

Real Information Status

Future directions

How can the cutoff bit recommended by the RIC be translated into a parameter setting such as relative error for a given compression algorithm?

We would like to further consider the question of how we can translate the cutoff bit recommended by the RIC to a parameter setting for a given compression algorithm. We want to more thoroughly explore the connection between the RIC and other quantities of interest such as the DSSIM as described in Chapter \ref{chapTwo}. Previous work in \cite{klower} performs a brute-force approach to matching the RIC bit cutoff with the median error obtained by the ZFP compression algorithm. This is related to the approach shown in \ref{FLNS}, where multiple compression algorithms and settings are applied, and the parameters that lead to median errors near the location of the bit cutoff line are a ``match" for that RIC cutoff. We will explore the parameter settings that align with the RIC and the corresponding levels of compression that can be obtained.

Future directions

What kinds of artifacts are produced by different compressors?

Another topic we want to investigate is to understand what kind of artifacts are produced by different compressors. So far we have looked at the artifacts introduced by ZFP and BitGrooming, but we will also compress data using SPERR \cite{sperr} and SZ \cite{sheng2016}, two other often used transform-based lossy compressors for climate data. We would answer the question of what is causing these artifacts, do they behave in similar ways across compressors and error modes including the precision and absolute error mode in ZFP (of which we only example the precision mode so far), and can adjustments to the compressors be made to remove these spatial artifacts?

Future directions

What are the shortcomings of the RIC, when do they apply, and how can the RIC be improved?

Finally, there are some situations in which the RIC is imperfect. The same numerical values stored in different ways, such as with exponent values that differ from neighboring points, causes alterations in the computed RIC. This will affect data that spans a large range of exponent values in a small spatial area, such as PRECT (precipitation rate) data. Adjustments to the process of computing RIC can be made to account for these shortcomings. Additionally, the RIC as defined in \cite{klower} operates 1-dimensionally, where the dataset is flattened and neighboring bits are considered to be only the bits adjacent in the resulting bit stream. The measure could be extended into two dimensions to more accurately capture the notion of spatial structure.