Collection 2 VIIRS Reservoir Product User's Guide
Version 3.0

Huilin Gao

Department of Civil & Environmental Engineering

Texas A&M University

September 2023

Document Change History Log

Revision	Date	Prepared by	Description
Draft (1.0)	8/2021	Huilin Gao	User Guide first draft for VNP28
Draft (2.0)	8/2022	Huilin Gao	Updated User Guide draft for VNP28/VJ128
Draft (3.0)	9/2023	Huilin Gao	Updated User Guide draft for algorithm refinements

Table of Contents

1. Introduction

This document provides the most current information about the NASA's Visible Infrared Imaging Radiometer Suite (VIIRS) Collection 2 Global Water Reservoir product from Soumi NPP (SNPP) and JPSS-1 satellite (also known as NOAA-20) satellites. The Global Water Reservoir (GWR) product data associated with SNPP and JPSS-1 are named as VNP28 and VJ128, respectively. This product provides data for 164 reservoirs (Figure 1), which includes 151 man-made reservoirs (2,672 km^3) and 13 regulated natural lakes (23,801 km^3) (Li et al., 2021). The total storage capacity of the 151 man-made reservoirs represents 45.82% of the global capacity (in its category) according to the Global Reservoir and Dam (GRanD) database (Lehner et al., 2011). More details about these reservoirs and their attributes are provided in Appendix A.

For each reservoir, the results are presented at two temporal resolutions: 8-day and monthly (Table 1). The surface area, elevation, and storage values are available at both 8-day and monthly intervals, while the evaporation rate and volumetric evaporation are only available at monthly intervals. This document is intended to provide the end user with practical information about how to use these products.

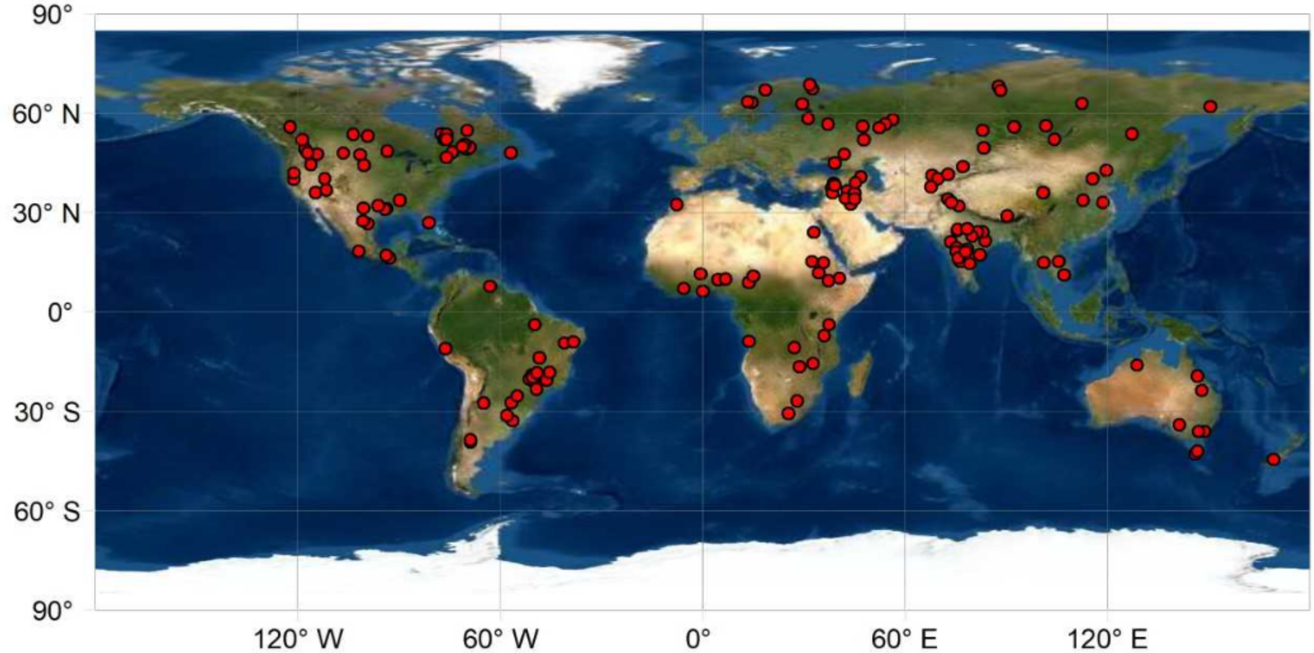


Figure 1. Locations of the 164 global reservoirs contained in this product.

Table 1. Summary of the VNP28/VJ128 Global Water Reservoir product

Terra/ Aqua Product ID	Temporal Resolution	Variables
VNP28C2/VJ128C2	8-day	Area, elevation, and storage

VNP28C3/VJ128C3	Monthly	Area, elevation, storage, evaporation rate, volumetric evaporation

2. VIIRS Reservoir Product (VNP28/VJ128) Algorithm Summary

2.1 VNP28C2/VJ128C2 Algorithm

Figure 2 shows the algorithm for generating the VNP28C2/VJ128C2 8-day product, which is described in more detail in Shah et al. (under review). First, the 8-day reservoir area values are extracted from the 500-m Near Infrared (NIR) bands of VIIRS reflectance (VNP09H1/VJ109H1) data. During the classification process, an area enhancement algorithm (after Zhao et al., 2020) is adopted to minimize the effects of various sources of contamination (e.g., cloud and snow/ice). Then, the area values are applied to the Area-Elevation (A-E) relationship—provided by the Global Reservoir Bathymetry Dataset (GRBD, Li et al., 2020)—to calculate the corresponding elevation value. Lastly, the reservoir storage can be estimated using Equation (1) (after Gao et al., 2012):

$$V_{VIIRS} = V_c - (A_c + A_{VIIRS})(h_c - h_{VIIRS})/2 \qquad (1)$$

where V_c, A_c, and h_c represent storage, area, and water elevation values at capacity; and V_{VIIRS}, A_{VIIRS}, and h_{VIIRS} are the estimated storage, area, and water elevation from VIIRS.

Figure 2. Flow chart of the algorithm for deriving the VNP28C2/VJ128C2 product, which contains 8-day area, elevation, and storage results for the 164 reservoirs. The product components are shown in green boxes.

2.2 VNP28C3/VJ128C3 Algorithm

Figure 3 shows the algorithm for generating the VNP28C3/VJ128C3 monthly product, which is described in detail in Shah et al. (under review). The monthly area values are first estimated based on the composite of the 8-day area classifications, and then converted to monthly elevation and storage results. In addition, monthly evaporation rates are calculated after the Lake Temperature and Evaporation Model (LTEM) (Zhao et al., 2020) using VIIRS LST product (VNP21/VJ121) and meteorological data from Global Land Data Assimilation System (GLDAS). Lastly, monthly evaporative volumetric losses are calculated as the product of evaporation rate and reservoir open water area values.

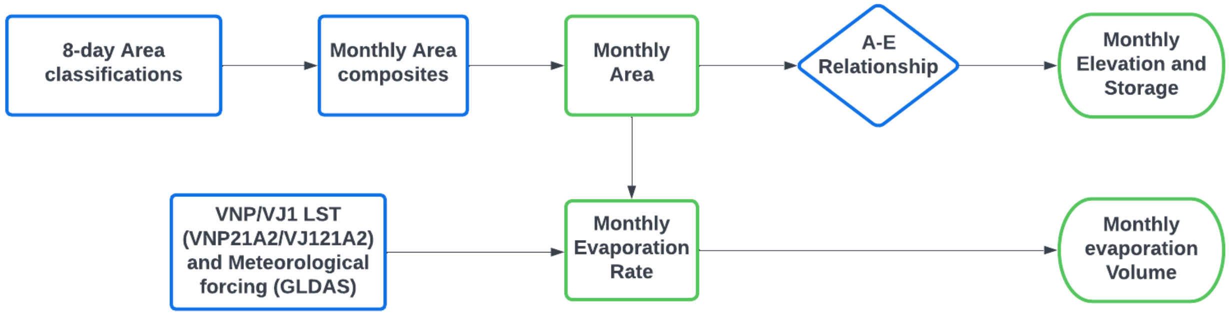


Figure 3. Flow chart of the algorithm for deriving the VNP28C3/VJ128C3 product, which contains monthly area, elevation, storage, evaporation rate, and volumetric evaporation loss results for the 164 reservoirs. The product components are shown in green boxes.

3. VNP28/VJ128 Global Water Reservoir Product Suite

3.1 Level 2 8-day Product (VNP28C2/VJ128C2)

The VNP28C2/VJ128C2 reservoir product contains the 8-day results of the reservoir area, elevation, and storage.

3.1.1 Naming Convention

The file naming convention is as follows:
VNP28C2.AYYYYDDD.002.YYYYDDDHHMMSS.h5
or
VJ128C2.AYYYYDDD.002.YYYYDDDHHMMSS.h5
where,
VNP stands for the SNPP product and VJ1 stands for JPSS-1 product.
YYYY= Year mapped
DDD = Start day of the year (Julian day) at the 8-day interval (see Table 2 for converting to calendar date)
002 = Collection 2
YYYYDDDHHMMSS = Production time

Example: The product file VNP28C2.A2012161.002.2020323115311.h5 contains the Collection 2 reservoir results based on VIIRS data (onboard of SNPP) started on June 9, 2012 (with VIIRS data collected during the 8-DAY period of June 9-16, 2012). The file was produced on November 18, 2020 at 11:53:11 UTC.

Table 2. Lookup table for converting the start date of the 8-day file from Julian Day to the calendar date.

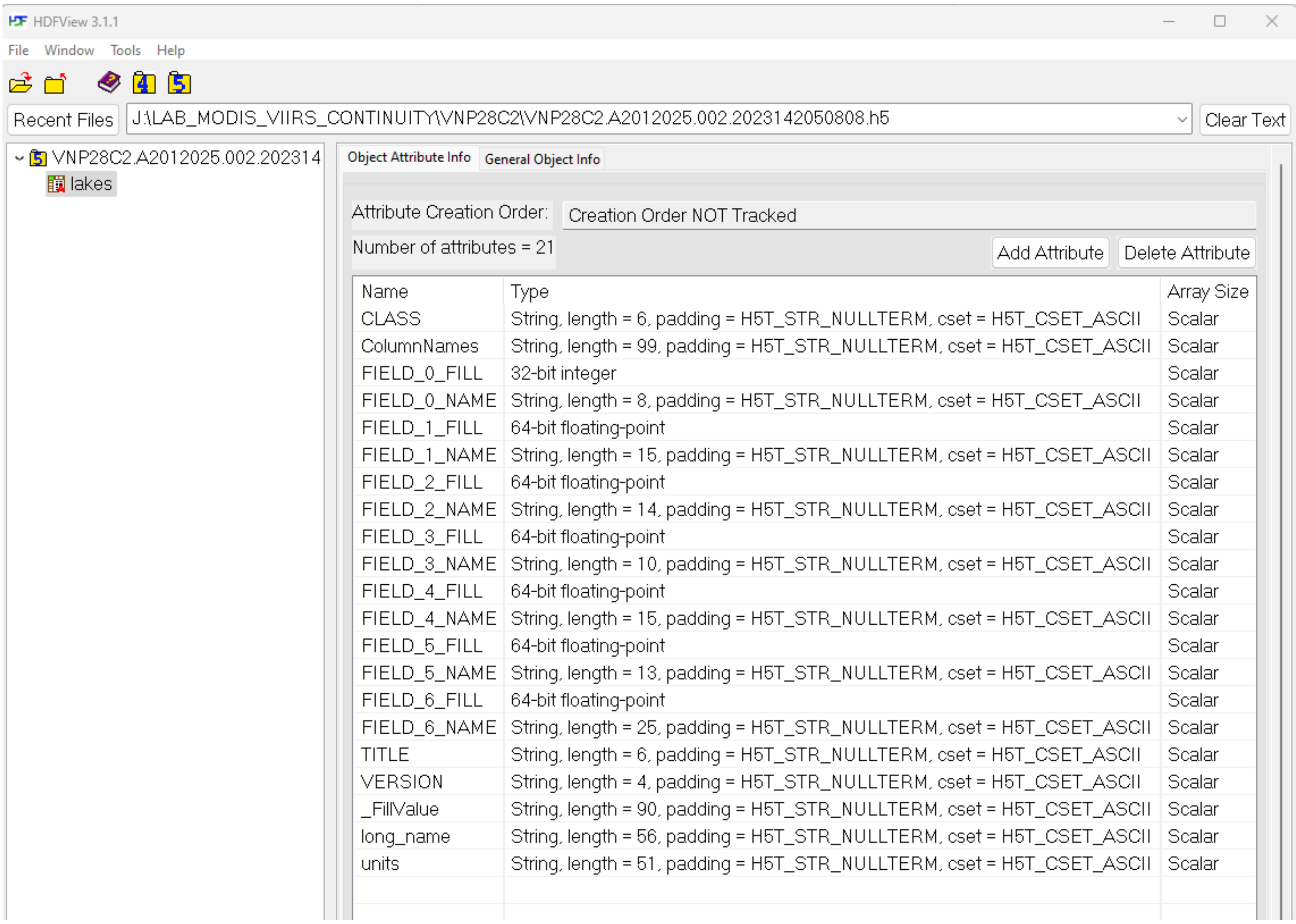
DDD (8-day interval*) **(Regular Years)**	**Corresponding Month and Day** **(Regular Years)**	**DDD (8-day interval)** **(Leap Years)**	**Corresponding Month and Day** **(Leap Years)**
001, 009, 017, 025	Jan. 1, Jan. 9, Jan 17, Jan 25	001, 009, 017, 025	Jan. 1, Jan. 9, Jan 17, Jan 25

033, 041, 049, 057	Feb. 2, Feb. 10, Feb. 18, Feb. 26	033, 041, 049, 057	Feb. 2, Feb. 10, Feb. 18, Feb. 26
065, 073, 081, 089	Mar. 6, Mar. 14, Mar. 22, Mar. 30	065, 073, 081, 089	Mar. 5, Mar. 13, Mar. 21, Mar. 29
097, 105, 113	Apr. 7, Apr. 15, Apr. 23	097, 105, 113, 121	Apr. 6, Apr. 14, Apr. 22, Apr. 30
121, 129, 137, 145	May 1, May 9, May 17, May 25	129, 137, 145	May 8, May 16, May 24
153, 161, 169, 177	Jun. 2, Jun. 10, Jun. 18, Jun. 26	153, 161, 169, 177	Jun. 1, Jun. 9, Jun. 17, Jun. 25
185, 193, 201, 209	Jul. 4, Jul. 12, Jul. 20, Jul. 28	185, 193, 201, 209	Jul. 3, Jul. 11, Jul. 19, Jul. 27
217, 225, 233, 241	Aug. 5, Aug. 13, Aug. 21, Aug. 29	217, 225, 233, 241	Aug. 4, Aug. 12, Aug. 20, Aug. 28
249, 257, 265, 273	Sep. 6, Sep. 14, Sep. 22, Sep. 30	249, 257, 265, 273	Sep. 5, Sep. 13, Sep. 21, Sep. 29
281, 289, 297	Oct. 8, Oct. 16, Oct. 24	281, 289, 297, 305	Oct. 7, Oct. 15, Oct. 23, Oct. 31
305, 313, 321, 329	Nov. 1, Nov. 9, Nov. 17, Nov. 25	313, 321, 329	Nov. 8, Nov. 16, Nov. 24
337, 345, 353, 361[1]	Dec. 3, Dec. 11, Dec. 19, Dec. 27[1]	337, 345, 353, 361[2]	Dec. 2, Dec. 10, Dec. 18, Dec. 26[2]

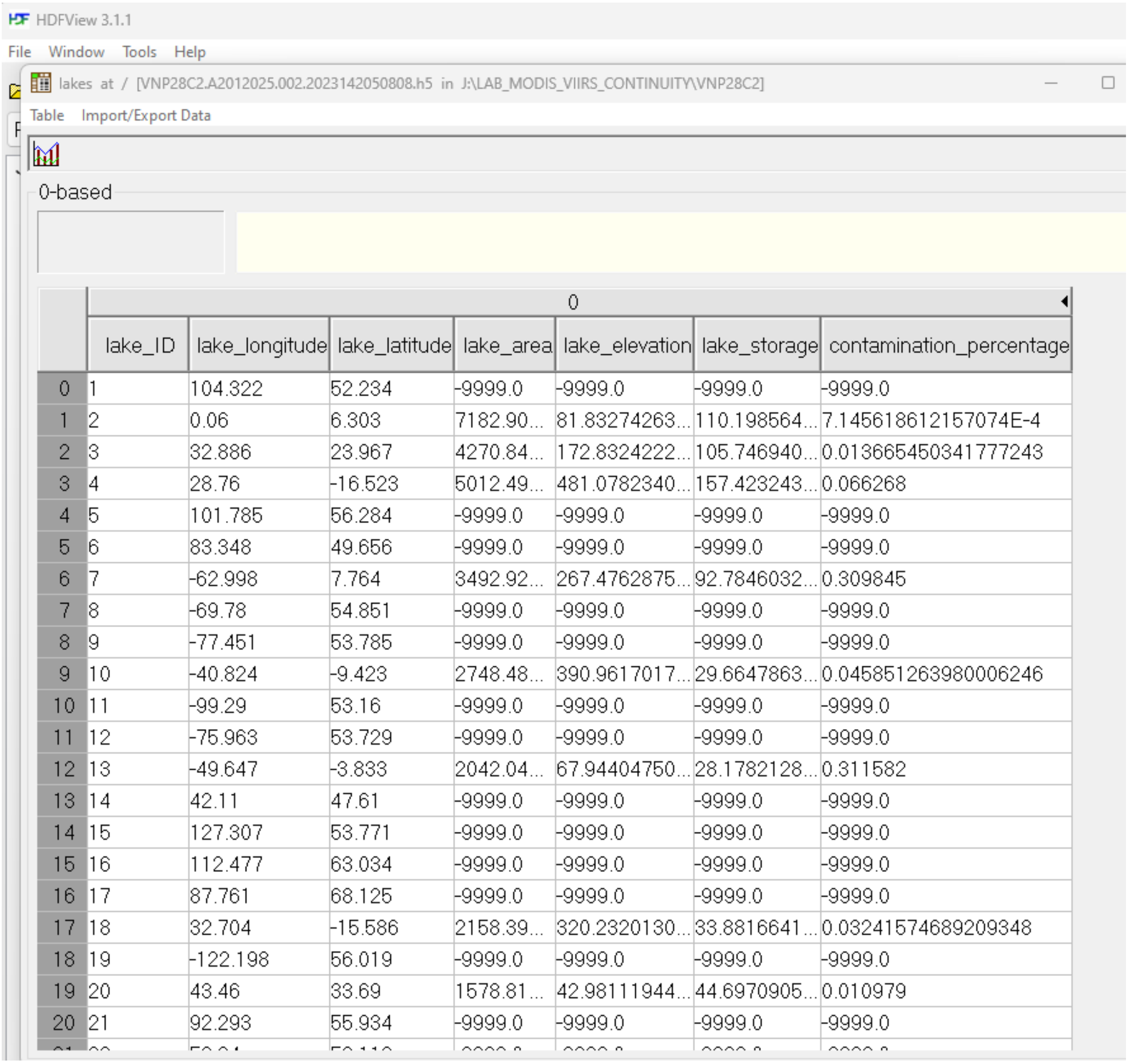
[1]File with this start date is only based on the VIIRS data from the last 5 days of a normal year;
[2] File with this start date is only based on the VIIRS data from the last 6 days of a leap year.

3.1.2 Data Layers

The product contains one single layer named 'lakes'. This layer contains information about the reservoir id, dam location (longitude and latitude), reservoir area, elevation, storage, and contamination percentage. Figure 4 shows the general information about the 'lakes' layer displayed using HDFview.



Name	Type	Array Size
CLASS	String, length = 6, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar
ColumnNames	String, length = 99, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar
FIELD_0_FILL	32-bit integer	Scalar
FIELD_0_NAME	String, length = 8, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar
FIELD_1_FILL	64-bit floating-point	Scalar
FIELD_1_NAME	String, length = 15, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar
FIELD_2_FILL	64-bit floating-point	Scalar
FIELD_2_NAME	String, length = 14, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar
FIELD_3_FILL	64-bit floating-point	Scalar
FIELD_3_NAME	String, length = 10, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar
FIELD_4_FILL	64-bit floating-point	Scalar
FIELD_4_NAME	String, length = 15, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar
FIELD_5_FILL	64-bit floating-point	Scalar
FIELD_5_NAME	String, length = 13, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar
FIELD_6_FILL	64-bit floating-point	Scalar
FIELD_6_NAME	String, length = 25, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar
TITLE	String, length = 6, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar
VERSION	String, length = 4, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar
_FillValue	String, length = 90, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar
long_name	String, length = 56, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar
units	String, length = 51, padding = H5T_STR_NULLTERM, cset = H5T_CSET_ASCII	Scalar



	lake_ID	lake_longitude	lake_latitude	lake_area	lake_elevation	lake_storage	contamination_percentage
0	1	104.322	52.234	-9999.0	-9999.0	-9999.0	-9999.0
1	2	0.06	6.303	7182.90...	81.83274263...	110.198564...	7.145618612157074E-4
2	3	32.886	23.967	4270.84...	172.8324222...	105.746940...	0.013665450341777243
3	4	28.76	-16.523	5012.49...	481.0782340...	157.423243...	0.066268
4	5	101.785	56.284	-9999.0	-9999.0	-9999.0	-9999.0
5	6	83.348	49.656	-9999.0	-9999.0	-9999.0	-9999.0
6	7	-62.998	7.764	3492.92...	267.4762875...	92.7846032...	0.309845
7	8	-69.78	54.851	-9999.0	-9999.0	-9999.0	-9999.0
8	9	-77.451	53.785	-9999.0	-9999.0	-9999.0	-9999.0
9	10	-40.824	-9.423	2748.48...	390.9617017...	29.6647863...	0.045851263980006246
10	11	-99.29	53.16	-9999.0	-9999.0	-9999.0	-9999.0
11	12	-75.963	53.729	-9999.0	-9999.0	-9999.0	-9999.0
12	13	-49.647	-3.833	2042.04...	67.94404750...	28.1782128...	0.311582
13	14	42.11	47.61	-9999.0	-9999.0	-9999.0	-9999.0
14	15	127.307	53.771	-9999.0	-9999.0	-9999.0	-9999.0
15	16	112.477	63.034	-9999.0	-9999.0	-9999.0	-9999.0
16	17	87.761	68.125	-9999.0	-9999.0	-9999.0	-9999.0
17	18	32.704	-15.586	2158.39...	320.2320130...	33.8816641...	0.03241574689209348
18	19	-122.198	56.019	-9999.0	-9999.0	-9999.0	-9999.0
19	20	43.46	33.69	1578.81...	42.98111944...	44.6970905...	0.010979
20	21	92.293	55.934	-9999.0	-9999.0	-9999.0	-9999.0

Figure 4. General information about the 'lakes' layer displayed using HDFView. Top: sample file imported and metadata attributes; and Bottom: data information contained in the sample file. The Reservoir ID for the last row (not shown) is 164.

3.1.3 Metadata

Detailed description about the metadata can be found in Table 3.

Table 3. List of metadata for VNP28C2/VJ128C2

Column Name	lake_ID	lake_longitude	lake_latitude	lake_area	lake_elevation	lake_storage

Fill Value	-9999.0	-9999.0	-9999.0	-9999.0	-9999.0	-9999.0
Long name	8-day aggregated global lake area					
Unit	None	Decimal Degree	Decimal Degree	km^2	m	km^3

3.1.4 Example Codes

Example codes (in MATLAB and Python) for reading the VNP28C2/VJ128C2 hdf files are provided in Appendix B.

3.2 Level 3 Monthly Reservoir Product (VNP28C3/VJ128C3)

The VNP28C3 product contains the monthly results of the reservoir area, elevation, storage, evaporation rate, and evaporation volume.

3.2.1 Naming Convention

The file naming convention is as follows:
VNP28C3.AYYYYDDD.002.YYYYDDDHHMMSS.h5
OR
VJ128C3.AYYYYDDD.002.YYYYDDDHHMMSS.h5
where,
VNP stands for the SNPP product and VJ1 stands for JPSS-1 product.
YYYY= Year mapped
DDD = Day-of-year (DOY) of the first day of each calendar month (please see Table 4 for more details)
002 = Collection 2
YYYYDDDHHMMSS = Production time

Example: The product file VNP28C3.A2012001.002.2020333100640.h5 contains the Collection 2 reservoir results based on VIIRS data (onboard of SNPP) in January 2012. The file was produced on November 28, 2020, at 10:06:40 UTC.

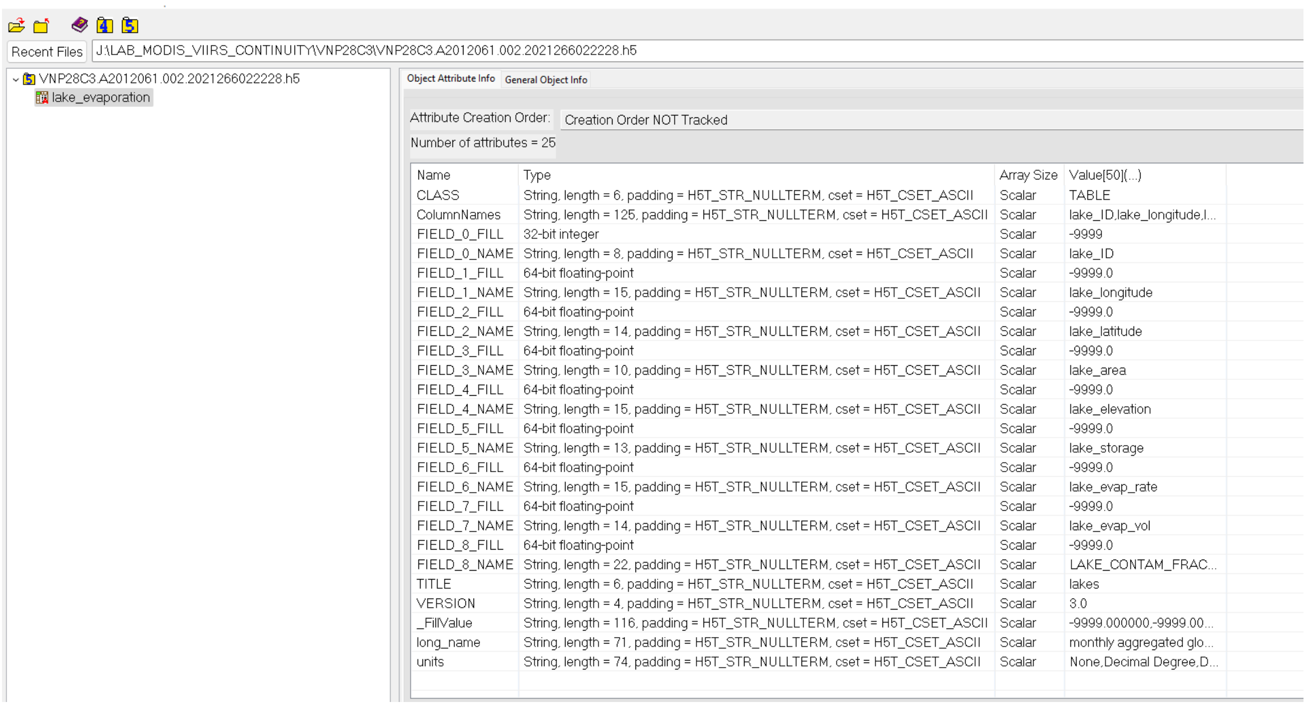
Table 4. Day-of-year (DOY) of the first day of each calendar month.

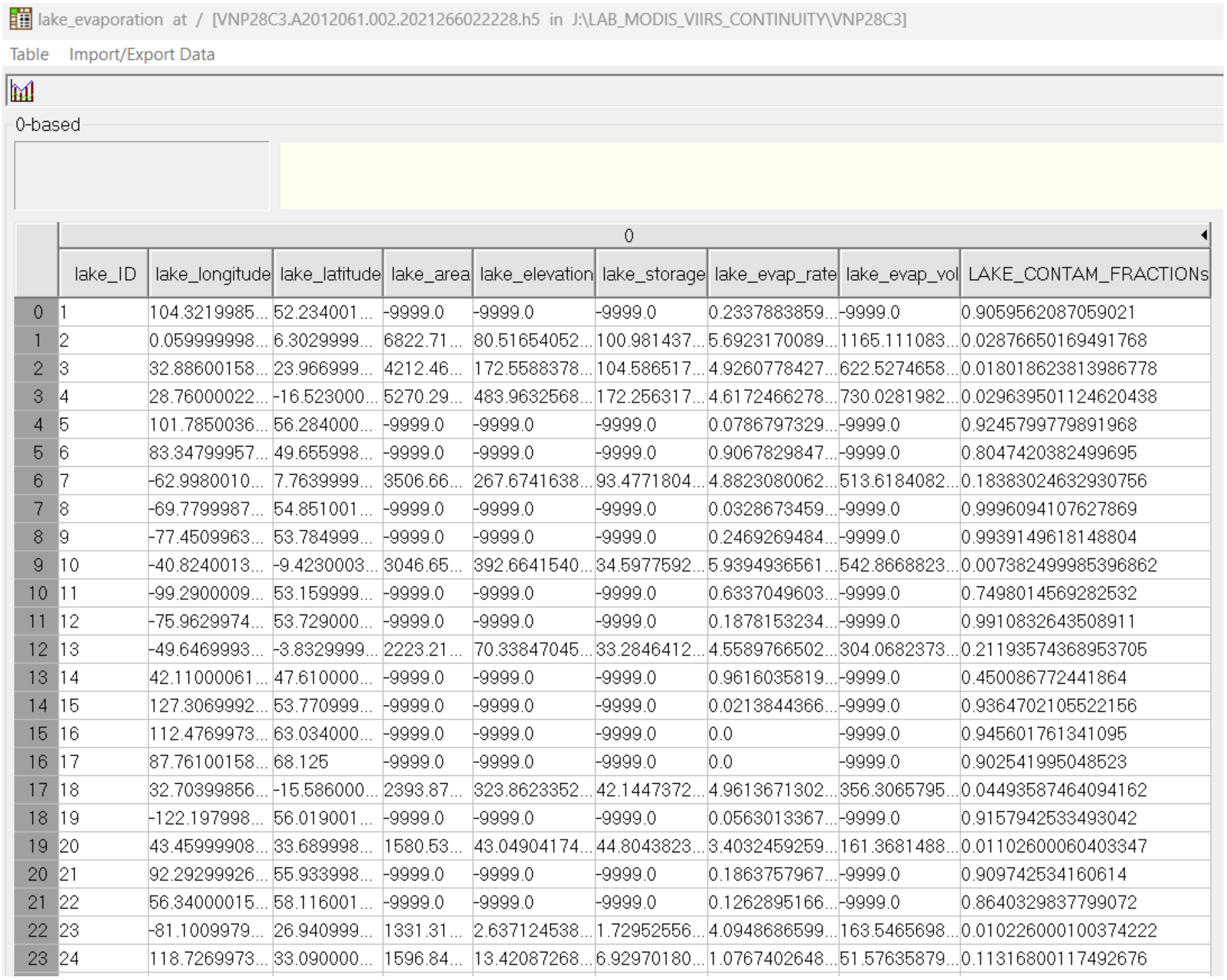
Month	Non-Leap Year Start DOY	Leap Year Start DOY
January	1	1
February	32	32
March	60	61
April	91	92
May	121	122
June	152	153

July	182	183
August	213	214
September	244	245
October	274	275
November	305	306
December	335	336

3.2.2 Data Layers

The product contains one single layer named 'lake_evaporation'. This layer contains information about the reservoir id, dam location (longitude and latitude), monthly reservoir area, elevation, storage, evaporation rate, evaporation volume, and contamination percentage. Figure 5 shows the general information about the 'lake_evaporation' layer displayed using HDFView.





	lake_ID	lake_longitude	lake_latitude	lake_area	lake_elevation	lake_storage	lake_evap_rate	lake_evap_vol	LAKE_CONTAM_FRACTIONs
0	1	104.3219985...	52.234001...	-9999.0	-9999.0	-9999.0	0.2337883859...	-9999.0	0.9059562087059021
1	2	0.059999998...	6.3029999...	6822.71...	80.51654052...	100.981437...	5.6923170089...	1165.111083...	0.02876650169491768
2	3	32.88600158...	23.966999...	4212.46...	172.5588378...	104.586517...	4.9260778427...	622.5274658...	0.018018623813986778
3	4	28.76000022...	-16.523000...	5270.29...	483.9632568...	172.256317...	4.6172466278...	730.0281982...	0.029639501124620438
4	5	101.7850036...	56.284000...	-9999.0	-9999.0	-9999.0	0.0786797329...	-9999.0	0.9245799779891968
5	6	83.34799957...	49.655998...	-9999.0	-9999.0	-9999.0	0.9067829847...	-9999.0	0.8047420382499695
6	7	-62.9980010...	7.7639999...	3506.66...	267.6741638...	93.4771804...	4.8823080062...	513.6184082...	0.18383024632930756
7	8	-69.7799987...	54.851001...	-9999.0	-9999.0	-9999.0	0.0328673459...	-9999.0	0.9996094107627869
8	9	-77.4509963...	53.784999...	-9999.0	-9999.0	-9999.0	0.2469269484...	-9999.0	0.9939149618148804
9	10	-40.8240013...	-9.4230003...	3046.65...	392.6641540...	34.5977592...	5.9394936561...	542.8668823...	0.007382499985396862
10	11	-99.2900009...	53.159999...	-9999.0	-9999.0	-9999.0	0.6337049603...	-9999.0	0.7498014569282532
11	12	-75.9629974...	53.729000...	-9999.0	-9999.0	-9999.0	0.1878153234...	-9999.0	0.9910832643508911
12	13	-49.6469993...	-3.8329999...	2223.21...	70.33847045...	33.2846412...	4.5589766502...	304.0682373...	0.21193574368953705
13	14	42.11000061...	47.610000...	-9999.0	-9999.0	-9999.0	0.9616035819...	-9999.0	0.450086772441864
14	15	127.3069992...	53.770999...	-9999.0	-9999.0	-9999.0	0.0213844366...	-9999.0	0.9364702105522156
15	16	112.4769973...	63.034000...	-9999.0	-9999.0	-9999.0	0.0	-9999.0	0.945601761341095
16	17	87.76100158...	68.125	-9999.0	-9999.0	-9999.0	0.0	-9999.0	0.902541995048523
17	18	32.70399856...	-15.586000...	2393.87...	323.8623352...	42.1447372...	4.9613671302...	356.3065795...	0.04493587464094162
18	19	-122.197998...	56.019001...	-9999.0	-9999.0	-9999.0	0.0563013367...	-9999.0	0.9157942533493042
19	20	43.45999908...	33.689998...	1580.53...	43.04904174...	44.8043823...	3.4032459259...	161.3681488...	0.01102600060403347
20	21	92.29299926...	55.933998...	-9999.0	-9999.0	-9999.0	0.1863757967...	-9999.0	0.909742534160614
21	22	56.34000015...	58.116001...	-9999.0	-9999.0	-9999.0	0.1262895166...	-9999.0	0.8640329837799072
22	23	-81.1009979...	26.940999...	1331.31...	2.637124538...	1.72952556...	4.0948686599...	163.5465698...	0.010226000100374222
23	24	118.7269973...	33.090000...	1596.84...	13.42087268...	6.92970180...	1.0767402648...	51.57635879...	0.11316800117492676

Figure 5. General information about the 'lake_evaporation' layer displayed using HDFView. Top: sample file imported and metadata attributes; and Bottom: data information contained in the sample file. The Reservoir ID for the last row (not shown) is 164.

3.2.3 Metadata

A detailed description of the metadata can be found in Table 5.

Table 5. List of metadata for VNP28C3/VJ128C3

Column Name	lake_ID	lake_longitude	lake_latitude	lake_area	lake_elevation	lake_storage	lake_evap_rate	lake_evap_vol
Fill Value	-9999.0	-9999.0	-9999.0	-9999.0	-9999.0	-9999.0	-9999.0	-9999.0
Long	monthly aggregated global lake area							

nam e								
Unit	None	Decimal Degree	Decimal Degree	km^2	m	km^3	mm/d	million m^3/month

3.2.4 Example codes

Example codes (in Matlab and Python) for reading the VNP28C3/VJ128C3 hdf files are provided in Appendix C.

4. Known Problems

Although the image enhancement algorithm after Zhao et al. (2020) can significantly improve the accuracy of the 8-day area classifications, there are still some biased estimations, especially over the high latitudes. To address this issue in time series analysis, an "outlier removal and gap-filling" script (Appendix D) is applied on the 8-day surface area time series. The outliers are identified using the biases between the area values and their moving averages (with a window size of 7). If the bias of an area value falls more than 3 standard deviations from the average bias, that area value is considered an outlier. When an area outlier is removed, its corresponding elevation and storage values are also removed. The area value data gaps are then filled in via interpolation. Afterwards, the missing elevation and storage data are calculated using the interpolated areas.

Another problem is that the storage values are sometimes calculated to be negative—in which case they are reassigned to zero in the product. This is attributed to the following: First, some reservoir capacity values only include the live capacity (and not the dead storage capacity—i.e. the portion of the reservoir capacity that is not used for operational purposes). Therefore, the calculated storage value will be negative when the water surface level is below the dead storage level. The second reason is attributed to the uncertainty of the A-E relationship. If the estimated slope is steeper than the actual condition, negative values can happen when the water area values decrease beyond a certain threshold. It should be noted, however, that these zero values only occur occasionally (when the water level is very low) in a small number of specific reservoirs.

5. Relevant Web and FTP Sites

• VIIRS Reservoir Product Overview: General information about the VIIRS water reservoir products, with user guide and Algorithm Theoretical Basic Document (ATBD) is provided.
https://viirsland.gsfc.nasa.gov/Products/NASA/GWR.html

• VIIRS Reservoir Product Downloading: Websites where VIIRS Land products are distributed.
https://lpdaac.usgs.gov/ and https://ladsweb.modaps.eosdis.nasa.gov/

• Scripts for post-processing VIIRS/MODIS Reservoir Product: Scripts from the science team to assist the users to further process the product (e.g., removing outliers from time series).
https://dataverse.tdl.org/dataverse/HGao

• VIIRS Reservoir Product Validation:
(Note: link for publication will be provided later)

• VIIRS LDOPE Tools: A collection of programs, written by members of the Land Data Operational Product Evaluation (LDOPE) group, to assist in the analysis and quality assessment of VIIRS Land products.
https://ladsweb.modaps.eosdis.nasa.gov/tools-and-services/#ldope

6. References

Gao, H., Birkett, C., & Lettenmaier, D. P. (2012). Global monitoring of large reservoir storage from satellite remote sensing. Water Resources Research, 48(9).

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., ... & Nilsson, C. (2011). High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. Frontiers in Ecology and the Environment, 9(9), 494-502.

Li, Y., Zhao, G., Shah, D., Zhao, M., Sarkar, S., Devadiga, S., Zhao, B., Zhang, S., & Gao, H. (2021). NASA's MODIS/VIIRS Global Water Reservoir Product Suite from Moderate Resolution Remote Sensing Data. Remote Sensing, 13(4), 565.

Li, Y., Gao, H., Zhao, G., & Tseng, K. H. (2020). A high-resolution bathymetry dataset for global reservoirs using multi-source satellite imagery and altimetry. Remote Sensing of Environment, 244, 111831.

Shah, D., Zhang, S.,….,Gao, H., (Under Review). Transitioning from NASA's MODIS to VIIRS Global Water Reservoir (GWR) Products: A Perspective of Continuity. Scientific Data.

Zhao, G., Gao, H., & Cai, X. (2020). Estimating lake temperature profile and evaporation losses by leveraging MODIS LST data. Remote Sensing of Environment, 251, 112104.

Appendix A - List of the 164 Reservoirs and their Attributes

Lake ID	GRanD_id[1]	Reservoir Name	lon, lat (◦)	Country	Continent	A-E Coefficients[2] a, b	Storage capacity (km3)	Area capacity (km2)	Elevation capacity (m)	Capacity Source[3]
1[*]	5058	Baikal	104.32, 52.24	Russia	Asia	0.00447, 312.77026	23615.39	32265.61	456.88	GRanD
2	3667	Volta	0.06, 6.3	Ghana	Africa	0.00365, 55.58562	148	8502	86.65	wikipedia
3	4478	Nasser	32.89, 23.97	Egypt	Africa	0.00469, 152.81994	162	6500	183.28	literature
4	4056	Kariba Reservoir	28.76, -16.52	Zambia	Africa	0.01119, 424.98467	180	5400	485.41	wikipedia
5	5055	Bratsk Reservoir	101.78, 56.29	Russia	Asia	0.00657, 367.92163	169.27	5470	403.85	wikipedia
6	4787	Zaysan	83.35, 49.66	Kazakhstan	Asia	0.00465, 370.20585	49.8	5490	395.74	GRanD
7	2294	Guri Reservoir	-63, 7.77	Venezuela	South America	0.0144, 217.16716	135	4250	278.38	wikipedia
8	1995	Caniapiscau Reservoir	-69.78, 54.85	Canada	North America	0.01218, 488.99841	53.79	4275	541.08	GRanD
9	1394	Robert Bourassa Reservoir	-77.45, 53.79	Canada	North America	0.0111, 143.99061	61.7	2905	176.24	Hydro-Québec
10	2516	Sobradinho Reservoir	-40.82, -9.42	Brazil	South America	0.00571, 375.26816	34.1	3017.9	392.5	GRanD
11	712	Cedar	-99.29, 53.16	Canada	North America	0.00217, 250.49224	9.64	2668.46	256.29	GRanD
12	1396	La Grande 3 Reservoir	-75.96, 53.73	Canada	North America	0.02539, 195.25843	60	2451	257.48	Hydro-Québec
13	2365	Tucurui Reservoir	-49.65, -3.83	Brazil	South America	0.01322, 40.95573	45.5	2606	75.4	GRanD
14	4375	Tsimlyanskoye Reservoir	42.11, 47.61	Russia	Euro	0.01177, 7.63989	23.86	2702	39.44	literature
15	5834	Zeyskoye Reservoir	127.31, 53.77	Russia	Asia	0.02065, 266.43675	68.4	2420	316.41	wikipedia
16	5180	Vilyuy Reservoir	112.48, 63.03	Russia	Asia	0.02852, 182.74156	35.9	2170	244.62	wikipedia
17	4783	Khantayskoye Reservoir	87.81, 68.16	Russia	Asia	0.00445, 49.76375	23.5	2221.61	59.64	GRanD

18	4505	Cahora Bassa Reservoir	32.7, -15.58	Mozambique	Africa	0.01542, 286.9568	55.8	2739	329.18	wikipedia
19	6	Williston	-122.2, 56.02	Canada	North America	0.0529, 580.99344	39.47	1773	674.79	literature
20	4472	Buhayrat ath Tharthar	43.46, 33.69	Iraq	Asia	0.03955, -19.46261	85.59	2135.54	65	literature
21	5056	Krasnoyarsk Reservoir	92.29, 55.93	Russia	Asia	0.03863, 162.77316	73.3	2000	240.04	wikipedia
22	4623	Kama Reservoir	56.34, 58.12	Russia	Euro	0.00744, 96.07894	12.2	1915	110.32	wikipedia
23	1957	Okeechobee	-81.1, 26.94	United States of America	North America	0.00617, -5.57499	3.546	1536.8	3.9	wikipedia
24	5295	Hungtze	118.73, 33.09	China	Asia	0.00749, 1.45816	13.5	2074.61	17	literature
25	4474	Razazah	43.89, 32.7	Iraq	Asia	0.01457, 11.06852	25.75	1621	34.69	literature
26	2023	Gouin Reservoir	-74.1, 48.36	Canada	North America	0.00068, 402.90611	8.57	1570	403.98	GRanD
27	4789	Qapshaghay Bogeni Reservoir	77.1, 43.92	Kazakhstan	Asia	0.00897, 467.10974	28.1	1850	483.71	GRanD
28	753	Fort Berthold Reservoir	-101.43, 47.51	United States of America	North America	0.02467, 528.64792	29.38	1477.4	565.1	wikipedia
29	2445	Aperea Reservoir	-56.63, -27.39	Paraguay	South America	0.02242, 48.84199	21	1600	84.71	literature
30	870	Oahe	-100.4, 44.46	United States of America	North America	0.02172, 462.72715	28.35	1429.57	493.78	wikipedia
31	2390	Ilha Solteira Reservoir	-51.38, -20.37	Brazil	South America	0.03237, 290.94542	21.17	1200	329.78	GRanD
32	4629	Saratov Reservoir	47.76, 52.05	Russia	Euro	0.02563, -0.27741	12.9	1117.7	28.36	GRanD
33	4350	Imandra	32.55, 67.41	Russia	Euro	0.18726, -62.86735	10.8	1062.37	136.07	GRanD
34	3640	Kainji Reservoir	4.61, 9.87	Nigeria	Africa	0.03997, 93.99579	15	1071.23	136.81	wikipedia
35	4785	Novosibirskoye	83, 54.84	Russia	Asia	0.01419, 98.78019	8.8	1070	113.97	wikipedia
36	4625	Cheboksary	47.46, 56.14	Russia	Euro	0.02447, 39.29789	13.85	1080.38	65.73	literature
37	4359	Ilmen	31.28, 58.46	Russia	Euro	0.0083, 9.98411	12	1120	19.28	wikipedia

38	4480	Jebel Aulia Reservoir	32.48, 15.24	Sudan	Africa	0.00624, 375.01032	3.5	861.19	380.39	FAO
39	1397	Opinaca Reservoir	-76.58, 52.21	Canada	North America	0.02118, 194.07727	8.5	1040	216.1	wikipedia
40	2392	Furnas	-46.31, -20.67	Brazil	South America	0.0437, 720.07262	22.59	1127.07	769.32	wikipedia
41	2368	Serra da Mesa Reservoir	-48.3, -13.84	Brazil	South America	0.03356, 410.19963	54.4	1784	470.07	wikipedia
42	4624	Votkinsk Reservoir	54.08, 56.8	Russia	Euro	0.03892, 53.1356	9.4	850.82	86.25	wikipedia
43	6201	Argyle Reservoir	128.74, -16.12	Australia	Oceania	0.02806, 66.43617	10.76	981.21	93.97	wikipedia
44	731	Rainy	-93.36, 48.62	Canada	North America	0.00078, 336.08674	0.69	829.45	336.73	GRanD
45	307	Fort Peck	-106.41, 48	United States of America	North America	0.04376, 643.31691	22.77	969.86	685.76	wikipedia
46	2375	Tres Marias Reservoir	-45.27, -18.21	Brazil	South America	0.03553, 539.11233	21	1040	576.06	wikipedia
47	2012	Pipmuacan Reservoir	-69.77, 49.36	Canada	North America	0.0498, 360.46403	13.9	978	409.16	wikipedia
48	4679	Chardarinskoye	67.96, 41.25	Kazakhstan	Asia	0.01786, 238.24413	5.7	800.66	252.54	wikipedia
49	4626	Nizhnekamsk Reservoir	52.28, 55.7	Russia	Euro	0.0138, 50.37324	13.8	1084	65.34	wikipedia
50	2456	Negro Reservoir	-56.42, -32.83	Uruguay	South America	0.0194, 62.00777	8.8	1070	82.77	wikipedia
51	2343	Chocon Reservoir	-68.76, -39.27	Argentina	South America	0.01519, 365.74893	22	820	378.2	GRanD
52	4442	Ataturk Dam	38.32, 37.49	Turkey	Asia	0.10643, 454.25042	48.7	817	541.2	GRanD
53	2513	Itaparica Reservoir	-38.31, -9.14	Brazil	South America	0.03337, 279.33376	10.7	781.21	305.4	wikipedia
54	4464	Assad	38.55, 35.86	Syria	Asia	0.05942, 266.62629	11.7	610	302.87	wikipedia
55	3650	Lagdo Reservoir	13.69, 9.06	Cameroon	Africa	0.0374, 190.15542	7.7	691.12	216	FAO
56	1269	Toledo Bend Reservoir	-93.57, 31.18	United States of America	North America	0.02039, 39.45546	5.52	636.18	52.43	wikipedia
57	6922	Eastmain Reservoir	-75.89, 52.19	Canada	North America	0.06785, 245.91598	6.94	602.9	286.82	literature

58	2009	Outardes 4 Reservoir	-68.91, 49.71	Canada	North America	0.19049, 239.61011	24.5	640	361.53	Hydro-Québec
59	4349	Kovdozero	31.76, 68.6	Russia	Euro	0.00193, 78.17686	11.52	745	79.62	GRanD
60	2380	Sao Simao Reservoir	-50.5, -19.02	Brazil	South America	0.0523, 369.16877	12.5	703	405.94	wikipedia
61	610	Mead	-114.73, 36.02	United States of America	North America	0.13619, 288.76038	34.07	659.3	374.6	USBR
62	5087	Yamdrok	90.38, 29.1	China	Asia	0.01275, 4435.35521	14.6	638	4443.49	literature
63	1391	Angostura Reservoir	-92.78, 16.4	Mexico	North America	0.08079, 478.95889	18.2	640	530.67	wikipedia
64	4991	Srisailam Reservoir	78.9, 16.09	India	Asia	0.03079, 253.3044	8.29	534.05	269.75	CWC
65	2455	Grande Reservoir	-57.94, -31.27	Argentina	South America	0.03068, 16.88963	5	592.83	35.08	wikipedia
66	4843	Gandhisagar Reservoir	75.55, 24.7	India	Asia	0.03366, 379.03449	6.83	619.89	399.9	CWC
67	2397	Promissao Reservoir	-49.78, -21.3	Brazil	South America	0.08038, 342.73167	7.41	513.39	384	GRanD
68	282	Arrow	-117.78, 49.34	Canada	North America	0.17477, 351.0668	10.3	504.82	439.3	USACE
69	2382	Agua Vermelha Reservoir	-50.35, -19.87	Brazil	South America	0.05626, 351.61681	11.03	563.15	383.3	wikipedia
70	4898	Hirakud Reservoir	83.85, 21.52	India	Asia	0.02204, 177.26302	5.38	669.62	192.02	CWC
71	3041	Kossour Reservoir	-5.47, 7.03	Ivory Coast	Africa	0.03423, 169.77945	27.68	1058.2	206	GRanD
72	4784	Kureiskaya	88.29, 66.95	Russia	Asia	0.04971, 67.89284	9.96	558	95.63	literature
73	3071	Storsjon	14.47, 63.3	Sweden	Euro	0.00422, 291.0872	0.5	484.6	293.13	GRanD
74	316	Flathead Lake	-114.23, 47.68	United States of America	North America	0.13239, 816.09051	23.2	510	883.61	wikipedia
75	2004	Kempt	-70.53, 50.66	Canada	North America	0.03312, 478.60112	2.22	470.44	494.18	GRanD
76	6700	Kolyma dam	150.23, 62.05	Russia	Asia	0.13658, 390.9085	15.08	454.6	453	wikipedia

77	4501	Mtera Reservoir	35.98, -7.14	United Republic of Tanzania	Africa	0.02183, 688.04662	3.2	478.83	698.5	literature
78	4686	Kayrakkumskoye	69.82, 40.28	Tajikistan	Asia	0.02143, 335.23897	4.2	513	346.23	wikipedia
79	250	Kinbasket	-118.57, 52.08	Canada	North America	0.31717, 622.76738	24.76	430	759.15	wikipedia
80	4634	Mingechaurskoye	47.03, 40.8	Azerbaijan	Asia	0.07215, 42.01887	15.73	567.97	83	wikipedia
81	2431	Lago del Río Yguazú	-54.97, -25.37	Paraguay	South America	0.04517, 203.13232	8.47	620	231.14	wikipedia
82	4858	Govind Ballabah Pant	83, 24.2	India	Asia	0.06208, 241.75327	5.65	426.36	268.22	CWC
83	4422	Keban Baraji	38.76, 38.81	Turkey	Asia	0.11302, 772.50564	30.6	675	848.79	wikipedia
84	2340	Los Barreales	-68.69, -38.58	Argentina	South America	0.30759, 290.07305	27.7	413	417.11	literature
85	4859	Bansagar Lake	81.29, 24.19	India	Asia	0.05088, 317.64432	5.17	471.6	341.64	CWC
86	1275	Sam Rayburn Reservoir	-94.11, 31.07	United States of America	North America	0.0355, 35.65711	3.55	455.64	50.11	TWDB
87	2414	Barra Bonita	-49.23, -23.21	Brazil	South America	0.00228, 565.24837	7.01	542	566.48	GRanD
88	4739	Ukal	73.6, 21.26	India	Asia	0.04229, 83.59772	6.62	509.85	105.16	CWC
89	479	Utah Lake	-111.89, 40.36	United States of America	North America	0.02307, 1359.51211	1.07	380	1368.28	wikipedia
90	305	Pend Oreille Lake	-117, 48.18	United States of America	North America	0.22845, 541.65792	54.2	381.47	628.8	wikipedia
91	4994	Tungabhadra	76.33, 15.27	India	Asia	0.04122, 483.33699	3.28	349.42	497.74	CWC
92	4461	Mosul Dam Lake	42.83, 36.63	Iraq	Asia	0.16032, 273.38375	11.1	353.16	330	wikipedia
93	4470	Habbaniyah	42.35, 34.21	Iraq	Asia	0.07125, 114.61642	8.2	418.4	144.43	literature
94	4946	Sriramsagar Reservoir	78.34, 18.97	India	Asia	0.04005, 319.94975	2.3	314.38	332.54	CWC
95	2376	Lago das Brisas	-49.1, -18.41	Brazil	South America	0.08818, 471.03368	17	559.6	520.38	wikipedia

96	2356	Meelpaeg	-56.78, 48.17	Canada	North America	0.0041, 269.35893	2.16	314.9	270.65	GRanD
97	4260	Hendrik Verwoerd	25.5, -30.62	South Africa	Africa	0.06907, 1236.10289	5.34	374	1261.93	wikipedia
98	1387	Malpaso	-93.6, 17.18	Mexico	North America	0.30032, 89.06386	9.17	309.45	182	literature
99	1379	Inhernillo	-101.89, 18.27	Mexico	North America	0.14118, 116.65544	12	400	173.13	wikipedia
100	4184	Vaaldam	28.12, -26.88	South Africa	Africa	0.0358, 1472.81742	2.61	320	1484.27	wikipedia
101	5062	Longyangxia	100.92, 36.12	China	Asia	0.18321, 2518.97907	24.7	383	2589.15	wikipedia
102	3727	Hoytiainen	29.48, 62.83	Finland	Euro	0.0064, 86.17122	2.39	293	88.05	GRanD
103	1423	Baskatong	-75.98, 46.72	Canada	North America	0.05663, 207.28526	2.63	280	223.14	GRanD
104	5803	Tri An Lake	107.04, 11.11	Vietnam	Asia	0.07216, 39.48203	2.76	323	62.79	wikipedia
105	2007	Peribonka	-71.25, 49.9	Canada	North America	0.10611, 411.5385	5.18	270.72	440.26	GRanD
106	4942	Jayakwadi	75.37, 19.49	India	Asia	0.03201, 451.67121	2.17	382.39	463.91	CWC
107	3638	Shiroro	6.84, 9.97	Nigeria	Africa	0.08602, 350.89662	7	312	377.73	FAO
108	4379	Tshchikskoye	39.12, 44.99	Russia	Euro	0.06161, 16.03972	3.05	286.28	33.68	FAO
109	710	Tobin	-103.4, 53.66	Canada	North America	0.00897, 311.22766	2.2	263.86	313.59	GRanD
110	5796	Noi	105.43, 15.21	Thailand	Asia	0.05709, 129.50217	1.97	288	145.94	wikipedia
111	4483	Roseires Reservoir	34.39, 11.8	Sudan	Africa	0.02506, 475.84407	7.4	450	487.12	wikipedia
112	4675	Toktogul'skoye	72.65, 41.68	Kyrgyzstan	Asia	0.55471, 743.53409	19.5	284.3	901.24	wikipedia
113	6698	Gordon	145.98, -42.73	Australia	Oceania	0.37007, 208.53588	12.4	278	311.42	wikipedia
114	4964	Ujani	75.12, 18.07	India	Asia	0.05453, 482.16622	1.52	268.91	496.83	CWC
115	2312	Hondo	-64.89, -27.52	Argentina	South America	0.02922, 266.72004	1.74	330	276.36	WLDB

116	4362	Ivankovo Reservoir	37.12, 56.73	Russia	Euro	0.01794, 119.50914	1.17	220.57	123.47	GRanD
117	4702	Tarbela	72.69, 34.09	Pakistan	Asia	0.52839, 351.45663	13.69	250	483.55	wikipedia
118	4985	Nagarjuna	79.31, 16.57	India	Asia	0.29044, 100.77784	6.84	272.18	179.83	wikipedia
119	3070	Kallsjon	13.34, 63.43	Sweden	Euro	0.02782, 387.52135	0.45	189.74	392.8	GRanD
120	4431	Karakaya	39.14, 38.23	Turkey	Asia	0.22073, 631.76449	9.5	298	697.54	wikipedia
121	4792	Beas	75.95, 31.97	India	Asia	0.20473, 371.49329	6.16	254.85	423.67	CWC
122	4047	Tshangalele	27.24, -10.75	Democratic Republic of the Congo	Africa	0.03102, 1119.03312	1.267	225.65	1126.03	GRanD
123	4485	Finchaa	37.36, 9.56	Ethiopia	Africa	0.01891, 2216.55235	0.65	196.13	2220.26	FAO
124	4989	Almatti	75.89, 16.33	India	Asia	0.05275, 504.12335	3.11	293.42	519.6	CWC
125	4707	Mangla	73.64, 33.15	Pakistan	Asia	0.20109, 320.1312	9.12	251	370.6	wikipedia
126	4836	Rana Pratap	75.58, 24.92	India	Asia	0.1391, 324.74	1.44	197.66	352.81	CWC
127	3014	Bagre	-0.55, 11.47	Burkina Faso	Africa	0.05719, 223.53693	1.7	255	238.12	literature
128	1991	Junin	-76.19, -10.98	Peru	South America	0.02312, 4079.83703	1.08	206.71	4084.62	WLDB
129	4881	Bargi Dam Reservoir	79.93, 22.95	India	Asia	0.08518, 401.51078	3.18	236.24	422.76	CWC
130	6686	Great Lake	146.73, -41.98	Australia	Oceania	0.40346, 969.53157	3.36	176	1040.54	GRanD
131	6800	Hawea	169.25, -44.61	New Zealand	Oceania	0.14631, 323.54085	2.18	150	345.49	GRanD
132	3676	Albufeira da Quiminha	13.79, -8.96	Angola	Africa	0.13121, 34.99206	1.56	129.05	51.93	GRanD
133	6629	Eucumbene	148.62, -36.13	Australia	Oceania	0.46484, 1097.64507	4.8	145.42	1165.24	wikipedia
134	1320	Falcon Reservoir	-99.17, 26.56	United States of America	North America	0.06972, 71.73912	3.88	311.84	93.48	TWDB

135	597	Lake Powell	-111.49, 36.94	United States of America	North America	0.1406, 1047.2	30	609.38	1127.76	wikipedia
136	4463	Dukan	44.96, 35.96	Iraq	Asia	0.18893, 462.6788	6.97	270	513.69	wikipedia
137	1230	Cedar Creek Reservoir	-96.07, 32.18	United States of America	North America	0.09423, 85.91971	0.8	133.03	98.15	TWDB
138	4041	Lake Maga	15.05, 10.83	Cameroon	Africa	0.01933, 309.62551	0.68	148.72	312.5	literature
139	5157	Pasak Chonlasit	101.08, 14.85	Thailand	Asia	0.05295, 33.58769	0.79	158.87	42	literature
140	6594	Fairbairn	148.06, -23.65	Australia	Oceania	0.13, 186.48395	2.29	179.43	209.81	wikipedia
141	6628	Hume	147.03, -36.11	Australia	Oceania	0.15399, 161.81633	3.04	201.9	192	wikipedia
142	4500	Kikuletwa	37.47, -3.82	United Republic of Tanzania	Africa	0.1, 677.01366	0.6	126.33	689.65	wikipedia
143	4958	Nizam sagar	77.93, 18.2	India	Asia	0.0893, 419.95709	0.5	92.75	428.24	CWC
144	6606	Victoria	141.28, -34.04	Australia	Oceania	0.16558, 7.52685	0.68	122	27.73	GRanD
145	1869	Grenada Lake	-89.77, 33.82	United States of America	North America	0.12614, 49.34905	1.54	128.29	65.53	Lakes Online
146	138	Canyon	-121.09, 40.18	United States of America	North America	0.68749, 1300.93547	1.61	108.39	1373.12	wikipedia
147	4638	Aras Dam Lake	45.4, 39.09	Azerbaijan	Asia	0.11845, 762.76554	1.35	145	779.94	wikipedia
148	4481	Khashm el-Girba	35.9, 14.93	Sudan	Africa	0.09342, 463.08227	1.3	125	474.76	wikipedia
149	370	Lake Cascade	-116.05, 44.52	United States of America	North America	0.16232, 1455.02068	0.85	101.98	1471.57	wikipedia
150	3695	Seitevare	18.57, 66.97	Sweden	Euro	0.62918, 419.1852	1.68	81	470.15	GRanD
151	4484	Yardi	40.54, 10.23	Ethiopia	Africa	0.33044, 533.59442	2.32	104.87	568.25	GRanD
152	119	Clear Lake Reservoir	-121.08, 41.93	United States of America	North America	0.19968, 1345.80215	0.65	100.36	1365.84	wikipedia
153	5196	Guanting Shuiku	115.6, 40.23	China	Asia	0.10764, 465.09336	4.16	130	479.09	GRanD

154	2953	Barrage Al Massira	-7.64, 32.47	Morocco	Africa	0.33916, 241.40761	2.76	80	268.54	wikipedia
155	1319	Venustiano Carranza	-100.62, 27.51	Mexico	North America	0.09456, 252.29236	1.31	150.56	266.53	literature
156	4471	Lake Hamrin	44.97, 34.12	Iraq	Asia	0.11963, 80.22516	4.61	228	107.5	literature
157	4826	Matatila	78.37, 25.1	India	Asia	0.10028, 297.22136	0.71	112.07	308.46	CWC
158	1263	Twin Buttes	-100.52, 31.37	United States of America	North America	0.49517, 576.78181	0.23	29.47	591.37	TWDB
159	4997	Somasila	79.3, 14.49	India	Asia	0.17144, 74.32045	1.99	153.17	100.58	CWC
160	5183	Hongshan Reservoir	119.7, 42.75	China	Asia	0.23268, 422.07692	2.56	66.9	437.64	GRanD
161	6583	Lake Ross	146.74, -19.41	Australia	Oceania	0.11178, 32.60497	0.417	82	41.77	wikipedia
162	4978	Yeleru Reservoir	82.08, 17.3	India	Asia	0.58856, 57.51076	0.51	49.36	86.56	CWC
163	4696	South Surkhan Reservoir	67.63, 37.83	Uzbekistan	Asia	0.33795, 397.79868	0.8	40.26	411.41	GRanD
164	5287	Zhaopingtai Reservoir	112.77, 33.73	China	Asia	0.35804, 157.62551	0.71	46.5	174.27	GRanD

[*] The 164 reservoirs include 13 regulated natural lakes, whose IDs are 1, 6, 20, 23, 33, 37, 44, 62, 73, 102, 131, 150, and 151.

[1] GRanD: Global Reservoir and Dam Database (Lehner et al., 2011). All of the geographical location information is adopted from GRanD.

[2] a and b are the coefficients used in the A-E relationship equation: h=a*A+b, where h and A are elevation (m) and area (km^2), respectively.

[3] List of full names: FAO- Food and Agriculture Organization; USBR- United States Bureau of Reclamation; CWC- Central Water Commission of India; USACE- United States Army Corps of Engineers; TWDB- Texas Water Development Board; WLDB- World Lake Database.

Appendix B- Example codes (in MATLAB and Python) for reading VNP28C2/VJ128C2 hdf files

1) MATLAB code for reading the VNP28C2/VJ128C2 hdf file

```
%%%   To get information about the file  %%%

    file = ('VNP28C2.A2012017.002.2021217150026.h5'); OR file =
    ('VJ128C2.A2020017.002.2021217150026.h5');

    h5disp(file);  %% To get information about the file %%

    x=h5read(file,'/lakes/');

    %%% Read the data %%%

    lake_id=x.lake_ID;
    longitude= x.lake_longitude;
    latitude=x.lake_latitude;
    area=x.lake_area;
    elevation=x.lake_elevation;
    storage=x.lake_storage;
```

2) Python code for reading the VNP28C2/VJ128C2 data from the hdf file

```
import numpy as np
import h5py
import pandas as pd

##input the H5 filename and open it
filename = "VNP28C2.A2012017.002.2021217150026.h5" OR “VJ128C2.A2020017.002.2021217150026.h5”

with h5py.File(filename, "r") as f:
   ##read the datasets
   a_group_key = list(f.keys())[0]
   data = pd.DataFrame(np.array(f[a_group_key]))
   ##set Lake_id as the ID
   data=data.set_index('lake_ID')
   print (data)
```

Appendix C- Example codes (in MATLAB and Python) for reading VNP28C3/VJ128C3 hdf files

1) MATLAB code for reading the VNP28C3/VJ128C3 hdf file

```
clc;clear all;

%% file name

file= ('VNP28C3.A2012001.002.2022138033759.h5'); OR file = (‘VJ128C3.A2020017.002.2021217150026.h5’);
x=h5read(file,'/lake_evaporation/');

%%% Read the data %%%

lake_id=x.lake_ID;
longitude= x.lake_longitude;
latitude=x.lake_latitude;
area=x.lake_area;
elevation=x.lake_elevation;
storage=x.lake_storage;
evaporation_rate=x.lake_evap_rate;
evaporation_volume=x.lake_evap_vol;
```

2) Python code for reading the VNP28C3/VJ128C3 data from the hdf file

```
import numpy as np
import h5py
import pandas as pd

##input the H5 filename and open it
filename = "VNP28C3.A2012017.002.2021217150026.h5" OR “VJ128C3.A2020017.002.2021217150026.h5”

with h5py.File(filename, "r") as f:
   ##read the datasets
   a_group_key = list(f.keys())[0]
   data = pd.DataFrame(np.array(f[a_group_key]))
   ##set Lake_id as the ID
   data=data.set_index('lake_ID')
   print (data)
```

Appendix D – Outlier removal and gap filling script

To use this attached script, copy the code below to a file (e.g., process.py). Change necessary paths and then run the script using python (with prerequisite packages installed). Output file (VNP28C2_outlier_removed.csv) contains values for all variables after removing outliers.

```
import matplotlib.pyplot as plt
plt.rcParams.update({'font.size': 20})

import numpy as np
import pandas as pd
import glob
from datetime import datetime

import h5py
import os

#read .h5 files
output = []
lake_num   = 164
lake_ids   = range(1,165)
key_word = 'VNP' #VNP or VJ1
files = glob.glob('./VNP28C2/VNP28C2*.h5', recursive=True) #look for input files

for fname in files:
    #print(fname)
    mdate = fname.split('/')[-1].split('.')[1]

    with h5py.File(fname, 'r') as hf:
        mdata=hf['lakes']
        for i in range(lake_num):
            output.append([mdate, lake_ids[i], mdata[i][3]])#lake area

output_df = pd.DataFrame(output, columns=['date', 'lake_id', 'pge698_area'])
display(output_df)
```

```
date_start = datetime.strptime('1/1/2012', "%m/%d/%Y")  ## any date before the start of VIIRS is good

def sum_area(group):

    if np.min(group['pge698_area'] == 0):
        c = 0
    else:
        c = np.sum(group['pge698_area'])

    return c

df = output_df.groupby(['date', 'lake_id']).apply(lambda grp: sum_area(grp)).reset_index()
df = df.rename({0: 'pge698_area'}, axis=1).replace({0: np.nan})
df['date'] = pd.to_datetime(df['date'], format='A%Y%j', errors='coerce')
df['date_int'] = df['date'].apply(lambda x: (x-date_start).days)
df = df.sort_values(['lake_id', 'date'])
display(df)
df.to_csv(key_word+'28C2.csv')
## Remove Outliers

def remove_bias(values, dates, wz):

    ## values: area or storage time series
    ## dates: corresponding dates (in integer)
    ## wz: windows size for calculating the moving average to determine the outliers

    weight = np.arange(1,wz,1)[int(wz/2):]
    weight = np.concatenate((weight, np.ones(len(values)-wz+1)*wz, np.flip(weight, 0)))/wz

    nol = 2
    limit = 50
    while nol > 0 and limit > 0:
        wt_ma = np.convolve(values, np.ones((wz,))/wz, mode='same')

        bias = values - wt_ma/weight
        b_avg = np.average(bias)
        b_sd = np.std(bias)

        outlier = [0 if (x >= b_avg+3*b_sd or x <= b_avg-3*b_sd) else 1 for x in bias]
        inreduce = nol - (len(outlier) - np.sum(outlier))
        nol = len(outlier) - np.sum(outlier)
        limit = limit - 1
        #print(nol)

        n_dates = [x for x,y in zip(dates, outlier) if y==1]
        n_enh = [x for x,y in zip(values, outlier) if y==1]
        if len(n_dates) >= 5:
            values = np.interp(dates, n_dates, n_enh)
        else:
            return dates, values

        if inreduce == 0 and nol <=2:
            break
    return n_dates, n_enh
```

```
lake_info_data = pd.read_table('./viirs_lake_info.txt', sep='  ')

output_intp = []
for lake_id, group in df.groupby('lake_id'):

   lake_info = lake_info_data[lake_info_data['LAKE_ID'] == lake_id].iloc[0]
   sto_cap_km3, area_cap_km2, elev_cap_m, a, b = lake_info[['sto_cap_km3','area_cap_km2', 'elev_cap_m', 'a', 'b']]

   group_valid = group[(group['pge698_area'] > 0)]

   if len(group_valid) > wz:
      areas = group_valid['pge698_area'].values
      dates = group_valid['date_int'].values
      dates_new, areas_new = remove_bias(areas, dates, wz)

      areas_intp = np.interp(group['date_int'], dates_new, areas_new)
   else:
      areas_new = group[(group['pge698_area'] > 0)]['pge698_area'].values
      dates_new = group[(group['pge698_area'] > 0)]['date_int'].values
      areas_intp = np.interp(group['date_int'], dates_new, areas_new)

   elevs_intp = [a*area+b for area in areas_intp]
   stors_intp = [sto_cap_km3 - (area_cap_km2+area)*(elev_cap_m-elev)/2000.
              for area, elev in zip(areas_intp, elevs_intp)]

   intp_flag = [0 if area==area_new else 1 for area, area_new in zip(group['pge698_area'], areas_intp)]

   group_new = pd.DataFrame({'date': group['date'], 'lake_area': areas_intp, 'lake_elevation': elevs_intp,
                  'lake_storage': stors_intp, 'intp_flag': intp_flag})
   group_new['lake_id'] = lake_id
   output_intp.append(group_new)

df_intp = pd.concat(output_intp, axis=0).reset_index(drop=True)
display(df_intp)
df_intp.to_csv(key_word+'28C2_outlier_removed.csv', index=False)
```