WaZuBay IT-Workshop

Ivana Jovanovic, Hai Nguyen, Wolfgang Kurtz, Tobias Neckel, Anton Frank

13.04.2021

Hydro-BITS

Basis IT-Struktur für die Wasser-Zukunft-Bayern

Agenda

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

1) Einführung LfU (10 Minuten)
2) Übersichtsvortrag “IT-Themen in WaZuBay aus Sicht von Hydro-BITS” (TUM/LRZ, 30 Minuten)
3) Feedback der WaZuBay-Partner (alle, 45 Minuten)
4) Wrap-up (10 Minuten)
5) Übersicht “Geplante IT-Workshops” und Feedback der WaZuBay-Partner (TUM/LRZ, 20 Minuten)
6) Nächste Schritte (alle, 15 Minuten)

Motivation

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

In den verschiedenen WaZuBay-Projekten wird eine Vielzahl von Datenprodukten und Werkzeugen entwickelt, z.B.

- Aufbereitung/ Weiterentwicklung von Geodaten des LfU
- Erzeugen von Modelleingabedaten
- Modellsimulationen (Desktop, HPC)
- Modellweiterentwicklungen (Parameterisierung, Kalibrierung, UQ)
- Analyse-Tools
- …

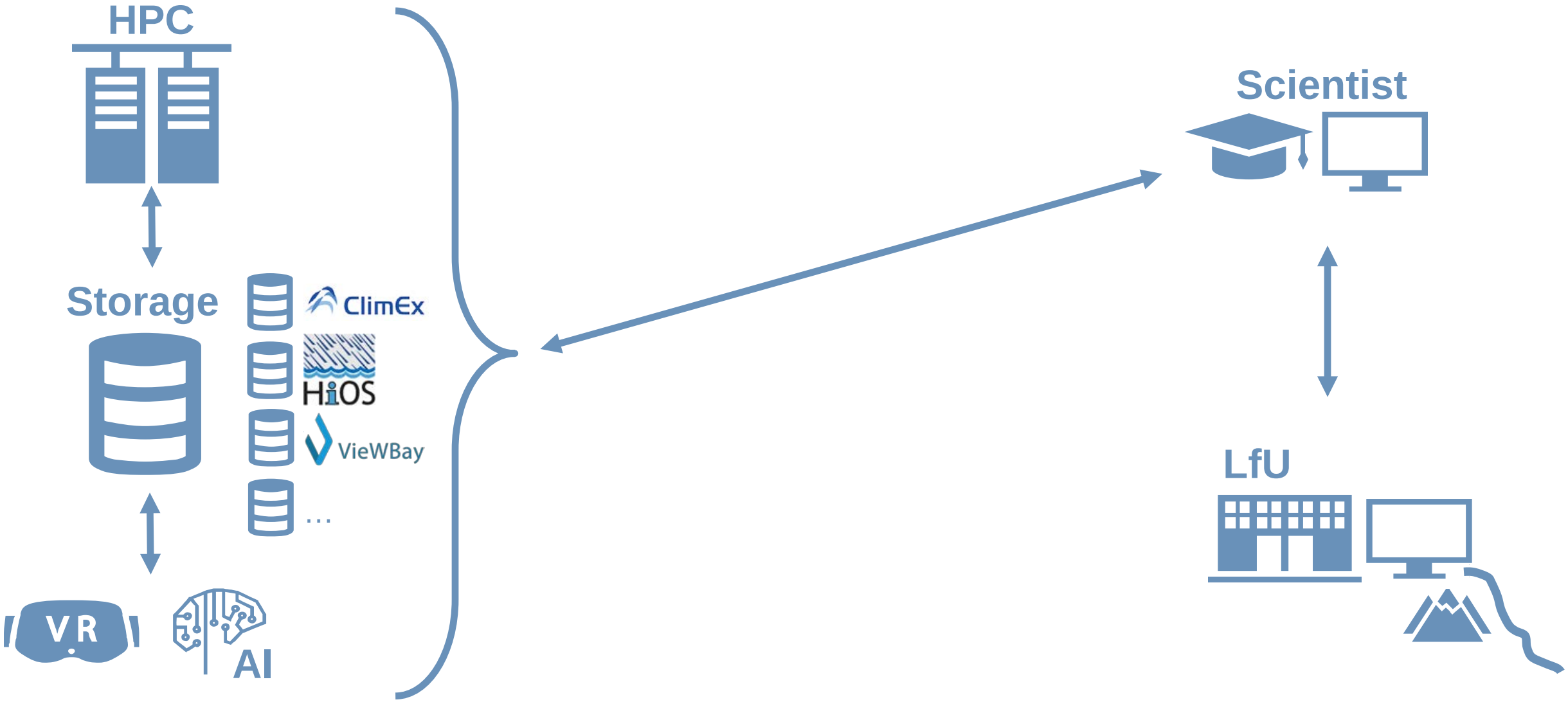
Motivation

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

- Synergien zwischen Projekten erzeugen (gemeinsam genutzte Tools und Daten)
- Bereitstellung von Infrastruktur zur Unterstützung der Datenprozessierung/-analyse
- Standardisierung von Datenformaten und Workflows
- Bereitstellung von Daten durch das LfU zu Projektbeginn
    - Einfache Bereitstellung für verschiedene Projekte/ Nutzer
    - Modifikationen/ Abgeleitete Daten versioniert zurückspielen
- Dissemination von Projektergebnissen nach Projektende
    - Zugriffsmöglichkeiten für LfU
    - Nachhaltigkeit z.B. Wiederverwendung in Folgeprojekte

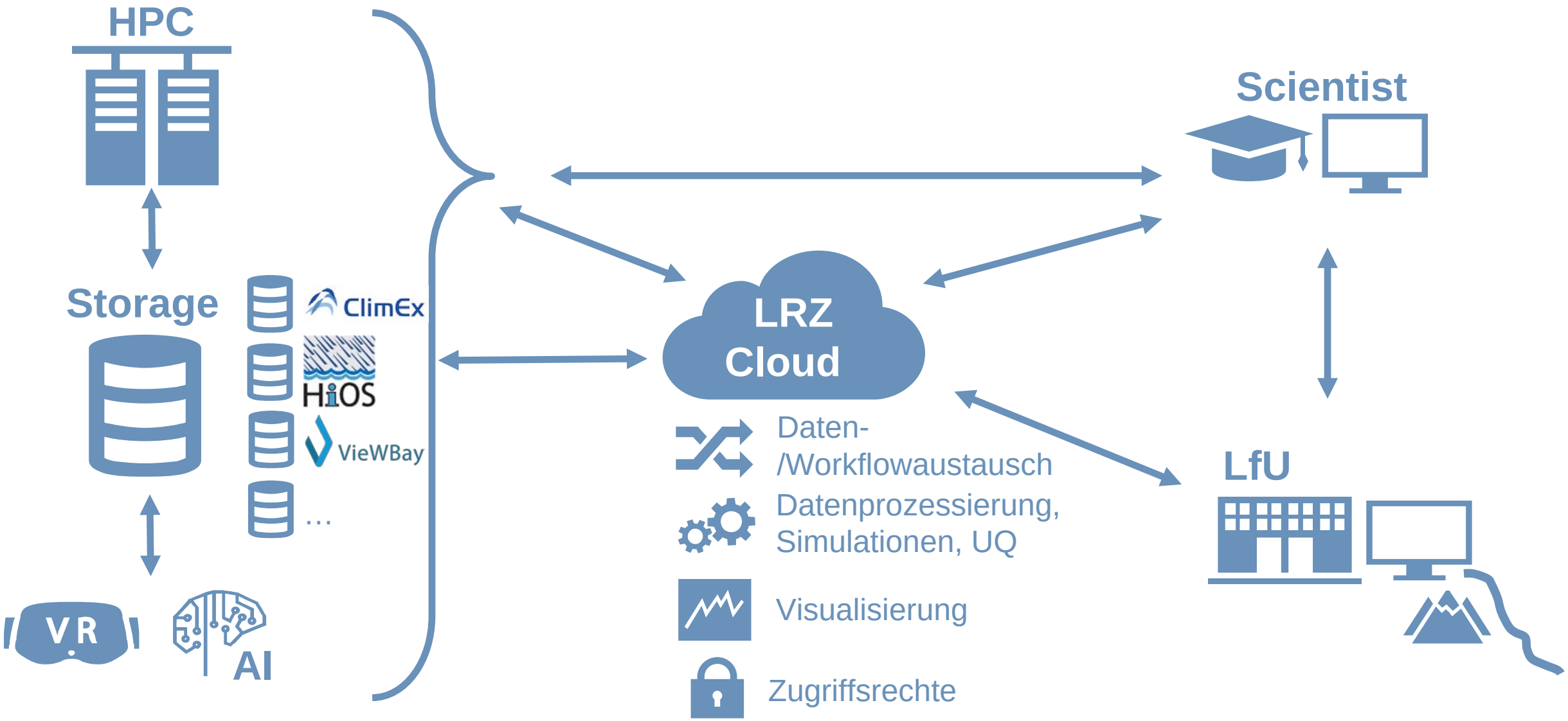
Datenaustausch

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern



Datenaustausch

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern



Datenaustauschplattform

- Datenaustauschplattform sollte nicht ‘from scratch’ entwickelt werden, sondern auf bereits vorhandenen etablierten Bausteinen aufbauen
- Es gibt keine ‘out-of-the-box’ Lösung für die spezifischen Anforderungen in WaZuBay
- Untersuchung bereits existierender Frameworks hinsichtlich der Anwendbarkeit in Hydro-BITS/ WaZuBay
  - OpenDataCube
  - PAVICS
  - HydroShare
  - Tethys

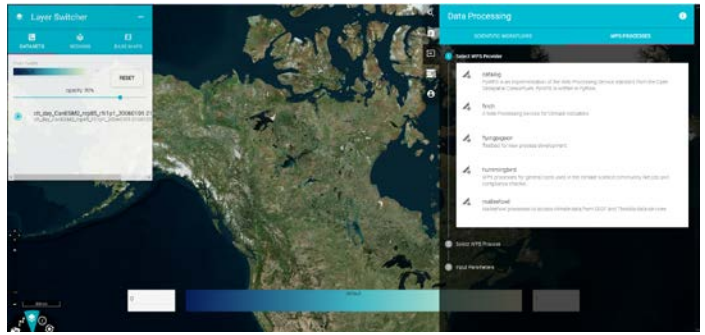
Schema PAVICS

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

Frontend

Backend

jupyter



Zugriffsmanagement

Magpie

Web Processing Services



Datenmanagement

GeoServer

Datensuche/-indexierung

Solr



Backend-Komponenten

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

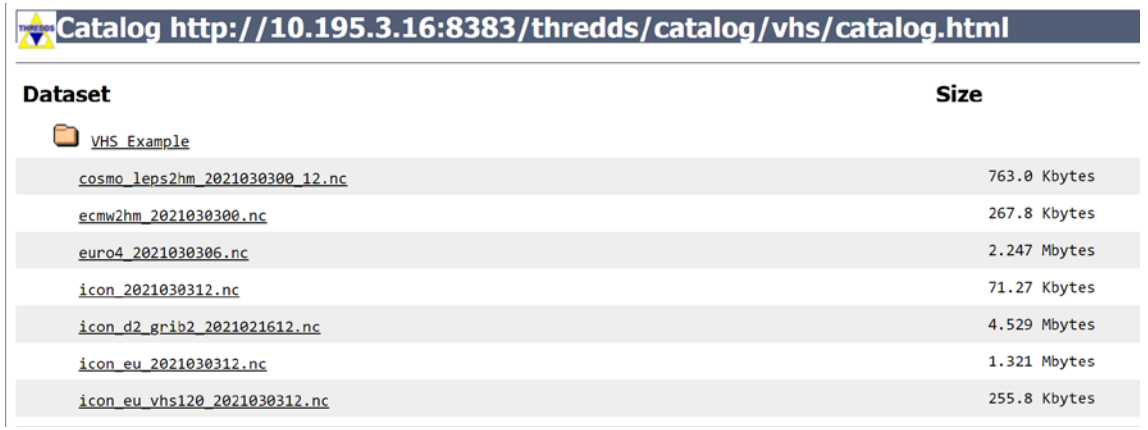
Geoserver:

- Standardkomponente für die Bereitstellung von Vektor- und Rasterdaten
- Bereitstellung von OGC-Services (z.B. Web Map Service)

GeoServer

THREDDS-Server:

- Webserver für die Bereitstellung von netCDF-Dateien
- Möglichkeit der Erstellung virtueller Datensätze (Aggregation, Modifikation der Metadaten)
- OpenDAP-Schnittstelle für Datenzugriff
- Web Map Service
- Subsetting Service



Backend-Komponenten

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

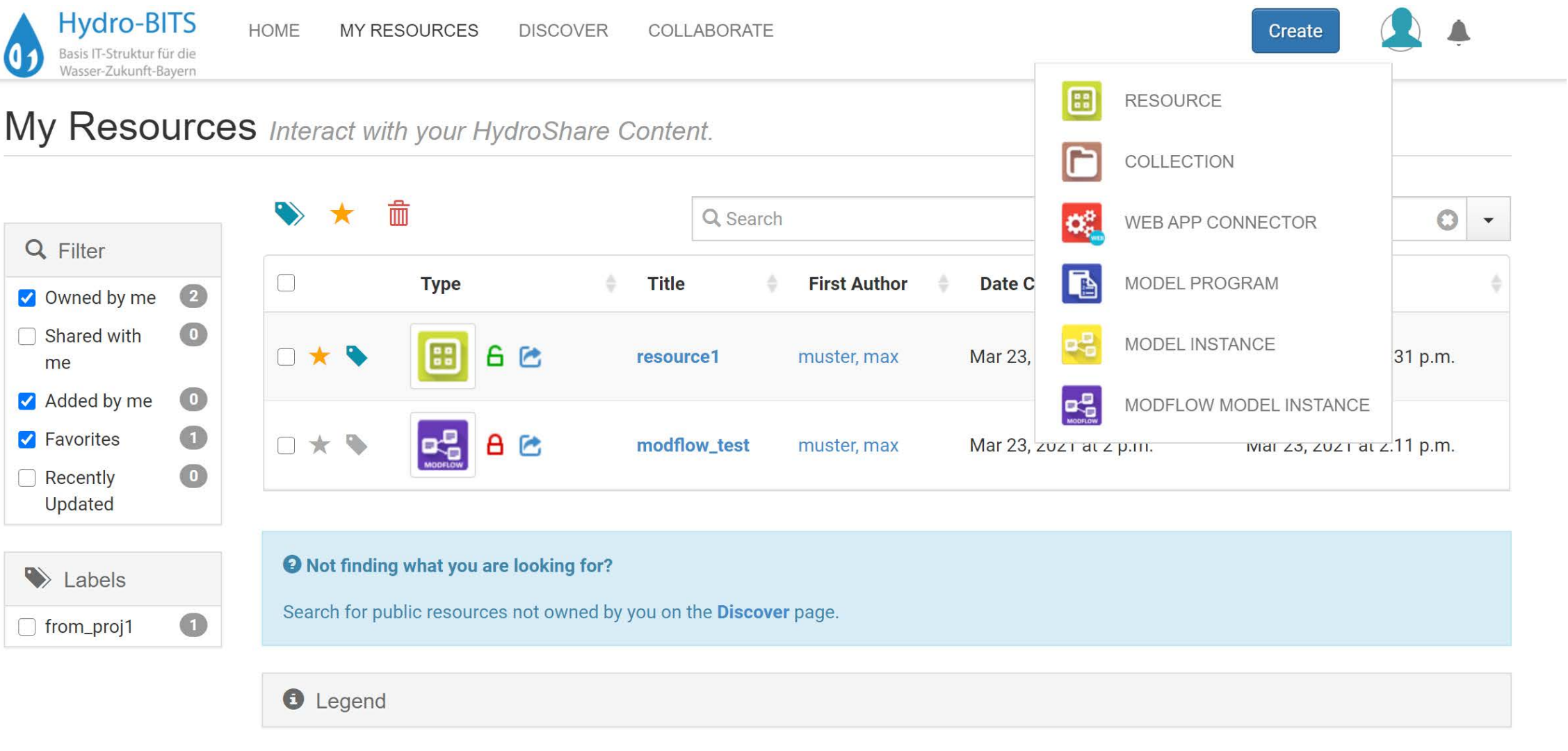
HydroShare:

- Dateibasierte Ablage unterschiedlicher Datensätze (Daten, Modellinput, Modellergebnisse, Skripte, Jupyter-Notebooks, Berichte, …)
- Metadaten nach Dublin-Core + Erweiterungen, z.B. für Modelle
- Datensätze können privat und öffentlich sein oder nur mit bestimmten Nutzern oder Nutzergruppen geteilt werde
- API für Datenzugriff z.B. in Skripten oder für Webapplikationen

Web Processing Services:

- Für Standard-Prozeduren mit vordefinierten Input-/Output-Parametern, z.B. für Datenextraktion, Datenanalyse, Simulationen
- Datenprozessierung erfolgt remote auf Cloud-/ HPC-Cluster-Infrastruktur

HydroShare


Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern
HOME
MY RESOURCES
DISCOVER
COLLABORATE
Create
RESOURCE
COLLECTION
WEB APP CONNECTOR
MODEL PROGRAM
MODEL INSTANCE
MODFLOW MODEL INSTANCE
My Resources
Interact with your HydroShare Content.
Search
Filter
Owned by me 2
Shared with me 0
Added by me 0
Favorites 1
Recently Updated 0
Labels
from_proj1 1
Type
Title
First Author
Date C
resource1
muster, max
Mar 23,
31 p.m.
modflow_test
muster, max
Mar 23, 2021 at 2 p.m.
Mar 23, 2021 at 2:11 p.m.
Not finding what you are looking for?
Search for public resources not owned by you on the Discover page.
Legend

HydroShare

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

Hydro-BITS HOME MY RESOURCES DISCOVER COLLABORATE

modflow_test

Authors:	max muster	**Sharing Status:**	Private
Owners:	max muster	**Views:**	11
Resource type:	MODFLOW Model Instance Resource	**Downloads:**	0
Storage:	The size of this resource is 42.7 KB	**+1 Votes:**	Be the first one to +1 this.
Created:	Mar 23, 2021 at 2 p.m.	**Comments:**	No comments (yet)
Last updated:	Mar 23, 2021 at 2:11 p.m. max muster		
Citation:	See how to cite this resource		

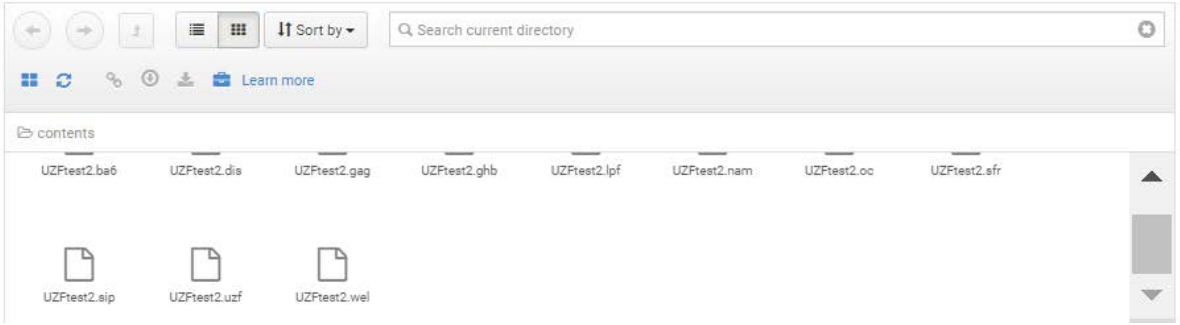
Abstract

Some ModFlow test resource

Subject Keywords

Groundwater ModFlow

Content



Resource Specific

Model Output

Includes output files?	No

Study Area

Total length in meters	22860.0
Total width in meters	15240.0

Stress Period

Type	Steady and Transient

Boundary Condition

Specified flux boundary package(s)	WEL
Head-dependent flux boundary package(s)	SFR, GHB, UZF

Executed By

Name	*Unspecified*

Grid Dimensions

Number of layers	1
Type of rows	Regular
Number of rows	15
Type of columns	Regular
Number of columns	10

Groundwater Flow

Flow package	LPF
Includes Unsaturated Zone Package (UZF)	Yes

Model Input

General

Model solver	SIP
Output control package	OC, GAGE

Frontend-Komponenten

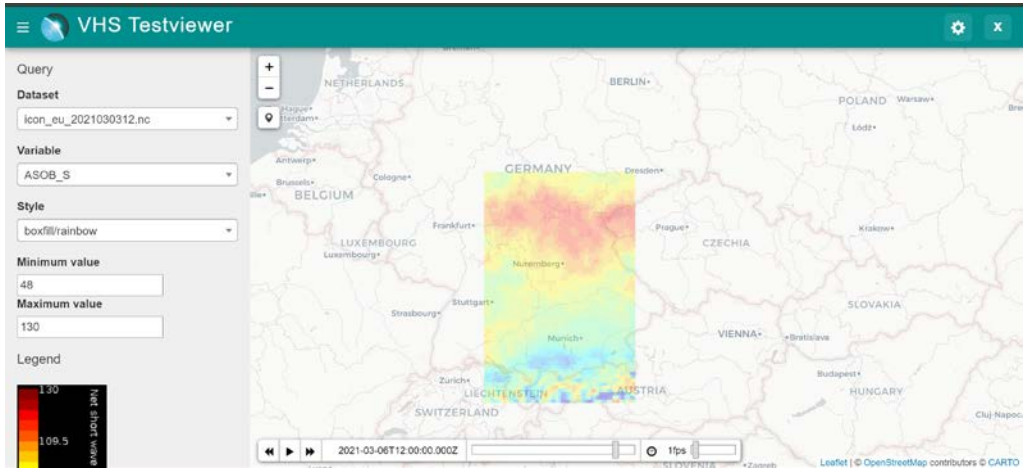
Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

Jupyter:

- Interaktive Datenbearbeitung für verschiedene Interpreter (Python, R, …)
- Flexibel erweiterbar z.B. für Git-Anbindung, Visualisierung, …
- Zugriff auf Datenmanagementkomponenten über APIs
- Einfacher Austausch von Workflows über Skripten

Webapplikationen:

- Interaktive Dashboards für Datenanalyse, Visualisierung, …
- Übersicht/ Management verfügbarer Daten (z.B. HydroShare)



ClimEx Dashboard

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

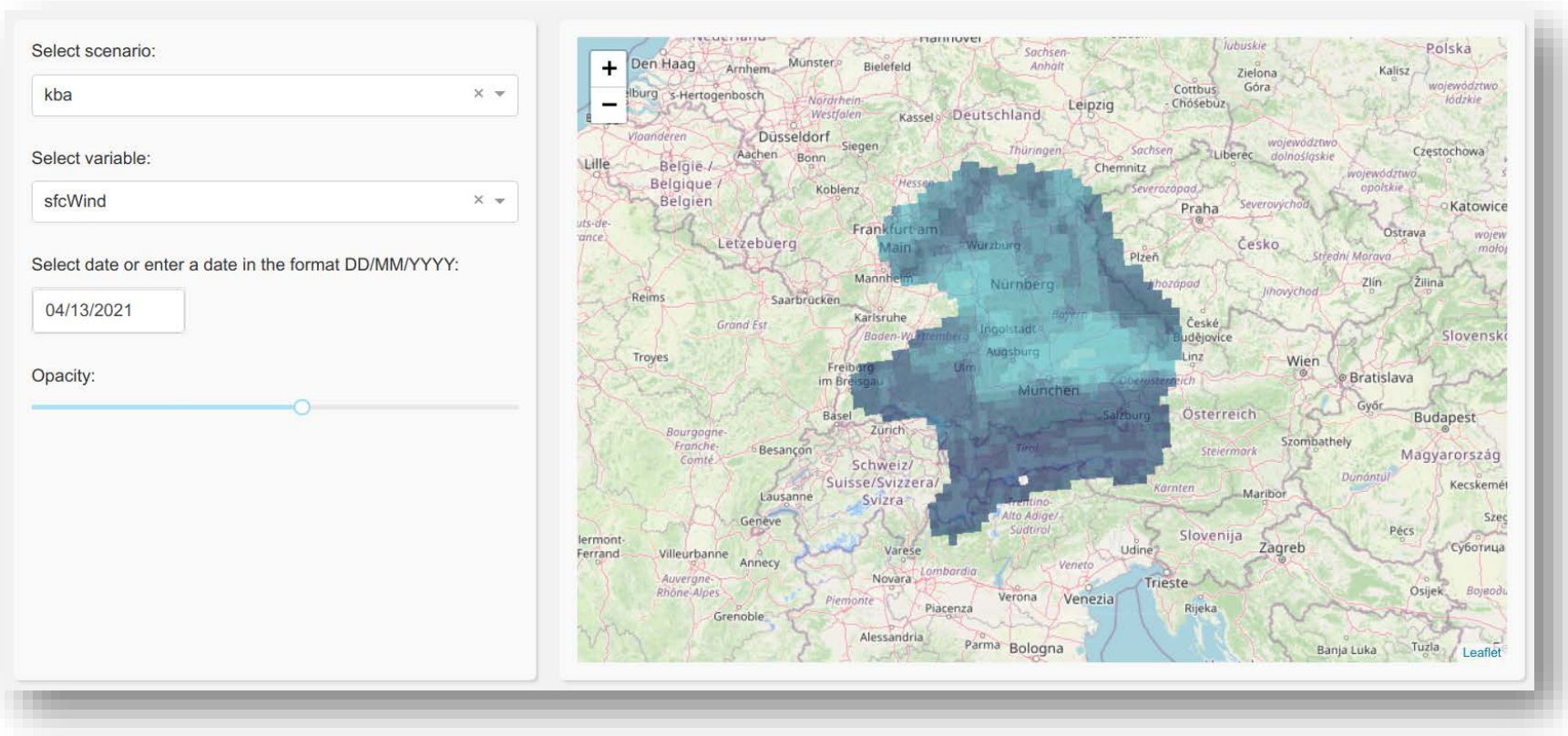
Ein beispielhafter Anwendungsfall

- Demonstrator für georeferenzierte Daten im netcdf Format
- Datenzugriff über eine Web-Oberfläche
- Darstellung auf einer web-basierten Karte
- Anwendung einfacher statistischer Analysewerkzeuge
- Räumlich-Zeitliches Zuschneiden des Datensatz vor dem Herunterladen

WebViewer

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

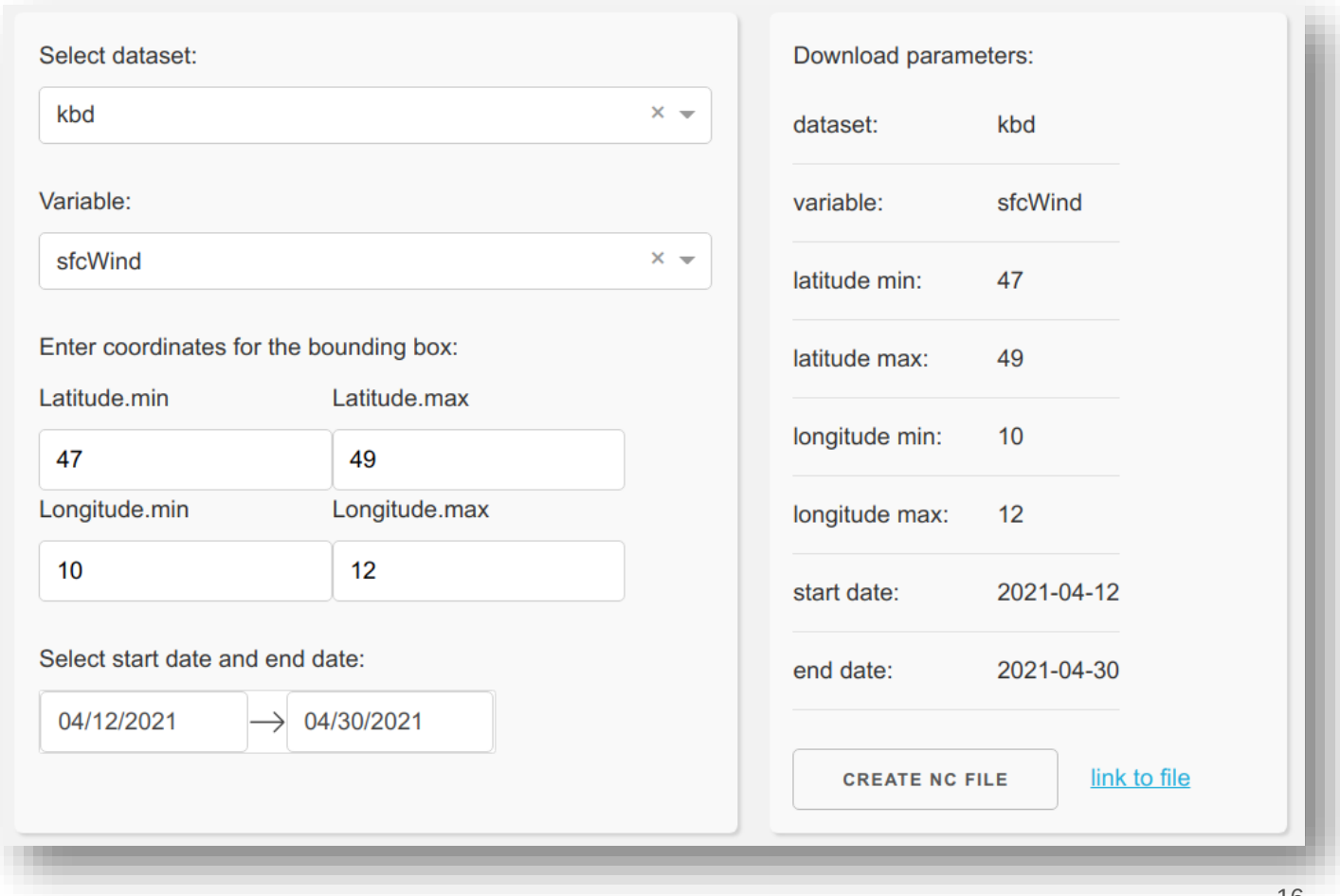
- Auswahl von Datensatz, Variable und Tag
- Darstellung auf einer Karte



Download ClimEx-Daten

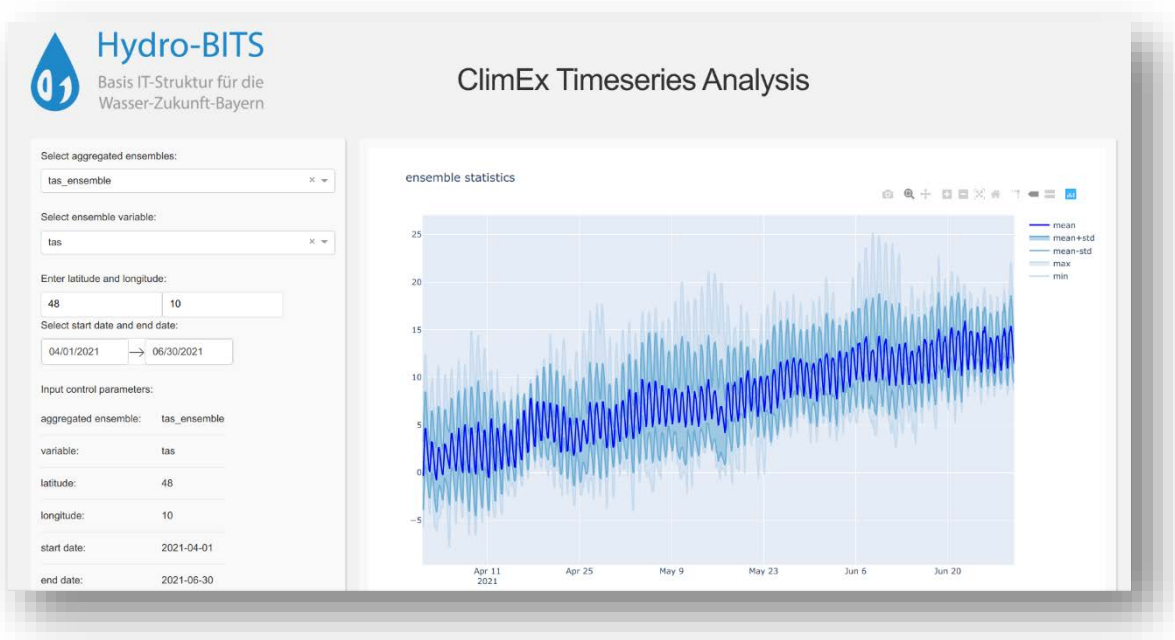
Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

- Auswahl von Datensatzen
- Zuschneiden auf ein Rechteck
- Einschränken des Zeitbereiches
- Bereitstellen einer netcdf Datei



Statistische Auswertung

- Auswahl einer Variable an einem Punkt (Breitengrad, Längengrad)
- Einschränken des Zeitbereiches
- Berechnung und Anzeige von Mittelwert, Standardabweichung der Ensembles in einem Graphen



Statistische Auswertung

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

Select variable:

tas_ensemble

Latitude: 48.1

Longitude: 10.1

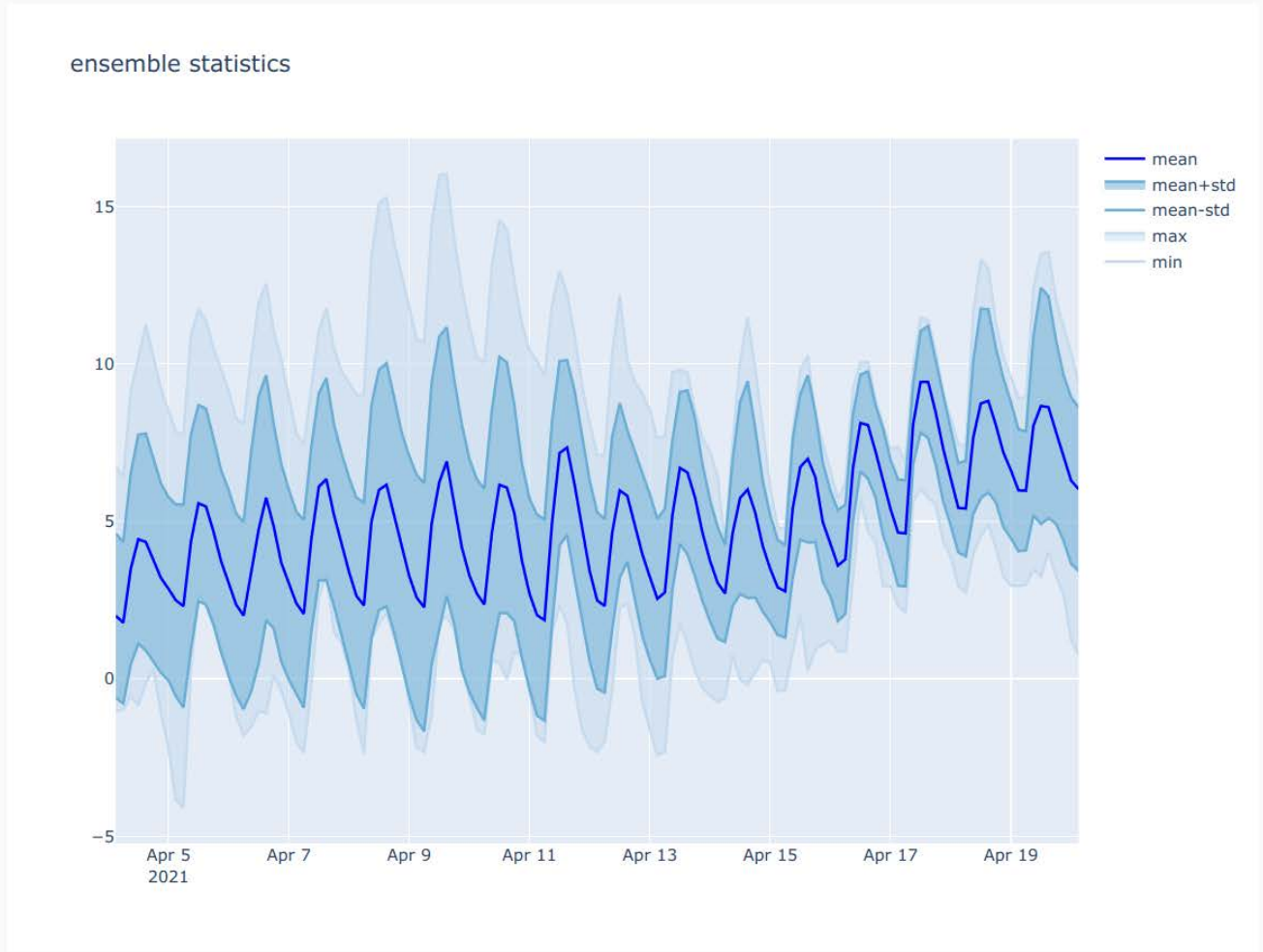
Select start date and end date:

04/04/2021 → 04/20/2021

Input control parameters:

variable:	tas_ensemble
latitude:	48.1
longitude:	10.1
start date:	2021-04-04
end date:	2021-04-20

PLOT GRAPH



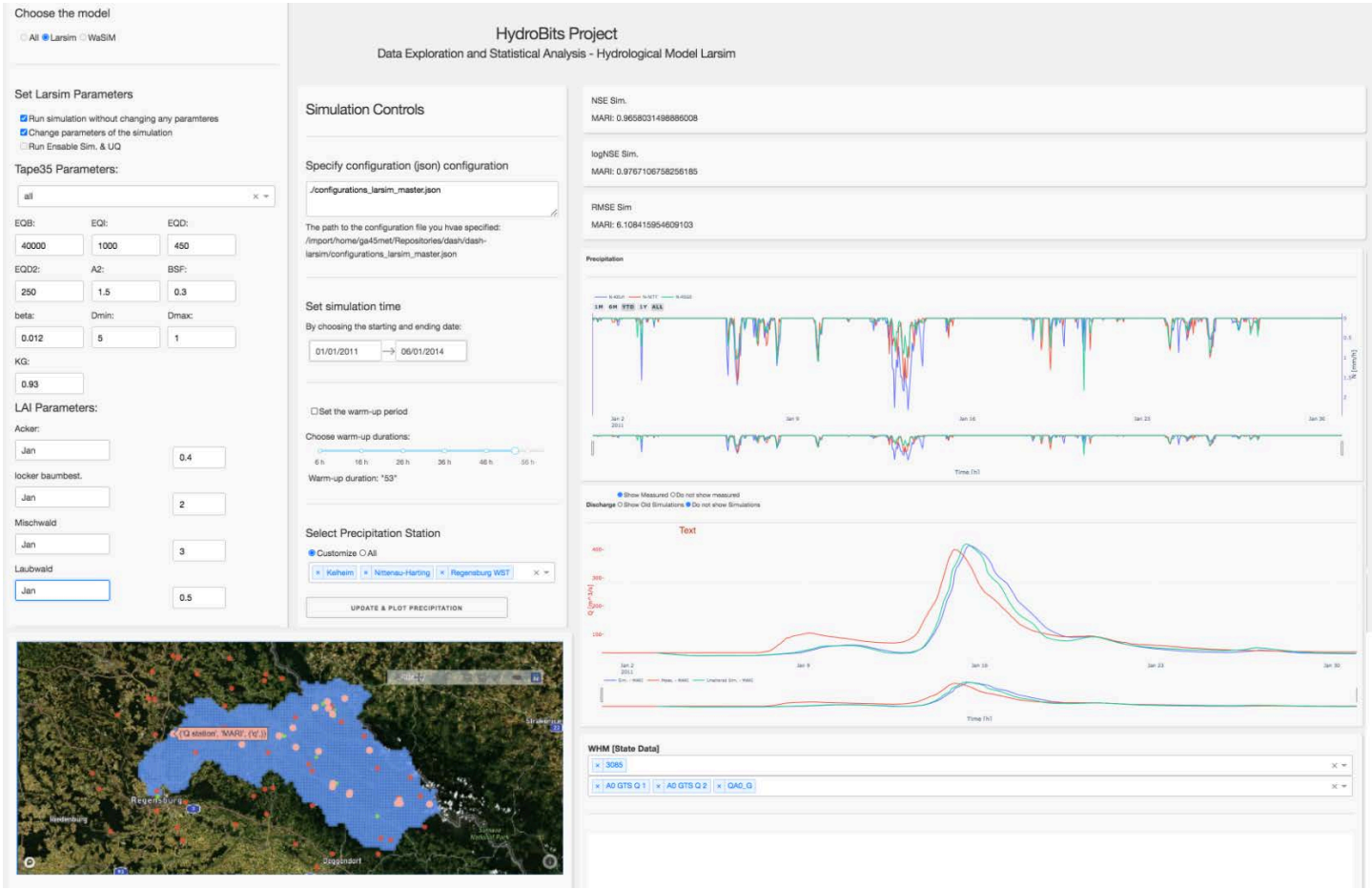
JupyterHub/Dashboard for Model Simulation(s)

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

The idea behind JupyterHub - an 'easy' way to:

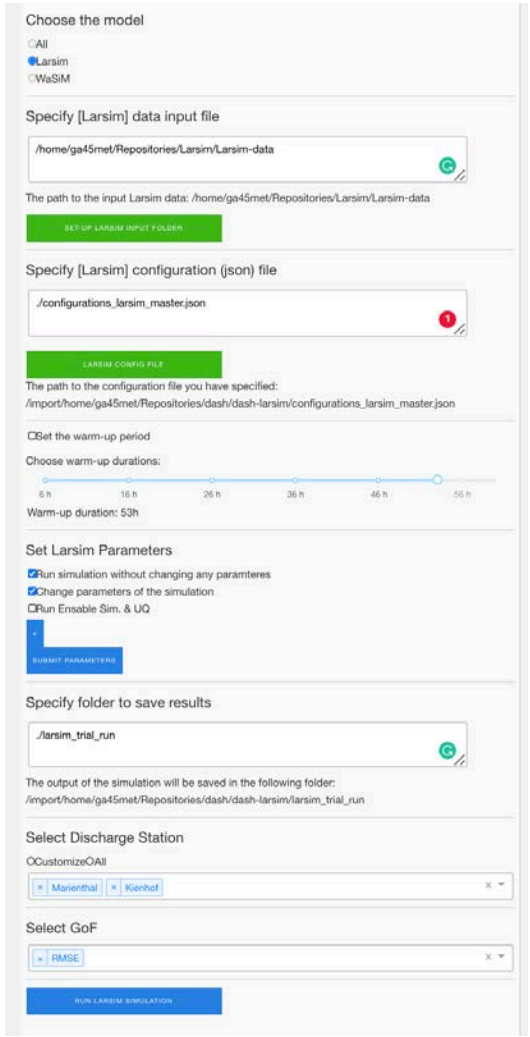
- Collect the code and **simulation/processing pipelines** among researchers, developers, etc.
- Share execution **environments** as well to facilitate the easy execution of the pipeline and reproduction of the results
- Share modest **datasets** necessary

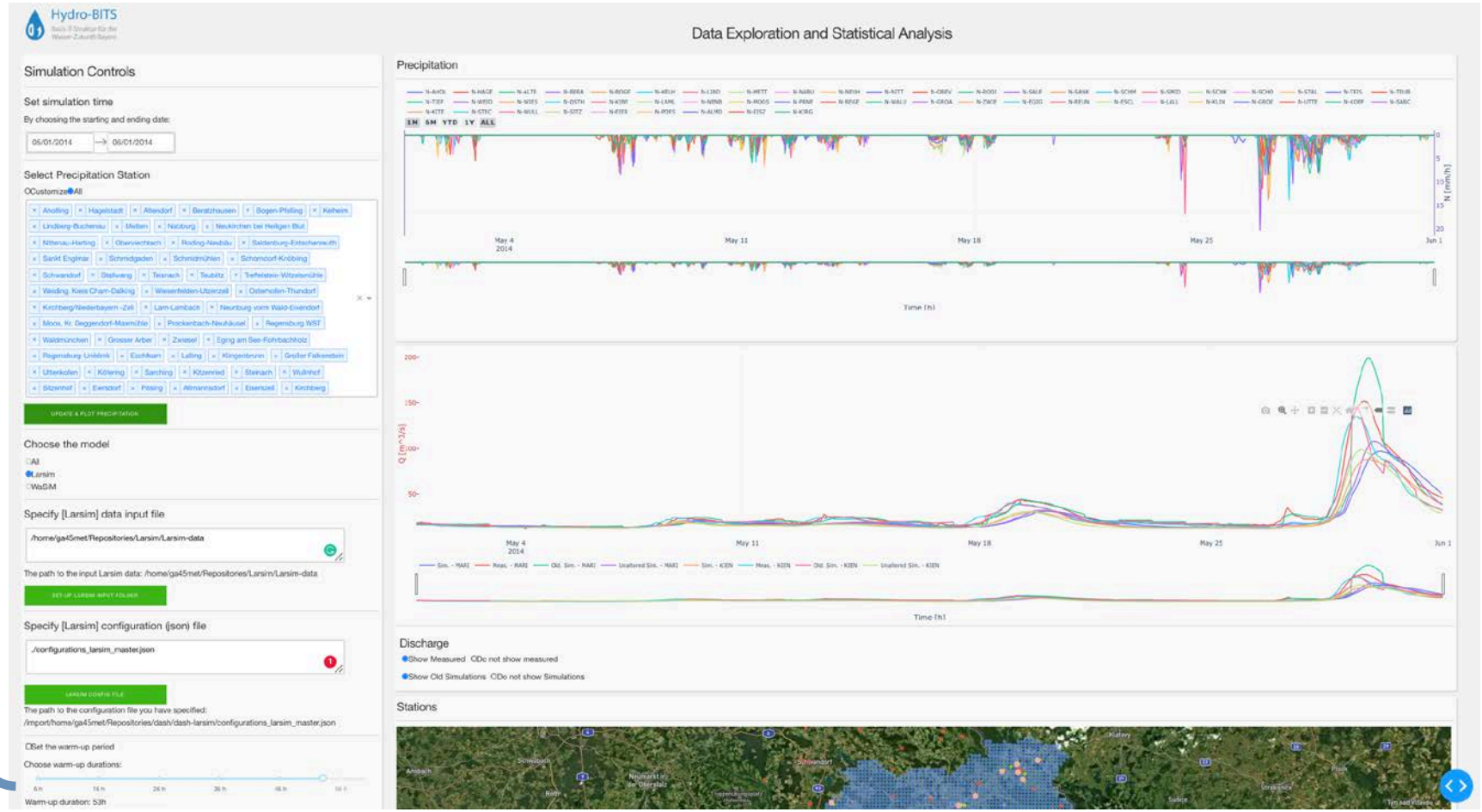
for the execution of the pipelines.



JupyterHub/Dashboard for Model Simulation(s)

Example: Automatic and adaptable Larsim simulation runs; featuring: change of parameters, computation of GoF, inspect the input/forcing and output data, etc.



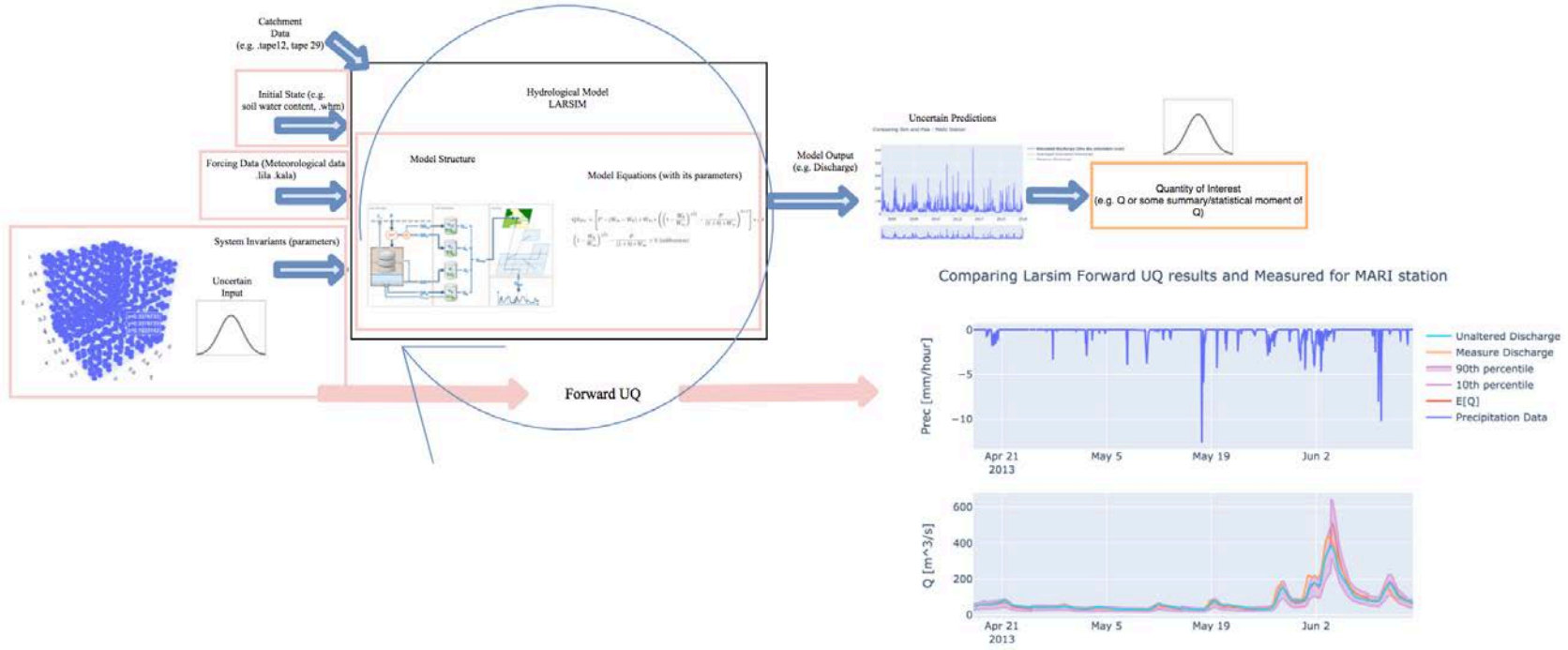


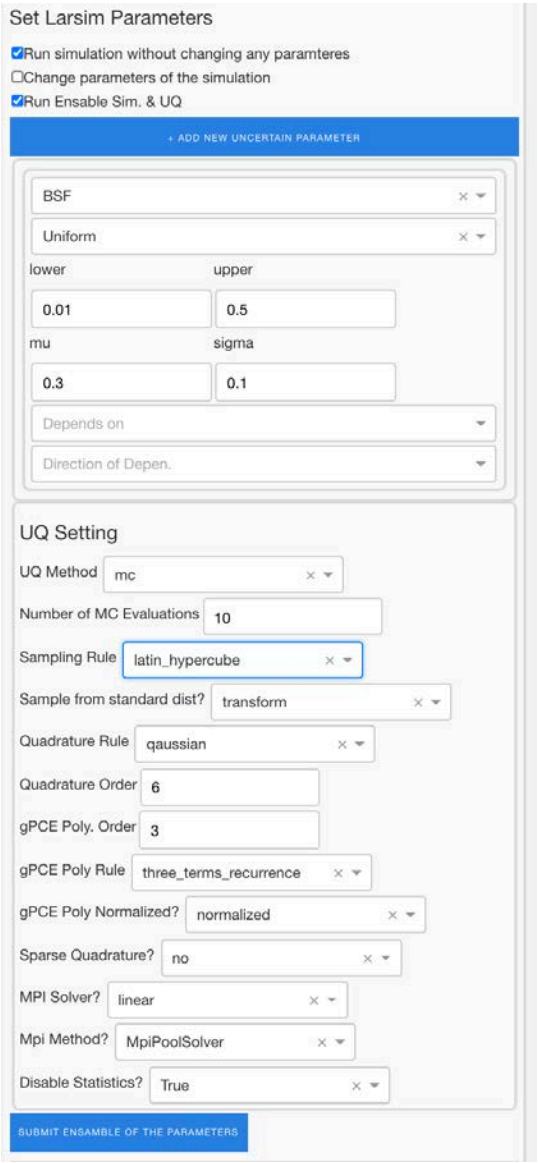
JupyterHub/Dashboard for forward Uncertainty Quantification (UQ)

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

- One of the options in the dashboard:
  - **Account for the variability in the output data/QoI due to the uncertainty in the model parameters in an ensemble manner**

  - In the background, relying on the algorithms and libraries for forward uncertainty quantification and global time-varying sensitivity analysis





Nächste Schritte

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

1) Feedback der WaZuBay-Projekte:

- Welche Funktionalitäten sind für die Projekte wünschenswert?
- Für welche Daten/ Workflows ist die Datenaustauschplattform sinnvoll?

2) Workshops:

- Detaillierte Diskussion zu spezifischen Themen: Datenaustauschplattform, Modelle, UQ/ Kalibrierung, Anwendungsmöglichkeiten am LfU

3) Implementierung:

- Fokussierung auf die Implementierung eine Untermenge von Services/ Funktionalitäten bis zu Projektende

Geplante IT-Workshops

Hydro-BITS

Basis IT-Struktur für die Wasser-Zukunft-Bayern

WaZuBay-Datenaustauschplattform

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

- Datenbereitstellung durch Behörden: Zugriffsrechte, Behandlung prozessierter Daten aus Projekten, Vorgehen bei Fehlerbereinigung
- Bestandsaufnahme existierender/ prospektiver Datenbestände in WaZuBay, welche über eine Datenaustauschplattform geteilt/ analysiert werden sollen
- Überblick über Standardisierung und Datenformate
- Verfügbarmachung von Prozessierungsworkflows und entsprechenden Datenprodukten zwischen verschiedenen Projekten und LfU
- Welche Datenanalyse- und Visualisierungsmöglichkeiten werden von den einzelnen Projektpartnern benötigt?

Hydrologische/ Hydrodynamische Modelle

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

- Bestandsaufnahme verwendeter Modelle, Modellregionen und zugehöriger IT-relevanter Aspekte (Rechenaufwand, Versionen, technische Probleme)
- Bestandsaufnahme verwendeter/ gewünschter Prä- und Postprozessierungs-werkzeuge. Welche Möglichkeiten und Wünsche bestehen für deren Zusammenführung und Wiederverwendung?
- Mögliche gemeinsame Modellweiterentwicklungen, z.B. in Bezug auf WaSiM

Unsicherheitsanalyse & Kalibrierung

- Überblick über bisherigen Arbeiten zu UQ in Hydro-BITS
- Bestandsaufnahme über Verwendung von UQ/ Kalibrierung in WaZuBay: Welche Algorithmen, welche Modelle, welche Parameter, welche Software?
- Welche Werkzeuge sind bereits verfügbar (z.B. für Larsim)?
- Welche UQ/Kalibierungs-Methoden und Werkzeuge sollen in Zukunft in WaZuBay verwendet werden?
- Wie könnte eine mögliche Zusammenführung und Erweiterung der verfügbaren Tools umgesetzt werden?

Anwendungsmöglichkeiten am LfU

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

- Weiterverwendung projektbezogener Daten am LfU
- Welche Datenzugriffs-/ Analysemöglichkeiten sind in der Praxis relevant?
- Wie werden Modelländerungen/ neue Modelle und Tools (z.B. im Bezug auf UQ) in die Praxis überführt?

Organisatorisches

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

- Weitere Vorschläge bzgl. der Inhalte der Workshops?
- Gibt es weitere Themen für die Workshopreihe?
- Zeitliche Planung:
  - Fixer Termin z.B. im 1-2 wöchentlichen Turnus?
  - Individuelle Terminfindung je Workshop?

Extra Slides

Hydro-BITS

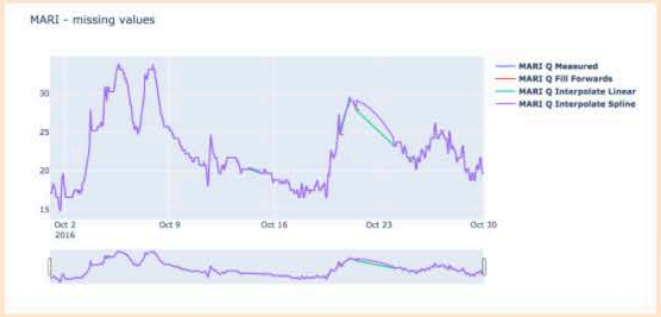
Basis IT-Struktur für die Wasser-Zukunft-Bayern

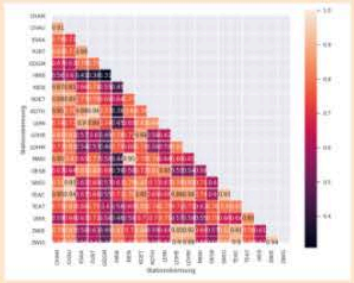
JupyterHub for Larsim

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

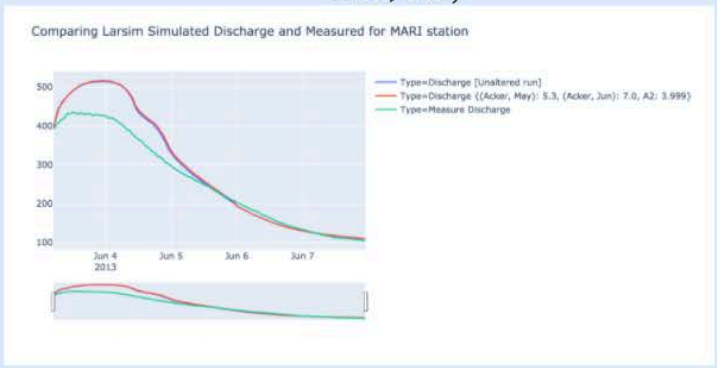
Pipelines in the form of Jupyter Notebooks for different purposes, e.g., processing, inspection, and statistical analysis of the input and output data

Processing, inspection and statistical analysis of the input and output data

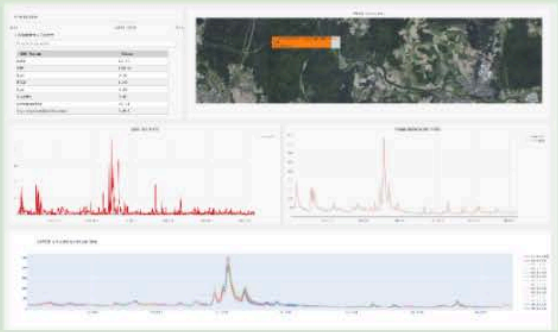




Automatic and adaptable Larsim simulation runs (featuring change of parameters, computation of GoF, etc.)



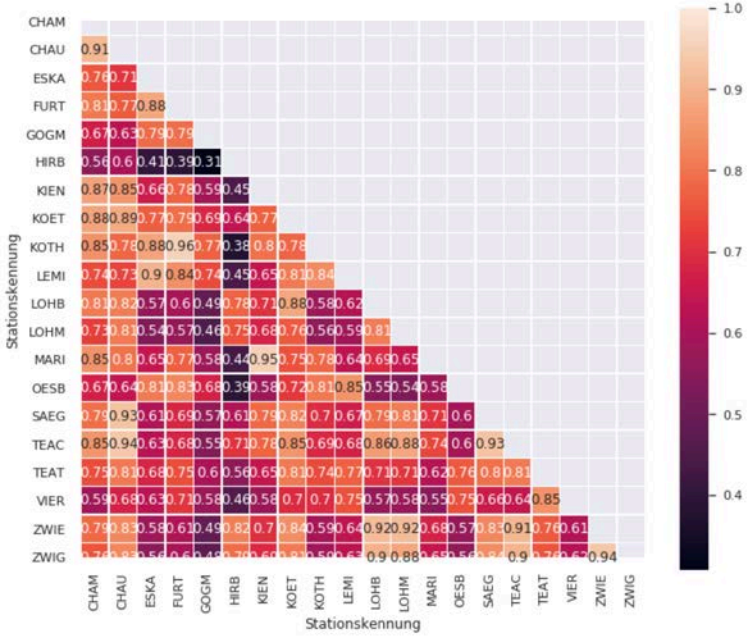
Detailed analysis of the Larsim simulation results (processing and visualization)

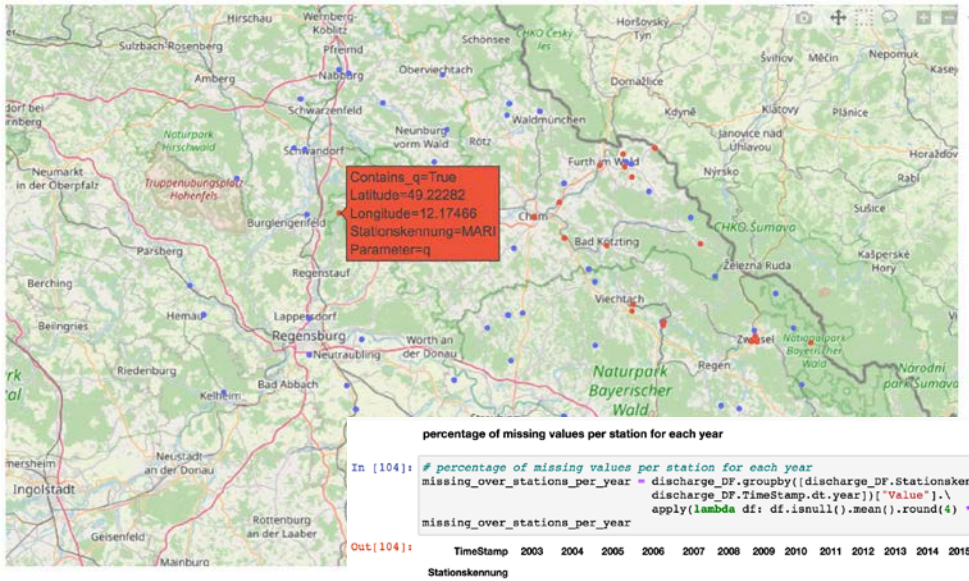


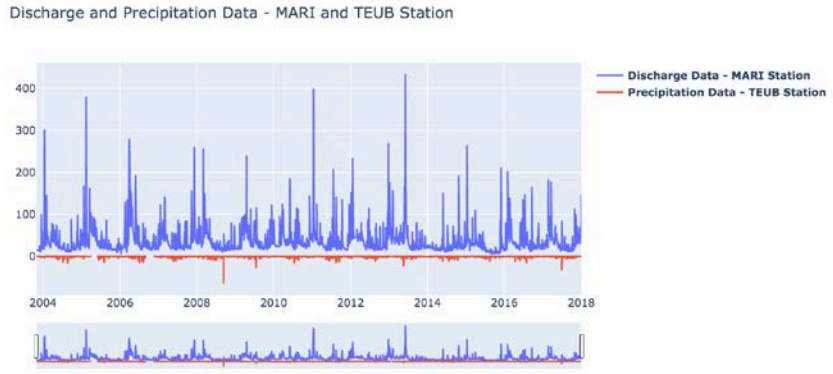
JupyterHub for Larsim

Hydro-BITS
Basis IT-Struktur für die Wasser-Zukunft-Bayern

Pipelines in the form of Jupyter Notebooks for different purposes, e.g., processing, inspection, and statistical analysis of the input and output data







percentage of missing values per station for each year

```
In [104]: # percentage of missing values per station for each year
missing_over_stations_per_year = discharge_DF.groupby([discharge_DF.Stationskennung,\
                                 discharge_DF.TimeStamp.dt.year])["Value"].\
                                 apply(lambda df: df.isnull().mean().round(4) * 100).unstack()
missing_over_stations_per_year
```

Out[104]:

TimeStamp	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Stationskennung																
CHAM	100.0	100.00	100.00	100.00	100.00	100.0	88.88	0.00	0.09	0.00	0.00	5.66	0.00	0.00	0.00	0.0
CHAU	0.0	2.81	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	7.12	0.0
ESKA	100.0	100.00	100.00	100.00	37.19	0.0	0.00	0.00	0.00	0.59	1.60	0.00	0.00	0.00	0.00	0.0
FURT	0.0	0.00	0.00	0.00	0.00	0.0	0.00	3.17	0.00	0.00	0.00	0.00	1.00	0.14	0.25	0.0
GOGM	0.0	0.00	0.00	0.00	0.00	0.0	0.00	3.55	14.29	0.00	0.00	0.00	0.00	36.54	18.89	0.0
HIRB	0.0	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.27	0.36	16.05	43.66	100.0
KIEN	0.0	0.00	0.00	0.00	0.00	0.0	0.00	0.35	0.80	0.82	0.00	0.00	0.00	0.00	0.00	0.0
KOET	0.0	0.00	0.00	0.00	0.00	0.0	0.00	8.52	0.00	0.00	0.00	0.00	0.00	0.14	0.00	0.0
KOTH	0.0	0.00	2.93	0.00	0.00	0.0	0.00	0.17	0.00	0.00	0.39	0.00	0.00	9.93	0.00	0.0
LEMI	0.0	0.00	3.00	5.06	0.00	0.0	0.00	0.00	0.38	4.14	0.00	0.00	0.00	0.00	0.00	0.0
LOHB	0.0	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	6.77	0.68	0.00	0.0
LOHM	0.0	0.00	0.00	8.89	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.03	0.0
MARI	0.0	0.00	0.00	2.88	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.42	0.00	2.32	0.19	0.0
OESB	0.0	0.00	0.00	5.39	0.00	0.0	0.00	0.00	0.86	7.02	0.00	0.00	0.00	0.00	0.00	0.0
SAEG	0.0	0.00	3.56	7.09	0.00	0.0	0.00	0.00	0.08	0.00	2.67	0.26	0.00	0.00	0.00	0.0
TEAC	0.0	0.00	3.55	6.99	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.27	10.14	0.0
TEAT	0.0	0.00	3.55	4.78	0.00	0.0	0.00	0.00	0.00	0.36	0.00	0.00	0.00	0.00	0.00	0.0
VIER	0.0	0.00	0.00	7.80	0.00	0.0	0.00	4.38	2.56	2.82	0.00	0.00	0.00	0.00	0.00	0.0
ZWIE	0.0	0.00	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.48	0.0
ZWIG	100.0	37.17	1.64	18.64	0.00	0.0	0.64	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0