

# Package ‘WaterML’

February 12, 2020

**Type** Package

**Title** Fetch and Analyze Data from 'WaterML' and 'WaterOneFlow' Web Services

**Version** 1.9.0

**Description** Lets you connect to any of the 'Consortium of Universities for the Advancement of Hydrological Science, Inc.' (CUAHSI) Water Data Center 'WaterOneFlow' web services and read any 'WaterML' time series data file. To see list of available web services, see <<http://hiscentral.cuahsi.org>>. All versions of 'WaterML' (1.0, 1.1 and 2.0) and both types of the web service protocol ('SOAP' and 'REST') are supported. The package has six data download functions: `GetServices()`: show all public web services from the HIS Central Catalog. `HISCentral_GetSites()` and `HISCentral_GetSeriesCatalog()`: search for sites or time series from the HIS Central catalog based on geographic bounding box, server, or keyword. `GetVariables()`: Show a data.frame with all variables on the server. `GetSites()`: Show a data.frame with all sites on the server. `GetSiteInfo()`: Show what variables, methods and quality control levels are available at the specific site. `GetValues()`: Given a site code, variable code, start time and end time, fetch a data.frame of all the observation time series data values. The `GetValues()` function can also parse 'WaterML' data from a custom URL or from a local file.

**License** MIT + file LICENSE

**Depends** R (>= 3.0)

**Imports** stats, XML, httr

**LazyData** true

**URL** <https://github.com/jirikadlec2/waterml>

**Suggests** knitr, qpdf

**VignetteBuilder** knitr

**RoxygenNote** 7.0.2

**NeedsCompilation** no

**Author** Jiri Kadlec [aut, cre]

**Maintainer** Jiri Kadlec <[jirikadlec2@gmail.com](mailto:jirikadlec2@gmail.com)>

**Repository** CRAN

**Date/Publication** 2020-02-12 10:50:05 UTC

## R topics documented:

WaterML-package . . . . .	2
GetServices . . . . .	3
GetSiteInfo . . . . .	3
GetSites . . . . .	5
GetValues . . . . .	6
GetVariables . . . . .	8
HISCentral_GetSeriesCatalog . . . . .	9
HISCentral_GetSites . . . . .	10
MakeSOAPEnvelope . . . . .	12
WaterMLVersion . . . . .	13
WaterOneFlowNamespace . . . . .	13
WaterOneFlowVersion . . . . .	14
<b>Index</b>	<b>15</b>

---

WaterML-package	<i>Fetch and analyze data from WaterML or CUAHSI WaterOneFlow web service</i>
-----------------	---

---

## Description

Package: WaterML  
 Type: Package  
 License: MIT  
 LazyLoad: yes

## Details

Fetch and analyze data from WaterML or CUAHSI WaterOneFlow web service.

Please see <https://github.com/jirikadlec2/waterml> for more information.

## Author(s)

Jiri Kadlec <jirikadlec2@gmail.com>

---

GetServices	<i>GetServices</i>
-------------	--------------------

---

**Description**

This function gets the table of web services from the HIS Central catalog

**Usage**

```
GetServices()
```

**Examples**

```
GetServices()
```

---

GetSiteInfo	<i>GetSiteInfo</i>
-------------	--------------------

---

**Description**

This function gets the table variables measured at a specific site from the WaterML web service

**Usage**

```
GetSiteInfo(server, siteCode)
```

**Arguments**

server	The URL of the web service ending with .asmx or .wsdl, for example: <a href="http://hydroportal.cuahsi.org/ipswic">http://hydroportal.cuahsi.org/ipswic</a>
siteCode	The full site code, for example: IRWA:FB-BV. To get a list of available site codes, see GetSites() function and use the FullSiteCode field.

**Value**

a data.frame of data values with the following columns:

Name	Type	Description
SiteID	character	The site ID in the original database
SiteName	character	The name of the site
SiteCode	character	A short unique code of the site
FullSiteCode	character	The complete unique code of the site
in the format NETWORK:CODE, for example SNOTEL:879.		
Latitude	numeric	The WGS84 latitude in decimal degrees
Longitude	numeric	The WGS84 longitude in decimal degrees
Elevation	numeric	The elevation of the site above sea level in meters
State	character	Only for sites in the USA: the state of the site

County	character	Only for sites in the USA: The county of the site
Comments (note: this field is often empty)	character	Additional comments about the sites
VariableCode	character	Short code of the variable
FullVariableCode Use this value as the variableCode parameter in GetValues().	character	The full variable code, for example: SNOTEL:S
VariableName	character	The name of the variable
ValueType Field Observation or Derived Value	character	the type of observation:
DataType Average, Continuous, Sporadic..	character	the aggregate data type:
GeneralCategory Climate, Water Quality..	character	the general category of the measurements:
SampleMedium for example water, atmosphere, soil..	character	the sample medium:
UnitName	character	The name of the measurement units
UnitType	character	the type of the measurement units
UnitAbbreviation (m, cm, in..)	character	The abbreviation of the measurement units
NoDataValue	numeric	The value that indicates missing data
IsRegular	boolean	TRUE if the measurements are regular, FALSE o
TimeUnitName	character	The name of the time units
TimeUnitAbbreviation	character	The time units abbreviation
TimeSupport one measurement is taken	character	The length of the time period over which
Speciation (as nitrogen, as phosphorus..)	character	The chemical sample speciation
methodID	character	The ID of the sensor or measurement method
methodCode Usually the same as methodID.	character	The code of the sensor or measurement method.
methodDescription of the data collection instrumentation / measurement method.	character	The description of the sensor or
methodLink of the sensor or measurement method.	character	The hyperlink of the website
sourceID	character	The ID of the data source or author
organization	character	The name of the organization collecting the data
sourceDescription	character	The description of organization collecting the da
citation	character	Instruction how to cite the data
qualityControlLevelID Usually 0 means raw data and 1 means quality controlled data.	character	The ID of the quality control level.
qualityControlLevelCode:	character	The code of the quality control level. Usually sa
qualityControlLevelDefinition:	character	The quality control level definition.
valueCount:	character	The number of observations in this time series
beginDateTime: observation in this time series.	POSIXct	The local date and time of the first available
endDateTime: observation in this time series.	POSIXct	The local date and time of the last available
beginDateTimeUTC:	POSIXct	The UTC date and time of the last available obse

endDateTimeUTC: POSIXct The UTC date and time of the last available observation

The output data.frame also has attributes with information about the status: download.time, parse.time, download.status, parse.status These attributes can be used for troubleshooting WaterOneFlow/WaterML server errors. If parse status is "NO\_SERIES\_FOUND", then this site doesn't have any available data.

### Examples

```
server <- "http://hydroportal.cuahsi.org/SNOTEL/cuahsi_1_1.asmx"
siteInfo <- GetSiteInfo(server, siteCode="SNOTEL:879")
```

---

GetSites

*GetSites*

---

### Description

This function gets the table of sites from the WaterML web service

### Usage

```
GetSites(server, west = NULL, south = NULL, east = NULL, north = NULL)
```

### Arguments

server	The URL of the web service ending with .WSDL, for example: <a href="http://icewater.usu.edu/MudLake/cuahsi_1_1.asmx">http://icewater.usu.edu/MudLake/cuahsi_1_1.asmx</a> alternatively this can be the REST URL to get the sites.
west	Optional parameter: The west longitude of the geographic bounding box in decimal degrees. Allowed values are between -180.0 and +180.0
south	Optional parameter: The south latitude of the geographic bounding box in decimal degrees. Allowed values are between -90.0 and +90.0
east	Optional parameter: The east longitude of the geographic bounding box in decimal degrees. Allowed values are between -180.0 and +180.0
north	Optional parameter: The north latitude of the geographic bounding box in decimal degrees. Allowed values are between -90.0 and +90.0

### Value

a data.frame of sites. The data.frame has the following columns:

- SiteID: The site ID in the original database
- SiteName: The name of the site
- SiteCode: A short unique code of the site
- FullSiteCode: The complete unique code of the site in the format NETWORK:CODE. Use this value in the GetSiteInfo and GetValues functions

- Latitude: The WGS84 latitude in decimal degrees
- Longitude: The WGS84 longitude in decimal degrees
- Elevation: The elevation of the site above sea level in meters
- State: Only for sites in the USA: the state of the site
- County: Only for sites in the USA: The county of the site
- Comments: Additional comments about the sites (note: this field is often empty)

The output data.frame also has attributes with information about the status: download.time, parse.time, download.status, parse.status These attributes can be used for troubleshooting WaterOneFlow/WaterML server errors.

### Examples

```
#Getting all sites from a service
sites <- GetSites("http://hydroportal.cuahsi.org/ipswich/cuahsi_1_1.asmx?WSDL")

#Getting a subset of sites restricted by geographical area
server <- "http://hydroportal.cuahsi.org/NEON/cuahsi_1_1.asmx?WSDL"
sites_subset <- GetSites(server, west=-115.0, south=40.0, east=-110.0, north=42.0)
```

---

GetValues

*GetValues*

---

### Description

This function gets the time series data values from the WaterML web service

### Usage

```
GetValues(
  server,
  siteCode = NULL,
  variableCode = NULL,
  startDate = NULL,
  endDate = NULL,
  methodID = NULL,
  sourceID = NULL,
  qcID = NULL,
  daily = NULL
)
```

### Arguments

- |          |   |
|----------|---|
| server   | The URL of the web service, for example: <a href="http://worldwater.byu.edu/interactive/rushvalley/services/index">http://worldwater.byu.edu/interactive/rushvalley/services/index</a> . This can be also a custom REST URL or the file name of the WaterML file. |
| siteCode | The site code. To get a list of available site codes, see GetSites() function and use the FullSiteCode field.   |

variableCode	The variable code. To get a list of possible variable codes, see GetVariables() function and use the FullVariableCode field
startDate	(optional) The start date in "yyyy-mm-dd" format
endDate	(optional) The end date in "yyyy-mm-dd" format
methodID	(optional) The ID of the observation method. To get a list of possible method IDs, see methodID column in the output of GetSiteInfo(). If methodID is not specified, then the observations in the output data.frame won't be filtered by method.
sourceID	(optional) The ID of the source. To get a list of possible source IDs, see sourceID column in the output of GetSiteInfo(). If sourceID is not specified, then the observations in the output data.frame won't be filtered by source.
qcID	(optional) The ID of the quality control level. Typically 0 is used for raw data and 1 is used for quality controlled data. To get a list of possible quality control level IDs, see QualityControlLevelID column in the output of GetSiteInfo(). If qcID is not specified, then the observations in the output data.frame won't be filtered by quality control level.
daily	(optional) If you set daily="max", daily="min" or daily="mean", then the data values are aggregated to daily time step.

## Value

a data.frame of data values with the following columns:

- time: The local date/time of the observation. The data type is POSIXct. POSIXct is a data type in R for storing time.
- DataValue: The observed data value
- UTCOffset: The difference between local time and UTC time in hours
- CensorCode: The code for censored observations. Possible values are nc (not censored), gt (greater than), lt (less than), nd (non-detect), pnq (present but not quantified)
- DateTimeUTC: The UTC time of the observation. The data type is POSIXct. POSIXct is a special data type in R for storing time.
- MethodCode: The code of the method or instrument used for the observation
- SourceCode: The code of the data source
- QualityControlLevelCode: The code of the quality control level. Possible values are -9999 (Unknown), 0 (Raw data), 1 (Quality controlled data), 2 (Derived products), 3 (Interpreted products), 4 (Knowledge products)

The output data.frame also has attributes with information about the status: download.time, parse.time, download.status, parse.status These attributes can be used for troubleshooting WaterOneFlow/WaterML server errors. If parse status is "NO\_VALUES\_FOUND", then this time series doesn't have any available data for the selected time period.

## Examples

```
#example 1: Get Values from a known site and variable from Ipswich River server.
server <- "http://hydroportal.cuahsi.org/ipswich/cuahsi_1_1.asmx?WSDL"
v1 <- GetValues(server, site="IRWA:FB-BV", variable="IRWA:DO",
               startDate="1999-01-01", endDate="1999-12-31")
#example 2: Get values from an external REST URL (in this case the Provo USGS NWIS site id 10163000)
url <- "http://waterservices.usgs.gov/nwis/dv/?format=waterml,1.1&sites=10163000&parameterCd=00060"
v2 <- GetValues(url)
#example 3: Get values from WaterML 2.0 file and show year, month, day
url2 <- "http://www.waterml2.org/KiWIS-WML2-Example.wml"
waterml2_data <- GetValues(url2)
waterml2_data$year <- strftime(waterml2_data$time, "%Y")
waterml2_data$month <- strftime(waterml2_data$time, "%M")
waterml2_data$day <- strftime(waterml2_data$time, "%d")
```

---

GetVariables

*GetVariables*

---

## Description

This function gets the table of variables from the WaterML web service

## Usage

```
GetVariables(server)
```

## Arguments

`server` The URL of the web service ending with ?WSDL, for example: <http://worldwater.byu.edu/app/index.php/>

## Value

a data.frame of variables with the following columns:

VariableCode	character	Short code of the variable
FullVariableCode	character	The full variable code, for example: SNOTEL:897. Use this value as the variableCode p
VariableName	character	The name of the variable
ValueType	character	the type of observation: Field Observation or Derived Value
DataType	character	the aggregate data type: Average, Continuous, Sporadic..
GeneralCategory	character	the general category of the measurements: Climate, Water Quality..
SampleMedium	character	the sample medium, for example water, atmosphere, soil..
UnitName	character	The name of the measurement units
UnitType	character	the type of the measurement units
UnitAbbreviation	character	The abbreviation of the measurement units (m, cm, in..)
NoDataValue	numeric	The value that indicates missing data
IsRegular	boolean	TRUE if the measurements are regular, FALSE otherwise
TimeUnitName	character	The name of the time units
TimeUnitAbbreviation	character	The time units abbreviation



TimeSupport	character	The length of the time period over which one measurement is taken
Speciation	character	The chemical sample speciation (as nitrogen, as phosphorus..)

### Examples

```
GetVariables("http://worldwater.byu.edu/app/index.php/rushvalley/services/cuahsi_1_1.asmx?WSDL")
```

---

```
HISCentral_GetSeriesCatalog
    HISCentral_GetSeriesCatalog
```

---

### Description

This function searches the table of time series from the HIS Central catalog

### Usage

```
HISCentral_GetSeriesCatalog(
    west,
    south,
    east,
    north,
    serviceID = NULL,
    keyword = NULL,
    beginDate = NULL,
    endDate = NULL
)
```

### Arguments

west	The west longitude of the geographic bounding box in decimal degrees. Allowed values are between -180.0 and +180.0
south	The south latitude of the geographic bounding box in decimal degrees. Allowed values are between -90.0 and +90.0
east	The east longitude of the geographic bounding box in decimal degrees. Allowed values are between -180.0 and +180.0
north	The north latitude of the geographic bounding box in decimal degrees. Allowed values are between -90.0 and +90.0
serviceID	(optional): The ID of the service on HIS Central. To get the service ID, use the id column in the output of the GetServices() function.
keyword	(optional): The concept keyword (common name of variable) for searching the sites on HIS Central. Examples include Temperature, Precipitation, Snow Depth,... If the Keyword is not specified then sites with any variable will be returned.
beginDate	(optional): The begin date of the observations in yyyy-mm-dd format.
endDate	(optional): The end date of the observations in yyyy-mm-dd format.

**Value**

a data.frame of series catalog entries. The data.frame has the following columns:

- ServiceCode: The code of the HydroServer
- ServiceURL: The URL of the server. Use this as the server parameter in GetValues() function.
- FullSiteCode: The complete unique code of the site in the format NETWORK:CODE. Use this value as the siteCode parameter in the GetValues function.
- FullVariableCode: The complete unique code of the site in the format VOCABULARY:CODE. Use this value as the variableCode parameter in the GetValues function.
- BeginDateTime: The local date/time of the first observation of the time series in POSIXct format.
- EndDateTime: The local date/time of the last observation of the time series in POSIXct format.
- ValueCount: The number of measurements in the time series
- SiteName: The name of the site.
- Latitude: The WGS84 latitude of the site in decimal degrees
- Longitude: The WGS84 longitude of the site in decimal degrees
- DataType: The data type of the variable
- ValueType: The type of the observation (field observation, sample, or derived value)
- SampleMedium: The sample medium (air, water or other)
- TimeUnits: The time units
- TimeSupport: The length of the time period of one measurement

**Examples**

```
#Getting all time series from the (14.1E, 49.9N, 14.3E, 50.1N) bounding box
series_catalog <- HISCentral_GetSeriesCatalog(west=14.1, south=49.9, east=14.3, north=50.1)
```

---

HISCentral\_GetSites     *HISCentral\_GetSites*

---

**Description**

This function gets the table of sites from the HIS Central catalog

**Usage**

```
HISCentral_GetSites(
  west = -180,
  south = -90,
  east = 180,
  north = 90,
  serviceID = NULL,
  keyword = NULL,
  IncludeServerDetails = TRUE
)
```

**Arguments**

west	The west longitude of the geographic bounding box in decimal degrees. Allowed values are between -180.0 and +180.0
south	The south latitude of the geographic bounding box in decimal degrees. Allowed values are between -90.0 and +90.0
east	The east longitude of the geographic bounding box in decimal degrees. Allowed values are between -180.0 and +180.0
north	The north latitude of the geographic bounding box in decimal degrees. Allowed values are between -90.0 and +90.0
serviceID	(optional): The ID of the service on HIS Central. To get the service ID, use the id column in the output of the GetServices() function.
keyword	(optional): The concept keyword (common name of variable) for searching the sites on HIS Central. Examples include Temperature, Precipitation, Snow Depth,... If the Keyword is not specified then sites with any variable will be returned.
IncludeServerDetails	(optional): If set to TRUE, then the output will include the servCode and servURL for each site. If set to FALSE, then we assume that all sites are from the same server and the servURL and servCode are not included.

**Value**

a data.frame of sites. The data.frame has the following columns:

- SiteName: The name of the site
- SiteCode: A short unique code of the site
- FullSiteCode: The complete unique code of the site in the format NETWORK:CODE. Use this value in the GetSiteInfo and GetValues functions
- Latitude: The WGS84 latitude in decimal degrees
- Longitude: The WGS84 longitude in decimal degrees
- ServCode: The code of the service in HIS Central. Same as the networkName in the output from GetServices() function. This column is only shown if IncludeServerDetails is TRUE.
- ServURL: The URL of the web service for this site as registered in HIS Central. This column is only shown if IncludeServerDetails is TRUE.

**Examples**

```
#Getting all sites from the (14.1E, 49.8N, 14.6E, 50.2N) bounding box from the GLDAS web service
sites <- HISCentral_GetSites(west=14.1, south=49.8, east=14.6, north=50.2, serviceID=262)
```

---

MakeSOAPEnvelope	<i>MakeSOAPEnvelope</i>
------------------	-------------------------

---

### Description

A helper function that makes a SOAP envelope to send to the CUAHSI WaterOneFlow SOAP web service. It is internally used by the GetSites, GetSiteInfo, GetVariables and GetValues functions.

### Usage

```
MakeSOAPEnvelope(CUAHSINamespace, MethodName, parameters = NULL)
```

### Arguments

CUAHSINamespace	The SOAP namespace. This must be either "http://www.cuahsi.org/his/1.0/ws" for WaterML 1.0, or "http://www.cuahsi.org/his/1.1/ws" for WaterML 1.1
MethodName	The name of the WaterOneFlow web service method. It can be one of the following values: "GetSites", "GetSitesObject", "GetSitesByBoxObject", "GetSiteInfoObject", "GetVariablesObject", "GetValuesObject"
parameters	An optional vector of named parameters for the web method. For GetSites, GetSitesObject and GetVariables no parameters are required. For GetSiteInfoObject you need the "site" parameter. For GetValuesObject you need the "location", "variable", "startDate" and "endDate" parameters.

### Value

A <soap:Envelope> text in XML format. This text is send in a HTTP POST body to the SOAP service. Two headers must be sent in the request: Content-Type="text/XML" and SOAPAction=paste(CUAHSINamespace, MethodName). For example if MethodName is GetSites and the WaterML version is 1.1, then SOAPAction="http://www.cuahsi.org/his/1.1/ws/GetSites".

### Examples

```
library(httr)
library(XML)
myEnvelope <- MakeSOAPEnvelope("http://www.cuahsi.org/his/1.1/ws/", "GetSitesObject")
SOAPAction <- "http://www.cuahsi.org/his/1.1/ws/GetSitesObject"
url <- "http://hydrodata.info/chmi-d/cuahsi_1_1.asmx"
response <- POST(url, body = myEnvelope,
                 add_headers("Content-Type" = "text/xml", "SOAPAction" = SOAPAction),
                 verbose())
status.code <- http_status(response)$category
WaterML <- xmlParse(response)
WaterML
```

---

WaterMLVersion	<i>WaterMLVersion</i>
----------------	-----------------------

---

**Description**

A helper function that finds out the WaterML version from the WaterML document. By default it checks for "http://www.opengis.net/waterml/2.0" Otherwise it tries to detect "http://www.cuahsi.org/waterML/1.1/" (for WaterML 1.1) or "http://www.cuahsi.org/WaterML/1.0/" (for WaterML 1.0)

**Usage**

```
WaterMLVersion(doc)
```

**Arguments**

doc	The XML document object
-----	-------------------------

**Value**

A character with the WaterML version: either 1.0, 1.1, or 2.0

**Examples**

```
library(httr)
library(XML)
url <- "http://www.waterml2.org/KiWIS-WML2-Example.wml"
response <- GET(url)
doc <- xmlParse(response)
version <- WaterMLVersion(doc)
```

---

WaterOneFlowNamespace	<i>WaterOneFlowNamespace</i>
-----------------------	------------------------------

---

**Description**

A helper function that finds out the WaterOneFlow namespace information based on the version number 1.0 or 1.1.

**Usage**

```
WaterOneFlowNamespace(version)
```

**Arguments**

version	The version of the WaterOneFlow XML namespace. Must be either "1.0" or "1.1"
---------	--

**Value**

A list with the namespaces and corresponding prefixes. This namespace information is important for correct parsing of the WaterML XML document.

**Examples**

```
ns <- WaterOneFlowNamespace("1.0")
ns <- WaterOneFlowNamespace("1.1")
```

---

WaterOneFlowVersion	<i>WaterOneFlowVersion</i>
---------------------	----------------------------

---

**Description**

A helper function that finds out the WaterOneFlow service version from the URL of the wsdl file. By default it checks for "cuahsi\_1\_0" or "cuahsi\_1\_1" in the url. If that is not found, then the function checks the version inside the WSDL file.

**Usage**

```
WaterOneFlowVersion(WSDL)
```

**Arguments**

WSDL	The URL of the WSDL, for example <a href="http://icewater.usu.edu/MudLake/cuahsi_1_0.asmx?WSDL">http://icewater.usu.edu/MudLake/cuahsi_1_0.asmx?WSDL</a>
------	--

**Value**

A list with two items: Version (either 1.0 or 1.1), and Namespace (either <http://www.cuahsi.org/his/1.0/ws/> or <http://www.cuahsi.org/his/1.1/ws/>)

**Examples**

```
versionInfo <- WaterOneFlowVersion("http://icewater.usu.edu/MudLake/cuahsi_1_0.asmx?WSDL")
versionInfo$Version
versionInfo$Namespace
```

# Index

- \*Topic **CUAHSI**
    - WaterML-package, [2](#)
  - \*Topic **HydroServer**
    - WaterML-package, [2](#)
  - \*Topic **Hydrology**
    - WaterML-package, [2](#)
  - \*Topic **WaterML**
    - GetVariables, [8](#)
    - MakeSOAPEnvelope, [12](#)
    - WaterML-package, [2](#)
    - WaterMLVersion, [13](#)
    - WaterOneFlowNamespace, [13](#)
    - WaterOneFlowVersion, [14](#)
  - \*Topic **waterml**
    - GetServices, [3](#)
    - GetSiteInfo, [3](#)
    - GetSites, [5](#)
    - GetValues, [6](#)
    - HISCentral\_GetSeriesCatalog, [9](#)
    - HISCentral\_GetSites, [10](#)
- [GetServices, 3](#)
- [GetSiteInfo, 3](#)
- [GetSites, 5](#)
- [GetValues, 6](#)
- [GetVariables, 8](#)
- [HISCentral\\_GetSeriesCatalog, 9](#)
- [HISCentral\\_GetSites, 10](#)
- [MakeSOAPEnvelope, 12](#)
- [WaterML-package, 2](#)
- [WaterMLVersion, 13](#)
- [WaterOneFlowNamespace, 13](#)
- [WaterOneFlowVersion, 14](#)