CMLh5

Commercial microwave link data format recommendations for hydrometeorological research

VERSION 0.2

1 Background and Motivation

[Write some nice introduction……..]

...Attenuation data from commercial microwave links (CMLs) can be used to derive rain rate in humidity information. Since CML networks are used a backhaul in all cell phone networks, there is a good coverage of most populated regions on the globe. This makes

...The feasability of the use of CMLs as remote sensing tool has been shown in numerous studies from several countries with very different orography, climate and economic development.

...Also suitable for purpose built research microwave links (RMLs)....

2 CML Terminology

Up to now there is no unified nomenclature of the basic components of hydrometeorological CML research. For instance, different abbreviations for “commercial microwave link” are used. MW link, MWL, LINK, ...

This sections introduces a standardized terminology for the the main components that are commonly used and described when CMLs are used for hydrometeorological research.

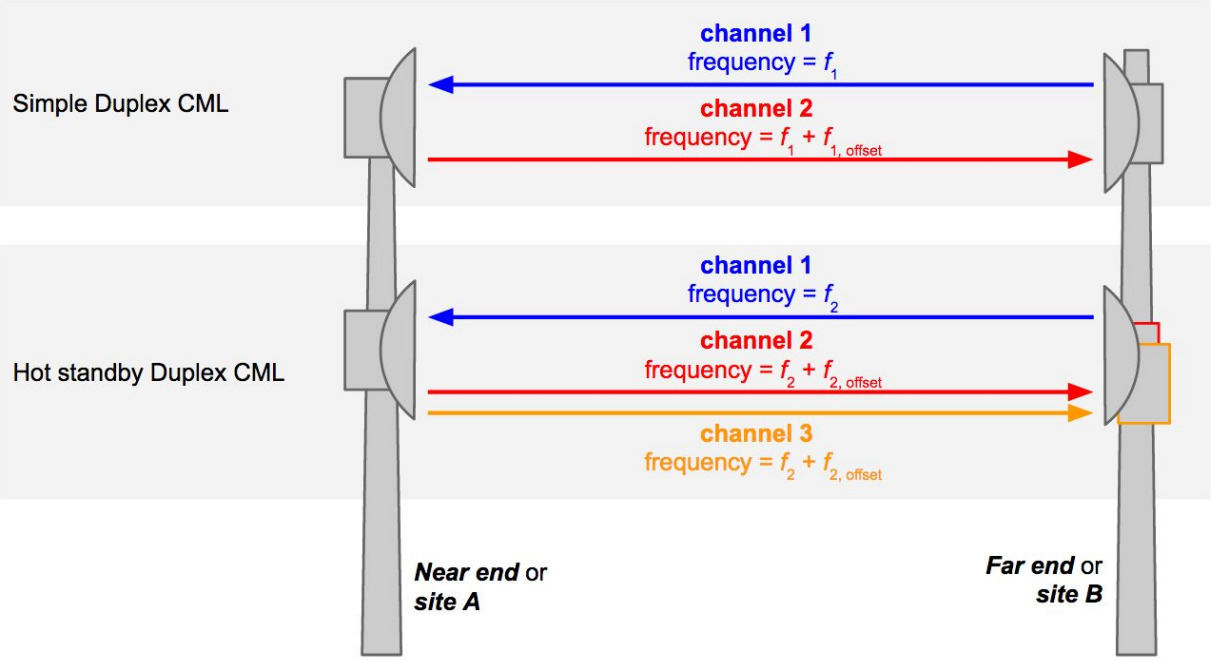

Figure 1: Overview of two CML configurations. The grey boxes indicate what *CMLh5* considers to be one CML (the figure can be found here).

Terminology	**Description**
CML	Commercial microwave link ● comprising one or more channels ● using one or two antennas per site (two antennas are used for space diversity configurations) ● using the same hardware type for all channels (i.e. dual frequency or dual polarization is allowed if the corresponding channels are the same hardware model.)

channel	One physical microwave transmission connection which is using: ● one frequency ● one polarization ● one direction, i.e. one transmitter and one receiver ● one antenna on each side
RX level, RSL,	Received signal level
TX level, TSL	Transmitted signal level
TXRX, TRSL	
ATPC	Automatic transmission power control

3 CMLh5 data format structure

For storage of the CML data and metadata HDF5 files are used. CMLh5 defines the structure which is used to store the data within the HDF5 container.

In CMLh5 every CML gets its own *HDF5 group* in the *HDF5 RootGroup.* Within the *CML group,* groups are generated for each channel of the particular CML. The actual data, records of transmitted and received signal level together with their time stamps, are stored as arrays in these *channel groups*. Additional records, e.g. error flags or bit error rate, if available at the same time stamps and if specific for the individual channel, should also be stored in the *channel subgroup* as a separate array.

Further additional records, e.g. of hardware temperature (which is site specific and not channel specific), can be optionally stored in a separate *auxiliary subgroup* of the CML *group*. They can have their own time index array or share one with one of the channels via a HDF5 linkage of the arrays.

[Write something about optional derived data…..]

3.1 Example structure of a CMLh5 file

```
/                          RootGroup
/cml_1                     Group for first CML
/cml_1/channel_1           Group for first channel
/cml_1/channel_1/rx        Array of RSL values in dBm
/cml_1/channel_1/tx        Array of TSL values in dBm
/cml_1/channel_1/time      Array of timestamps in POSIX time
/cml_1/channel_1/bit_errorrate  Additional data of the channel
/cml_1/channel_1/quality   Quality flag

/cml_1/channel_2           Group for second CML channel_2
/cml_1/channel_2/rx
/cml_1/channel_2/tx
/cml_1/channel_2/time

/cml_1/auxiliary_1/odu_temp_site_a Outdoor unit temperature
/cml_1/auxiliary_1/time          link to /time from channel_1

/cml_1/product_1/rain_rate
/cml_1/product_1/time
```

```
/cml_1/product_2/humidity
/cml_1/product_2/time

/cml_2                      Group for second CML
/cml_2/channel_1
/cml_2/channel_1/rx
/cml_2/channel_1/tx
/cml_2/channel_1/time
```

Things to note:

- If no **`TX`** record is supplied at the **`channel`** level the attribute TX_const must be set. (See table of metadata in next section.)

3.2 Specifications of the arrays in the channel subgroup

Single array for each “record” with own attributes…..

Attributes should contain information on averaging, etc.

Name	Units	Type	Description
rx	dBm	float16	
rx_min	dBm	float16	
rx_max	dBm	float16	
tx	dBm	float16	
tx_min	dBm	float16	
tx_max	dBm	float16	
time	POSIX time	float64	

3.3 Dimension scales

[to be written...]

4 CMLh5 metadata conventions

General note: In the tables below, bold text indicates obligatory metadata. Metadata in light font are optional but should follow the given naming convention.

4.1 Metadata at the root level

Metadata name	**Units**	**Type**	**Description**
file_format	-	string	This must always be set to ‘CMLh5’
file_format_version	-	string	examples: ‘0.1, ‘1.2’, ...
author_name	-	string	
author_email	-	string	

4.2 Metadata at the CML level

The following metadata is attached as attributes at the CML group of each CML entry in the CMLh5 file.

Metadata name	**Units**	**Type**	**Description**
site_a_latitude	Decimal degrees	float16	
site_a_longitude	Decimal degrees	float16	
site_a_altitude	Meter	float16	
site_a_antenna_above_ground	Meter	float16	
site_a_id	-	string	
site_b_latitude	Decimal degrees	float16	
site_b_longitude	Decimal degrees	float16	
site_b_altitude	Meter	float16	
site_b_antenna_above_ground	Meter	float16	

site_b_id	-	string	
length	Kilometer	float64	Strongly recommended
cml_id	-	string	
cml_owner	-	string	
cml_operator	-	string	
system_manufacturer	-	string	
system_model	-	string	
system_configuration	-	string	Options are: ● ‘1+0’ ● ‘1+1_SD’ ● ‘1+1_HSB’ ● (‘2+0’) HSB = Hot standby SD = Space diversity

4.3 Metadata at the channel level

The following metadata is attached as attributes at the channel group of each CML entry in the CMLh5 file.

Metadata name	Units/Values	Type	Description
frequency	GHz	float64	
polarization	[‘V’, ‘H’, ‘v’, ‘h’]	string	
tx_site	[‘site_a’, ‘site_b’]	string	
rx_site	-	string	
channel_id	-	string	
atpc	[‘on’, ‘off’]	string	
tx_quantization	dBm	float64	
tx_quantization_type	[‘rounded’, ‘truncated’]	string	

rx_quantization	dBm	float64	
rx_quantization_type	[‘rounded’, ‘truncated’]	string	
tx_const	dBm	float64	value of the TX level if ATPC is off
channel_name	-	string	
additional_info	-		
sampling_type	[‘mean’, ‘max’, ‘min’, ‘instantaneous’]	string	
time_stamp_alignment	[‘left’, ‘center’, ‘right’]	string	Info on the association of the time stamp regarding the aggregation period
temporal_resolution	[‘not_constant, ‘min’, ‘s’, ‘30s’, ‘H’]	string	

4.4 Metadata at the array level

The following metadata is attached as attributes at the individual data arrays of each CML entry in the CMLh5 file.

metadata name	**units/values**	**type**	**description**
quantity	e.g. ‘power’	string	
unit	e.g. ‘dBm’	string	
side?	[‘transmitter’, ‘receiver’]	string	
sampling?	[‘min’, ‘max’, ‘instant’, ‘mean’]	string	
array_id	-	string	

5 Missing data conventions

data type	**fill value**
float	NaN
int	-9999
string	“NA”