Snowcast Showdown

Model report

Ultimate Hydrology Team

Evgeniy Malygin, Maxim Kharlamov, Ivan Malygin, Maria Sakirkina, Ekaterina Rets

SUMMARY

Here we present combined physically-based and machine learning approaches to SWE (snow water equivalent) prediction based on a wide range of data (in-situ, remote sensing data, results of general circulation modeling) using different SOTA implementations of Gradient Boosting Machine algorithm: XGBoost, LightGBM, and CatBoost, and their ensembles. Here are the main results of our research. First of all, as a result of feature research and feature engineering, a dataset of 120 features was created, describing the physics of the process, as well as natural-climatic features of the region under study. Next, we trained and tested an ML model for snowpack forecast in the Colorado and Sierra Nevada Mountains (USA), using different implementations of the Gradient Boosting technique. Besides, the robustness of the model was checked by validation and hidden sampling and confirmed by low RMSE metric values (3.66) on the public leaderboard. Moreover, we developed an end-to-end solution and automated real-time forecast pipeline to reproduce the forecast for each week. Hence, the developed solution is universal and can be used both for operational monitoring and for forecasting in advance.

DATA PROCESSING

Data sources

Following data sources approved in the competition set up where used for this solution:

- SNOTEL
- CDEC
- MODIS Terra MOD10A1
- HRRR
- Copernicus DEM

Among approved snow cover data products we have opted to MODIS Terra MOD10A1 as the spatial resolution (500*500 m) was the closest to the projected swe data net resolution. That reduces computational costs and possible errors in aggregating the more detailed information from such data sources as Landsat and Sentinel. Secondly,, the MODIS data provides the best estimation for the given date, that reduces data noise due to the cloud cover. Thirdly, MODIS has a daily resolution that suits our solution the most. Moreover, MODIS provides pre-calculated characteristics and requires minimum data processing. The MODIS Aqua MYD10A1 was not included as the data source as it is mostly focused on polar regions.

The choice of Copernicus and HRRR dataset was based on the design of the solution and physically-based approach to feature set structure (see Feature engineering section).

Land Cover Gridded Map and FAO-UNESCO Global Soil Regions Map were tested on the first stage of development to characterize local conditions of snow redistribution (such as canopy interception) but was eliminated further due to low feature importance (see Feature engineering section)

Feature engineering

Feature set was designed including data directly characterizing snowcover properties (MODIS data, SNOTEL and CDEC, swe characteristics from HRRR) as well as a set of features serving as predictors of physical processes underlying snow cover evolution.

Snow evolution processes include snow accumulation, snow interception, redistribution (wind, avalanching), snowpack internal processes (including mass- and energy transfer, snow metamorphism), snow ablation and water yield.

Using the approved dataset a set of features was generated characterizing each of this processes:

1. Considered predictors of snow accumulation include:
    - precipitation sum
    - sum of solid precipitation (precipitation on days with mean daily negative temperature)
    - negative temperature sums
2. To characterize snow interception by vegetation landuse and soil data (Land Cover Gridded Map, FAO-UNESCO Global Soil Regions Map) was incorporated on the first steps of model training, and further removed due to low feature importance
3. No features to adress wind redistribution and avalanching directly were considered. However, "relief" features (different characteristics of altitude, slope, aspect, curvature etc) added to the feature set could indirectly reflect differences in swe distribution due to these processes
   The heat balance of the surface of melting snow can be expressed as:
   $w = S_{down}\,(1-A) + E_{lrd} - E_{lru} \pm LE \pm H \pm Q_m \pm Q_{snp} \pm Q_{act} + Q_{rain}$
4. where w is the net energy flux on the surface, W m^{-2}; S_{down} is the downward shortwave radiation flux, W m^{-2}; A – is surface albedo; E_{lru} is outgoing long-wave radiation, W m^{-2}; E_{lrd} is the counter radiation of the atmosphere, W m^{-2}; LE is the turbulent latent - heat flux density, W m^{-2}; H is the turbulent sensible - heat flux density, W m^{-2}, Q_{snp} – energy change due to snowpack processes and processes on the snow-ice boundary, Q_{rain}– energy flux due to liquid precipitation in snow

To characterize each of the fluxes following meteorological characteristics are added to feature set:

- downward radiation
- relief parameters characterizing spatial distribution of solar direct and diffuse radiation in mountainous areas (slope, aspect, elevation,)
- snow albedo (Modis data)
- air temperature, wind speed (characteristics of turbulent heat fluxes)
- liquid precipitation

Most of the features (attributes) can be classified into 4 main groups according to the different data sources. All of them are used in the training and inference code.

1. Ground Measures

Spatial-temporal interpolation based on Ground measure data (SNOTEL, CDEC) by Random Forest model (**rf_org_value_v2**). 4 parameters are used:

- lat
- lon
- alt
- time (day_of_year)

2. Meteorological parameters and their aggregates for 3 time periods from atmospheric model High-Resolution Rapid Refresh(HRRR NOAA):

Average for the day of the forecast

Aggregates by weekly period (week before the date of the forecast)

Aggregates by winter period (from December of the previous year)

The High-Resolution Rapid Refresh (HRRR) is a NOAA real-time 3km resolution, hourly updated, cloud-resolving, convection-allowing atmospheric model, initialized by 3km grids with 3km radar assimilation.

Data is updated weekly for historical and future predictions. We connected to the Approved data access location.

List of created features:

Daily period:

Name in the dataset	**Description**	**Unit**
si10	daily mean wind speed	m/s
dswrf	daily mean radiation	W/m²
t2m	daily mean temperature	°C
tp	daily sum precipitation	mm
tp_pls	daily sum of liquid precipitation	mm
tp_mns	daily sum of solid precipitation	mm
t2m_pls	daily sum of warm temperature	°C
t2m_mns	daily sum of cold temperature	°C
rain_nrg	daily sum of t2m_pls*tp_pls	mm · °C

Weekly period:

Name in the dataset	**Description**	**Unit**
temp_mean	mean of temperature over the last 7 days	°C
temp_sum	sum of temperature over the last 7 days	°C
temp_sum_cold	sum of cold temperature over the last 7 days	°C
temp_sum_warm	sum of warm temperature over the last 7 days	°C
temp_sum_cold_hours	count of cold hours over the last 7 days	°C
temp_sum_warm_hours	count of warm hours over the last 7 days	°C
tp_mean	mean of precipitation over the last 7 days	mm
tp_sum	sum of precipitation over the last 7 days	mm
tp_sum_liquid	sum of liquid precipitation over the last 7 days	mm
tp_sum_solid	sum of solid precipitation over the last 7 days	mm

Name in the dataset	Description	Unit
rain_enrg	last 7 days sum of t2m_pls*tp_pls	mm · °C
thaw_count	count of thaws over the last 7 days	-
dswrf_mean	mean of downward solar radiation over the last 7 days	W/m²
dswrf_sum	sum of downward solar radiation over the last 7 days	W/m²
si10_mean	mean of wind speed over the last 7 days	m/s
si10_sum	sum of wind speed over the last 7 days	m/s
sdwe_mean	mean of swe over the last 7 days	mm
sdwe_sum	sum of swe over the last 7 days	mm
sdwe_range	delta swe per week	mm
sdwe_last	swe 7 days ago	mm
sdwe_first	swe for real-time evaluation date	mm

Winter period:

Name in the dataset	Description	Unit
si10_cumsum	cumsum of wind speed from the beginning of winter	m/s
si10_mean_sws	cummean of wind speed from the beginning of winter	m/s
dswrf_cumsum	cumsum of radiation from the beginning of winter	W/m²
dswrf_mean_sws	cummean of radiation from the beginning of winter	W/m²
t2m_cumsum	cumsum of temperature from the beginning of winter	°C
t2m_mean_sws	cummean of temperature from the beginning of winter	°C
tp_cumsum	cumsum of precipitation from the beginning of winter	mm
tp_mean_sws	cummean of precipitation from the beginning of winter	mm
tp_pls_cumsum	cumsum of liquid precipitation from the beginning of winter	mm
tp_pls_mean_sws	cummean of liquid precipitation from the beginning of winter	mm
tp_mns_cumsum	cumsum of solid precipitation from the beginning of winter	mm
tp_mns_mean_sws	cummean of solid precipitation from the beginning of winter	mm

Name in the dataset	Description	Unit
t2m_pls_cumsum	cumsum of warm temperature from the beginning of winter	°C
t2m_pls_mean_sws	cummean of warm temperature from the beginning of winter	°C
t2m_mns_cumsum	cumsum of cold temperature from the beginning of winter	°C
t2m_mns_mean_sws	cummean of cold temperature from the beginning of winter	°C
rain_nrg_cumsum	cumsum of rain_nrg from the beginning of winter	mm · °C
rain_nrg_mean_sws	cummean of rain_nrg from the beginning of winter	mm · °C
si10_m7	daily mean wind speed 7 days ago	m/s
dswrf_m7	daily mean radiation 7 days ago	W/m²
t2m_m7	daily mean temperature 7 days ago	°C
tp_m7	daily sum precipitation 7 days ago	mm
tp_pls_m7	daily sum of liquid precipitation 7 days ago	mm
tp_mns_m7	daily sum of solid precipitation 7 days ago	mm
t2m_pls_m7	daily sum of warm temperature 7 days ago	°C
t2m_mns_m7	daily sum of cold temperature 7 days ago	°C
rain_nrg_m7	daily sum of t2m_pls*tp_pls 7 days ago	mm · °C
si10_cumsum_m7	cumsum of wind speed from the beginning of winter 7 days ago	m/s
si10_mean_sws_m7	cummean of wind speed from the beginning of winter 7 days ago	m/s
dswrf_cumsum_m7	cumsum of radiation from the beginning of winter 7 days ago	W/m²
dswrf_mean_sws_m7	cummean of radiation from the beginning of winter 7 days ago	W/m²
t2m_cumsum_m7	cumsum of temperature from the beginning of winter 7 days ago	°C
t2m_mean_sws_m7	cummean of temperature from the beginning of winter 7 days ago	°C
tp_cumsum_m7	cumsum of precipitation from the beginning of winter 7 days ago	mm
tp_mean_sws_m7	cummean of precipitation from the beginning of winter 7 days ago	mm
tp_pls_cumsum_m7	cumsum of liquid precipitation from the beginning of winter 7 days ago	mm
tp_pls_mean_sws_m7	cummean of liquid precipitation from the beginning of winter 7 days ago	mm

Name in the dataset	Description	Unit
tp_mns_cumsum_m7	cumsum of solid precipitation from the beginning of winter 7 days ago	mm
tp_mns_mean_sws_m7	cummean of solid precipitation from the beginning of winter 7 days ago	mm
t2m_pls_cumsum_m7	cumsum of warm temperature from the beginning of winter 7 days ago	°C
t2m_pls_mean_sws_m7	cummean of warm temperature from the beginning of winter 7 days ago	°C
t2m_mns_cumsum_m7	cumsum of cold temperature from the beginning of winter 7 days ago	°C
t2m_mns_mean_sws_m7	cummean of cold temperature from the beginning of winter 7 days ago	°C
rain_nrg_cumsum_m7	cumsum of rain_nrg from the beginning of winter 7 days ago	mm · °C
rain_nrg_mean_sws_m7	cummean of rain_nrg from the beginning of winter 7 days ago	mm · °C

3. MODIS Terra MOD10A1 satellite imagery product (Snow Cover Daily L3 Global 500m SIN Grid).

The snow cover algorithm calculates NDSI for all land and inland water pixels in daylight using MODIS band 4 (visible green) and band 6 (shortwave near-infrared). We used three main features:

Name in the product	Name in the dataset
NDSI_Snow_Cover	sc
NDSI	ndsi1
Snow_Albedo_Daily_Tile	sa1

Detailed information and descriptions of the values are given here https://nsidc.org/data/MOD10A1.

As the MODIS data is given in equal-area sinusoidal projection a task of reprojection was carried out. On the first step raster data was reprojected in WGS84, however it led to a substantial amount of erroneous values and “NoData” cells in the resulting raster. In order to solve that problem, centroids of the training and testing nets were reprojected in equal-area sinusoidal projection. The NDSI, snow albedo, snow cover values from the MODIS rasters were attributed to the centroids of the training and testing nets for each date. Gapes were filled with “missing” values according to the official User Guide (https://nsidc.org/data/MOD10A1). Additional engineering was not needed as the gradient boosting solution used in the study is sustainable to outliers, missing data and not-normalized data.

Data is updated daily, but we use only the necessary dates. We connected to Approved data access location: https://modissa.blob.core.windows.net/modis-006 .

4. Relief parameters, produced on Copernicus DEM (90 meter resolution).

Relief parameters were calculated by usage of Copernicus DEM for each grid cell, as well as their aggregates in 200 and 500 m geodetic buffers (min, max, average, median) with the opensource libraries (GDAL, RichDEM). These attributes were calculated in advance for the whole grid and do not change over time. We used open source GDAL and RichDEM libraries to calculate slope degrees taking into account the latitude, aspects, profile curvatures, curvatures, planform curvatures, TRI (topographic ruggedness index).

Additional features such as Vector Ruggedness Measure (calculated using open source GRASS GIS Application) and distance to nearest stream (calculated based on filled DEM flow accumulation raster by means of r.stream.extract) didn't contribute much to feature importance structure while being complicated to be executed outside GIS Applications. Accordingly, these features were eliminated from the feature set.
We used Approved data access location:

https://planetarycomputer.microsoft.com/api/stac/v1/collections/cop-dem-glo-90

List of created features:

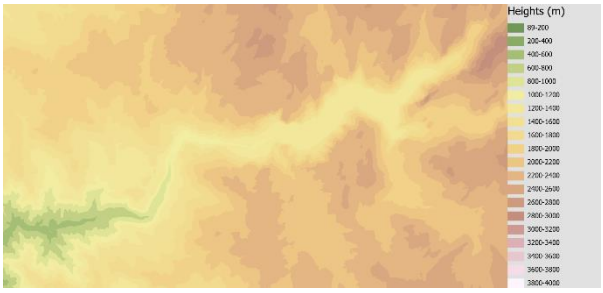
Name in the dataset	**Description**
alt* alt_min_200 alt_max_200 alt_mean_200 alt_median_200 alt_min_500 alt_max_500 alt_mean_500 alt_median_500	Height, meters alt_{agg}_{buffer}, where {agg} is aggregate function (min, max, mean, median), {buffer} is buffer radius in meters.
slope* slope_mean_200 slope_median_200 slope_mean_500 slope_median_500	Slope, degrees slope_{agg}_{buffer}, where {agg} is aggregate function (min, max, mean, median), {buffer} is buffer radius in meters. Without suffix value at point. https://richdem.readthedocs.io/en/latest/python_api.html
aspect* aspect_mean_200 aspect_median_200 aspect_mean_500 aspect_median_500	Aspect aspect_{agg}_{buffer}, where {agg} is aggregate function (min, max, mean, median), {buffer} is buffer radius in meters. https://richdem.readthedocs.io/en/latest/python_api.html
curv_prof* curv_prof_mean_200, curv_prof_median_200 curv_prof_mean_500 curv_prof_median_500	Profile curvature curv_prof_{agg}_{buffer}, where {agg} is aggregate function (min, max, mean, median), {buffer} is buffer radius in meters. https://richdem.readthedocs.io/en/latest/python_api.html
curv* curv_mean_200 curv_median_200	Curvature curv_{agg}_{buffer}, where {agg} is aggregate function (min, max, mean, median), {buffer} is buffer radius in

Name in the dataset	Description
curv_mean_500 curv_median_500	meters. https://richdem.readthedocs.io/en/latest/python_api.html
curv_plan* curv_plan_mean_200 curv_plan_median_200 curv_plan_mean_500 curv_plan_median_500	Planform curvature curv_plan_{agg}_{buffer}, where {agg} is aggregate function (min, max, mean, median), {buffer} is buffer radius in meters. https://richdem.readthedocs.io/en/latest/python_api.html
tri* tri_mean_200 tri_median_200 tri_mean_500 tri_median_500	TRI (Terrain Ruggedness Index) tri_{agg}_{buffer}, where {agg} is aggregate function (min, max, mean, median), {buffer} is buffer radius in meters. https://gdal.org/programs/gdaldem.html

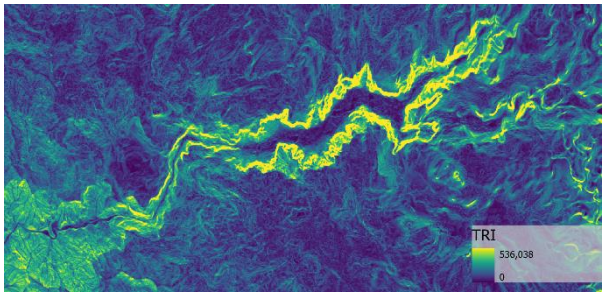
* without suffix -- value at the cell centoid

Visualization of created DEM features:

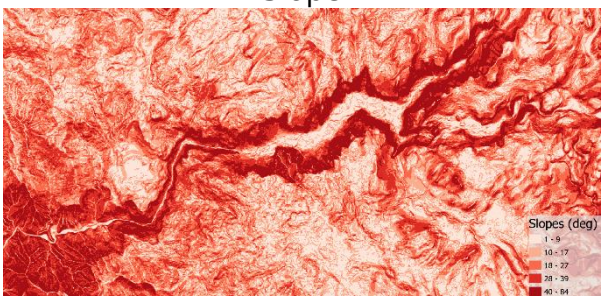
DEM



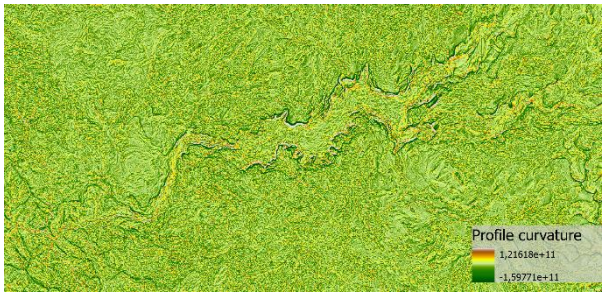
TRI



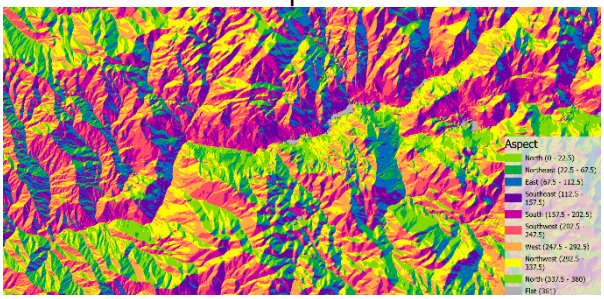
Slope



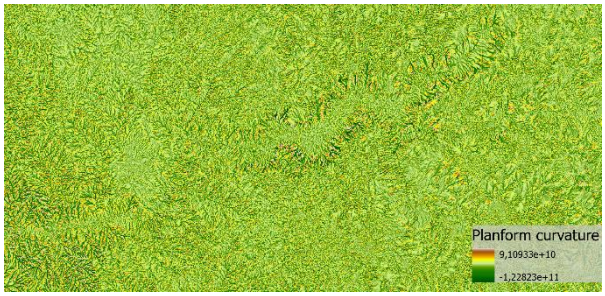
Profile curvature



Aspect



Planform curvature



5. Additional sources:
   - Coordinates of cell centroids
   - Height in the cell centroid
   - Ordinal date from the beginning of the year (Day of year)

MACHINE LEARNING MODEL

Algorithm selection

An accurate estimation of snow water equivalent (swe) is extremely important for a wide range of tasks from operational forecast of river runoff to assessment and prediction of climatically-induced changes in water cycle (Bair et al., 2018, IPCC, 2022, Rheinheimer et al., 2016, Şensoy and Uysal, 2012 etc.). Current models of snow cover evolution provide opportunities for detailed simulation of snow accumulation, vegetation interception, redistribution, internal processes of snowpack evolution and snowmelt (Krinner et al., 2018, Vionnet et al., 2021, Wever et al., 2014). However, while using for operational practice on the regional and global scale we face a trade off between accuracy of prediction for each point and time resources alongside with computational performance.

Here we aimed to combine physically-based and machine learning approaches to swe prediction using different sources of data (in-situ, remote sensing data, results of general circulation modeling).

The choice of the model's algorithm is determined by the characteristics of the SWE data. Target values are collected from two different sources: regular observations at the stations (time series data) and sporadic remote sensing measurements (time series snapshots). Observation points are located irregularly and sparsely, and snapshots are presented only for a short period before the snow melting. The dataset is heterogeneous, so the use of time-series appropriate neural networks (LSTM or CNN) would lead to reliable results only for a limited part of the dataset, being unsuitable for estimating values at the stations. Also, the choice of the model is determined by the dataset's content, which combines different categories of features. At the same moment, it has different features: temporal distributed, spatially immutable over time, with a distinguishing physical meaning described by different units of measurement.

In this competition, we applied gradient boosting models to predict the SWE values. Gradient boosting is an extremely popular machine learning algorithm that has proven successful across many domains. Gradient boosting has several advantages, for example, it is insensitive to scaling (and in general to any monotonic transformations) of feature values, and, therefore, valuable for the complex content of our dataset. Next, this algorithm is good at sensing differences in data composition and offers universal models for multiple sources, averaging model quality for the entire dataset without leading to overfitting. Also, there are a lot of methods to interpret gradient boosting results, which are important for understanding and explaining the processes under study. Clearly, we can recognize the importance and thresholds of every feature inside the model.

Also, we tried to use openAMUNDSEN snow and hydroclimatological modeling framework. openAMUNDSEN is a fully distributed model, designed primarily for resolving the mass and energy balance of snow and ice covered surfaces in mountain regions. Typically, it is applied in areas ranging from the point scale to the regional scale (i.e., up to some hundreds to thousands of square kilometers), using a spatial resolution of 10–100 m and a temporal resolution of 1–3 h.

First test has shown good level of representation of the SWE observational data for several sites (Fig. 1-2). However, these experiments were not included in the final solution due to a number of limitations. The model required creation of the separate dataset and calibration for each region (setting temperature and precipitation gradients and other physical parameters that affect the processes of SWE evolution), which complicates the calculations. openAMUNDSEN is written in Python and has poor optimization, which leads to low calculation speed. For example, one section of the Sierra Nevada region for one year was calculated about 9 hours. The inference time of the model was limited to 8 hours. Also, the source code of this framework contained bugs.

To sum it up, the developed solution could benefit from incorporation of physically-based model into the model framework. However, further study and development is needed for successful implementation.

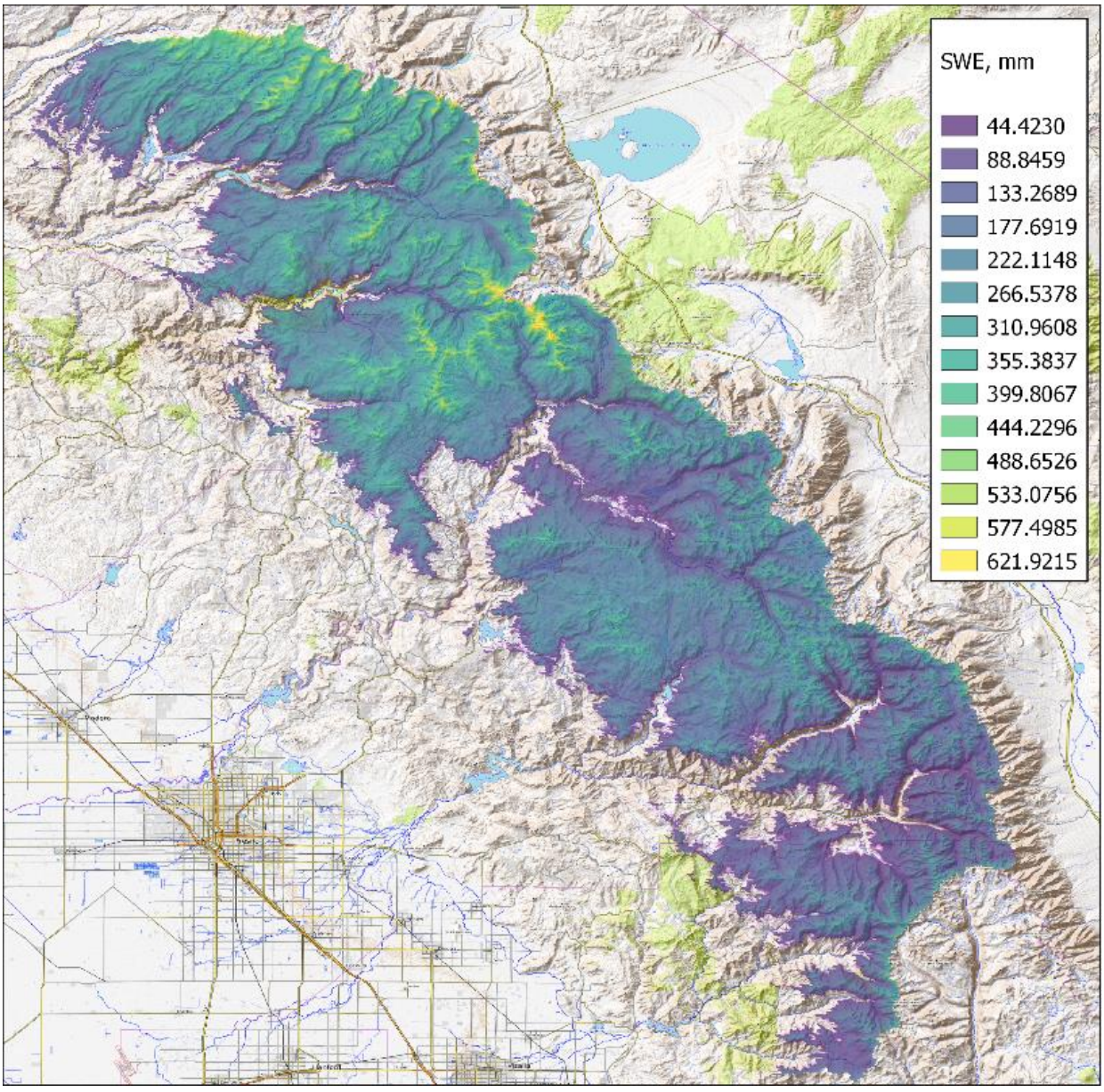


Figure 1 openAMUNDSEN predicted SWE

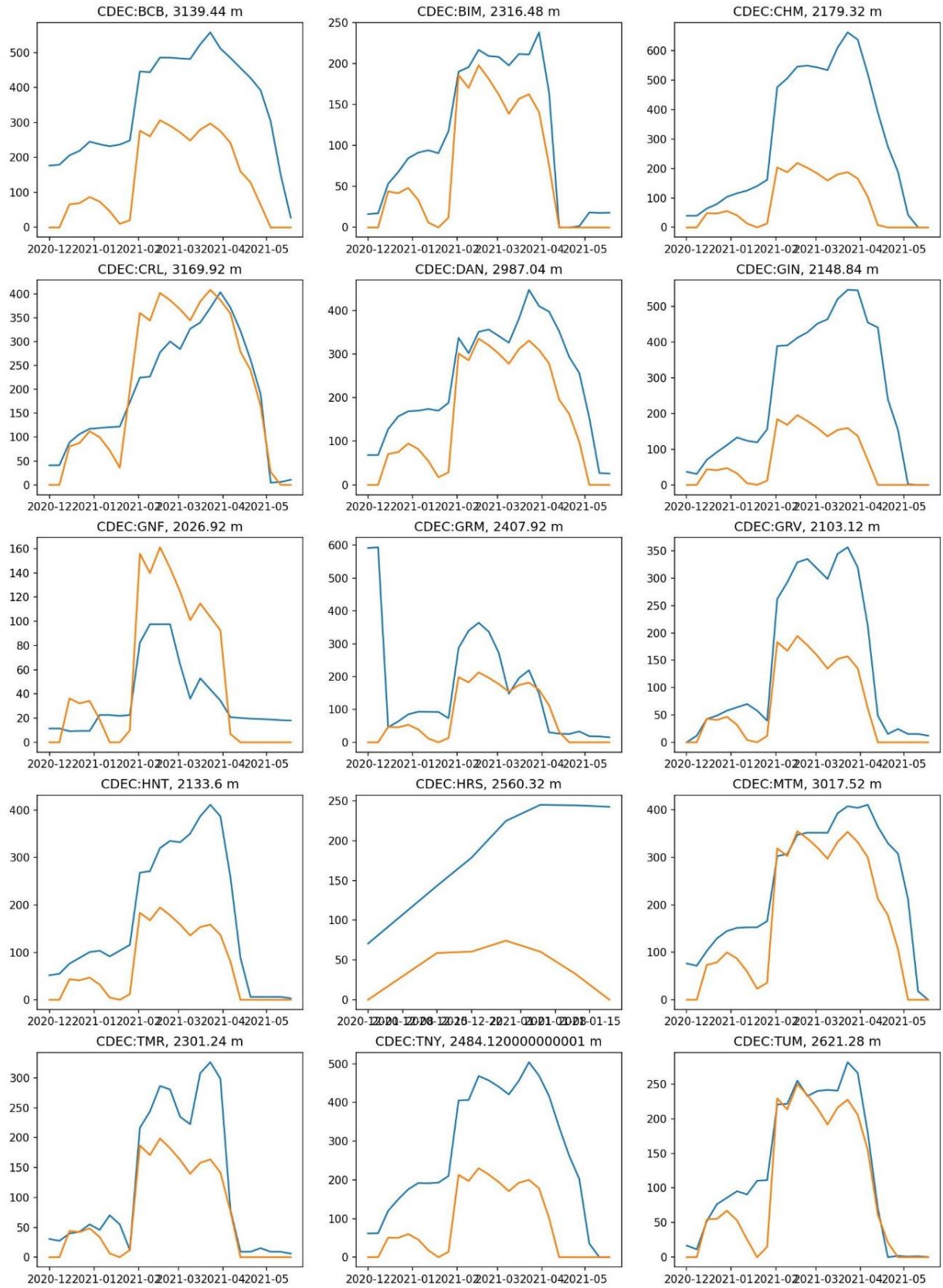


Figure 2 Comparing of the openAMUNDSEN simulated (orange) and CDEC observed (blue) SWE values, mm

Model description

Our final model is shown on Figure 3 below. It consists of several layers.

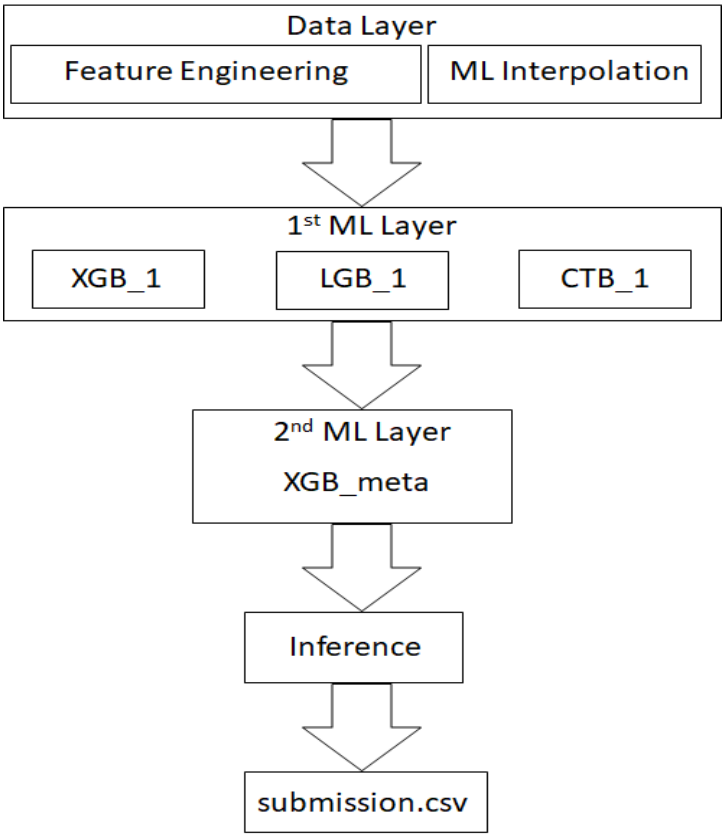


Figure 3 Model scheme

The first one includes data preparation and ML Interpolation process. As an ML model for interpolation initial SWE measurements from stations we use Random Forest and add this result as a feature in the main feature space. Next two layers are the main stacking model. At the first ML layer we use different SOTA implementations of Gradient Boosting Machine algorithm: XGBoost, LightGBM and CatBoost. Next ML layer is a meta-model which is based on predictions from the previous layer. Meta regressor is another instance of XGBoost. In order for the stacking process to be cross-validated at the second level, we use an OOF-approach. It's needed for meta-model hyperparameter tuning and for score calculation without overfitting effect at the second layer.

After all models complete it is possible to infer predictions and prepare a submission file.

Training, validation, and testing

Available observations were divided for training and validation by random sampling. To validate the model we used k-fold cross-validation whis 5 folds. To evaluate model performance against validation RMSE metric was used. We used data for 2020-2021 years for testing. Summarize validation results are presented in the table below:

	XGBoost	CatBoost	LightGBM	Meta-Model
CV Score (RMSE)	3.75	3.65	3.78	3.49

Explainability

The stability and consistency of the top of the feature importance structure for all three Gradient Boosting models and from one experiment set up to another indicates potential interpretability of the developed solution (Fig. 4-6).

The top features include first of all those features directly characterizing snow cover properties: **rf_org_value_v2** (spatial-temporal interpolation based on Ground measure data (SNOTEL, CDEC), **ndsi1** and **sc** (MODIS Terra MOD10A1). However, first 1-4 places in terms of feature importance in all cases include as well indirect predictors of swe: first of all, seasonal cumulative sum and average solid precipitation (**tp_mns_mean_sws, tp_mns_cumsum**) (calculated from HRRR) characterizing snow accumulation process. At the same time swe characteristics from HRRR show significantly less importance than solid precipitation sums that outlines low quality of swe characteristics prediction by HRRR, that was additionally confirmed by the comparison of HRRR swe data to ground measurements.

Among features characterizing snow ablation processes the most important are seasonal cumulative sum and average values of air temperature and the mean seasonal value of solar radiation (t2m_pls_mean_sws, t2m_pls_cumsum, dswrf_meam_sws). Relative higher importance of the air temperature characteristics indicates an important role of turbulent heat fluxes in snowmelt processes. It can as well be a better characteristic of seasonal effects in snow evolution, given a distinct two-phase nature of the process (accumulation and ablation periods). This assumption is supported as well by the relatively high importance of the feature **dayoftheyear.**

In all cases cumulative sums and long-term averages of solid precipitation, solar radiation and air temperature characteristics give a more important signal than daily values, that reflects the inertness of swe dynamics both during accumulation and ablation period.

Higher importance of *latitude* and *longitude* compared to altitude characteristics outlines prevalence of spatial differences in geographic and climatic conditions over the study area compared to altitude effects. Moreover, latitude is always more important than longitude, which can be interpreted as a more intense gradient of change in snow cover forming conditions in the south-north direction compared to west-east.

As the study is carried out for the mountainous region, the influence of vertical gradients of air temperature and precipitation is taken in account by different characteristics of altitude, among which those with a 200 or 500 m buffer produce a more distinct signal (**alt_mean_200, alt_median_200, alt_min_200, alt_max_200, alt_mean_200, alt_median_200, alt_min_200, alt_max_500**). The combination of these different characteristics of altitude can as well account for conditions for local snow redistribution processes (lake snow drift due to local winds and avalanching).

Features characterizing spatial differences in incoming solar radiation on the surface of snow cover (**aspect_median_500, aspect_median_200, slope_mean_200**) display substantial level of importance, especially the characteristics of the surface aspect.

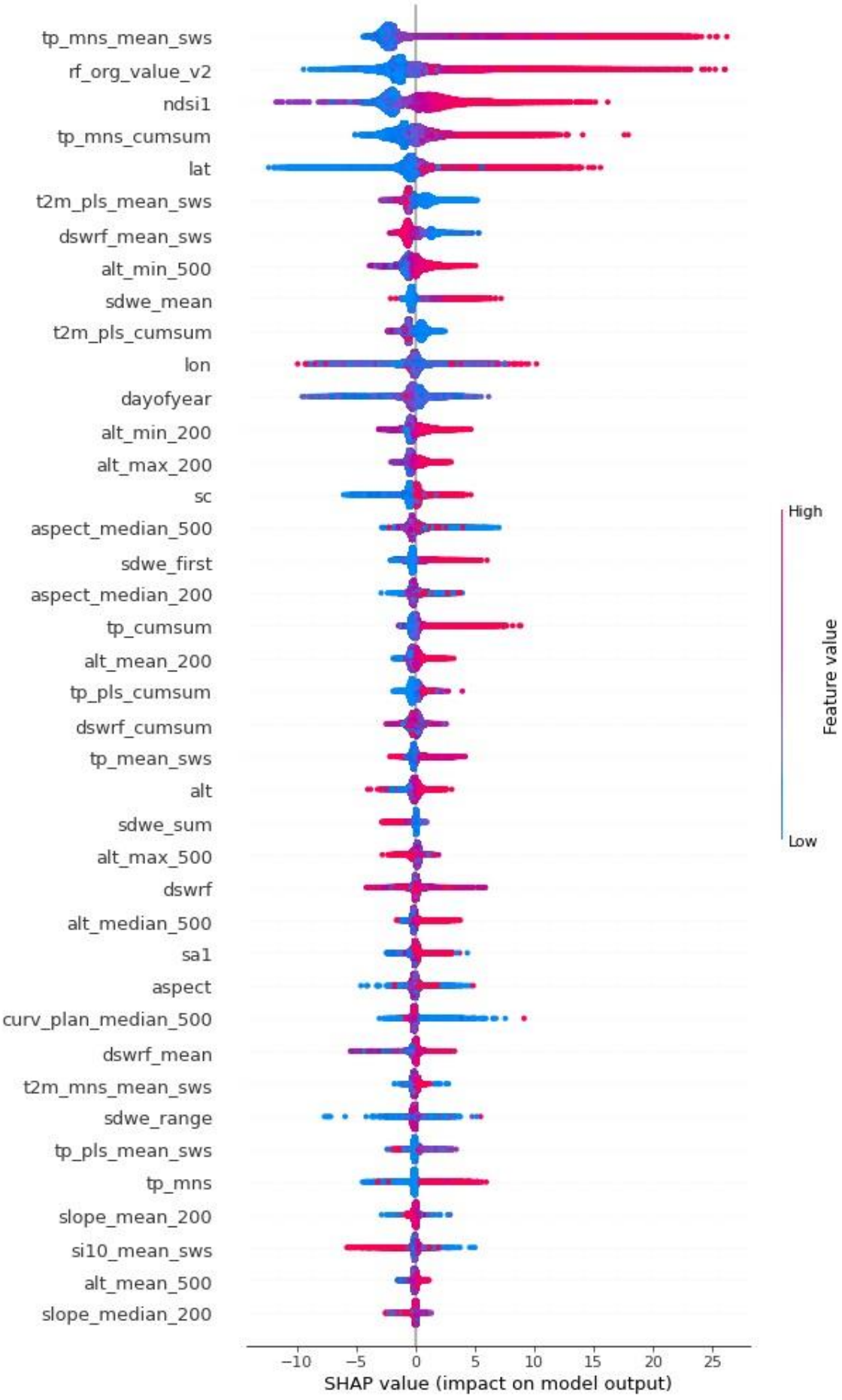


Figure 4 Feature importance structure in XGBoost

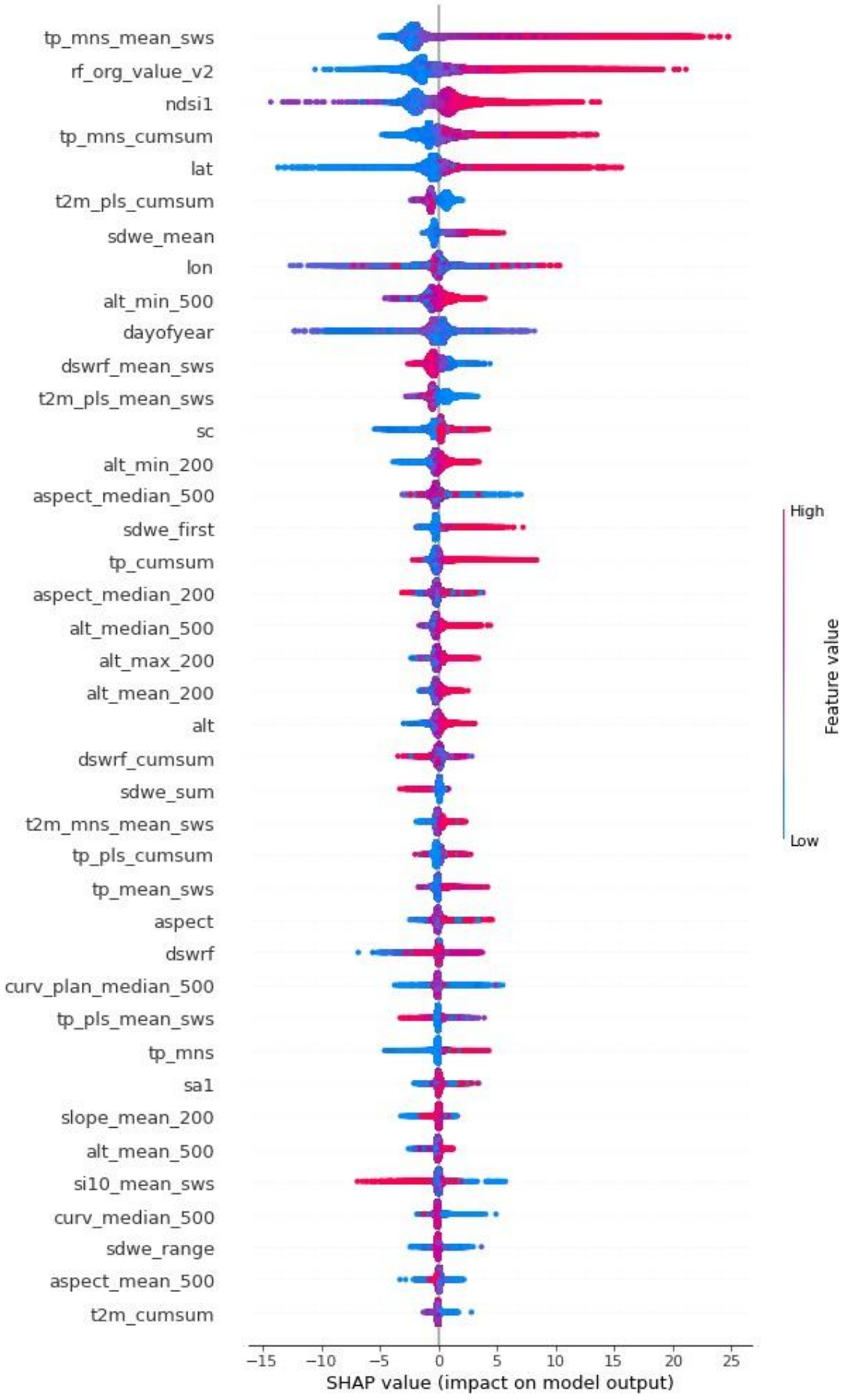


Figure 5 Feature importance structure in LightGBM

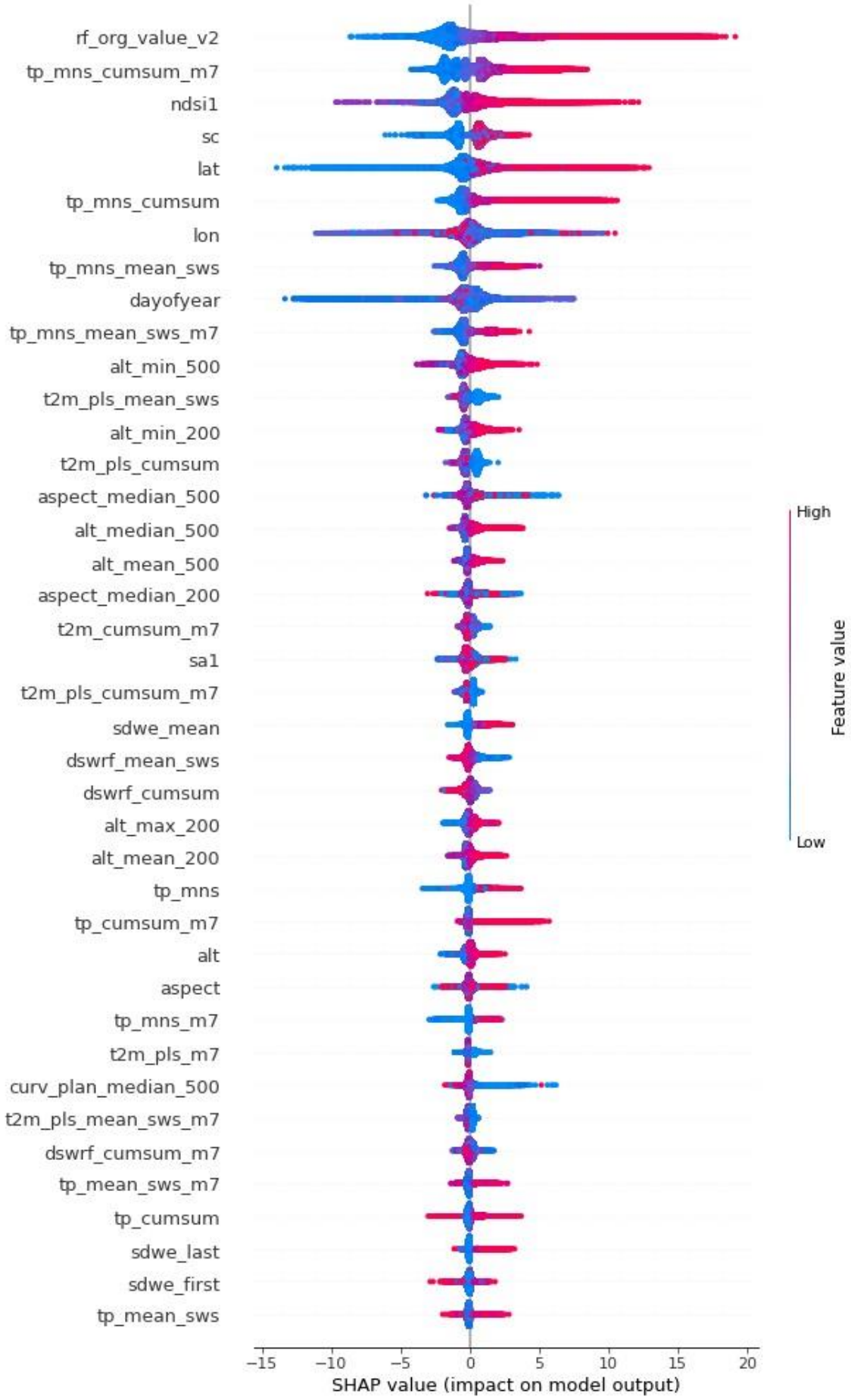


Figure 6 Feature importance structure in CatBoost

Model results

The developed model is an end-to-end pipeline that allows you to predict SWE for the desired date and grid cell. There are the following steps in the pipeline:

2. Downloading and processing the meteorological data for the previous week and winter period
3. Downloading and processing of MODIS space images for the previous week
4. Data collection from new (meteo, space images) and stable(relief, etc.) resources
5. Loading model weights from the pretrained model
6. Making predictions and preparing a weekly submission for the appropriate week.

The result of the model is a SWE forecast map for the current date for the studied regions (Fig. 7-1).

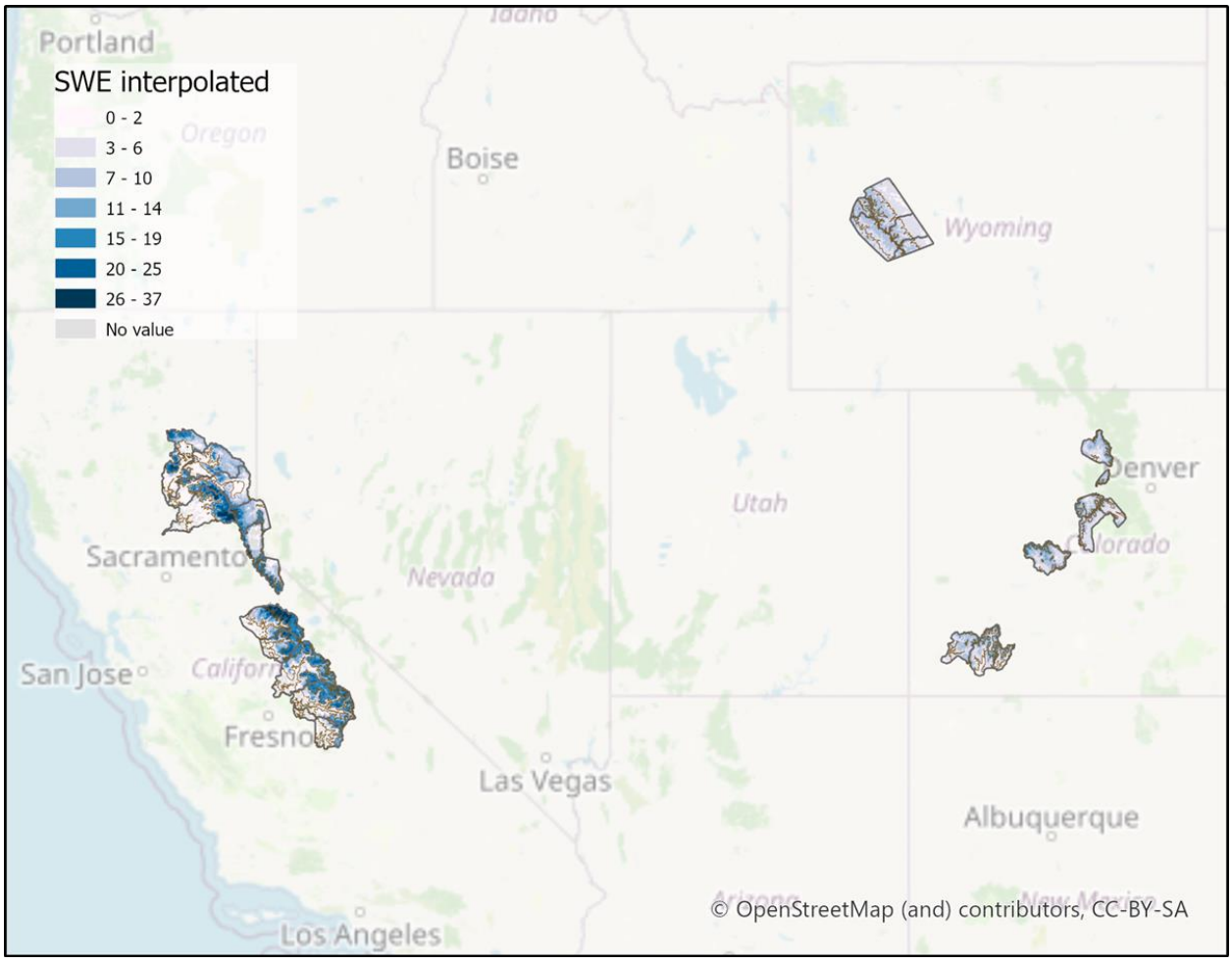


Figure 7 Model results

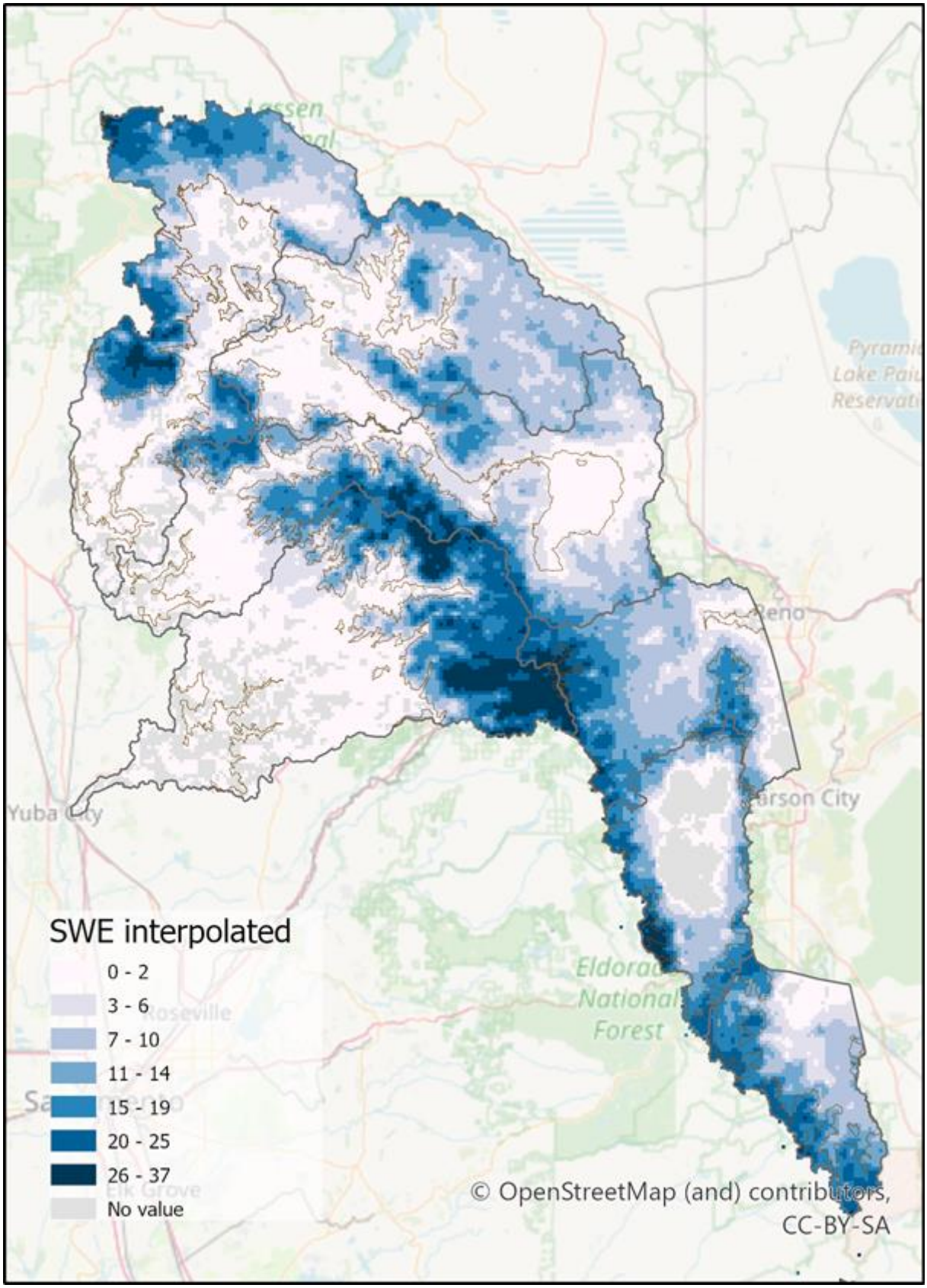


Figure 8 Model results

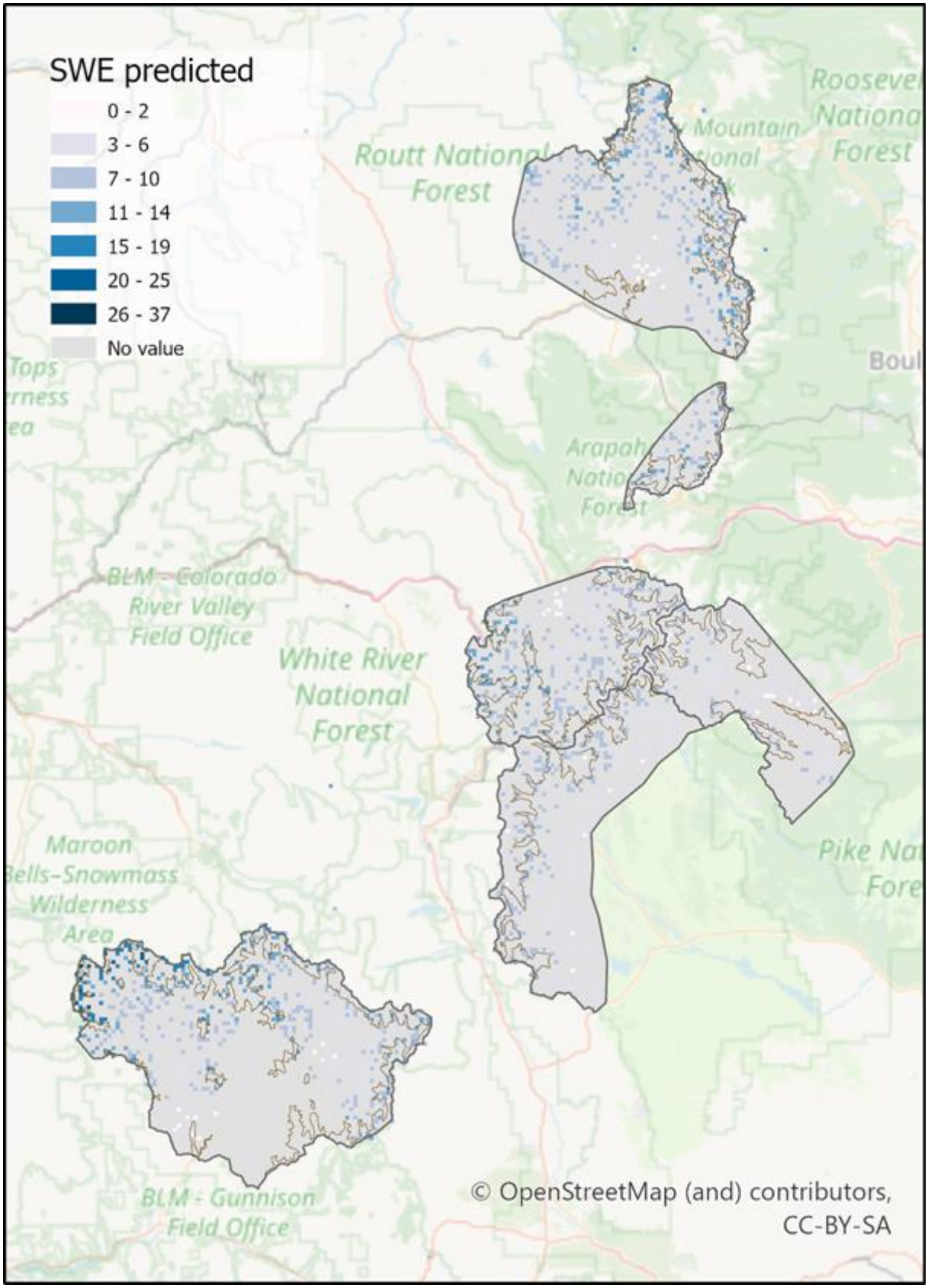


Figure 9 Model results

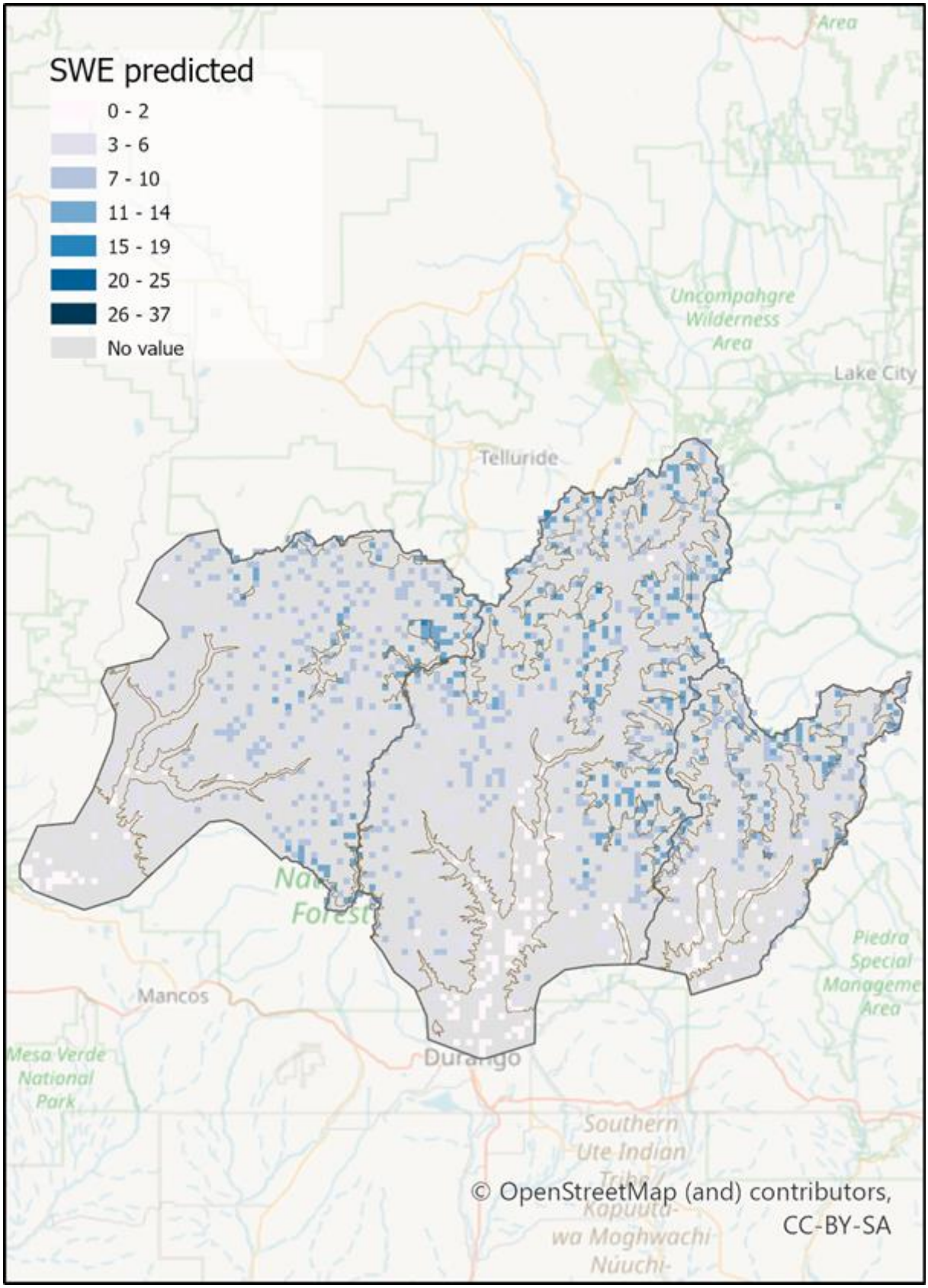


Figure 10 Model results

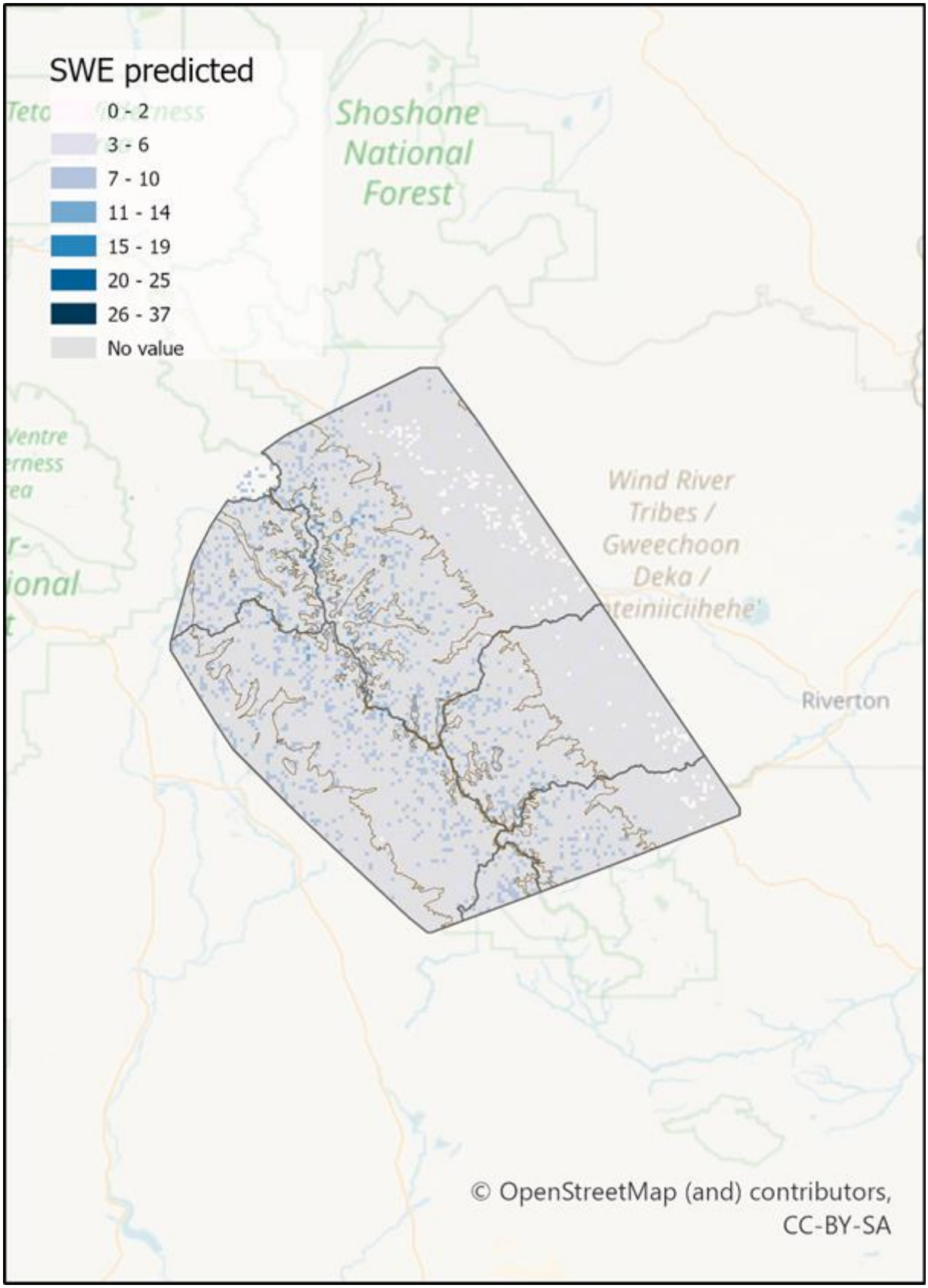


Figure 11 Model results

Model performance

The robustness of the model was assessed by several methods.

First, a comparison of the simulated SWE values and snow coverage according to MODIS satellite images data was carried out in Figure 12. The model accurately reproduces the spatial distribution of the snow cover. There is a high spatial correlation between the predicted SWE values and the actual snow cover in the study area. As we can see in the figure, the bulk of snow accumulates at high altitudes in the mountains, while there is no snow in the lower reaches of rivers and river valleys.

Secondly, the accuracy of the model was estimated at regular measurement points. Figure 13 presents the results of comparing the measured and simulated values for the best 24 stations according to the RMSE score. As we can see, the model also faithfully reproduces the seasonal trends in the process of snow accumulation and ablation.

Analysis of the results and model's metrics shows that it has a fine spatiotemporal generalizing ability.

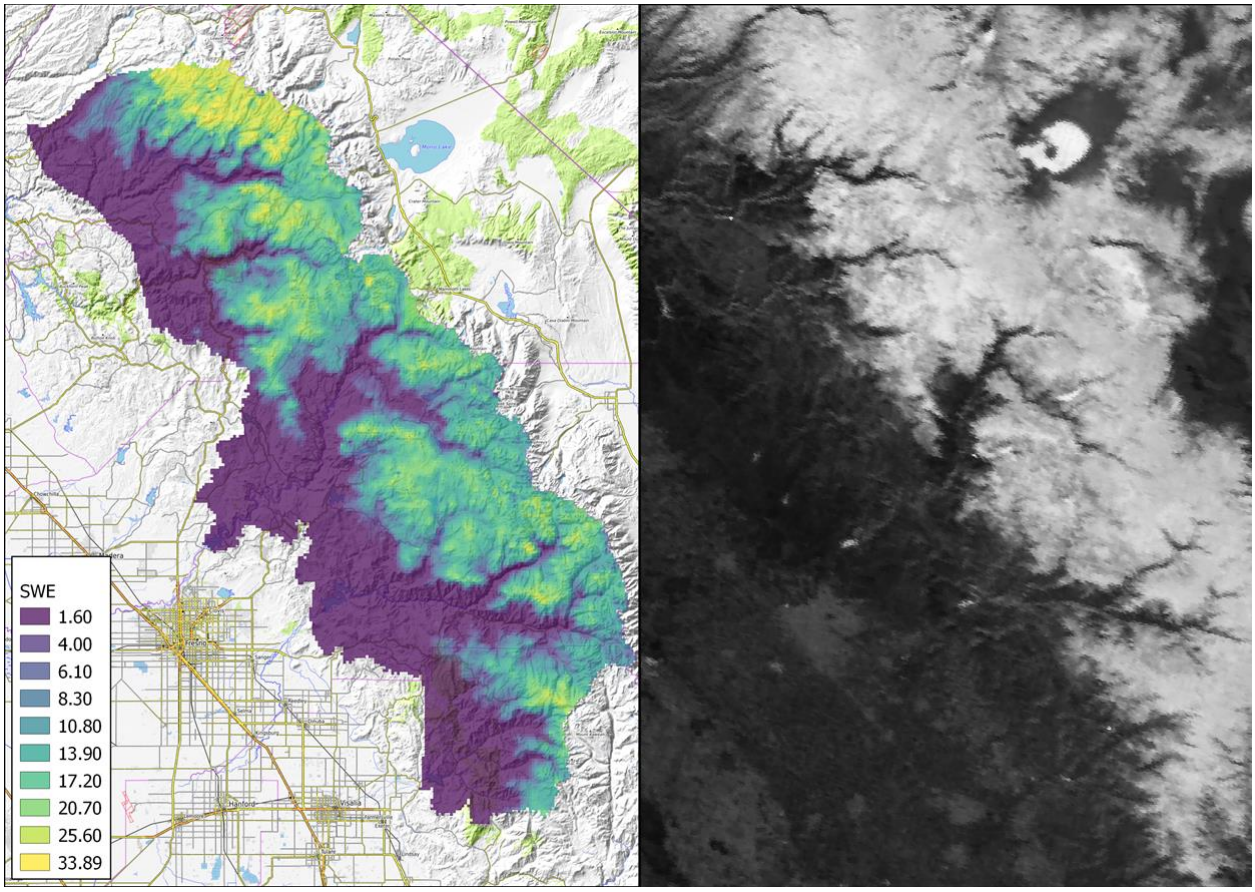


Figure 12 , Comparing of the simulated SWE values and snow coverage according to MODIS satellite images data

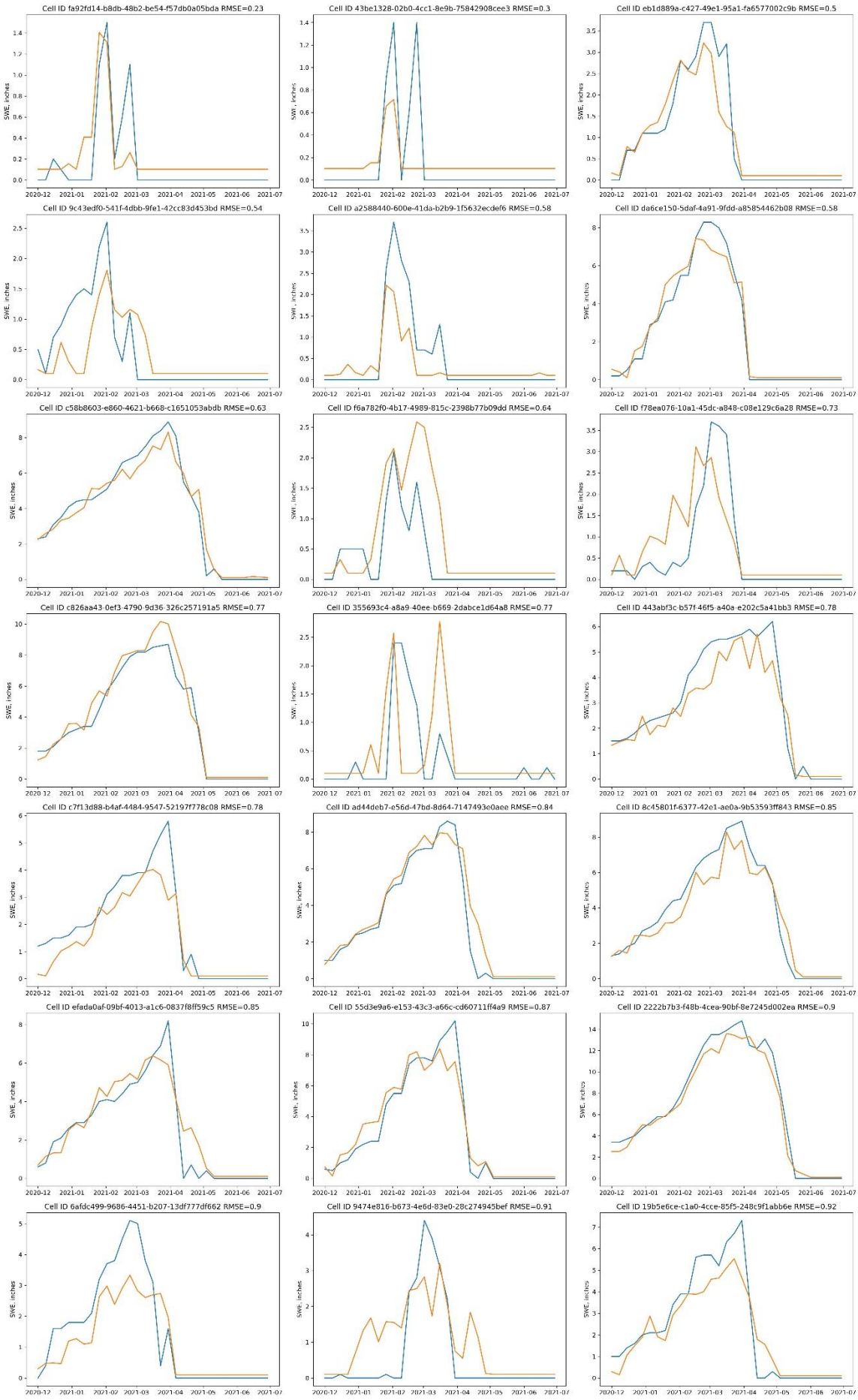


Figure 13 , Comparing of the simulated (orange) and observed (blue) SWE values

MACHINE SPECIFICATIONS

We used 3 workstations for data analysis, feature engineering, model training and experiments. All calculations were performed on the CPU, the GPU was not used. Specifications are in the table below:

	Workstation 1	Workstation 2	Workstation 3
CPU	AMD Ryzen Threadripper 1950X	AMD Ryzen 9 5950X	Intel i7 11800H
CPU cores	16	16	8
CPU threads	32	32	16
RAM, gb	64	32	16
GPU	2x Nvidia GeForce 1080Ti 11gb	Nvidia GeForce RTX 2070 8 gb	Nvidia GeForce RTX 3060 6 gb

REFERENCES

Bair, E.H., Abreu Calfa, A., Rittger, K. and Dozier, J., 2018. Using machine learning for real-time estimates of snow water equivalent in the watersheds of Afghanistan. The Cryosphere, 12(5), pp.1579-1594.

IPCC, 2022: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press. In Press.

Krinner, G. et al., 2018: ESM-SnowMIP: Assessing snow models and quantifying snow-related climate feedbacks. Geosci. Model Dev., 11, 5027–5049

Rheinheimer, D. E. , R. C. Bales, C. A. Oroza, J. R. Lund and J. H. Viers, "Valuing year-to-go hydrologic forecast improvements for a peaking hydropower system in the sierra nevada", Water Resour. Res., vol. 52, pp. 3815-3828, 2016.

Şensoy, A. and Uysal, G., 2012. The value of snow depletion forecasting methods towards operational snowmelt runoff estimation using MODIS and Numerical Weather Prediction Data. Water resources management, 26(12), pp.3415-3440.

Vionnet, V., Marsh, C.B., Menounos, B., Gascoin, S., Wayand, N.E., Shea, J., Mukherjee, K. and Pomeroy, J.W., 2021. Multi-scale snowdrift-permitting modelling of mountain snowpack. The Cryosphere, 15(2), pp.743-769.

Wever, N., Fierz, C., Mitterer, C., Hirashima, H., and Lehning, M.: Solving Richards Equation for snow improves snowpack melt- water runoff estimations in detailed multi-layer snowpack model, The Cryosphere, 8, 257–274, doi:10.5194/tc-8-257-2014, 2014.