**Allgemeines**

* Ziel: Konvertierung des TK3 Outputs (csv-Datei) zu NetCDF.
* Dabei orientieren wir uns, soweit es möglich ist und uns sinnvoll erscheint, an den CF Conventions, versuchen aber nicht, diese vollständig umzusetzen.
* Zur “Darstellung” unserer Daten orientieren wir uns an der in den CF Conventions beschriebenen [orthogonal multidimensional array representation of time series](https://cfconventions.org/Data/cf-conventions/cf-conventions-1.8/cf-conventions.html#_orthogonal_multidimensional_array_representation_of_time_series).
* Nützliche Links:
  + [NetCDF Users Guide](https://www.unidata.ucar.edu/software/netcdf/documentation/NUG/index.html).
  + [Aktuelle Version der CF Conventions](https://cfconventions.org/Data/cf-conventions/cf-conventions-1.8/cf-conventions.html).
    - [Appendix A > List of CF attributes](https://cfconventions.org/Data/cf-conventions/cf-conventions-1.8/cf-conventions.html#attribute-appendix).
    - [Appendix H.2 > Examples for time-series data](https://cfconventions.org/Data/cf-conventions/cf-conventions-1.8/cf-conventions.html#time-series-data).
  + [CF standard name table](https://cfconventions.org/Data/cf-standard-names/76/build/cf-standard-name-table.html).

Ein Versuch, die unten beschriebene Struktur und die Zuordnung der TK3 Variablen zu den netCDF-Variablen und deren Attributen abzubilden, findet sich in TK3\_to\_NetCDF.xlsx. Ob eine Exceldatei dafür das geeignete Mittel ist, sei mal dahingestellt… 😊

* Im Tabellenblatt *global attributes* sind alle von uns vorgeschlagenen globalen Attribute mit entsprechenden Beispielwerten aufgeführt.
* Im Tabellenblatt *station variables* sind alle von uns vorgeschlagenen Stationsvariablen mit zugehörigen Attributen und mit entsprechenden Beispielwerten aufgeführt.
* Im Tabellenblatt *TK3 variables to netCDF* findet sich unser Versuch der Zuordnung der TK3 Variablen zu den netCDF Variablen und deren Attributen.
  + Die obersten 4 Zeilen beziehen sich auf den TK3 Output (Variablenname in TK3; Beispielwert; Einheit, in der TK3 Variable angegeben ist; Umrechnung, die nötig ist, um die Werte von den „TK3 Einheiten“ zu den von uns vorgeschlagenen „netCDF Einheiten“ umzurechnen).
  + Die Zeilen 7 bis 17 beschreiben die zugehörigen netCDF Variablen und deren Attribute.
  + Nur die in der obersten Zeile grün hinterlegten TK3 Variablen werden in netCDF überführt.
  + Ist ein Feld ausgegraut, heißt das, das entsprechende Attribut wird der entsprechenden Variable nicht zugewiesen.
* Das Tabellenblatt *TK3 variables explanation* ist einfach nochmal eine Beschreibung des TK3 Outputs.

**Global Attributes**

* Globale Attribute dienen zur Beschreibung des Dateiinhalts (siehe [CF Conventions > Description of file contents](https://cfconventions.org/Data/cf-conventions/cf-conventions-1.8/cf-conventions.html#description-of-file-contents)).
* Alle in TK3\_to\_NetCDF.xlsx (Blatt *global attributes*) aufgeführten Attribute sind obligatorisch und sind auch weitestgehend in den CF Conventions vorgesehene Attribute.
* Zusätzliche globale Attribute können definiert werden, sind aber nicht obligatorisch.

**Dimensionsvariablen**

* Über die Dimensionsvariablen werden die Datenwerte in Zeit und Raum verortet. (*A data variable’s dimensions are used to locate data values in time and space.*)
* Entsprechend haben wir die Dimensionsvariablen time und station definiert.
* Data Variables haben die Dimension time, station.
* Stationsvariablen (siehe nächster Punkt) haben die Dimension station.

**station**

* Dient der Beschreibung der Messstation (über entsprechende „Stationsvariablen“).
* Bekommt die Länge 1. Das heißt, es wird nur eine Messtation definiert, da alle Data Variables von der selben Station kommen.
* Alle in TK3\_to\_NetCDF.xlsx (Reiter *station variables*) aufgeführten Variablen und zugehörigen Attribute sind obligatorisch und sind auch weitestgehend in den CF Conventions vorgeschlagene Variablen und Attribute.
* Zusätzliche Stationsvariablen können definiert werden, sind aber nicht obligatorisch.

**time**

* Die TK3-Variable T\_mid (Mitte des TK3-Mittelungsintervalls) wird die Dimensionsvariable time.
* Die TK3-Variablen T\_begin und T\_end (Anfang und Ende des TK3-Mittelungsintervalls) werden zu einer bounds-Variable time\_bnds mit den Dimensionen (time x 2) zusammengefasst (1. „Spalte“ in time\_bnds ist T\_begin, 2. „Spalte“ ist T\_end).
* Die Variable time\_bnds wird über das Attribut bounds der Dimensionsvariable time zugeordnet, wodurch der gesamte Zeitraum des entsprechenden TK3-Mittelungsintervalls definiert ist.

**Data Variables**

* Sind unsere “eigentlichen” Messwerte bzw. die Variablen, die von TK3 ausgegeben werden.
* Data Variables haben die Dimension time, station (siehe Abschnitt Dimensionsvariablen).
* Data Variables haben einen Namen.
* Data Variables haben verschiedene Attribute, welche einerseits die Variablen beschreiben und andererseits auch Beziehungen zwischen verschiedenen Variablen abbilden.
* Unsere Vorschläge zu den nachfolgenden Punkten finden sich TK3\_to\_NetCDF.xlsx (Reiter *TK3 variables to netCDF*).

**Variablennamen**

* Können grundsätzlich frei definiert werden (d.h. es gibt keine Vorgaben in den CF Conventions).
* Unsere Vorschläge sind möglichst eingängige bzw. uns geläufige Namen.
* Die schlussendlich festgelegten Variablennamen sind obligatorisch.

**Attribute**

* Aufgenommen haben wir die in den CF Conventions vorgesehen Attribute standard\_name, long\_name, units, \_FillValue, ancillary\_variables, bounds, cell\_methods, coordinates und comment. Nicht jede Variable hat aber unbedingt alle diese Attribute (siehe unten)!
* Außerdem haben wir zusätzlich das Attribut instrument aufgenommen.
* Weitere Attribute können eingeführt werden, sind aber nich obligatorisch.
* Im Folgenden sind die einzelnen Attribute näher erklärt.

standard\_name

* Sind in den CF Conventions genau definierte Namen (siehe [CF standard name table](https://cfconventions.org/Data/cf-standard-names/76/build/cf-standard-name-table.html)), die sozusagen zur exakten Definition der entsprechenden Variable dienen (z.B. ist durch den standard\_name klar definiert, welche Richtung ein positiver / negativer Fluss hat).
* Der standard\_name ist nicht zwingend gleich dem Variablennamen und verschiedene Variablen können auch den selben standard\_name haben!
* Das Attribut standard\_name wird zwingend zugewiesen, wenn es einen entsprechenden CF Standardnamen gibt und weggelassen, wenn es keinen gibt.

long\_name

* Dienst als “langer” beschreibender Name der Variable (z.B. zur Beschriftung von Diagrammen).
* Können frei definiert werden (d.h. es gibt keine Vorgaben in den CF Conventions).
* Unsere Vorschläge siehe TK3\_to\_NetCDF.xlsx.
* Die schlussendlich festgelegten long\_name sind obligatorisch.

units

* Definiert die Einheit, in der die Werte angegeben sind.
* Unser Vorschlag ist, alles in SI Basiseinheiten anzugeben.
* Die schlussendlich festgelegten units sind obligatorisch.

\_FillValue

* Ist der Wert, der Fehlwerten zugewiesen wird.
* Wert sollte (deutlich) außerhalb des Wertebereichs der entsprechenden Variable sein.

ancillary\_variables

* Das Attribut ancillary\_variables bietet die Möglichkeit, Variablen die miteinander assoziiert sind bzw. von denen eine Variable die andere beschreibt, miteinander zu verlinken.
* Die „beschriebene“ Variable (z.B. sensibler Wärmefluss) bekommt das Attribut ancillary\_variables zugewiesen, in dem wiederum die „beschreibende“ Variable (z.B. zugehöriger QA/QC Flag) oder auch mehrere „beschreibende“ Variablen als Wert(e) stehen.
* Dieses Attribut wird nur Variablen zugewiesen, bei denen es entsprechende Beziehungen gibt (siehe TK3\_to\_NetCDF.xlsx).

bounds

* Dieses Attribut ist nur für die Dimensionsvariable time relevant und wird nur dort zugewiesen (siehe Abschnitt Dimensionsvariablen weiter oben).

cell\_methods

* Das Attribut cell\_methods bietet die Möglichkeit genauer zu definieren, um was für einen Wert es sich genau handelt (z.B. das es sich um einen Mittelwert über ein Zeitintervall handelt).
* Dieses Attribut haben wir überall dort zugewiesen, wo in den CF Conventions eine entsprechende Methode definiert ist (siehe [CF cell methods](https://cfconventions.org/Data/cf-conventions/cf-conventions-1.8/cf-conventions.html#appendix-cell-methods)) und uns eine entsprechende Zuordnung möglich erschien.

coordinates

* Dient zur Zuordnung der Datenvariablen zu den entsprechenden Stationsvariablen.

comment

* Über das Attribut comment können weitere Information eingetragen werden, die durch keines der anderen Attribute abgedeckt sind, aber wichtig erscheinen.
* Sollte sich jedoch herausstellen, dass ein entsprechender comment bei sehr Variablen eingetragen wird, ist zu überlegen, ob nicht ein entsprechendes neues Attribut eingeführt wird.
* Dieses Attribut wird allen Variablen zugewiesen. Ein Eintrag ist jedoch nicht obligatorisch und es gibt keine Vorgaben zum Format der Einträge.

instrument

* Nicht von den CF Conventions definiertes, sondern von uns zusätzlich eingefügtes Attribut zur Beschreibung mit welchem Instrument der entsprechende Wert gemessen wurde.
* Dieses Attribut wird nur Variablen zugewiesen, die eindeutig einem Instrument zugeordnet werden können.
* Es gibt keine Vorgaben zum Format der Einträge.