

## Data download manual Anemos@OZ System Documentation

### ABSTRACT

This document describes the time series data storage concepts used in the Anemos platforms and gives instructions on how to download this data.

### GENERAL INFORMATION

<b>Author</b>	Felix Dierich
<b>Affiliation</b>	Overspeed GmbH & Co. KG
<b>Address</b>	Marie-Curie-Str. 1, 26129 Oldenburg, Germany
<b>Telephone</b>	+49-441-36116-301
<b>Email address</b>	f.dierich@overspeed.de
<b>Further authors</b>	Igor Waldl
<b>Reviewer</b>	Dr. Hans-Peter Waldl
<b>Approver</b>	

### DOCUMENT INFORMATION

<b>Document type</b>	Memo
<b>Document name</b>	Anemos@OZ.data-download-manual.20090907.odt
<b>Version</b>	0.2
<b>Revision number</b>	2
<b>Revision date</b>	2009-12-10
<b>Status</b>	Draft for comments
<b>Confidentiality</b>	Restricted to project members

### STATUS, CONFIDENTIALITY AND ACCESSIBILITY

Status			Confidentiality			Accessibility	
<b>S0</b>	Approved/Released		<b>R0</b>	General public		Local file	X
<b>S1</b>	Reviewed		<b>R1</b>	Restricted to vendor/customer		Private web site	
<b>S2</b>	Pending for review		<b>R2</b>	Restricted to project members	X	Public web site	
<b>S3</b>	Draft for comments	X	<b>R3</b>	Restricted to WP members/PL		Paper copy	
<b>S4</b>	Under preparation		<b>R4</b>	Restricted to task members/WPL/PL			

**PL:** Project leader, **WPL:** Work package leader, **WP:** Work package

# Table of Contents

Title (from properties) .....	1
Subject (from properties).....	1
1. Introduction .....	3

## CHANGE HISTORY

Version	Date	Author	Changes
0.1	07.09.09	F. Dierich	initial document version
0.2	09.09.09	Igor Waldl	Corrections and review

# 1. Introduction

This document describes the time series data storage concepts used in the Anemos platform and gives instructions on how to download this data using a command line client or a web interface.

## 2. Time series data storage concepts

The Time Series Data Repository (TSDR) is a central part of the Anemos platform and used for the storage of all time series data including NWP, SCADA-derived wind farm production and meteorological data and wind power predictions (interim and final). It has a standardized SOAP web-service interface, which is used for all data exchange. Internally the data is stored in a database, which resides on Oracle database servers. The web service is hosted on an Apache Tomcat J2EE application server.

The term “TSDR” is generally used for the whole component consisting of the web-service with its SOAP interface and the encapsulated database tables. From a platform module’s perspective the term denotes the service and not the database, as the modules never access the databases directly.

While the SOAP interface can be addressed directly by using standard SOAP libraries, ready-to-use clients have been developed to aid the integration of platform components. They are available in the form of a Java library or a command-line client and internally handle all SOAP processing. The TSDR has the built in ability to process and produce time series data in the DEPRI file format.

The TSDR maintains meta information on the time series data. It keeps internal write operation logs in addition to sending messages to the Anemos UMA system.

### 2.1. Data contexts

As time series data of different natures and for different sites is kept in a TSDR installation, a concept has been designed to identify a certain data stream. The streams are referenced by so called data contexts. A context has the following selection parameters:

- **major type:** *nwp* for weather prediction data, *wpp* for wind power prediction data, *measured* for SCADA data etc.
- **subtype:** used to distinguish between different prediction models, model parameters etc.
- **site:** used to distinguish between different wind farms, aggregations of wind farms or other sites (existing DUIDs)

For instance, the data of the single point numerical weather prediction data of the weather model BoM LAPS125 for the wind farm “Cathedral Rocks” is stored in the context *nwp bom-laps125 CATHROCK*.

The repository can also be accessed via a web-based GUI. It allows viewing time series data in tabular form and TSDR administration. The Anemos.Live GUI reads data from the TSDR.

## 2.2. Channel naming

The channels used in the TSDR are named using a standardized naming system. The channel name consists of two parts, the so called PhysSymbol and the GlobName, separated by underscore. The PhysSymbols are fixed, while the GlobNames can be used to distinguish between variants. E.g. the wind speed is stored in channels called WSpdNWP\_xxx, where xxx specifies the weather model and the height of the prediction wind speed, e.g. WSpdNWP\_ECMWF10m.

### 2.2.1. Time channels

For all output data of WPP models the a certain system of timestamps is used.

"PTimePred" is the scheduled time of a model run, which is the first scheduled target time in the output – formally this was the NWP base time.

The time channels to be used are hence the following:

- TimePred: the point in time for which the power value is predicted
- PTimePred: the scheduled time of a model run, which is usually the first scheduled target time in the output.
- UTimePred: the time of the last SCADA value used by the model – if the model does not use online SCADA data or the time is not known, set to the same value as PTimePred
- NTimePred: the base time of the NWP run that was used to produce this value

Note that for combined WPPs like the exported products, NTimePred and UTimePred are often not unique and not provided.

Example:

<i>TimePred</i>	<i>PTimePred</i>	<i>UTimePred</i>	<i>NTimePred</i>	<i>PowerPred</i>	...
200704021200	200704021200	200704020647	200704020000	119.26	
200704021300	200704021200	200704020647	200704020000	145.32	
200704021400	200704021200	200704020647	200704020000	163.23	
200704021500	200704021200	200704020647	200704020000	181.77	

...

### 2.2.2. NWP Data channels

Typical NWP data channel names include:

- WSpdNWP wind speed
- WDirNWP wind direction
- etc.

### 2.2.3. SCADA Data channels

Typical SCADA data channel names include:

- TimeMeas      timestamp
- PowerMeas    measured active power
- PowerControlMeas downregulation
- WSpdMeas    wind speed
- WDirMeas    wind direction
- etc.

### 2.2.4. WPP Data channels

Typical WPP data channel names include:

- timestamps: see above
- PowerPred    predicted power
- etc.

### 2.2.5. Quality channels

For all time series data used within the system the data quality is marked by quality channels. These have two functions:

- tracing the “history” of a data value
- flagging the data, how it can be used by subsequent modules

To keep these two functions apart, two different channels are used:

1. *Data Status* for the “history” of a data value
2. *Data Quality* for the usage flags

For data contexts with one data status and one data quality channel per normal data channel, e.g. for SCADA data, the status and quality channels are named by appending “DS” and “DQ” to the channels physSymbol, i.e. the first part of the channel name. E.g. a channel “PowerMeas\_PI” would have two channels “PowerMeasDQ\_PI” and “PowerMeasDS\_PI” for Data Quality and Data Status information. In WPP data a channel e.g. named “PowerPred\_WPP” would have the extra channels “PowerPredDQ\_WPP” and “PowerPredDS\_WPP”.

For data contexts with only one data status and data quality channel for all normal data channels, i.e. one status and one quality value per data record, the combined extra channels are named “RecordDQ\_DQ” and “RecordDS\_DS”.

### 3. The DEPRI file format

The DEPRI format is a standardized data format for storing time series data. It is used for storing measured data, numerical weather prediction data (NWP) and wind power prediction data in the Anemos system. It is a plain ASCII format. The DEPRI format is used for all file representations of time series data in the Anemos system. The data download facilities described below all provide the data in DEPRI format. To understand the format, it is recommended to inspect files produced that way.

In the DEPRI file format meta information describing the time series data is included in the same file as the data itself, so that it is always possible to interpret the data without further external information.

The document “DEPRI data format 1.2A with Anemos specialities” is provided as a reference guide introducing the file syntax and giving some examples.

### 4. TSDR web access

The TSDR provides a web-based GUI, which allows to inspect and download time series and meta data among other functions. Take the following steps to use this interface:

- Use a web browser to open the URL <http://hostname:8080/anemos/>, where hostname needs to be the name of one of the Anemos APP servers, e.g. NORPWAPP2.
- Enter your login data, where Anemos username and password is the individual login provided to you and “HTTP authentication” username and password is a common login also provided.
- To view time series data select “Data – view” in the menu on the left. This will give you a selection of the data contexts that can be read with your login. (See above for data context principle.)
- Select a context, e.g. “wpp-comb standard CATHROCK” and click “Select”.
- On the next screen you can select further data retrieval options (see also screenshot): You can select the timespan to load or an individual model run, select a display type or reduce the number of channels to fetch. Click “Fetch data” to load and display the actual data.

*Note: You should never retrieve too large timespans on a production system to avoid compromising system performance. The maximum timespans depend on the data context, please consult someone who knows.*

- The second screenshot shows an example of data display in an HTML table. Other display options are in a textarea on the page or as sole page content. Via the “Data – download” option you can also retrieve Depri or CSV files.

view data - retrieval options

Data context: wpp-comb - standard - CATHROCK

Times selection:

Use ISO format with the complete year e.g. "2003-01-01 12:00:00" or "2003-01-01" or "200301011200" or "2003". If you want to include only the data of a specific model run, enter the time of the model run below. Alternatively, if you want the newest predictions only, check the box below.

start time:

>= 2000-01-01

2008-04-01 00:00:00.000

end time:

<= 2015-12-31

2008-04-03 00:00:00.000

model run time:

2008-04-01 06:30:00.000

☐ all

only the newest?

☐ Yes ☒ No

minimum horizon

0

hours

Fetch data

Data display format:

☒ display in table

☐ display in textarea

☐ display in page

☐ use commas as decimal markers

Channel selection:

Uncheck the channels you do not want to include in the result

☒ TimePred.UTC

☒ PTimePred.UTC

☒ UTimePred.UTC

☒ NTimePred.UTC

☒ PowerPred.WPP

☒ PowerPredDQ.WPP

☒ PowerPredDS.WPP

Fetch data

Figure 1: TSDR web access data retrieval detail selection.

Data						
vvv scroll to retrieval options vvv						
TimePred_UTC	PTimePred_UTC	UTimePred_UTC	NTimePred_UTC	PowerPred_WPP	PowerPredDQ_WPP	PowerPredDS_WPP
2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	7546.5	0.0	35.0
2008-04-01T07:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	8069.1	0.0	35.0
2008-04-01T07:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	8723.7	0.0	35.0
2008-04-01T08:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	9226.4	0.0	35.0
2008-04-01T08:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	9501.9	0.0	35.0
2008-04-01T09:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	9676.8	0.0	35.0
2008-04-01T09:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	8478.8	0.0	35.0
2008-04-01T10:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	8117.8	0.0	35.0
2008-04-01T10:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	10988.2	0.0	35.0
2008-04-01T11:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	16219.1	0.0	35.0
2008-04-01T11:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	22719.7	0.0	35.0
2008-04-01T12:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	27188.4	0.0	35.0
2008-04-01T12:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	31306.3	0.0	35.0
2008-04-01T13:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	35355.5	0.0	35.0
2008-04-01T13:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	39400.5	0.0	35.0
2008-04-01T14:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	43397.4	0.0	35.0
2008-04-01T14:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	47387.5	0.0	35.0
2008-04-01T15:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	51289.7	0.0	35.0
2008-04-01T15:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	52877.7	0.0	35.0
2008-04-01T16:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	53771.2	0.0	35.0
2008-04-01T16:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	54159.0	0.0	35.0
2008-04-01T17:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	54136.6	0.0	35.0
2008-04-01T17:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	53582.9	0.0	35.0
2008-04-01T18:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	51896.8	0.0	35.0
2008-04-01T18:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	51366.9	0.0	35.0
2008-04-01T19:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	50886.5	0.0	35.0
2008-04-01T19:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	50425.2	0.0	35.0
2008-04-01T20:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	49915.3	0.0	35.0
2008-04-01T20:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	49460.3	0.0	35.0
2008-04-01T21:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	48972.5	0.0	35.0
2008-04-01T21:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	49056.2	0.0	35.0
2008-04-01T22:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	49198.8	0.0	35.0
2008-04-01T22:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	49273.2	0.0	35.0
2008-04-01T23:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	49367.4	0.0	35.0
2008-04-01T23:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	49489.4	0.0	35.0
2008-04-02T00:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	49539.5	0.0	35.0
2008-04-02T00:30:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	49274.1	0.0	35.0
2008-04-02T01:00:00.000	2008-04-01T06:30:00.000	2008-04-01T06:30:00.000	2000-01-01T00:00:00.000	49000.0	0.0	35.0

Figure 2: TSDR web access data display in HTML table.

## 5. TSDR command line client

To load time series data from the TSDR in form of DEPRI files or to inspect meta data associated with data contexts, a command line client is provided, which accesses the TSDR via SOAP over HTTP. The following is a quick start to using the client:

Use DEPRI client to download TSDR data to files:

- Get TSDRClient package and extract to a convenient location. Note: You need to install a Java runtime environment of version 1.5 or newer and need to add the Java executable to the Windows PATH.
- Open a Windows command line in the extracted directory (or a Cygwin shell).
- Edit tsdrclient-login.bat and set the login. (Note: This may already be prepared.)
- Run “tsdrclient-login.bat check” to see if the login and the connectivity is working.
- Chose a context to get data of. You can inspect a list of available contexts in the web interface (see above) or by running “tsdrclient-login.bat contextlist”. E.g. for Starfish Hill SCADA data chose “measured filtered STARHLWF”



- Execute the batch file 'tsdrclient-login.bat' with the command "getdepri" and add information for the context (majorType, subType and farmName) and the timespan you want to download: `tsdrclient-login.bat getdepri majorType subType farmName start-Time endTime`
  - Note: You should never retrieve too large timespans on a production system to avoid compromising system performance. The maximum timespans depend on the data context, please consult someone who knows.
  - Example: `tsdrclient-login.bat getdepri measured filtered STARHLWF 2009-09-07T00:00 2009-09-07T23:59 -O output-directory/`
  - Note: The -O flag specifies the path and optional filename to write the DEPRI file to. If only an existing directory is specified, the filename is automatically build including relevant information.
- For more detail and all available commands, consult the help function of the client by executing 'tsdrclient-login.bat help' or see the reference copy in the appendix.

## 6. Appendix 1: TSDRClient usage information

ANEMOS TSDR Client V1.58

List of possible commands, their usage and optional flags:

GETDEPRI: Get a Depri file with timeseries data

Usage: getdepri majorType subType farmName [startTime endTime] [chan1 [chan2 ...]]

Optional flags:

- M modelRunTime include run with this time as PTime only
- D updateTime include run with this time as UTime only
- T runTime get latest run with PTime <= runTime
- N include only newest data for each time
- L fillup missing data lines
- Z hours minimum horizon to retrieve
- B use blanks instead of tabs in output
- S split file output in separate Depri files for each run
- O dir/file.dep use file or directory for output

SENDEPRI: Send one Depri file with timeseries data

Usage: senddepri [majorType subType farmName]

Optional flags:

- F force write mode, write even incomplete sets
- C comment comment to save in writing log
- H check Depri header for correctness
- I file.dep use file as input, optionally with path
- X don't write data to the failover TSDR

SENDEPRIMULTI: Send multiple Depri files with timeseries data

Usage: senddeprimulti file1.dep [file2.dep [...]]

Optional flags:

- F force write mode, write even incomplete sets
- C comment comment to save in writing log
- H check Depri header for correctness
- X don't write data to the failover TSDR

CHECK: Check connection to TSDR (including HTTP auth)

Usage: check

CHECKFO: Check connection to failover TSDR (including HTTP auth)

Usage: checkfo

CONTEXTLIST: Get a list of contexts readable with current username

Usage: contextlist

CHANNELLIST: Get a list of the channels of a context

Usage: channellist majorType subType farmName aTime

CHANNELINFO: Get meta information about a channel

Usage: channelinfo majorType subType farmName aTime channelName

CHANNELPROPERTY: Get one channel meta info property

Usage: channelproperty majorType subType farmName aTime channelName propertyName

METAINFO: Get context meta info property (one/all)

Usage: metainfo majorType subType farmName aTime [keyName]

AVAILABLE: Query time series lines available

Usage: available majorType subType farmName startTime endTime

Optional flags:

- M modelRunTime include run with this time as PTime only
- D updateTime include run with this time as UTime only

WRITEABLE: Query if context can be written to with login

Usage: writeable majorType subType farmName aTime

DATALINES: Load data lines, display on stdout

Usage: datalines majorType subType farmName [startTime endTime] [chan1 [chan2 ...]]

Optional flags:

- M modelRunTime include run with this time as PTime only
- D updateTime include run with this time as UTime only
- N include only newest data for each time
- L fillup missing data lines
- Z hours minimum horizon to retrieve

SINGLELINE: Get a single line (record) of data

Usage: singleline majorType subType farmName [time] [chan1 [chan2 ...]]

Optional flags:

- M modelRunTime include run with this time as PTime only
- D updateTime include run with this time as UTime only

RUNTIMES: Get the model run times in timespan

Usage: runtimes majorType subType farmName startTime [endTime]

WRITEDATA: Write data via console to TSDR

Usage: writedata majorType subType farmName startTime chan1 [chan2 ...]

Optional flags:

- F force write mode, write even incomplete sets
- C comment comment to save in writing log
- X don't write data to the failover TSDR

MULTI: Execute several commands in multi-query mode

Usage: multi

SETTINGS: Print current settings and flags

Usage: settings

HELP: Help/usage information

Usage: help

Global options:

- l soapURL TSDR SOAP service URL
- o soapURL failover TSDR SOAP service URL
- U tsdrUsername username for TSDR login (Anemos URM user)
- P tsdrPassword password for TSDR login (Anemos URM user)
- u httpUsername username for HTTP auth login
- w httpPassword password for HTTP auth login
- t timeout for all SOAP calls in msecs
- V be verbose
- n no check of TSDR interface version
- ? get help / usage message

## **7. Appendix 2: Anemos DEPRI Description**

(see PDF file)