Glacier Covered Area for the State of Alaska, 1985-2020, Version 1

USER GUIDE

How to Cite These Data

As a condition of using these data, you must include a citation:

Roberts-Pierel, B.M., Kirchner, P.B., Kilbride, J.B., Kennedy, R.E. 2022. *Glacier Covered Area for the State of Alaska, 1985-2020, Version 1*. [Indicate subset used]. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. https://doi.org/10.7265/8esq-w553. [Date Accessed].

FOR QUESTIONS ABOUT THESE DATA, CONTACT NSIDC@NSIDC.ORG

FOR CURRENT INFORMATION, VISIT https://nsidc.org/data/G10040

National Snow and Ice Data Center

TABLE OF CONTENTS

1 DATA DESCRIPTION

1.1 Summary

This data set captures changes in glacier covered area across the state of Alaska for the period 1985 to 2020. The data set includes 18 biannual shapefiles each for overall glacier covered area, supraglacial debris area, and debris-free glacier covered area. Within each shapefile, attributes for the polygons include information from the Randolph Glacier Inventory (RGI Consortium, 2017) such as RGI glacier ID, latitude, longitude, minimum and maximum elevation, slope, and aspect, to name a few. Shapefiles for overall glacier covered area are inclusive of the entire state of Alaska while those for supraglacial debris area and debris-free glacier area do not include the Brooks Range. These data were derived from a time series of Landsat data, supplemented with elevation data from the Alaska 5 m IfSAR digital elevation model (DEM). The spatial extent of the elevation data and the project objectives focused the study area on the state of Alaska, so glacier outlines do not extend into Canada.

1.2 Parameters

The main parameters consist of a time series of area for the state of Alaska for three glacier categories: overall glacier coverage, supraglacial debris coverage, and debris-free glacier coverage.

1.3 File Information

1.3.1 Format

The data are provided in 54 shapefiles for the three glacier categories (18 shapefiles per category): overall glacier covered area, supraglacial debris area, and debris-free glacier area. Each shapefile is composed of four component files:

- .shp – primary geometry file for shapefile features
- .dbf – dBASE table with information for feature attributes
- .prj – coordinate system information for features
- .shx – index file with feature geometry

1.3.2 File Contents

Each shapefile contains the attributes listed in Table 1. Note that attributes marked with an asterisk (*) are metadata that were copied from the Randolph Glacier Inventory (RGI) database for each of the corresponding polygons. These attributes are simply ancillary information used to describe

glaciers in RGI 6.0. All 22 RGI 6.0 attributes are included here. The two main attributes that are applicable to the analysis done for this data set are GN_area and GN_comp_ye.

Table 1. Shapefile Attributes.

Shapefile Attribute	Description
RGIId*	Randolph Glacier Inventory Glacier Id: A 14-character identifier of the form RGIvv-rr.nnnnn, where vv is the version number, rr is the first-order region number and nnnnn is an arbitrary identifying code that is unique within the region.
GLIMSId*	Global Land Ice Measurements from Space Glacier Id: A 14-character identifier in the GLIMS format GxxxxxxEyyyyycoord, where xxxxxx is longitude east of the Greenwich meridian in millidegrees, yyyyy is north or south latitude in millidegrees, and coord is N or S depending on the hemisphere.
BgnDate*	Beginning date from the RGI database of the survey or image from which the RGI outline was taken, in the form yyyymmdd, with missing dates represented by -9999999. **Note**: The RGI date was included here for completeness to indicate when the original RGI outline was from. For the date of the input data used as part of the Glacier Covered Area for the State of Alaska data set, use the GN_comp_ye attribute.
EndDate*	End date of a range of dates in the form yyyymmdd from the RGI database. This value is only populated if the data provider supplied a range of dates. The two codes together (BgnDate and EndDate) give the date range for the RGI outline. Otherwise, this is set to missing (-9999999). **Note**: The RGI date was included here for completeness to indicate when the original RGI outline was from. For the date of the input data used as part of the Glacier Covered Area for the State of Alaska data set, use the GN_comp_ye attribute.
CenLon*	Longitude, in degrees, of a point representing the geometric center of the glacier polygon.
CenLat*	Latitude, in degrees, of a point representing the geometric center of the glacier polygon.
O1Region*	The codes of the first-order region to which the glacier belongs. See Table 2 in the Randolph Glacier Inventory – A Dataset of Global Glacier Outlines Technical Document for a list of these regions.
O2Region*	The codes of the second-order region to which the glacier belongs. See Table 2 in the Randolph Glacier Inventory – A Dataset of Global Glacier Outlines Technical Document for a list of these regions.
Area*	Area of the glacier in km^2 from the RGI database. Note that this is *not* the area calculated during the processing of the Glacier Covered Area for the State of Alaska data set. For that value, see the GN_area attribute.
Zmin*	Minimum elevation (m above sea level) of the glacier.
Zmax*	Maximum elevation (m above sea level) of the glacier.

Shapefile Attribute	Description
Zmed*	Median elevation (m) of the glacier.
Slope*	Mean slope of the glacier surface (deg).
Aspect*	The aspect (orientation) of the glacier surface (deg) presented as an integer azimuth relative to 0° at due north.
Lmax*	Length (m) of the longest surface flowline of the glacier
Status*	The Status attribute flags glaciers that have outlines that await subdivision or are nominal circles: 0 - Glacier or ice cap 1 - Glacier complex 2 - Nominal glacier 9 - Not assigned
Connect*	N/A. This attribute is only applicable to glaciers in Greenland because it describes the connectivity level of a glacier to the Greenland ice sheet.
Form*	Contains information on the form of the ice body: 0 - Glacier 1 - Ice cap 2 - Perennial snowfield 3 - Seasonal snowfield 9 - Not assigned
TermType*	Contains information on terminus type: 0 - Land-terminating 1 - Marine-terminating 2 - Lake-terminating 3 - Dry calving 4 - Regenerated 5 - Shelf-terminating 9 - Not assigned
Surging*	Contains information on evidence for surging: 0 - No evidence 1 - Possible 2 - Probable 3 - Observed 9 - Not assigned
Linkages*	Indicates whether there is a link to mass-balance measurements in the World Glacier Monitoring Service Fluctuations of Glaciers database: 0 - Not in FG 1 - In FG 9 - Not assigned
Name*	Name of the glacier, if available/known.

Shapefile Attribute	Description
GN_area	GlacierCoverNet glacier area in km^2.
GN_comp_ye	GlacierCoverNet composite label year. When image composites are created, a biannual composite includes images from the label year and the previous year, i.e., the year 1986 in the data refers to a medoid (Multi-Dimensional Median) composite (Flood, 2013) of Landsat images from 1985 and 1986. Additional information on compositing is included in section 2.3 Processing Steps.

1.3.3 Directory Structure

All data reside via HTTPS: https://noaadata.apps.nsidc.org/NOAA/G10040/. Within that directory are three subdirectories as described in Table 2.

Table 2. Directory Structure

Directory Name	Description
debris_free_area	Shapefiles of areas classified as free of supraglacial debris for Alaska (excluding the Brooks Range).
overall_area	Shapefiles of all glacier covered area with no distinction with respect to debris for the entire state of Alaska.
supraglacial_debris_area	Shapefiles of supraglacial debris areas for Alaska (excluding the Brooks Range).

1.3.4 Naming Convention

The files are named according to the following convention and as described in Table 3:

`AK_YYYY_AREATYPE.ext`

Table 3. File Naming Convention

Variable	Description
AK	Identifies this file as containing Alaskan glacier outlines.
YYYY	4-digit year of the outlines. This year matches the GN_comp_ye attribute in the data (see Table 1 for description).
AREATYPE	Area type refers to three categories corresponding to the three parameters: debris_free_area, overall_glacier_covered_area, and supraglacial_debris_area
.ext	File extension for the four shapefile components: .shp, .dbf, prj, and .shx

1.3.5 File Size

The complete data set totals approximately 2.6 GB.

1.4 Spatial Information

1.4.1 Coverage

The following are the approximate latitude/longitude bounding coordinates:

Northernmost Latitude: 69° N
Southernmost Latitude: 53° N
Easternmost Longitude: 130° W
Westernmost Longitude: 167° W

1.4.2 Resolution

The vector polygons were created from satellite imagery with a resolution of 30 m. See section 2 Data Acquisition and Processing for full details.

1.4.3 Geolocation

Table 4 provides information for geolocating this data set.

Table 4. Geolocation Details

Geographic coordinate system	NAD83
Projected coordinate system	EPSG:3338 – Alaska Albers
Longitude of true origin	-154
Latitude of true origin	50
Scale factor at longitude of true origin	N/A
Datum	North American Datum 1983
Ellipsoid/spheroid	GRS 1980
Units	meter
False easting	0
False northing	0
EPSG code	3338
PROJ4 string	+proj=aea +lat_1=55 +lat_2=65 +lat_0=50 +lon_0=-154 +x_0=0 +y_0=0 +ellps=GRS80 +datum=NAD83 +units=m +no_defs

1.5 Temporal Information

1.5.1 Coverage and Resolution

The data span 1985 to 2020 in biannual composites (one every two years). Specifically, the data are derived from Landsat images collected between 1985 and 2020. The temporal resolution of the dataset is two years. This temporal resolution is defined by the repeat interval of the medoid composites (Flood, 2013) created from the Landsat archive (see section 2.3 Processing Steps for more information).

2 DATA ACQUISITION AND PROCESSING

Material in Section 2 was provided by Roberts-Pierel and describes the processing that was carried out to create these data.

2.1 Background

This data set was created to fill a need for accurate, high-temporal granularity data of glacier area for the state of Alaska. The temporal component is aimed at improving the ability to analyze periods of change at a scale relevant to the physical processes observed in situ. Emphasis was placed on creating a method that was transferable and updatable without extensive resource-intensive, manual editing of glacier outlines. These objectives were fulfilled by employing a neural network (ResNet101-PSPNet) (Zhang et al., 2020; Zhao et al., 2017) informed by research in computer vision, deep learning (He et al., 2015; Szegedy et al., 2014; Xie et al., 2017), and glacier inventory creation. This neural network is referred to here as GlacierCoverNet (Roberts-Pierel et al., 2022). These data provide high spatial and temporal resolution maps of debris-free ice and debris cover. The data set is intended for use by National Park Service scientists and interpretive staff in support of science-based management decisions, education, and outreach. The dataset is also intended for public research use. Methods used here are adaptable to future work generating glacier inventories globally.

2.2 Acquisition

Primary data were collected from Landsat Collection 1 Tier 1 surface reflectance (SR) from the Landsat 4, 5, 7, and 8 missions (30 m spatial resolution) and the 5-meter Alaska DEMs – USGS National Map 3DEP Downloadable Data Collection (U.S. Geological Survey, 2020). All available Landsat images in the USGS archive for Tier 1 SR over the state of Alaska between 1985 and 2020 collected between July 20 and September 20 of each year were utilized. The Landsat data were accessed on the Google Earth Engine platform (Gorelick et al., 2017) where they are

collocated with the computer resources used to produce this dataset. The late-summer window was selected to minimize conflation of glacier ice and seasonal snow cover, while maximizing the number of possible repeat passes of the satellites.

2.3 Processing Steps

At a high level, the workflow for this project consisted of: 1) Creation of a map of probability of glacier presence to which areas of supraglacial debris were identified and added based on RGI 6.0. This map of debris free and supraglacial debris covered glacier is referred to as the class label, 2) Training a deep neural network (GlacierCoverNet) using optical and topographic predictors and the class label, and 3) Vector post-processing. What follows is a brief overview of these steps. A more complete description can be found in Roberts-Pierel et al. (2022).

1. Acquired and processed Landsat data through the LandTrendr algorithm (Kennedy et al., 2010, 2018) on the Google Earth Engine platform (GEE) (Gorelick et al., 2017) to create landcover maps (Figure 1 – GlacierCoverNet Training).
    a. Created a time series of images with consistent temporal, spatial, and spectral characteristics. Image processing, training, and classification were all conducted at a 30 m spatial resolution. The primary inputs to the LandTrendr algorithm were a time series of medoid composites (Flood, 2013). There were 18 medoid composites with each acting as a representative image for a two-year interval in the study time period.
    b. Calculated five optical indices using the homogenized time series data (output of Step 1a) including: Normalized Difference Snow Index (Hall, 2010), Normalized Difference Vegetation Index, Normalized Burn Ratio (Key & Benson, 1999), Tasseled Cap Wetness, and Tasseled Cap Brightness (Kauth & Thomas, 1976). The output optical imagery time series was 18 five-channel images, each representing a biannual late summer period (July 20th to September 20th) from 1985 to 2020.
    c. Calculated four topographic indices using elevation data from the 5 m Alaska IfSAR product (U.S. Geological Survey, 2020) including: two aspect intensity variables (Kirchner et al., 2014), curvature (Zevenbergen & Thorne, 1987), and a DEM.
    d. Created a landcover map using the nine aforementioned indices (five optical and four topographic) as predictor variables within a Spatiotemporal Exploratory Model (STEM) (Hooper & Kennedy, 2018). The output of the STEM model produced a map of probability of glacier presence, capturing primarily debris-free, clean-ice glaciers. This step produced an output analogous to the widely used band-ratioing method of glacier delineation but without the necessity of selecting thresholds. Like existing approaches to mapping supraglacial debris, the resulting STEM output of debris-free ice was combined with outlines from the RGI 6.0 boundaries, assuming

the difference in area to be supraglacial debris. This combined product is referred to as the class label and was used as the target layer to inform the deep learning model.

2. Developed a deep neural network (more specifically ResNet101-PSPNet), referred to hereafter as GlacierCoverNet, to improve capture of supraglacial debris without manual editing and to automate classification of glacier covered area over large geographic extents and fine temporal cadence to create biannual classification maps (Figure 1 – GlacierCoverNet Training).
    a. Combined the class label (discussed in step 1d) with the nine predictor variables – the five optical processed indices from the LandTrendr/Landsat time series and the four topographic indices from the elevation data – to build a sufficiently complex data set for training and evaluating GlacierCoverNet.
    b. Subsetted class label using a “fishnet” with 256 x 256 pixel partitions (i.e., laid a 256 x 256 grid over the optical and topographic images), with 80 percent of grid cells allocated to training and 20 percent to validation.
    c. To train GlacierCoverNet, employed and used systematic sub-sampling of the training set (80 percent) into 128 x 128 pixel image chips (n = 655,452) as model training inputs. The outputs of the GlacierCoverNet model were 18 biannual classification maps (every two years for 1985-2020) including three classes: no glacier, supraglacial debris, and debris free glacier.
3. Migrated the raster outputs to vectors by moving RGI 6.0 labels and metadata to the GlacierCoverNet data set (Figure 1 – Vector Processing).
    a. Labeled GlacierCoverNet grid cells that were spatially coincident with RGI 6.0.
    b. Employed a novel region growing approach to conservatively apply RGI labels to adjacent GlacierCoverNet pixels not included in the original labeling. In all cases, the RGI 6.0 boundaries had no limitations on the GlacierCoverNet outputs, i.e. GlacierCoverNet outputs were allowed to extend beyond or stop short of the RGI boundaries. The only case where RGI boundaries were inherited to the GlacierCoverNet outputs was internally on ice fields (e.g. Bagley Icefield) where the GlacierCoverNet model did not automatically delineate boundaries between glaciers.
    c. Converted labeled rasters to vectors, removed any unlabeled polygons, and applied the RGI recommended minimum mapping unit of 0.01 km^2.
    d. Used a table join to bring the RGI attribute information into the GlacierCoverNet outputs (Figure 1 – Final Data set). These are the attributes marked with a * in Table 1.

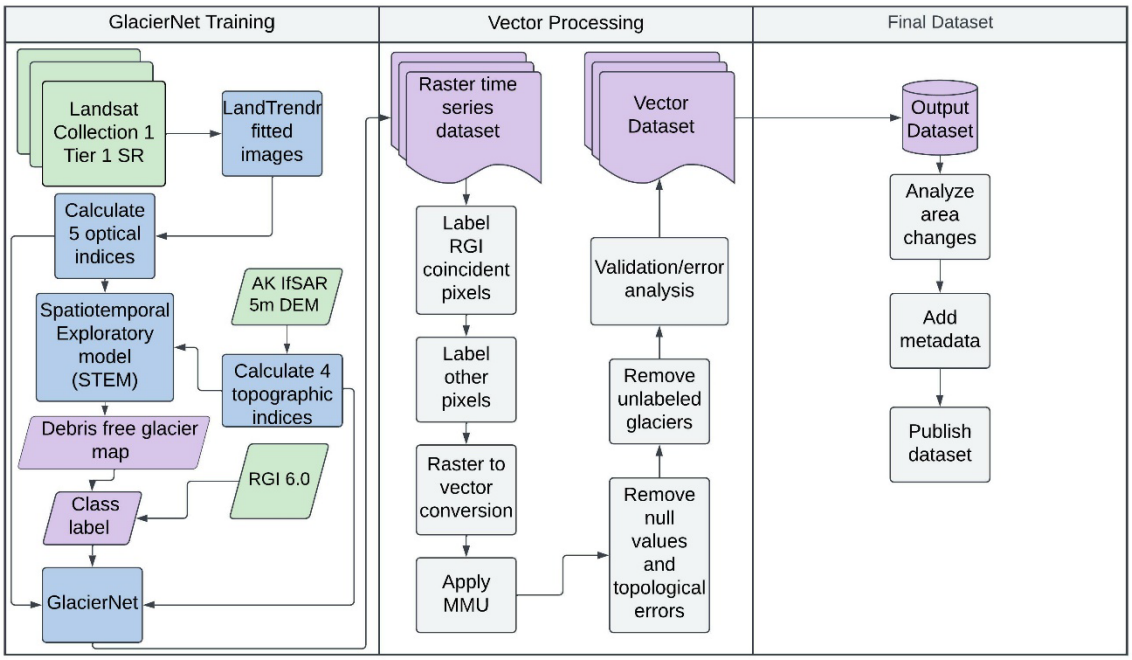

Figure 1. Conceptual diagram of overall project workflow

2.4 Quality, Errors, and Limitations

This data set has higher spatial reliability when capturing glaciers that are greater than approximately 1 km^2 and when assessing landscape-scale changes as compared to the RGI 6.0. A comprehensive error assessment is included in the accompanying publication (Roberts-Pierel et al., 2022). However, it does not perform as well for mapping small glaciers (<0.5 km^2 - 1 km^2) due to the model design; and therefore, the investigators caution users against using the data set solely for analyzing glaciers less than or equal to 1 km^2. The data set also has an overall higher error in the Brooks Range due to the preponderance of small glaciers, morphological characteristics, and the GlacierCoverNet model design.

3 VERSION HISTORY

Table 5. Version History Summary

Version	Release Date	Description of Changes
Version 1	March 2022	Initial release

4 RELATED DATA SETS

Echelmeyer, K. A., V. B. Valentine, and S. L. Zirnheld. 2002, updated 2004. *Airborne Surface Profiling of Alaskan Glaciers, Version 1*. [Indicate subset used]. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. doi: https://doi.org/10.7265/N5RF5RZJ. [Date Accessed].

National Park Service Special Collections within
the Glacier Photograph Collection (https://nsidc.org/data/glacier_photo/)

5 RELATED WEBSITES

https://www.nps.gov/im/swan/index.htm

6 CONTACTS AND ACKNOWLEDGMENTS

Ben Roberts-Pierel
robertsb@oregonstate.edu or brobertspierel@gmail.com
Ocean Administration Building, 104
101 SW 26th St
Corvallis, OR 97331

Peter Kirchner
peter_kirchner@nps.gov
National Park Service, 240 W. 5th Avenue
Anchorage, Ak 99501

6.1 Acknowledgments

This study was supported by the National Park Service, Inventory & Monitoring Program, Southwest Alaska Network and Focused Condition Funds, through a Cooperative Ecosystems Studies Unit agreement awarded to R. Kennedy (P20AC00176). Publication of these data was made possible by NCEI support to NOAA@NSIDC thorough the NOAA Cooperative Agreement with CIRES, NA17OAR4320101.

7 REFERENCES

Flood, N. (2013). Seasonal Composite Landsat TM/ETM+ Images Using the Medoid (a Multi-Dimensional Median). *Remote Sensing* 5(12), 6481–6500. doi: 10.3390/rs5126481.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* 202, 18–27. doi: 10.1016/j.rse.2017.06.031.

Hall, D.K. (2010). Normalized Difference Snow Index (NDSI). In *Encyclopedia of Snow, Ice and Glaciers*. NASA. https://ntrs.nasa.gov/citations/20100031195.

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep Residual Learning for Image Recognition. *ArXiv:1512.03385 [Cs]*. doi: 10.48550/arXiv.1512.03385.

Hooper, S., and Kennedy, R.E. (2018). A spatial ensemble approach for broad-area mapping of land surface properties. *Remote Sensing of Environment* 210, 473–489. doi: 10.1016/j.rse.2018.03.032.

Kauth, R.J., and Thomas, G.S. (1976). The Tasselled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by LANDSAT. *The Laboratory for Applications of Remote Sensing*, Purdue University. https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1160&context=lars_symp.

Kennedy, R.E., Yang, Z., Gorelick, N., Braaten, J., Cavalcante, L., Cohen, W.B., and Healey, S. (2018). Implementation of the LandTrendr Algorithm on Google Earth Engine. *Remote Sensing* 10(5), 691. doi: 10.3390/rs10050691.

Kennedy, R.E., Yang, Z., and Cohen, W.B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. *Remote Sensing of Environment* 114, 2897–2910. doi: 10.1016/j.rse.2010.07.008.

Key, C H., and Benson, N.C. (1999). The Normalized Burn Ratio (NBR): A Landsat TM radiometric measure of burn severity. United States Geological Survey, Northern Rocky Mountain Science Center.

Kirchner, P.B., Bales, R.C., Molotch, N.P., Flanagan, J., and Guo, Q. (2014). LiDAR measurement of seasonal snow accumulation along an elevation gradient in the southern Sierra Nevada, California. *Hydrology and Earth System Sciences* 18(10), 4261-4275. doi: 10.5194/hess-18-4261-2014.

Pfeffer, W.T., Arendt, A.A., Bliss, A., Bolch, T., Cogley, J.G., Gardner, A.S., Hagen, J.O., Hock, R., Kaser, G., Kienholz, C., Miles, E.S., Moholdt, G., Mölg, N., Paul, F., Radić, V., Rastner, P., Raup, B.H., Rich, J., Sharp, M.J., … Wyatt, F.R. (2014). The Randolph glacier inventory: A globally complete inventory of glaciers. *Journal of Glaciology* 60(221), 537–552. doi: 10.3189/2014JoG13J176.

RGI Consortium. (2022). *Randolph Glacier Inventory - A Dataset of Global Glacier Outlines, Version 6*. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. doi: 10.7265/4m1f-gd79.

RGI Consortium. (2017). Randolph Glacier Inventory – A Dataset of Global Glacier Outlines: Version 6.0: Technical Report, Global Land Ice Measurements from Space, Colorado, USA. Digital Media. doi: 10.7265/N5-RGI-60.

Roberts-Pierel, B.M., Kirchner, P.B., Kilbride, J.B., Kennedy, RE. (2022). Changes over the Last 35 Years in Alaska's Glaciated Landscape: A Novel Deep Learning Approach to Mapping Glaciers at Fine Temporal Granularity. *Remote Sensing* 14(18), 4582. doi: 10.3390/rs14184582.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2014). Going Deeper with Convolutions. *ArXiv:1409.4842 [Cs]*. doi: 10.48550/arXiv.1409.4842.

U.S. Geological Survey. (2020). 5-meter Alaska Digital Elevation Models (DEMs)—USGS National Map 3DEP Downloadable Data Collection: U.S. Geological Survey. https://data.usgs.gov/datacatalog/data/USGS:e250fffe-ed32-4627-a3e6-9474b6dc6f0b.

Xie, S., Girshick, R., Dollár, P., Tu, Z., and He, K. (2017). Aggregated Residual Transformations for Deep Neural Networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. doi: 10.48550/arXiv.1611.05431.

Zevenbergen, L.W., and Thorne, C.R. (1987). Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms* 12(1), 47–56. doi: 10.1002/esp.3290120107.

Zhang, H., Wu, C., Zhang, Z., Zhu, Y., Lin, H., Zhang, Z., Sun, Y., He, T., Mueller, J., Manmatha, R., Li, M., and Smola, A. (2020). ResNeSt: Split-Attention Networks. *ArXiv:2004.08955 [Cs]*. doi: 10.48550/arXiv.2004.08955.

Zhao, H., Shi, J., Qi, X., Wang, X., and Jia, J. (2017). Pyramid Scene Parsing Network. *ArXiv:1612.01105 [Cs]*. doi: 10.48550/arXiv.1612.01105.

8 DOCUMENT INFORMATION

8.1 Author

This document was compiled and edited by A. Windnagel with significant text contribution from B. Roberts-Pierel.

8.2 Publication Date

March 2022

8.3 Revision History

No revisions