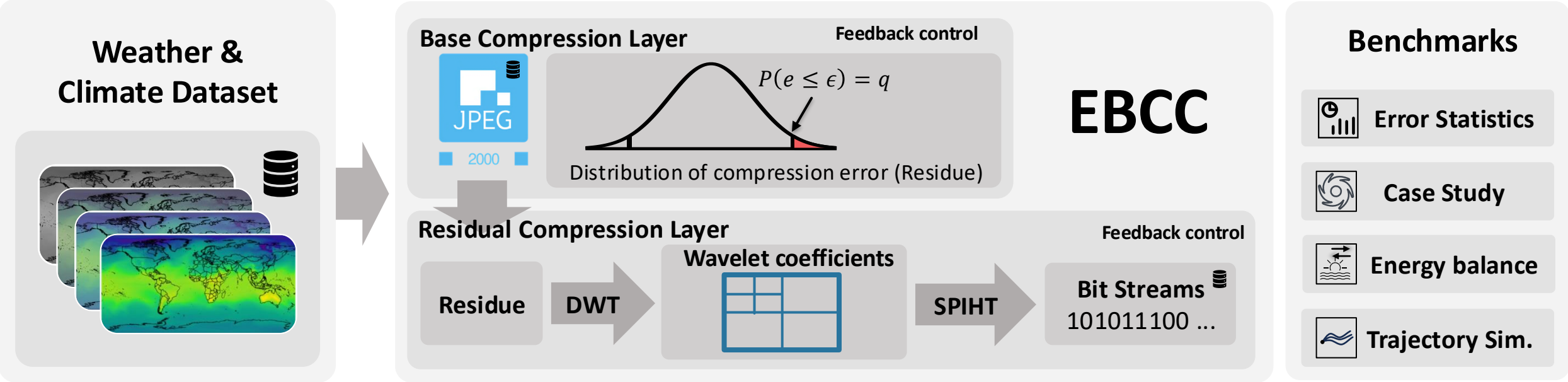


EBCC: Error Bounded Compressor for Climate data

LANGWEN HUANG, LUIGI FUSCO, FLORIAN SCHEIDL, JAN ZIBELL, MICHAEL ARMAND
SPRENGER, SEBASTIAN SCHEMM, TORSTEN HOEFLE

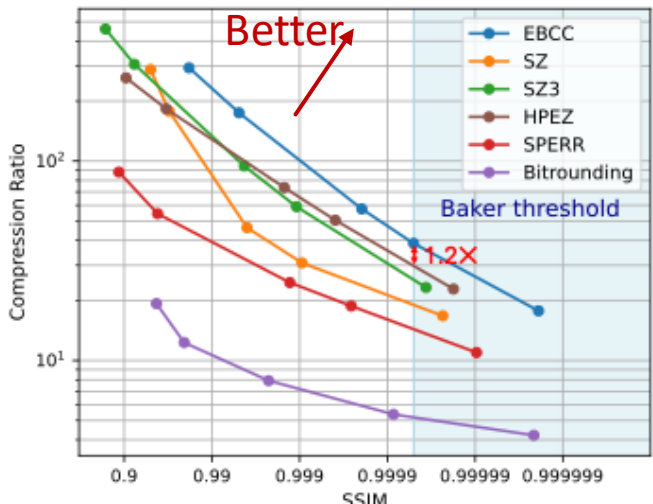
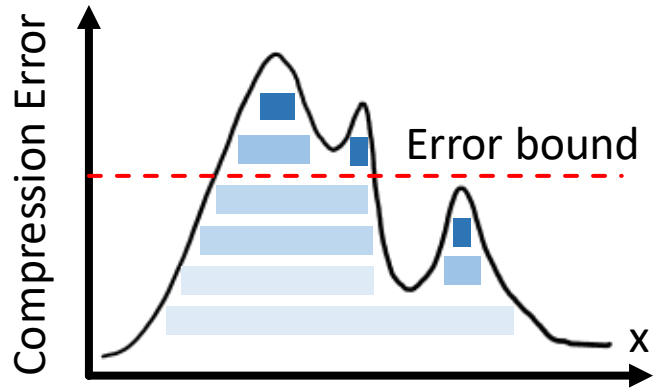


- ### Benchmarks
-  **Error Statistics**
 -  **Case Study**
 -  **Energy balance**
 -  **Trajectory Sim.**

SPIHT compressed significant bits of wavelet coefficients

MSB
↓
LSB

011100000101110000011000011011100
111010101110011001000010101100100
111100110010101111111100110111111
001111100111010111100101101010101
11001000010100101011111111100



Integration with existing toolchains

Paper



GitHub



github.com/spcl/ebcc

Implementation

- Pure C
- Efficient SPIHT implementation

Compression with oneliner

- HDF5/netCDF: filter API
- Zarr: numcodecs compression plugin
- CDO


```
cdo --filter FILTER_ID, `python filter_wrapper.py --base_cr=1000 --height=721 --width=1440` copy temperature.nc comp.nc
```
- Xarray

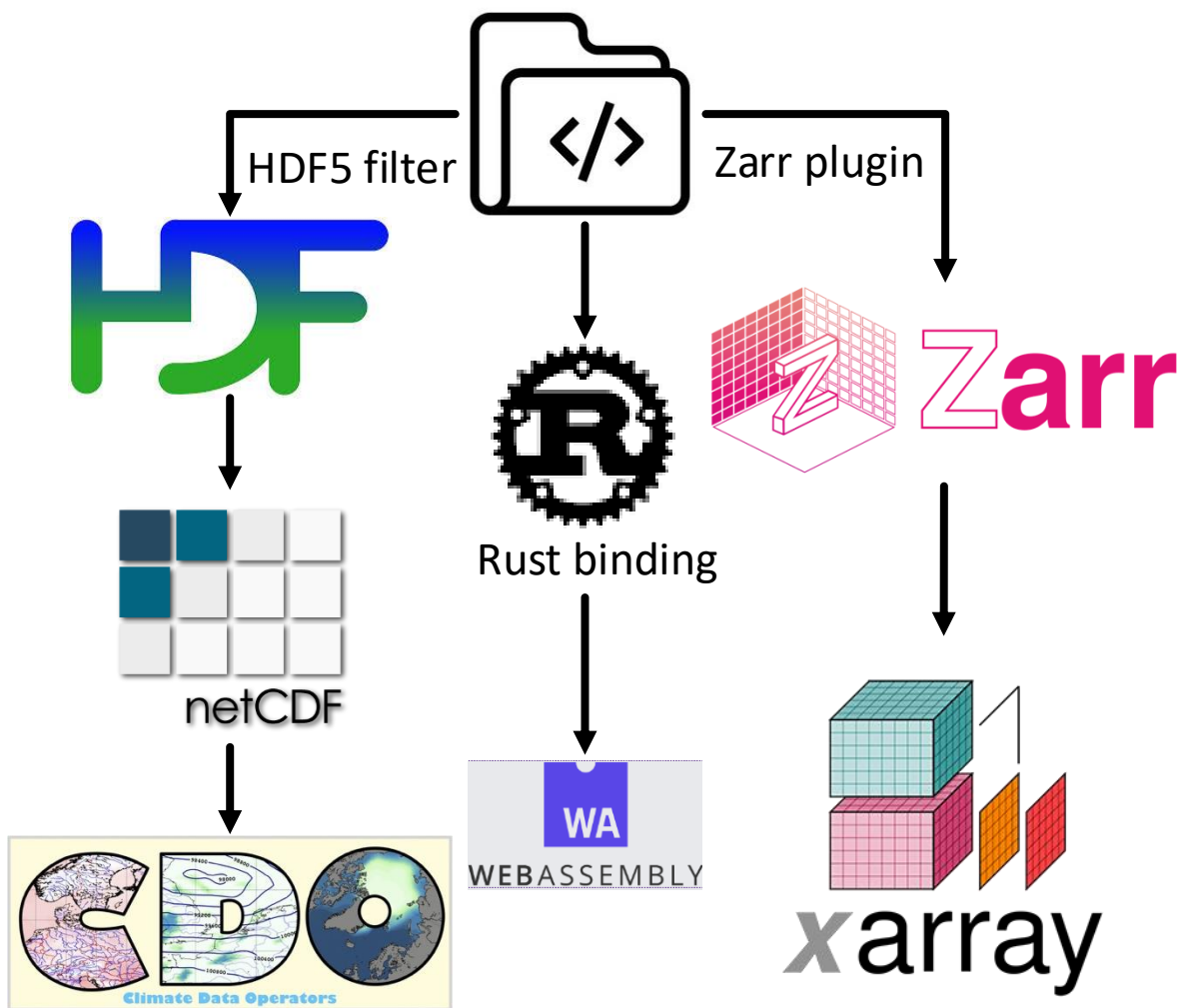

```
ds.to_zarr("comp.zarr", encoding={"foo": {"compressor": J2KFilter(JP2SPWV_FilterOpts(base_cr=1000,height=721,width=1440).hdf_filter_opts)})
```

Transparent decompression

HDF5_PLUGIN_PATH=<path/to/filter>

- HDF5: `h5dump -d /temperature comp.nc`
- netCDF: `ncdump -v temperature comp.nc`
- CDO: `cdo info comp.nc`
- Zarr: `zarr.open('comp.zarr', mode='r')`
- Xarray: `xarray.open_zarr('comp.zarr')`

Standalone C library



Appendix

Installation (HDF Filter, Zarr, Xarray integration)

Instructions can be found in: <https://github.com/spcl/EBCC>

- Go to <https://github.com/spcl/EBCC/releases> and download whl file that matches your platform.
- Install the downloaded whl package via `pip install ebcc-*.whl`

Support of non-finite values (NaN/Inf): EBCC expects NaN/Inf values to be prefilled.

Point-wise absolute/relative error bound: EBCC only supports chunk-wise absolute/relative error bound.

Supported datatypes: float32, other data types have to be casted into float32.

Chunk-size requirement: #dimension ≥ 2 , last two dimensions $\geq (32, 32)$, values spanning last two dimensions should be spatially correlated.