

# File IO Documentation

## 1 Introduction

The File IO portion of the DASH code is designed to quickly funnel gridded climate data into a format optimized for use with dash, which I refer to as gridfiles. This format permits partial loading and partial saving to allow efficient use of computational memory. All formatting is handled by utility functions, so a user does not need to worry about the format. The user functions are designed to mimic several basic netCDF operators, and preserve dimensional metadata.

## 2 User Responsibilities

1. Converting data to a Matlab array.
2. Knowing the name and order of the dimensions in this array.
3. Providing metadata for each index of each data dimension. NaN metadata are permitted but strongly discouraged.

## 3 User Functions

The following is a summary of file IO functions intended for the user:

**buildMetadata:** Converts metadata to a structure optimized for use with dash.

**newGridfile:** Creates a new gridded data file.

**extendGridfile:** Writes data to a gridfile and extends the length of a specified data dimension.

**indexGridfile:** Writes data to specified existing indices in a grid file.

**fillGridfile:** Overwrites specified indices with NaN.

**deleteGridfile:** Deletes specified indices in a gridfile.

**metaGridfile:** Returns metadata for a gridfile.

## 4 Dimension Names

A list of all recognized dimension name IDs is specified in the **getKnownIDs** function. By default, this list consists of lon, lat, lev, time, and run. Essentially, the x, y, z, time, and model ensemble coordinates. Gridded data is not required to have all listed dimensions, but all gridded data dimensions must exist within this list. The recognized dimensions may be renamed by editing the list, and new data dimensions may be created by adding them to the list. Defining an exhaustive list of all possible dimensions is recommended before building gridfiles.