DHM product specification v1.0.0

Release 1.0.0

Danish Agency for Data Supply and Efficiency

Oct 27, 2020

CONTENTS:

CHAPTER

ONE

ABOUT THIS SPECIFICATION

1.1 Scope

This document serves to authoritatively specify the technical details of the core data products of the DHM elevation model.

Except where noted, this document specifies normative requirements for the publicly available DHM data products, starting from the 2018 LiDAR data collection. The released DHM data products will include pre-2018 LiDAR data at least until data from the 2022 collection season is released. While the specification aims to cover this pre-2018 data to the highest extent possible, it is only possible to do so on a best-effort basis.

1.2 Nomenclature

Danmarks Højdemodel, abbreviated **DHM**, is the collective name for the nationwide LiDAR-based elevation model data products produced by the Danish Agency for Data Supply and Efficiency (Styrelsen for Dataforsyning og Effektivisering, SDFE). DHM is unofficially translated into English as the "Danish Elevation Model" or "DK-DEM".

1.3 Normative references

DHM relies on the following data formats for distribution:

- LAS specification 1.4, specified by American Society for Photogrammetry and Remote Sensing (ASPRS)
- The LAZ format produced by the LASzip software
- TIFF revision 6.0, specified by Adobe Systems Incorporated
- GeoTIFF version 1.1, specified by Open Geospatial Consortium (OGC)

Note: While the LAZ format does not appear to have a publicly available formal specification as of June 2020, the source code of the LASzip reference implementation may in practice serve as the LAZ specification.

DHM relies on the following specifications and datasets for georeferencing and tiling:

- The EPSG Geodetic Parameter Dataset, maintained by the International Association of Oil & Gas Producers (IOGP)
- Det danske Kvadratnet, specified by Statistics Denmark

DHM relies on supporting datasets specified in:

- GeoDanmark Specifikation 6.0 by GeoDanmark

1.4 Stability and versioning

This specification follows Semantic Versioning 2.0.0. For the purpose of applying Semantic Versioning, the DHM data products shall be considered "software", and the normative parts of the DHM product specification shall serve as its "public API".

New major versions of this specification may be released without further notice.

CHAPTER

TWO

DATA PRODUCTS

The fundamental DHM data products are:

- **DHM/Punktsky** (pointcloud)
- **DHM/Terræn** (DTM)
- **DHM/Overflade** (DSM)
- **DHM/Højdekurver** (contours)
- **DHM/Oprindelse** (data origin)
- **DHM/Korrektion** (data corrections)

2.1 DHM/Punktsky

DHM/Punktsky is a pointcloud dataset, derived from internal master pointcloud data and provided in units of tiles.

2.2 DEMs

DHM provides two raster-based digital elevation models (DEMs), **DHM/Terræn** and **DHM/Overflade**, in units of tiles.

The DEMs are computed on a per-tile basis from internal master pointcloud data. Each tile is produced using a Delaunay triangulation of a particular filtered subset of the points located within the tile extent plus a spatial buffer. The amount of spatial buffering is unspecified.

Very long triangles may be removed from the triangulation prior to rasterization. The precise criteria for triangle removal are unspecified.

The Delaunay triangulation is rasterized by point sampling the triangles in the pixel centers. The resulting raster is subject to raster-based postprocessing in areas of water bodies, see *Water bodies*.

2.2.1 DHM/Terræn

DHM/Terræn is a digital terrain model (DTM). It is based on the following point classes from the pointcloud:

- 2 (ground)
- 9 (water)
- 17 (bridge deck)

2.2.2 DHM/Overflade

DHM/Overflade is a digital surface model (DSM). It is based on the following point classes from the pointcloud:

- 2 (ground)
- 3 (low vegetation)
- 4 (medium vegetation)
- 5 (high vegetation)
- 6 (building)
- 9 (water)
- 17 (bridge deck)

2.3 DHM/Højdekurver

DHM/Højdekurver is the collective name of several vector datasets of elevation contours. A contour dataset is provided for at least each of the following contour intervals:

- 0.25 m
- 0.5 m
- 2.5 m

The contour datasets are generated from a digital terrain model. The DTM used for contour generation may deviate in unspecified ways from the one provided as DHM/Terræn in order to facilitate e.g. omission of bridges and to apply smoothing. The DTM processing may differ between the individual contour datasets, as the amount of smoothing appropriate for producing aesthetically pleasing contours is dependent on the interval size.

2.4 DHM/Oprindelse

DHM/Oprindelse is a vector dataset of polygons containing metadata describing the origin of the point data in the corresponding DHM/Punktsky. The polygon features contain attributes describing planar and vertical accuracy, production identifier, sensor type and date of data capture.

The polygons are intended as a guide, and may be generalized and/or simplified and should not be used for analysis. If the exact timestamp of a given point is desired, it should be obtained by querying the given point in the pointcloud.

2.5 DHM/Korrektion

DHM/Korrektion is a polygon dataset indicating where and how data has been manipulated. DHM/Korrektion consists of two parts:

- Corrections performed systematically (countrywide), specifically where lake polygons from GeoDanmark has been used to flatten lakes.
- Corrections performed sporadically, e.g. where old point cloud data has been inserted to fill gaps. Such gaps will generally have been caused by seasonal surface water during the data capture.

The attribute "rettelsestype" (type of correction) conveys the action taken within the polygon.

Note: DHM/Korrektion takes precedence over DHM/Oprindelse. DHM/Oprindelse should ideally be overlaid with DHM/Korrektion when displayed in a GIS system.

CHAPTER

THREE

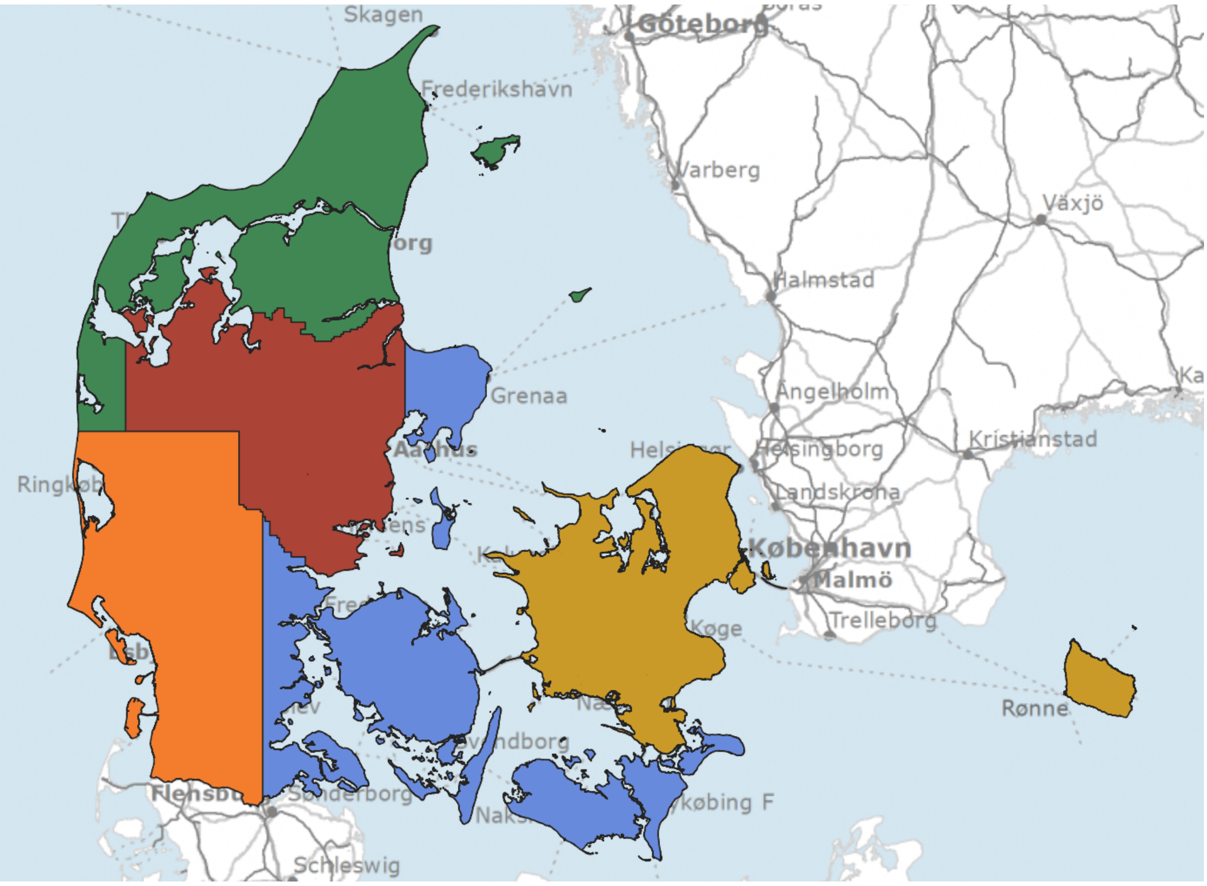
RELEASES

3.1 Regular releases

Starting with the 2018 data collection, the DHM data products are updated once every five years in any given location. In a given year, LiDAR data will be collected for a number of GeoDanmark blocks (cf. *Blocks and seams*) corresponding to approximately 1/5 of the land area of Denmark. Such a set of blocks is referred to as a *lot*.

New versions of the DHM data products are released once a year after SDFE's quality control is completed, incorporating the past year's data collection.

The lots are shown for illustrative purposes below:



The five-year cycle covers the lots in the following order, starting with 2018 as year 1:

- Year 1: blue
- Year 2: brown
- Year 3: orange

- Year 4: red
- Year 5: green

Note: Prior to the 2018 collection, new data were collected for the entire country twice: once in 2007, and once in 2014–15.

3.2 Corrections

Corrections to errors discovered in the DHM products are published by the time of the following regular release. As a general rule, such corrections will largely prioritize ensuring the quality of hydrological data products based on DHM.

3.3 Supplementary data sources

Where deemed relevant, such as for areas of known major construction projects, the DHM products may incorporate data collected outside of the regular cycle, and outside of the usual data collection channels. Such supplementary data will be quality controlled by SDFE on a best-effort basis.

The presence of such supplementary source data will be indicated by the DHM/Oprindelse (data origin) dataset.

CHAPTER

FOUR

FORMATS

4.1 Pointcloud format

The pointcloud data is published as tiles in the LAZ (LASzip-compressed ASPRS LAS) format, version 1.3 or newer.

4.1.1 Classification

Point classification follows a subset of the standard LAS classes, see section *Point classification*.

4.1.2 Coloring

The pointcloud data may or may not include RGB color. If available, such colorization is not subject to any particular requirements regarding geometric or colorimetric quality, completeness or time of data collection, and is provided only as a rough guide for visualization purposes.

4.1.3 Extra bytes

Beginning with the 2018 data collection, the raw pointcloud data from SDFE's suppliers has been delivered with LAS "extra bytes" describing *amplitude* and *pulse width* for each point.

Note: The availability of extra bytes is not stable and is subject to change without notice.

4.2 Raster format

The raster data products are provided in tiles in GeoTIFF format.

4.2.1 Resolution

The DHM raster products are provided with a pixel size that, in georeferenced coordinates, is square and measures 40cm×40cm or smaller.

4.2.2 Data type

The GeoTIFF files store data in 32-bit floating point.

4.2.3 Compression

The GeoTIFF files use lossless DEFLATE compression.

4.3 Vector format

The DHM vector-format data products are guaranteed to be provided in one or more formats readable by OGR, but the format is otherwise unspecified.

CHAPTER

FIVE

TILING AND GEOREFERENCING

5.1 Reference systems

The horizontal reference system in all DHM products is UTM projection, zone 32N, using datum ETRS89 (**EPSG:25832**). The tiling scheme is based on this same reference system.

The vertical reference system in all DHM products is DVR90 (**EPSG:5799**).

The DHM data products may report either EPSG:25832 or the compound reference system **EPSG:7416**.

5.2 Tiling scheme

The DHM data products are divided into 1km×1km tiles.

The tiling follows the one-kilometer "Kvadratnet" convention, i.e. with tile names of the form `1km_NNNN_EEE`, where `NNNN` is northing (in kilometers, truncated to integer) and `EEE` is easting (in kilometers, truncated to integer).

CHAPTER

SIX

GEOMETRY

6.1 Blocks and seams

Since the 2018 collection, data collection and pointcloud delivery from suppliers for DHM is done in units of GeoDanmark blocks. This ensures a manageable size of deliverables and sufficient availability of ground control points within each deliverable. The actual extent of data coverage in deliverables may exceed the nominal block geometries, in part to allow flexibility in choice of seamlines.

Quality control is performed in units of these blocks. The blocks are inspected for internal consistency as part of this QC. The nationwide DHM data products are produced by mosaicking the blocks of pointcloud data approximately according to the block geometries, with the precise seamlines chosen to minimize visible discontinuities, on a best-effort basis as the available data allows.

Note: As of December 2019, the block and seamline geometries are not publicly available for download. A non-normative overview of the blocks can be viewed at SDFE's status maps for the current year's data collection.

Note: Data collected prior to 2018 was handled in a different block scheme. Further details are available in the quality assessment reports for the 2007 version and the 2014/15 version.

6.2 Geometric quality

The geometric quality of the pointcloud is specified in terms of its precision and accuracy.

The guarantees specified below apply only to point data acquired through SDFE's regular data collection cycle. That is, extraordinary/supplementary data (as described in *Supplementary data sources*) are not covered by these guarantees, and such data will generally only be subject to a best-effort quality check before release.

6.2.1 Precision

As a measure of its internal geometric consistency, the precision of the pointcloud describes how closely data from one flight line fits with other flight lines, and how closely measurements of the same object fit.

Vertical precision is determined by comparing points on road surfaces from overlapping flightlines, while horizontal precision is determined by comparing points on gable building roofs.

For data collected in 2018 and later, the precision of the pointcloud is guaranteed to be within the following RMS values, subject to availability of sufficient road surfaces and buildings:

Precision (RMS)	Intra-flight line	Inter-flight line
Vertical	3 cm	5 cm
Horizontal	3 cm	7 cm

For data collected before 2018, the precision is unspecified.

6.2.2 Accuracy

The accuracy of the pointcloud describes how closely the pointcloud as a whole fits to known real-world coordinates. Accuracy determination for the pointcloud relies on ground control points (GCPs) whose accuracy is known to exceed that of the pointcloud. For DHM, the GCPs fall into two distinct categories: GCPs for vertical accuracy, and GCPs for horizontal accuracy.

Vertical accuracy is determined using GCPs shared with other SDFE photogrammetry and remote sensing projects. Typical examples of such GCPs include road markings and manhole covers.

Horizontal accuracy is determined using roof ridges of known rectangular houses with gable roofs.

For data collected in 2018 or later, the DHM pointcloud is guaranteed to fit the available GCPs to a vertical accuracy better than 6 cm RMS (1 sigma), and to a horizontal accuracy better than 15 cm RMS (1 sigma).

For data collected before 2018, the accuracy is unspecified.

CHAPTER

SEVEN

POINT CLASSIFICATION

To produce DHM, SDFE procures a classified pointcloud from a supplier, and that classification is quality controlled by SDFE as part of DHM's QC procedure.

The DHM/Punktsky data product is a pointcloud comprised of points classified as described in this document, and the raster products (DHM/Terræn and DHM/Overflade) are based on a pointcloud with that same classification, filtered to their respective criteria.

The point classes in DHM are generally determined algorithmically using supporting datasets and the geometry of the pointcloud, in combination with human judgement. Due to the tradeoff between cost and classification granularity, the point data in DHM does not use the full gamut of classes specified in the LAS format. **As such, the classification of points is done on a best-effort basis, with no guarantees as to the correctness of the assigned classes.**

7.1 Ground

Points on terrain and other firm, ground-like surfaces such as roads or pavements are assigned LAS class 2 ("Ground").

7.2 Water

Points on water surfaces are assigned LAS class 9 ("Water").

Note: Water bodies are given extensive special treatment when computing the DTM and DSM, see *Water bodies*.

7.3 Vegetation

Points on vegetation are assigned a LAS vegetation class (class 3, 4 or 5). The specific classification is assigned based on estimated height above ground as follows:

- Class 3: "Low Vegetation", up to 0.3 m above ground
- Class 4: "Medium Vegetation", 0.3–2 m above ground
- Class 5: "High Vegetation", more than 2 m above ground

Note: In addition to their use for vegetation, the vegetation classes are used as catch-all classes for unknown surfaces, see section *Other surfaces*.

7.4 Buildings

Points on building surfaces are assigned LAS class 6 ("Building"). The classification of buildings is verified using a snapshot of the GeoDanmark "Bygning" (building) layer. In order to correctly handle unfinished or demolished buildings, any given building in the GeoDanmark dataset may be omitted from DHM's QC and production routines.

Note: Points deemed to fall within building interiors, including ground and vegetation points located within the building footprint, are not included when computing the DTM and DSM. The classification of building interior points in DHM/Punktsky is unspecified.

7.5 Roads

Points on road surfaces, pavements and paths are assigned LAS class 2 ("Ground").

Note: LAS class 11 ("Road Surface") is not used in DHM.

7.6 Railways

Points on railway tracks and trackbeds are assigned LAS class 2 ("Ground"), with the possible exception of rails and other protruding objects, which may be assigned a vegetation class (see *Other surfaces*).

Note: LAS class 10 ("Rail") is not used in DHM.

7.7 Bridges

Points on bridge decks are assigned LAS class 17 ("Bridge Deck"). Other parts of bridges are classified according to the general rules described in this document.

Note: As of 2020, there is no supporting dataset available to substantially assist with bridge classification, and thus bridge classification largely relies on estimates from the available point data.

7.8 Power lines

Points on power lines and associated infrastructure are assigned a class that is not included in the DSM (DHM/Overflade). The target class and precise criteria for power line classification are unspecified, and are primarily aimed at avoiding a conspicuous presence of power lines in the DSM.

The handling of power lines is subject to change without notice.

Note: Since 2020, power line classification has been based on the GeoDanmark "Højspændingsledning" (power line) objects for point data collected in 2018 or later.

Point data collected before 2018 has generally had power lines filtered out using a voxel-based algorithm.

7.9 Construction cranes

For point data collected in 2018 or later, construction cranes are assigned an unspecified LAS class not included in DHM/Overflade (the DSM).

The handling of construction cranes is subject to change without notice.

Note: As of 2020, no supporting dataset is available to assist in construction crane classification. Due to the inherently ephemeral nature of construction cranes, classification of these will generally rely on a best-effort judgement from the available point data.

7.10 Other surfaces

Points believed to correspond to an actual physical surface, but which do not correspond directly to a category otherwise specified in this document, are assigned a vegetation class (class 3, 4 or 5 depending on estimated height above ground). Thus, the vegetation classes are used as a catch-all classification and will not necessarily correspond to vegetation.

7.11 Noise and other artifacts

Points in the point cloud believed to not correspond to a physical surface, such as noise or other scanning artifacts, are assigned a class not among the classes included in the DSM (DHM/Overflade). The specific classification of non-surface points is not quality controlled by SDFE, but will generally fall within one of the noise-like LAS classes (1, 7 or 18).

CHAPTER

EIGHT

WATER BODIES

For the DTM and DSM, **sea**, **lakes**, and **rivers** are burned into the raster.

Lidar data coverage is generally poor on specular surfaces such as water. Various processing is therefore performed to ensure hydrologically plausible water surfaces.

8.1 Sea

Within sea areas, a DVR90 elevation of zero is burned into the raster, except possibly where the raster exceeds a DVR90 elevation of 0.8 meters. This vertical tolerance is intended to help preserve any physical features that may not be included in the supporting vector data.

Sea areas are defined by an extract of GeoDanmark data. As of 2020, this is done through a custom data extract.

8.2 Lakes

Within each lake geometry as defined by the GeoDanmark "Sø" (lake) objects, a constant elevation appropriate for the lake surface is burned into the raster if reasonably possible. If it is deemed impossible to determine an appropriate constant elevation for a given lake object, raster values in that lake area are determined by the usual DTM or DSM interpolation between points.

The data source of these constant elevations is unspecified. The precise policy for deciding whether to burn in a constant elevation is also unspecified.

Note: Previous DHM releases have used various techniques for estimating lake elevation, such as using a best-effort estimate based on the available pointcloud data within buffered GeoDanmark lake objects. In case there were too few points to make this estimate, no constant elevation would be burned in.

8.3 Rivers

To ensure monotonic elevation changes along river courses, the geometry from some GeoDanmark "Vandløb-smidte" (river centerline) objects are buffered and burned into the raster.

Only river objects with both a GeoDanmark width class (`midtebredde`) of at least 2.5 meters and a visible center (`synligVandløbsmidte` attribute equal to true) are burned into the rasters. The elevation values burned into the raster originate directly from the Z coordinates of the GeoDanmark river geometries. The geometries are buffered horizontally by a distance of 3 meters before being burned into the raster.