Supplementary Material

Francisco Zambrano et al.

2024-04-05

1 Validation of ERA5L variables

1.1 Methods

We compared the ERA5L (Muñoz-Sabater et al. 2021) variables for monthly mean temperature, total precipitation, and volumetric soil water content against values retrieved by weather stations. For temperature and precipitation, we used data from the weather network from the Ministry of Agriculture of Chile (www.agromet.com) between 2015 and 2023. We used 277 stations located throughout Chile. For in-situ soil moisture, we used a private soil network that is owned by the agricultural enterprise Garces Fruit, which has 99 stations in Central Chile, located in fields with cherry fruit crops. The sensors are installed at 30, 60, and 90cm and are a Advanced Soil Moisture Sensor model Teros 12 from METER Group (https://metergroup.com/products/teros-12/). To avoid comaring ERA5L with in-situ soil moisture levels caused by irrigation, which are not captured by ERA5L, we used daily data for the year 2022 and the months outside the growing season, May to September.

We selected the following metrics:

$$MAE = \frac{1}{n}\sum |E - S|$$

$$Bias = \frac{\sum E}{\sum S}$$

$$ubRMSE = \sqrt{\frac{\sum \left[(E_i - \overline{E}) - (S_i - \overline{S})\right]^2}{n}}$$

$$CC = \frac{\sum (S_i - \bar{S})(E_i - \bar{E})}{\sqrt{(Si - \bar{S})^2 (E_i - \bar{E})^2}}$$

MAE: mean absolute error
$bias$: bias
$ubRMSE$: unbiassed root mean squared error
CC: coefficient of correlation
S: value of the variable measure by the weather station
E: value of the variable measure by ERA5L

1.2 Results

The average performance metrics of ERA5L over the 266 weather stations were in the case of monthly temperature: $ubRMSE = 1.06\,°C$, $MAE = 1.131\,°C$, and $CC = 0.963$, showing a good agreement, low error, and low overestimation. For cumulative monthly precipitation, $MAE = 28.1\,mm$, $bias = 1.93$, and $CC = 0.845$, showing a high correlation and a 93% bias and being overestimated by ERA5L. In the case of the 97 soil moisture stations, we averaged for the three depths (30, 60, and 90cm) and then compared it with the volumetric water content in the top 100cm of the soil derived from ERA5L. For this case, we made a daily comparison to determine the performance metrics per station, then we averaged over all stations, having a $CC = 0.71$, $RMSE = 0.174$, $MAE = 0.167$, and $bias = 1.74$. The ERA5 soil moisture overestimate is 74%, but there is some good correlation.

2 Land cover macroclasses and validation

2.1 Methods

To analyze the LULCC, we use the IGBP scheme from the MCD12Q1 Collection 6.1 from MODIS. This product has a yearly frequency from 2001 to 2022. The IGBP defines 17 classes; we regrouped these into ten macroclasses, as follows: classes 1-4 to forest, 5-7 to schrublands, 8-9 to savannas, 10 as grasslands, 11 as wetlands, 12 and 14 to croplands, 13 as urban, 15 as snow and ice, 16 as barren, and 17 to water bodies.

To validate the land cover obtained, we compare the macroclasses with those of a more detailed land cover map made by Zhao et al. (2016) for Chile with samples acquired in the years 2013–2014 (LCChile). The latter has a spatial resolution of 30 m and three hierarchy levels of defined classes; from those, we used level 1, which fits with the macroclass land cover. We chose the years 2013 (IGBP2013) and 2014 (IGBP2014) from the land cover macrolcasses to compare with LCChile. For this comparison, we used the following procedure:

i) we resampled LCChile to the spatial resolution (500m) of the land cover macroclasses using the majority method,

ii) we took a random sample of 1000 points within continental Chile and extracted the classes that fell within each point for LCChile, IGBP2013, and IGBP2014; we considered the point extracted from LCChile as the truth and the values from the other two years as predictions.

iii) we derivede a confusion matrix with the classes extracted from the 1000 points for LCChile, IGBP2013, and IGBP2014; and calculate the performance metrics of accuracy and F1.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{correct\ classifications}{all\ classifications}$$

$$F1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}$$

where TP and FN refer to true positive and false negative, correctly classified classes; TN and FP to true negative and false positive, wrongly classified classes.

2.2 Results

For vegetation, we obtained and use hereafter five macroclasses of land cover from IGBP MODIS: forest, shrubland, savanna, grassland, and croplands (Table 1).

3 Relationship between drought indices and land cover change

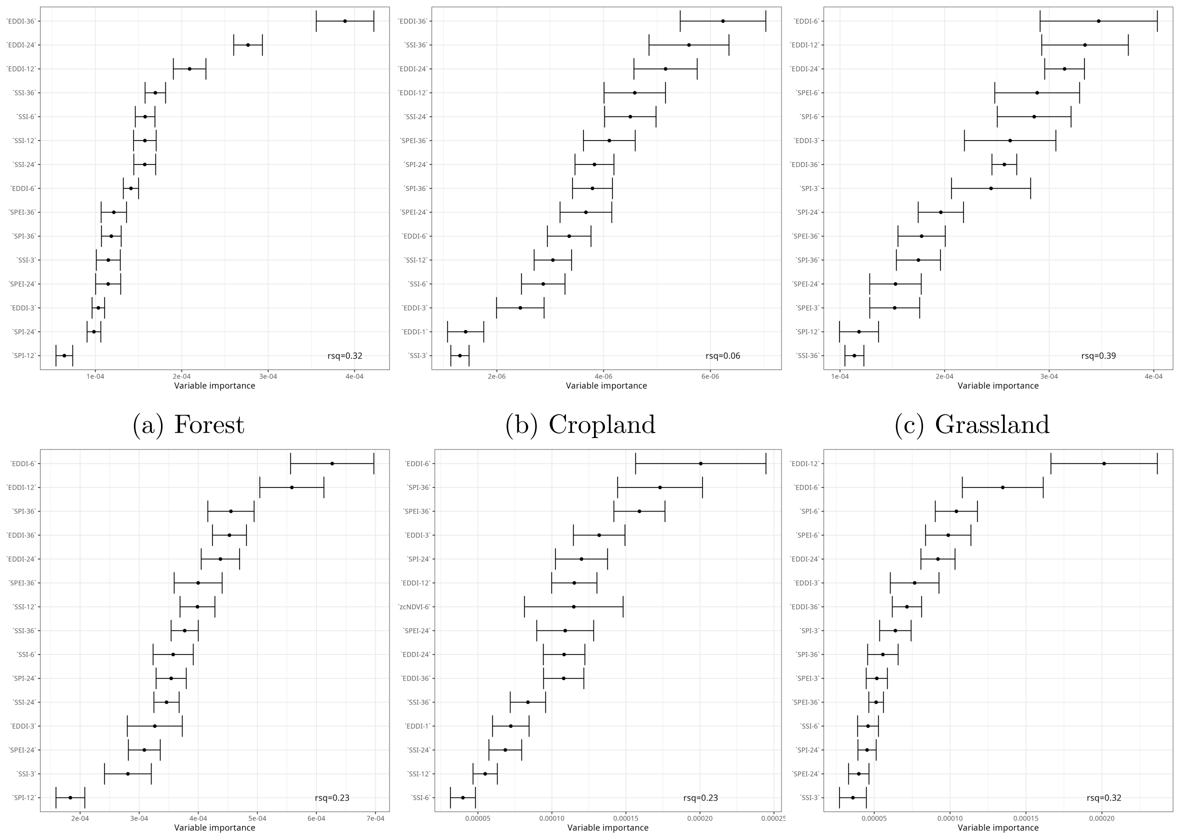
Figure 1 shows the ranking of variable importance obtained with the ten folds used in the resampling of the Random Forest model per land cover type.

4 Trend of vegetation productivity

Figure 2 shows the average trend for zcNDVI for 2000-2023 per macrozone and landcover macroclass.

5 Vegetation productivity

We analyzed the correlation of zcNDVI for time scales of 1, 3, 6, and 12 months versus net primary production (NPP). We obtained both the zcNDVI from MOD13A3.061 and the NPP from MOD17A3HGF.061, using MODIS products. We used the zcNDVI in December to correlate with the annual NPP. Figure 3 shows a map of the r-squared correlation between zcNDVI and NPP, and Figure 4 shows the aggregated values per macrozone.



(a) Forest (b) Cropland (c) Grassland

(d) Savanna (e) Shrubland (f) Barren land

Figure 1: Error bar for the variables importance obtained from the 10 resamples folded by the Random Forest used to model land cover trend using as predictors the trend in drought indices of water supply, water demand, soil moisture, and vegetation productivity.

Table 1: Landcover clases from IGBP MODIS and the corresponding macroclasses.

Name	Value	Description	Macroclass
Evergreen Needleleaf Forests	1	Dominated by evergreen conifer trees (canopy >2m). Tree cover >60%.	Forest
Evergreen Broadleaf Forests	2	Dominated by evergreen broadleaf and palmate trees (canopy >2m). Tree cover >60%.	Forest
Deciduous Needleleaf Forests	3	Dominated by deciduous needleleaf (larch) trees (canopy >2m). Tree cover >60%.	Forest
Deciduous Broadleaf Forests	4	Dominated by deciduous broadleaf trees (canopy >2m). Tree cover >60%.	Forest
Mixed Forests	5	Dominated by neither deciduous nor evergreen (40-60% of each) tree type (canopy >2m). Tree cover >60%.	Forest
Closed Shrublands	6	Dominated by woody perennials (1-2m height) >60% cover.	Shrublands
Open Shrublands	7	Dominated by woody perennials (1-2m height) 10-60% cover.	Shrublands
Woody Savannas	8	Tree cover 30-60% (canopy >2m).	Savanna
Savannas	9	Tree cover 10-30% (canopy >2m).	Savanna
Grasslands	10	Dominated by herbaceous annuals (<2m).	Grassland
Permanent Wetlands	11	Permanently inundated lands with 30-60% water cover and >10% vegetated cover.	Wetland
Croplands	12	At least 60% of area is cultivated cropland.	Cropland
Urban and Built-up Lands	13	At least 30% impervious surface area including building materials, asphalt, and vehicles.	Urban
Cropland/Natural Vegetation Mosaics	14	Mosaics of small-scale cultivation 40-60% with natural tree, shrub, or herbaceous vegetation.	Cropland
Permanent Snow and Ice	15	At least 60% of area is covered by snow and ice for at least 10 months of the year.	Snow/Ice
Barren	16	At least 60% of area is non-vegetated barren (sand, rock, soil) areas with less than 10% vegetation)	Barren land
Water Bodies	17	At least 60% of area is covered by permanent water bodies	Water
Unclassified	255	Has not received a map label because of missing inputs	Unclassified

6 References

Muñoz-Sabater, Joaquín, Emanuel Dutra, Anna Agustí-Panareda, Clément Albergel, Gabriele Arduini, Gianpaolo Balsamo, Souhail Boussetta, et al. 2021. "ERA5-Land: A State-of-the-Art Global Reanalysis Dataset for Land Applications." *Earth System Science Data* 13 (9): 4349–83. https://doi.org/10.5194/essd-13-4349-2021.

Zhao, Yuanyuan, Duole Feng, Le Yu, Xiaoyi Wang, Yanlei Chen, Yuqi Bai, H. Jaime Hernández, et al. 2016. "Detailed Dynamic Land Cover Mapping of Chile: Accuracy Improvement by Integrating Multi-Temporal Data." *Remote Sensing of Environment* 183 (September): 170–85. https://doi.org/10.1016/j.rse.2016.05.016.

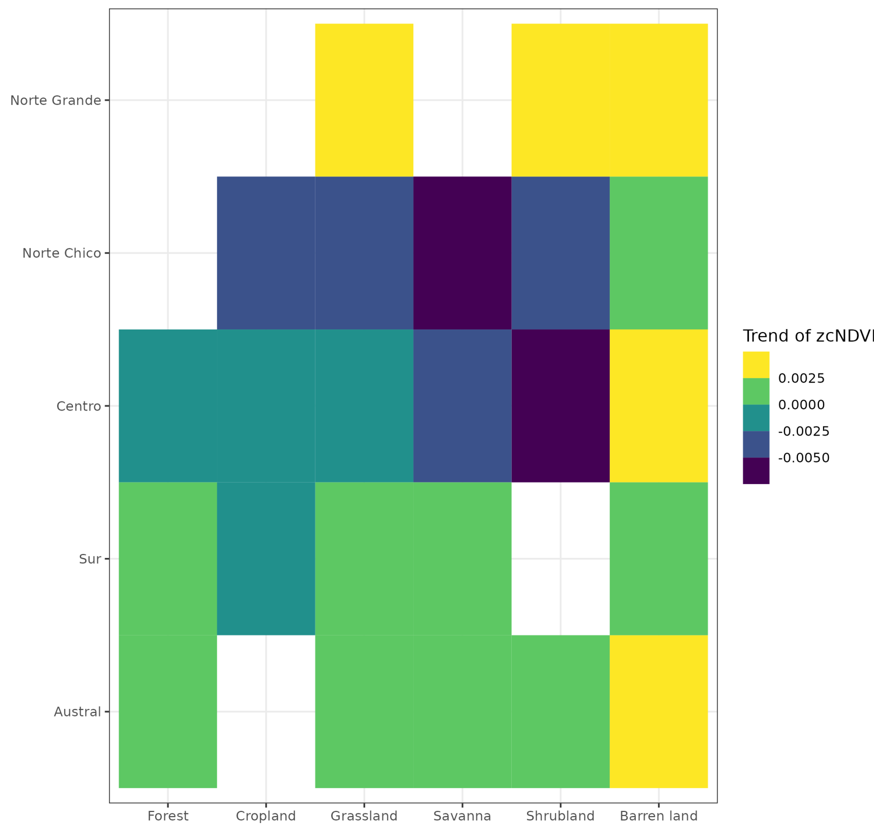


Figure 2: Heatmap of trends in zcNDVI for 2000 to 2023 per macrozone and landcover macro-class.

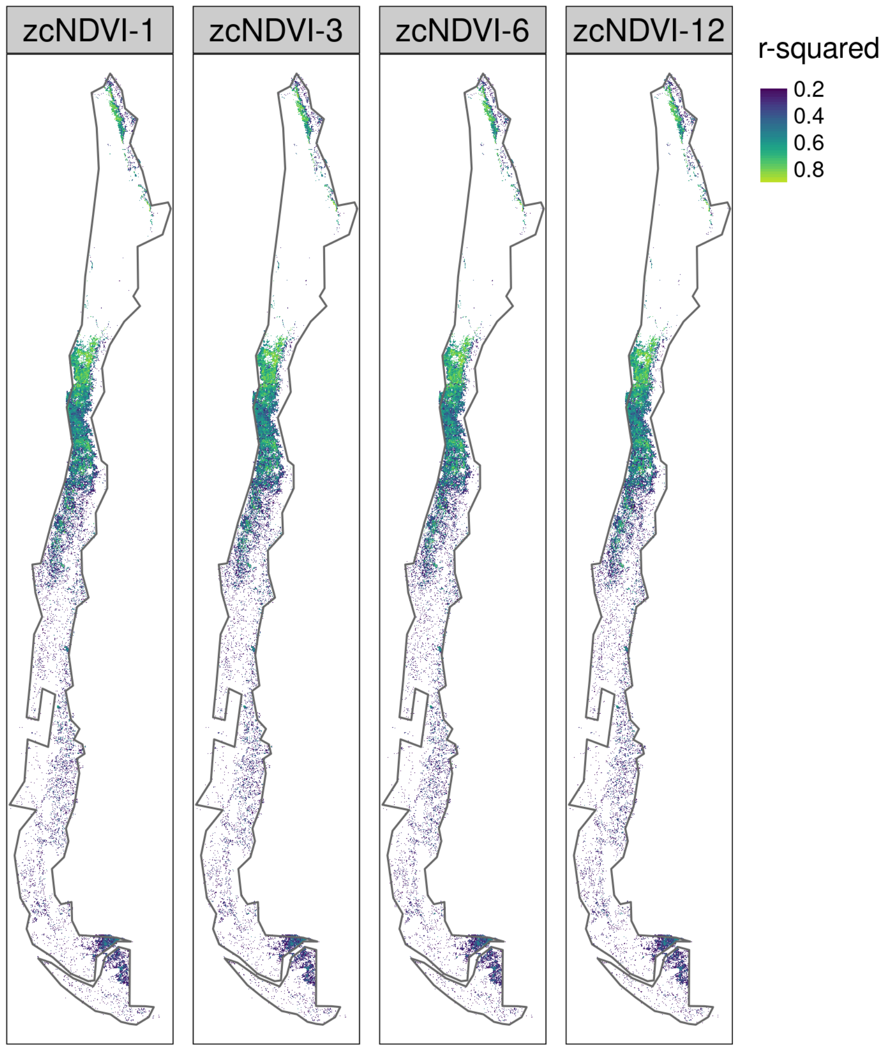


Figure 3: Spatial variation of the r-squared values obtained from the yearly correlation of zcNDVI of 1, 3, 6, and 12 months with the net primary productivity (NPP) for continental Chile.

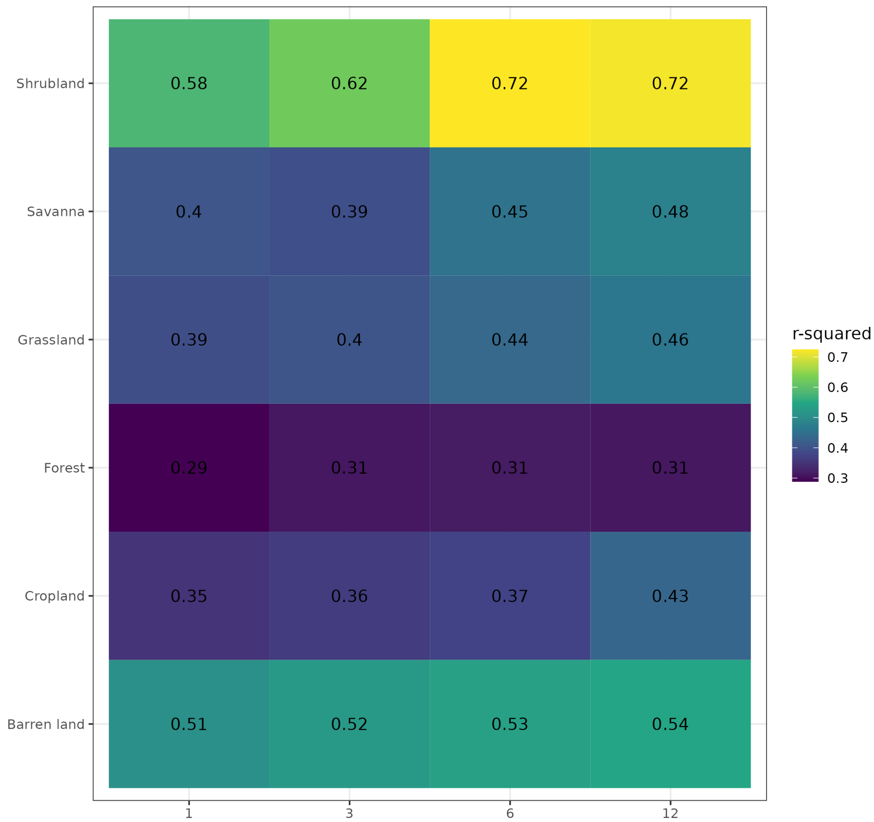


Figure 4: A heatmap showing the r-squared values obtained from the yearly correlation of zcNDVI of 1, 3, 6, and 12 months with the net primary productivity (NPP) for continental Chile.