**Assessing climate change impacts on snow cover area by using a hybrid NARX-LSTM machine learning model**

**1. Introduction**

1) Importance of snow dominated basins and sensitivity of mountains to climate change

2) Remote sensing products to monitor snow cover area; problem of gaps🡪 MOD10A1F

3) Models to simulate snow cover area. Importance of models that only use climate variables to use them to propagate climate change scenarios

4) Objective a novelty of the work. To highlight the use of the method in several case studies around the world.

**2. Method**

Figure 1. Flow chart of the proposed methodology

**2.1. Generation of climate change scenarios**

**2.2. Hybrid** **NARX-LSTM machine learning model**

**3. Case studies and data**

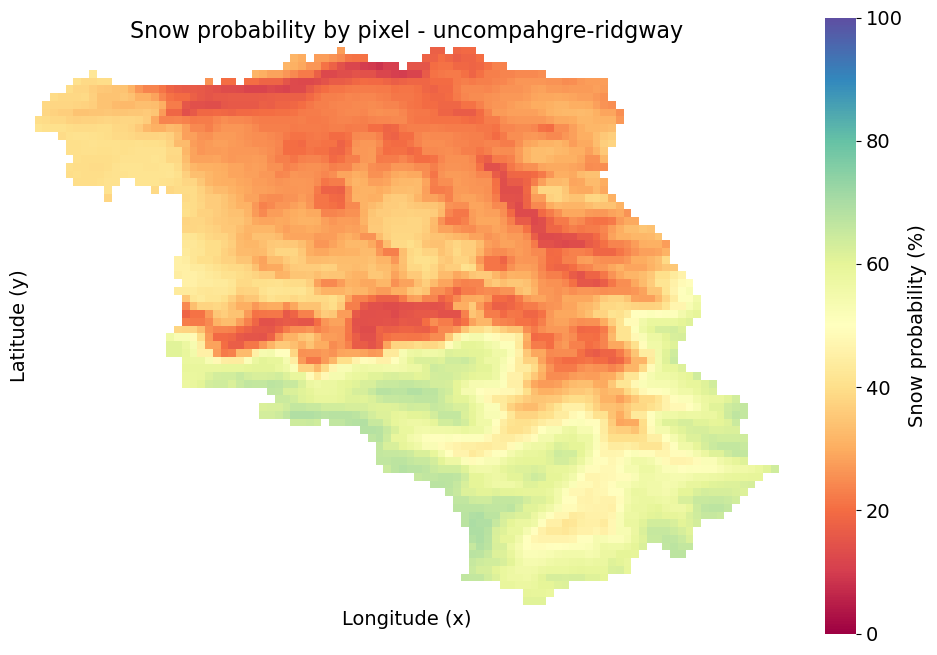
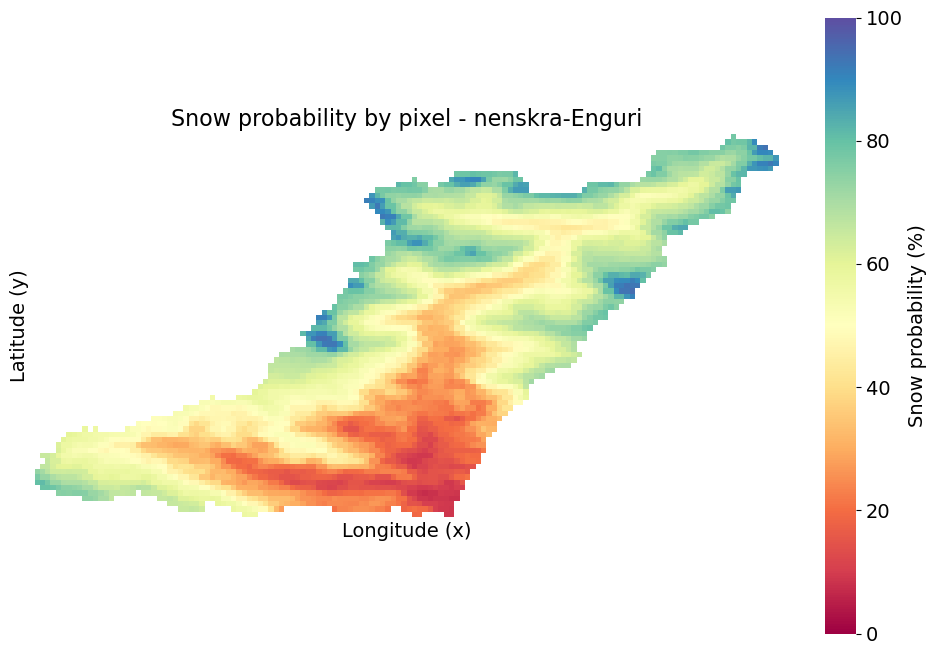
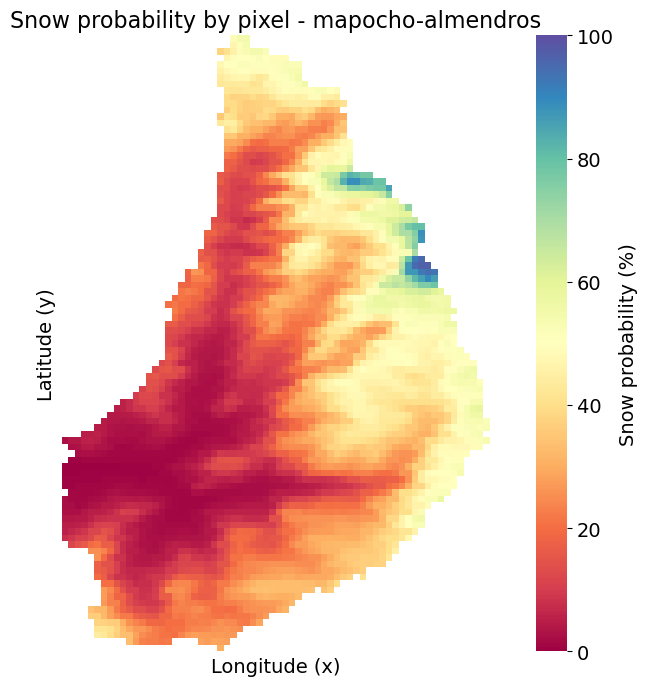
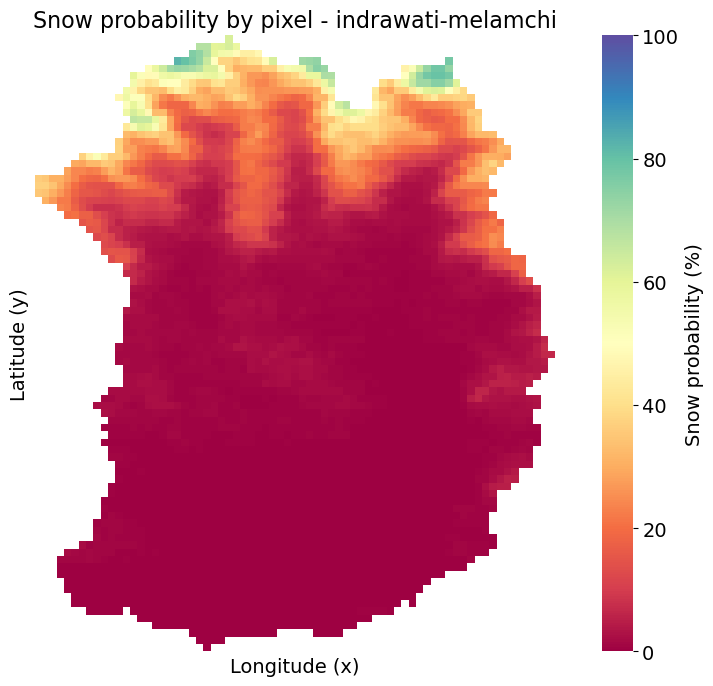
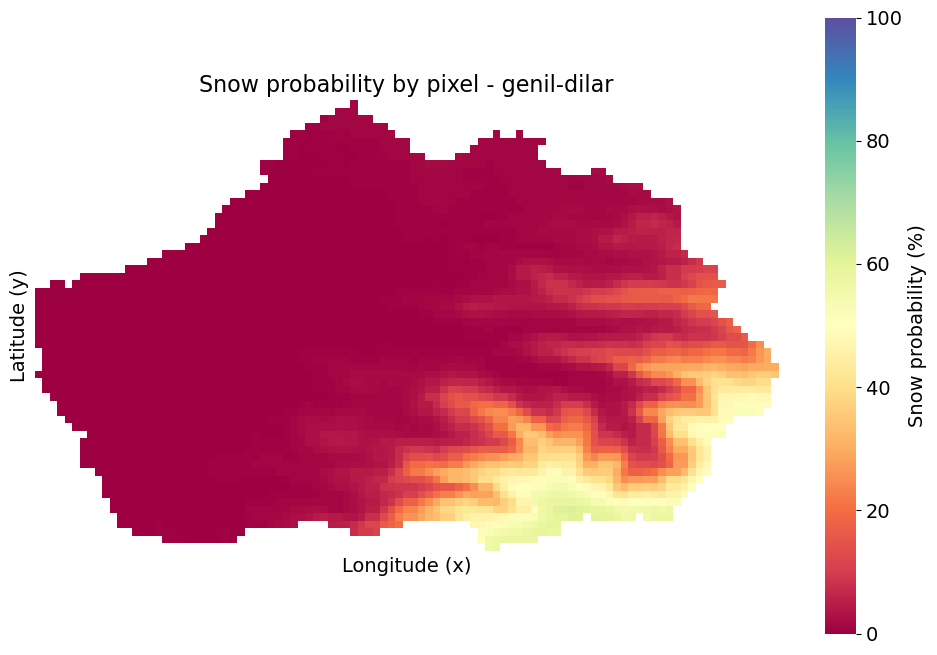
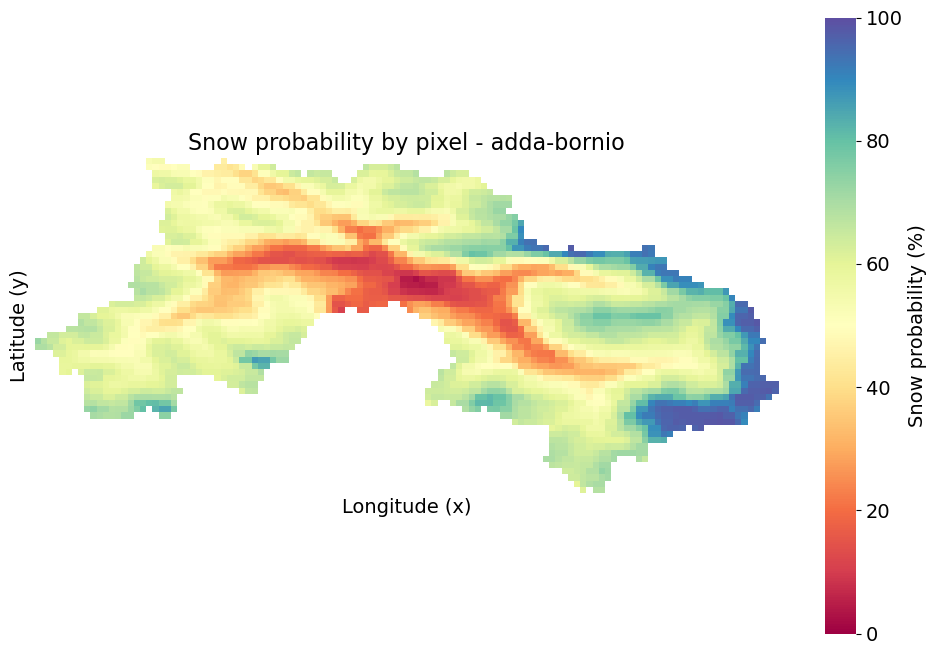
**3.1. Case studies**

Figure 2. Location of the case studies.

**3.2. Data**

Figure 3. Elevation maps for each basin.

Figure 4. The probability that a pixel is covered by snow during the historical period (2000-2023) for each basin.



This figure illustrates the long-term spatial distribution of snow cover probability across six distinct basins, derived from historical data spanning 2000-2023 through Kriging interpolation. A general trend observed across most basins (e.g., 'adda-bornio', 'iracheos-almendros', 'rercise-liquin', 'suwampalyu-ralgayu') is the **strong influence of topography**, with higher probabilities of snow cover consistently found in **elevated and mountainous regions**, and gradually diminishing probabilities in lower-lying areas. This altitudinal gradient is particularly evident in basins like 'rercise-liquin', which shows a clear north-south decrease in snow probability aligning with decreasing elevation.

Conversely, the 'imduwael-malanahi' basin (Figure 1c) presents a unique case, characterized by a remarkably **high and spatially uniform snow probability across its entire extent**, suggesting consistently favorable conditions for snow accumulation or high average elevations throughout the basin. The 'gqali-allar' basin (Figure 1b) also displays generally high probabilities over a large area, though with some internal variations. These maps collectively provide critical insights into the regional snow regimes, identifying persistent snow zones and highlighting the diverse hydro-climatic characteristics that govern snow presence across the studied basins.

**4. Results**

**4.1. Historical and future climate change scenarios**

Figure 5. Monthly average precipitation and temperature for the mean year over the historical scenario (1995-2014) and future climate change scenarios (2051-2070) considering the SSP 2 and RCP 4.5 and the different GCMs for each basin.

Figure 6. Monthly average precipitation and temperature for the mean year over the historical scenario (1995-2014) and future climate change scenarios (2081-2100) considering the SSP 2 and RCP 4.5 and the different GCMs for each basin.

Figure 7. Monthly average precipitation and temperature for the mean year over the historical scenario (1995-2014) and future climate change scenarios (2051-2070) considering the SSP 5 and RCP 8.5 and the different GCMs for each basin.

Figure 8. Monthly average precipitation and temperature for the mean year over the historical scenario (1995-2014) and future climate change scenarios (2081-2100) considering the SSP 5 and RCP 8.5 and the different GCMs for each basin.

**4.2. NARX-LSTM machine learning models performance**

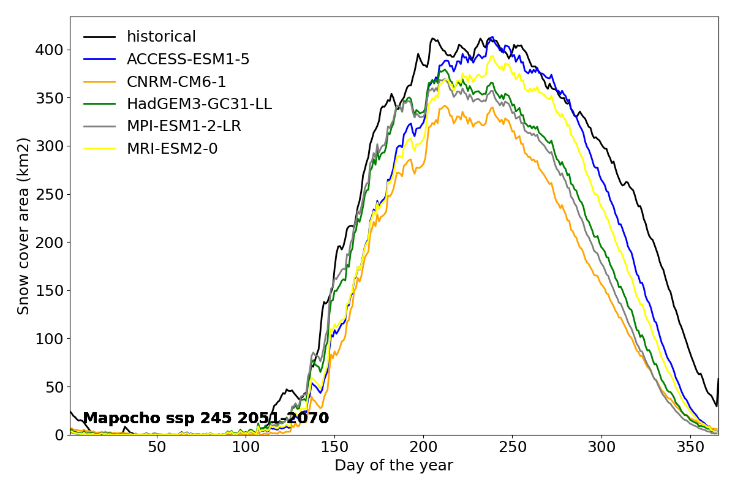
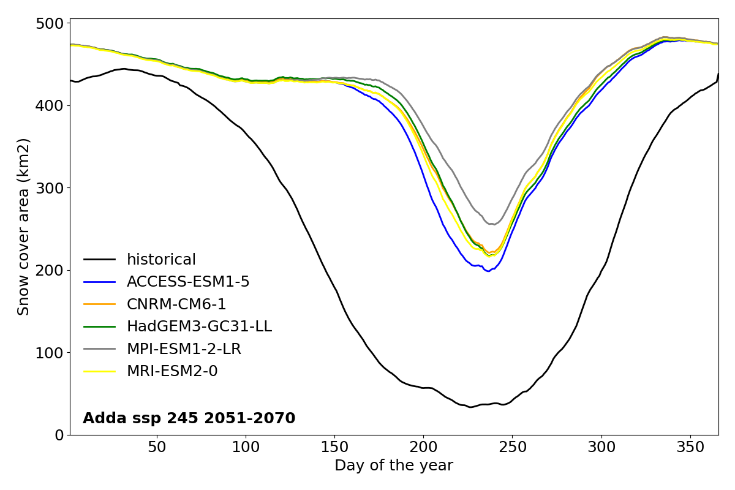
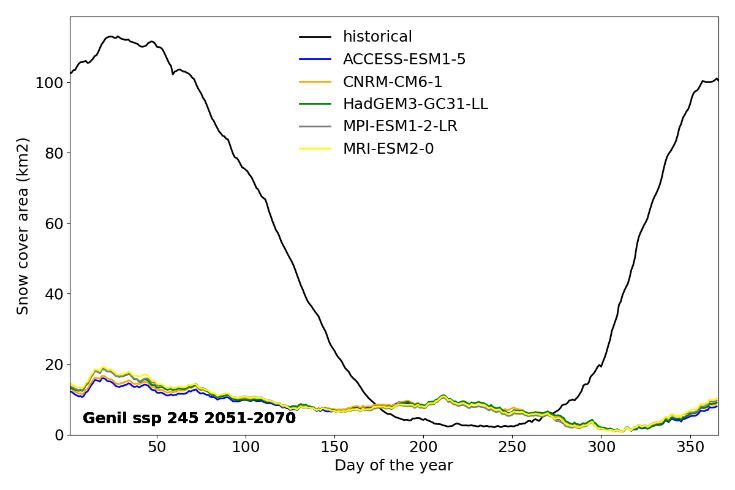
Table 1. Best configurations of the NARX-LSTM machine learning models for each basin.

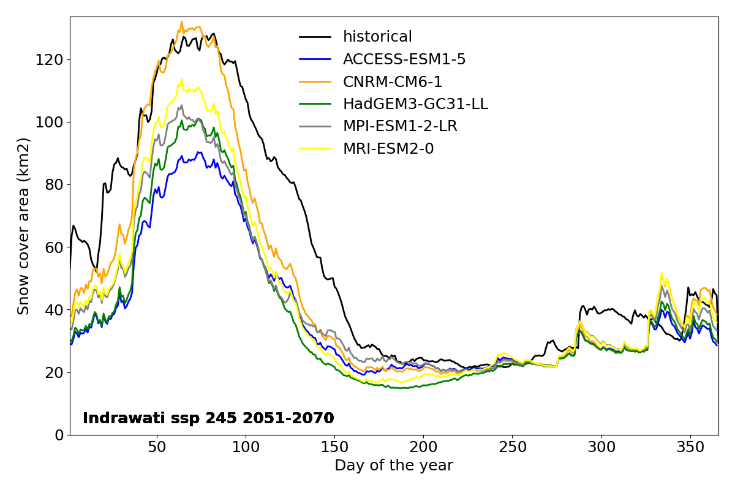
In order to find these best hyperparameters, the Optuna library from Python has been used to try several configurations and get the one that yields the best metrics. The objective of the optimization was to maximize the Nash-Sutcliffe Efficiency (NSE) metric calculated on the full historical dataset:

* **n\_lags\_area:** number of steps backwards of the target variable, since it is an autoregressiveneural network, it predicts the next day, based on what happened in the n\_lags\_area days before. Optimal values vary between 4 and 8 days, suggesting that recent past conditions (up to approximately one week) are critical for snow prediction across these basins.
* **n\_layers:** this parameter defines the deph of the neural network. Deeper networks can learn more complex patterns from the input data. Most basins achieved optimal performance with 1 to 2 LSTM layers, while Uncompahgre required 3, indicating varying levels of complexity needed to model snow dynamics across different basins.
* **n\_units:** number of neurons within each layer determines the learning capacity and memory of that layer. More units allow the layer to process and retain a greater amount of information. In the table we can see that optimal units vary from 10 to 29.
* **learning\_rate:** important hyperparameter in the optimization process. It controls the step size at which the model’s internal weights are adjusted during the training base on the calculated loss.
* **dropout\_rate:** this is meant to be a regularization parameter in order to prevent overfitting, a common challenge in neural networks. Optimal dropout rates are found within the 0.16 to 0.30 range, emphasizing the importance of regularization to prevent overfitting in these time-series prediction tasks.
* **epochs:** to define the maximum number of times the entire training dataset is passed forward and backward through the neural network. This mechanism automatically halts the training process if it does not improve over a specific number of epochs. In the project it varies from 21 to 99 demonstrating the difference between basins and how to train them without overfitting

Figure 9. Historical and simulated (using the forecasting mode) daily snow cover area series for the period 2000-2023 and performance metrics for each basin.

**4.3. Propagation of climate change scenarios to snow cover area**

****Figure 10. Daily average snow cover area for the mean year over the historical scenario (1995-2014) and future climate change scenarios (2051-2070) considering the SSP 2 and RCP 4.5 and the different GCMs for each basin.

****

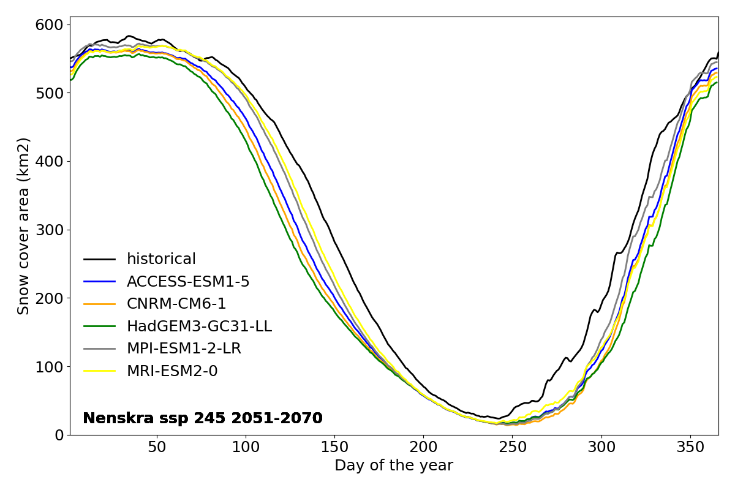
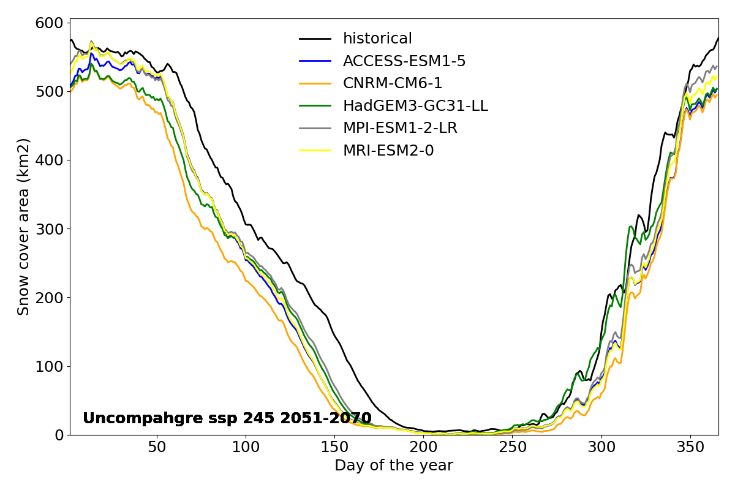
****Figure 11. Daily average snow cover area for the mean year over the historical scenario (1995-2014) and future climate change scenarios (2081-2100) considering the SSP 2 and RCP 4.5 and the different GCMs for each basin.

Figure 12. Daily average snow cover area for the mean year over the historical scenario (1995-2014) and future climate change scenarios (2051-2070) considering the SSP 5 and RCP 8.5 and the different GCMs for each basin.

Figure 13. Daily average snow cover area for the mean year over the historical scenario (1995-2014) and future climate change scenarios (2081-2100) considering the SSP 5 and RCP 8.5 and the different GCMs for each basin.

**5. Discussion**

**6. Conclusion**

**References**