00 PM
2003_FEN_Summary.csv	2406 KB	14-10-08	02:09:00 PM
2004_FEN_Summary.csv	2478 KB	14-10-08	02:09:00 PM
2005_FEN_Summary.csv	2449 KB	14-10-08	02:09:00 PM
2006_FEN_Summary.csv	2446 KB	14-10-08	02:09:00 PM
2007_FEN_Summary.csv	2433 KB	14-10-08	02:09:00 PM
2008_FEN_Summary.csv	2460 KB	14-10-08	02:09:00 PM
2009_FEN_Summary.csv	2479 KB	14-10-08	02:09:00 PM
2010_FEN_Summary.csv	2476 KB	14-10-08	02:09:00 PM
2011_FEN_Summary.csv	1866 KB	14-10-08	02:09:00 PM
2012_FEN_Summary.csv	1871 KB	14-10-08	02:09:00 PM
2013_FEN_Summary.csv	2351 KB	14-10-08	02:08:00 PM
2014_FEN_Summary.csv	2212 KB	14-10-08	02:11:00 PM

KiWIS

- Provides access to Wiski data through http requests
- Database queries
- Can be used by many types of programs
- Modules have been written for MATLAB and R

Time series IDs

- Each time series is identified by a unique ID number
- You can't get the ID number from the Wiski website
- Have to query the KIWIS server to get the ID number, then extract the data for the specified time series

Instructions

- Available at

<http://giws.usask.ca:8080/KiWIS/KiWIS?datasource=0&service=kisters&type=queryServices&request=getrequestinfo>

KISTERS QueryServices - Request Information

All commands are available as HTTP GET and POST (add ?kvp=true to URL!) with key-value

Version: 1.4.8-SNAPSHOT

Datasource Type: WDP

getStationList	retrieves a list of stations with metadata
getTimeseriesList	retrieves a list of timeseries with metadata
getTimeseriesValues	retrieves timeseries data
getTimeseriesValueLayer	retrieves a layer of timeseries values for a given date including location data
getSiteList	retrieves a list of sites with metadata
getParameterList	retrieves a list of parameters available at sites and/or stations
getParameterTypeList	retrieves the system defined parameter type list
getTimeseriesTypeList	retrieves a list of timeseries types
getGroupList	retrieves a list of timeseries and station groups
getGraph	retrieves a graph image of timeseries data
getGraphTemplateList	retrieves a list of available graph templates
Show All Commands	shows all command descriptions at once

Please select a request to show more information...

R and MATLAB

- MATLAB code developed by Alan Barr and Warren Helgason
- R package WISKIr developed by me
 - Functions:

```
find_WISKI_timeseries(stationName)
```

```
get_WISKI_timeseries_metadata(timeSeriesID)
```

```
get_WISKI_timeseries_values(timeSeriesID)
```

Acknowledgements

- Branko Zradovic
- Alan Barr
- Warren Helgason

Models come and go but a good data set lasts forever. Paul Quay, Science, 2002