Revista Internacional de Ciencia y Tecnología de la Información Geográfica
International Review of Geographical Information Science and Technology

Trojer, F., Téllez, L., Pesquer, L., Ninyerola M. (2024). Evaluating the contribution of auxiliary observations for climate mapping: a case study in the Guadalquivir region. *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica* (Articles), 33, 77-128. http://dx.doi.org/10.21138/GF.830

EVALUATING THE CONTRIBUTION OF AUXILIARY OBSERVATIONS FOR CLIMATE MAPPING: A CASE STUDY IN THE GUADALQUIVIR REGION

[1a]Fabian Trojer, [1b]Lluís Téllez, [2c]Lluís Pesquer, [3d]Miquel Ninyerola
[1] Universitat Autònoma de Barcelona. 08193, Bellaterra, Barcelona, Catalonia, Spain
[2] Grumets Research Group. Departament de Geografia. Edifici B. Universitat Autònoma de Barcelona. 08193, Bellaterra, Barcelona, Catalonia, Spain
[3] Grumets Research Group. Departament de Biologia Animal, Biologia Vegetal i Ecologia. Edifici C. Universitat Autònoma de Barcelona. 08193, Bellaterra, Barcelona, Catalonia, Spain
[a]fabian_trojer@hotmail.com, [b]lluistto8@gmail.com, [c]l.pesquer@creaf.uab.cat, [d]miquel.ninyerola@uab.cat

ABSTRACT

In order to improve the detection of climate variability patterns, we assess the impact of incorporating auxiliary meteorological networks into the AEMET network to produce temperature and precipitation maps in the Guadalquivir basin. We employ multiple linear regression and residual spatial interpolation to create monthly precipitation and air temperature (mean minimum, mean and mean maximum) maps using two different datasets: Official (only AEMET stations) and Extended (both AEMET and auxiliary network data). Comparison of the performance of both datasets focuses on three key indicators: adjusted R^2, cross-validation measured with RMSE, and the percentage of significant independent variables.

Overall, the results indicate that the inclusion of auxiliary networks did not consistently or significantly enhance regression models or reduce map inaccuracies. The extended dataset shows a slight decline in adjusted R^2 for most of the variables, with a maximum decrease of 0.082 in R^2. However, it allowed for the inclusion of more independent variables in the regression models. Notably, altitude, distance to the Atlantic Ocean, and distance to the Mediterranean Sea emerged as crucial predictor variables for both precipitation and temperature. The impact of auxiliary networks on the error metric lacked a consistent pattern. They led to decreased RMSE values for only half of the variables, with a maximum improvement of 1.24 d°C for temperature models and 6.27 dmm for precipitation models when using the extended dataset.

Keywords: multiple linear regression; spatial interpolation; temperature mapping; precipitation mapping; meteorological stations networks.

Received: 12/10/2023
Academic Editor: Dr. Cristina Domingo
Accepted: 29/04/2024
www.geofocus.org
Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)

EVALUACIÓN DE LA CONTRIBUCIÓN DE OBSERVACIONES AUXILIARES EN CARTOGRAFÍA CLIMÁTICA: ESTUDIO EN LA REGIÓN DEL GUADALQUIVIR

RESUMEN

Con el objetivo de mejorar la detección de los patrones de variabilidad climática, se evalúa el impacto de incorporar redes meteorológicas auxiliares a la red de AEMET para la generación de mapas de temperatura y precipitación en la cuenca del Guadalquivir. Empleamos regresión lineal múltiple con interpolación espacial de residuos para generar mapas mensuales de precipitación y temperatura del aire (media de las mínimas, media y media de las máximas) utilizando dos conjuntos de datos diferentes: Oficial (sólo estaciones de AEMET) y Extendido (tanto datos de AEMET como de redes auxiliares). La comparación del rendimiento de ambos conjuntos de datos se centra en tres indicadores: R^2 ajustado, validación cruzada con RMSE y porcentaje de variables independientes significativas.

Globalmente, los resultados indican que las redes auxiliares no mejoraron de forma consistente o significativa los modelos de regresión ni redujeron la imprecisión de los mapas. El conjunto de datos ampliado muestra una ligera disminución del R^2 ajustado para la mayoría de las variables, con una disminución máxima de 0.082 en R^2. Sin embargo, permitió la inclusión de más variables independientes explicativas en los modelos de regresión. En particular, la altitud, la distancia al Océano Atlántico y la distancia al Mar Mediterráneo fueron las variables predictoras cruciales tanto para la precipitación como para la temperatura. El impacto de las redes auxiliares sobre la métrica del error careció de un patrón consistente. Sólo disminuyeron los valores de RMSE para la mitad de las variables, con una mejora máxima de 1.24 d°C para los modelos de temperatura y de 6.27 dmm para los de precipitación en el conjunto de datos ampliado.

Palabras clave: regresión lineal múltiple; interpolación espacial; cartografía de temperatura; cartografía de precipitación; redes de estaciones meteorológicas.

1. Introduction

Climate mapping and modeling are vital tools in unraveling the intricate dynamics of our changing climate. A robust understanding of climatic spatial patterns relies on accurate and comprehensive meteorological datasets. The Agencia Estatal de Meteorología (AEMET) in Spain has established an official network of weather stations that serves as a cornerstone for climate research and applications. The AEMET network provides well-calibrated, long-term datasets that contribute to analyses like the Climate Atlas of the Iberian Peninsula (Ninyerola *et al*., 2005). However, in order to capture the full spectrum of climate variability, there is a growing interest in exploring the integration of supplementary meteorological data sources into the official network.

While the official AEMET network provides valuable insights, its density may not fully capture variations along topographic gradients and diverse land cover types. The availability of additional meteorological data sources, often maintained by various institutions, provides opportunities to enhance spatial coverage and capture local conditions. These sources, when integrated into the AEMET network, expand the heterogeneity of the dataset and offer a more comprehensive representation of climate patterns. The integration of diverse meteorological data sources, while promising, introduces challenges related to data quality, series continuity, and harmonization. The unique characteristics of each data source and the lack of data quality assurance can contribute to uncertainties, just as discontinuities in the time series can lead to problems in climatological modeling. These issues require a careful assessment of the spatial distribution and reliability of the integrated data set. This assessment is essential for the generation of accurate and robust climate maps.

In the context of this research, the present study focuses on two key meteorological variables: monthly accumulated precipitation and monthly air temperature (mean minimum, mean and mean maximum). These four variables play a pivotal role in climate studies as they influence various ecological processes and human activities and previous studies such as Ninyerola *et al*., 2007a and

Ninyerola *et al*., 2007b have successfully modeled them using similar methodologies and data sets as those presented in this study.

The availability of precipitation and temperature data is limited to specific ground points, recorded by weather station networks, so values at other point locations on the ground must be inferred from those nearby or from their relationship with other geographic variables (Marquínez *et al*., 2003). This leads to the main issue in climatological modeling: data availability. Weather stations in data networks are often not adequately distributed to capture the variability of temperature and precipitation patterns (Moral, 2010).

In this study, both climate elements share a common methodology that combines multiple linear regression with spatial interpolation of the residuals. This approach effectively integrates geographic information to interpolate climate variables. The residuals from the regression, reflecting unexplained variability, are then interpolated across the study area. This unified approach enables us to assess the contributions of supplementary networks, investigate their impact on climate variable patterns, and evaluate the possible improvements introduced by their integration.

This research delves into this integration of complementary meteorological networks with the AEMET network. It undertakes a comparative analysis of the two datasets: the official dataset, consisting of the AEMET network alone, and an extended dataset that incorporates additional meteorological networks, focusing on the evaluation of the regression model and the quantification of uncertainty in the resulting spatial maps. This evaluation uses the adjusted coefficient of determination (R^2) to measure model performance, considering the number of predictors in the model (Harel, 2009), the selection of predictor variables to assess the captured variability and the root mean square error (RMSE) to measure map uncertainty. The results provide insight into model effectiveness and map reliability and highlight the benefits and challenges of integrating additional data sources, to determine the extent to which these supplementary networks positively influence the accuracy and reliability of climate models.

2 Study Area

The study area is located in the south of Spain and consists of the hydrographic basin of the Guadalquivir River. It is defined by the following UTM-30 N coordinates (in km and datum ETRS89): 165 (minimum X coordinate), 573 (maximum X coordinate), 4066 (minimum Y coordinate), and 4289 (maximum Y coordinate) with a total area of approximately 58 000 km^2, see Figure 1. The Guadalquivir River basin extends across 12 provinces belonging to four autonomous communities, with Andalusia accounting for more than 90 % of the basin's area (Confederación Hidrográfica del Guadalquivir, n.d.). This area experiences a Mediterranean climate characterized by its warm and temperate nature (average annual air temperature of 16.8°C) and highly irregular precipitation (annual average of 550 mm). Due to its exposure to the Atlantic Ocean, intense southwesterly winds of subtropical character easily penetrate the territory, associated with storms that bring heavy rainfall (García de Pedraza, 1976). These winds result in a southwest to northeast distribution of precipitation, peaking in the highest peaks that delimit the watershed. These factors create a dry summer climate, with up to five months experiencing minimal or even no rainfall, in stark contrast to the winter months, which are characterized by torrential rains.

Although the study area focuses on this specific region, a buffer of 50 km around the basin has also been considered. This buffer ensures that the interpolation process performed is not affected by data gaps at the boundaries of the study area.

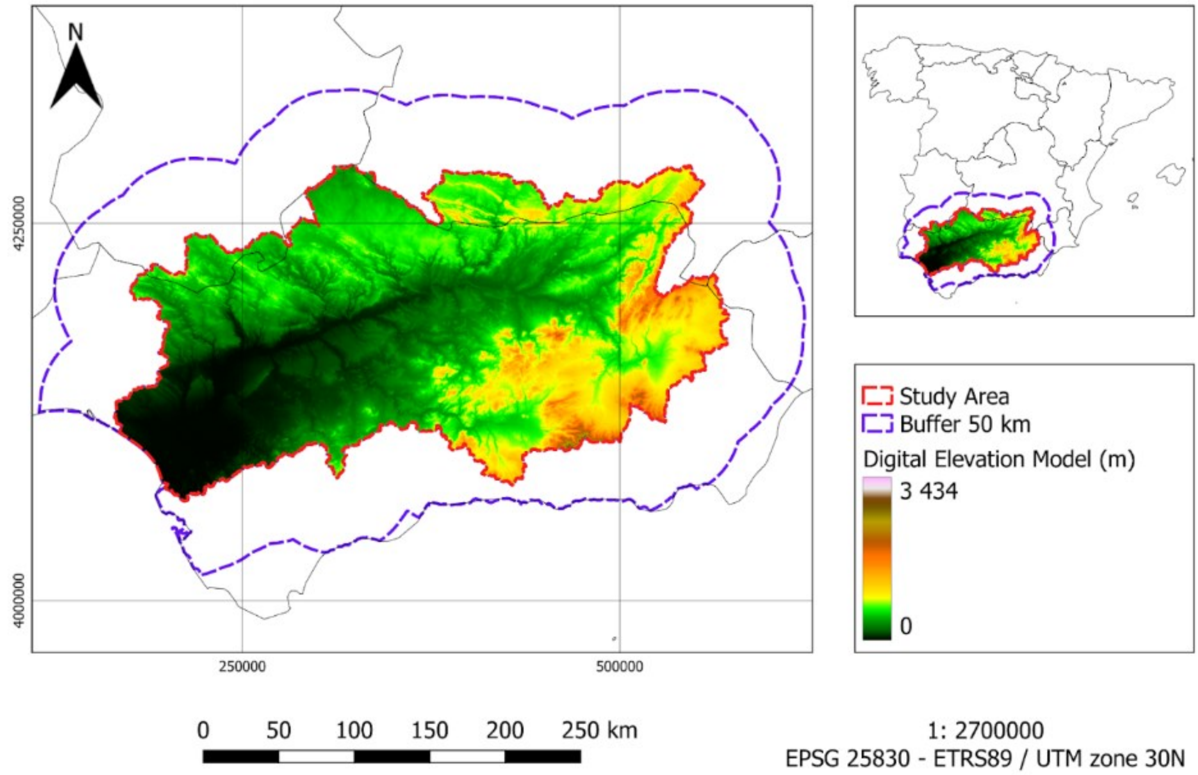


Figure 1: DEM of the Study Area in the Guadalquivir basin and a 50 km buffer zone.

3. Material

3.1 Meteorological stations

Based on SPEI drought monitoring service (Consejo Superior de Investigaciones científicas, n.d.) and statistical yearbook data on precipitation and temperature in the area of interest (Ministerio de Agricultura, P. y A., n.d.), three years have been selected for the study. The selection was made by identifying years wherein these parameters exhibited contrasting behaviors. The year 2003 was chosen as a year without significant anomalies, 2006 as a slightly dry and hot year and 2011 as a slightly wet and cold year. For each year data from four specific months corresponding to the solstices and equinoxes, namely March, June, September and December was processed. A representative random sampling could be sufficient to compare the results obtained with different datasets. However, we selected evenly spaced samples throughout each year that cover a wide range of climatic situations in order to test the results under different climatic conditions, while ensuring that samples contain auxiliary stations spatially distributed throughout the whole study area and the buffer area.

Monthly meteorological data have been obtained from AEMET for the corresponding months and years. The downloaded AEMET network for temperature consists of 433 weather stations covering an altitude range between 3 and 1800 m, with a density in the study area of one station per 134 km^2. For precipitation, the downloaded AEMET network consists of 846 stations covering the same altitude range but with a density of one station per 68.5 km^2. The records of these stations were already filtered to ensure homogeneity with neighboring stations and continuous time series of at least 60 observations.

Additional meteorological data have been obtained from a total of 11 auxiliary meteorological networks (Table 1). For temperature, a total of 283 meteorological stations belonging to auxiliary networks have been obtained, resulting in a spatial density of one station every 205 km^2 with altitudes ranging between 1 and 1603 m. For precipitation, the complementary networks provided 308 stations between 1 and 2002 m, with a density of one station every 188.3 km^2.

In total the spatial density of the Extended dataset for temperature is one station every 81 km^2 and for precipitation one station every 50 km^2.

Table 1: Name and acronym of the meteorological networks for temperature and precipitation

Network	Acronym
Agencia Estatal de Meteorología https://www.aemet.es/es/portada	AEMET
Instituto de Educación Secundaria Los Pedroches https://ieslospedroches.com/	Pedroches
Red de Estaciones Automáticas y Remotas de Meteorología https://www.juntadeandalucia.es/medioambiente/servtc5/WebClima/	EARM
Red de Estaciones Meteorológicas https://www.juntadeandalucia.es/medioambiente/servtc5/WebClima/	REM
Red de Alerta e Información Fitosanitaria https://www.juntadeandalucia.es/agriculturapescaaguaydesarrollorural/raif/	CAPDR
Red del Servicio de Calidad Ambiental https://www.miteco.gob.es/es/calidad-y-evaluacion-ambiental/	CMAOT
Instituto Andaluz de Investigación y Formación Agraria, Pesquera, Alimentaria y de la Producción Ecológica https://www.juntadeandalucia.es/agriculturaypesca/ifapa/web/	IFAPA
Red Sistema de Información Agraria de Murcia http://siam.imida.es/apex/f?p=101:1:8443957909132639	SIAM
Red Sistema de Información Agraria para el Regadío https://eportal.mapa.gob.es/websiar/Inicio.aspx	SIAR
Red SAIH de la cuenca mediterránea andaluza https://www.chguadalquivir.es/saih/	SAIH
Red de la Cuenca mediterránea Andaluza http://www.redhidrosurmedioambiente.es/saih/	CMA
Red de Estaciones no Oficiales https://www.juntadeandalucia.es/medioambiente/servtc5/WebClima/	PAART

In general, these additional data sources contain temporal gaps and often imprecisely calibrated instruments, so data filtering was an essential step prior to processing. Months with less than 29 days of data availability for temperature and more than one NODATA value for precipitation have been dismissed from the analysis to reduce potential sources of error and maintain representativeness, taking into account that precipitation is more sensitive to missing values than temperature.

The auxiliary dataset also underwent a filtering process to remove data records considered to be outliers, as in clearly erroneous values, specifically excluding temperature values that fell outside the range of the absolute maximum and minimum temperatures, historically recorded by the AEMET network in the region from 1920 to 2023 (AEMET, n.d.), plus a small margin of 1.5 ºC according to the IPCC AR15 report. Precipitation values outside this historical range of the AEMET network records in the region (AEMET, n.d.) have also been excluded without any margin, as there were no values recorded by auxiliary networks close enough above this boundary to be considered potentially real rather than outliers.

The total number of stations, including those for temperature and precipitation, amounts to 1371 stations belonging to AEMET (1002) and 11 auxiliary networks (369) between 1 and 2002 m, with a spatial density of one station every 42.3 km^2 (Figure 2).

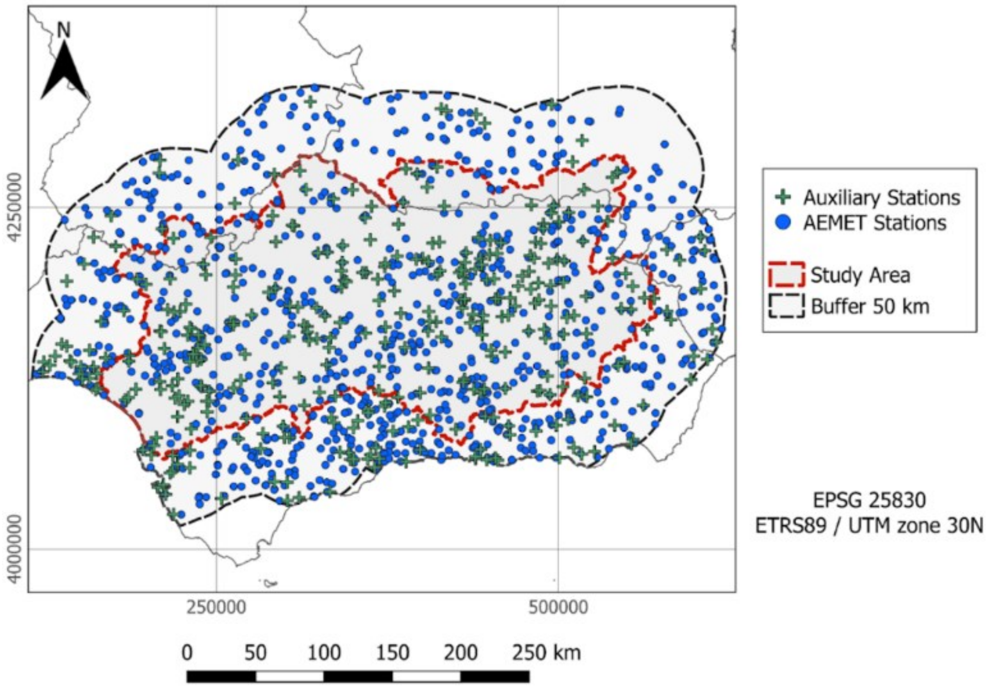


Figure 2: Geographical location of the AEMET and Auxiliary meteorological stations.

3.2 Independent variables

The geographical variables used in the multiple linear regression model were raster maps, generated from a Digital Elevation Model derived from official PNOA data (Plan Nacional de Ortofotografía Aérea, n.d.), with a spatial resolution of 100 m for altitude, potential solar radiation, cosine of aspect, cosine of slope, quadratic distance to the Mediterranean Sea and quadratic distance to the Atlantic Ocean. Altitude has a substantial influence on weather (Schermerhorn, 1967), particularly on temperature, which typically decreases with altitude, making it an excellent statistical predictor of this variable (Barry & Chorley, 1968). In complex terrain, slope and aspect can significantly improve the climate model (Daly *et al*., 1993) by providing a more detailed picture of the topographic conditions and can also influence cloud formation or wind circulation (Tullot, 2000). In addition, the climate regime can be strongly influenced by the proximity of a significant body of water (Daly *et al*., 2002). Finally, solar radiation can serve as an additional predictor due to its direct relationship with air temperature (Shrestha *et al*., 2019) and its topographic information related to cloud formation (Ninyerola *et al*., 2007b).

4. Methodology

This study aims to determine whether the description of the patterns of spatial variability of the variables is improved by comparing the same modeling method with two different datasets. Therefore, we have adopted the methodology proposed by Ninyerola *et al*., 2007a and Ninyerola *et al*., 2007b, which involves performing a multiple regression analysis with residual interpolation using meteorological station data and independent geographic variables, see Figure 3. First, a multiple linear regression analysis (MLR) has been conducted with 100 % of the meteorological stations and, in a second step, a multiple linear regression analysis with residual interpolation (MLR+I) using only 75 % of the AEMET stations.

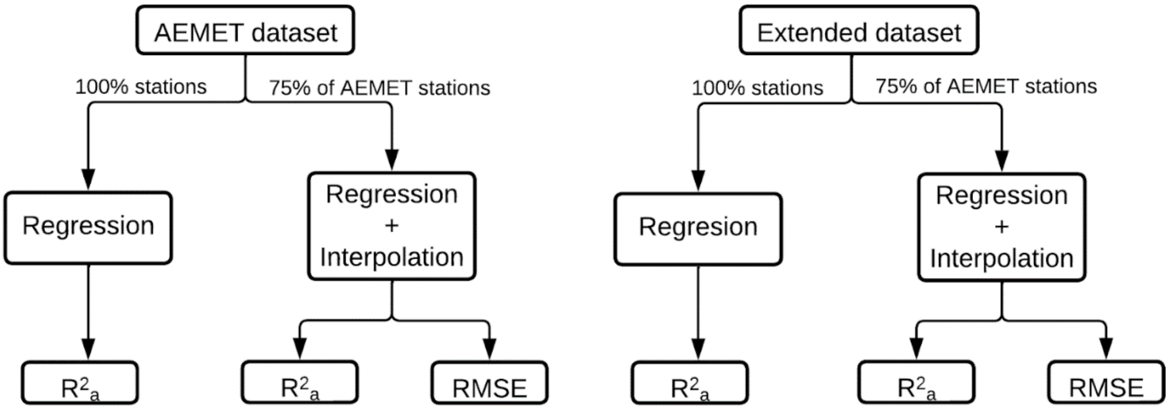


Figure 3: Flowchart of the methodology proposed by Ninyerola et al., 2007a and Ninyerola et al., 2007b adopted in this study. R^2_a signifies adjusted regression coefficient.

4.1 Multiple Linear Regression

The multiple linear regression analysis has been conducted in the GIS MiraMon using the *RegMult* tool (Pesquer *et al*., 2007) twice: once using the official AEMET network as the dependent variable and once using the extended dataset with the official AEMET network combined with the auxiliary networks as the dependent variable (for each one: precipitation and min, mean and max. temperatures). These analyses focused on the selected years and months. The resulting adjusted R^2 values, which represent the adjusted coefficient of determination, are presented in the 'Results' section.

In addition, a comparison was made to determine which independent variables were deemed statistically significant in both the official and extended models. This analysis helped to identify the contribution of the auxiliary networks into the variability of the stations in the different independent geographical variables.

4.2 Interpolation of residuals

In a second phase, the residuals, which represent the differences between the predicted values generated by the multivariate regression model and the observed values, were calculated. These residuals were then subjected to interpolation using the inverse distance weighting model with a cubic exponent in the case of monthly temperature variables. For monthly precipitation the splines interpolation technique (Mitasova & Mitas, 1993) with a tension of 800 was applied. This selection was made following an evaluation of interpolation method outcomes, which showed that the use of splines was able to reduce the RMSE values compared to the inverse of the distance technique only for the precipitation variable.

Unlike traditional approaches that directly interpolate climatological values, the *RegMult* tool leverages the predictive capabilities of the regression model by interpolating the residuals. Only the unexplained variation obtained from the multiple regression analysis is incorporated into the spatial interpolation process (Pons, 2004). After the calculation of the residuals, meteorological stations of the auxiliary networks with systematic errors, resulting in consistently extremely high residuals, were removed from the calculation.

4.3 Validation

In order to validate the developed models, a random subset of 25 % of the ground meteorological stations within the study area from the AEMET network was set aside. RMSE have been computed for each model, providing a quantitative measure of the uncertainty. The 'Results' section presents the RMSE analysis for each model, providing insight into the accuracy and precision of the generated models.

5. Results

For each run of the *RegMult* tool (flowchart Fig 3.), raster maps of monthly accumulated precipitation or monthly temperature with a spatial resolution of 100 m were obtained. For each of these maps, the effect of the auxiliary networks on the explained variability was studied by comparing the adjusted R^2 values obtained by each dataset, the effect on the interpolation of residuals with the RMSE values and the captured variability by consulting the predictor selection.

Upon examination of the results, it was found that there were no significant visual differences between the maps generated using the official and extended datasets (Figure 4). However, the extended dataset did introduce slightly more contrast in certain areas and subtle circular artifacts in some maps.

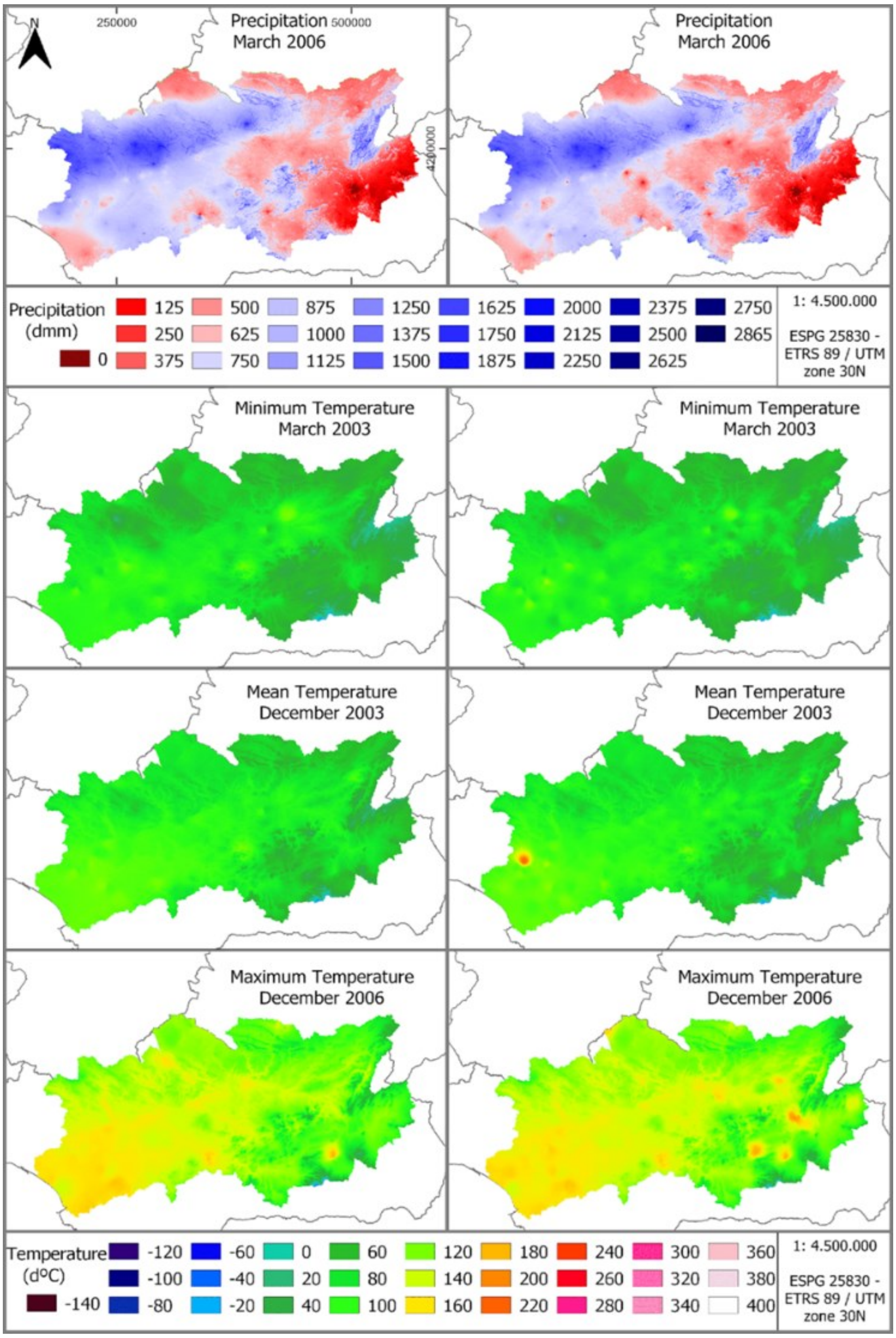
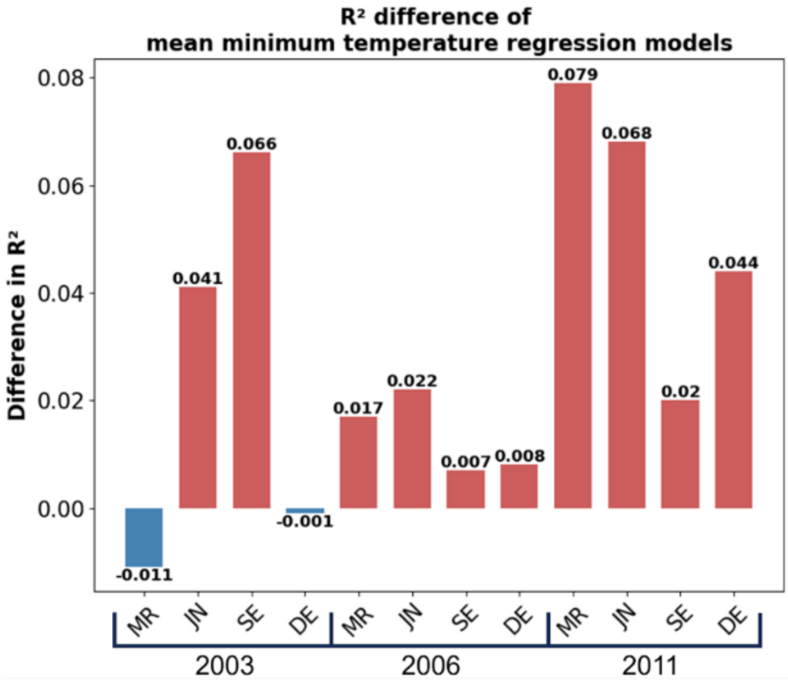


Figure 4: Comparison between official (left) and extended (right) regression and interpolation results for each variable.

5.1 Adjusted R^2 of the Regression Models

5.1.1 Temperature data

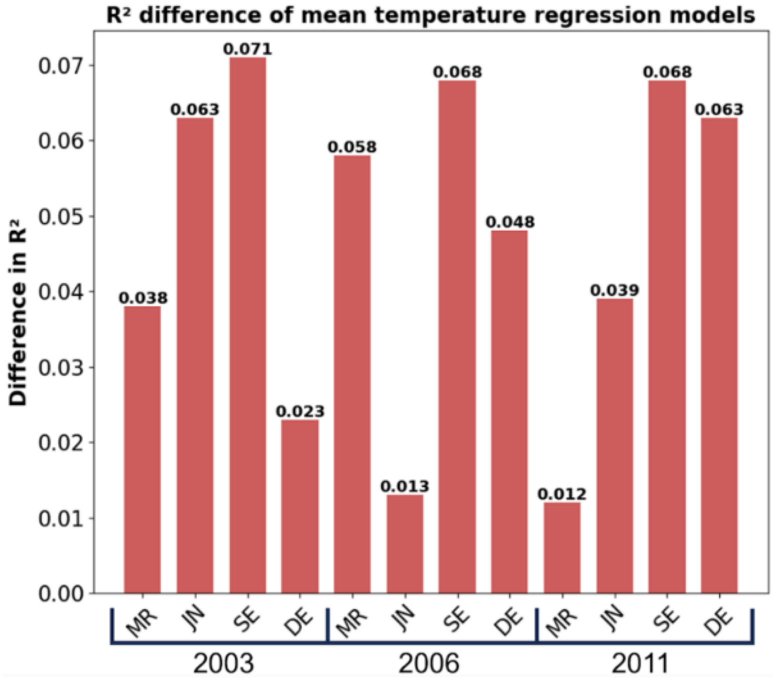
Based on the results of the multiple regression analysis, it was observed that the official dataset consistently achieved higher adjusted R^2 values compared to the extended dataset. However, there were certain cases where the inclusion of auxiliary networks improved model performance, specifically in modeling mean minimum and mean maximum temperatures, see Figures 5 to 7. The improvements in adjusted R^2 range from 0.001 to 0.041. In the months of March and December 2003, as well as March 2011, the mean minimum temperature models showed slight improvements when incorporating the auxiliary networks alongside the official AEMET data. Similarly, for the March 2003, June 2006 and March 2011 mean maximum temperature models, the inclusion of auxiliary data resulted in slight enhancements.



MODEL	OFFICIAL adjusted R^2	EXTENDED adjusted R^2
MR 2003	0.585	0.596
JN 2003	0.436	0.395
SE 2003	0.549	0.482
DE 2003	0.601	0.602
MR 2006	0.528	0.511
JN 2006	0.747	0.725
SE 2006	0.203	0.196
DE 2006	0.333	0.325
MR 2011	0.436	0.358
JN 2011	0.364	0.295
SE 2011	0.443	0.423
DE 2011	0.487	0.443

MR: March **JN**: June
SE: September **DE**: December

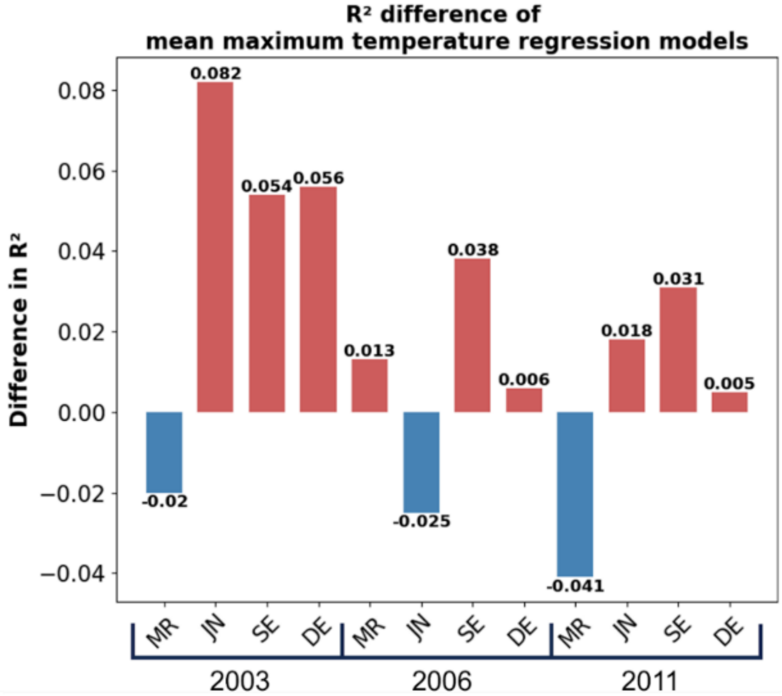
Figure 5: Adjusted R^2 difference of mean minimum temperature MLR models between the official models and the extended models (left graphic) and adjusted R^2 values of each model (right table).



MODEL	OFFICIAL adjusted R^2	EXTENDED adjusted R^2
MR 2003	0.701	0.663
JN 2003	0.685	0.622
SE 2003	0.726	0.655
DE 2003	0.749	0.726
MR 2006	0.690	0.632
JN 2006	0.848	0.835
SE 2006	0.249	0.182
DE 2006	0.396	0.348
MR 2011	0.398	0.386
JN 2011	0.576	0.537
SE 2011	0.623	0.555
DE 2011	0.549	0.486

MR: March **JN**: June
SE: September **DE**: December

Figure 6: Adjusted R^2 difference of mean temperature MLR models between the official models and the extended models (left graphic) and adjusted R^2 values of each model (right table).



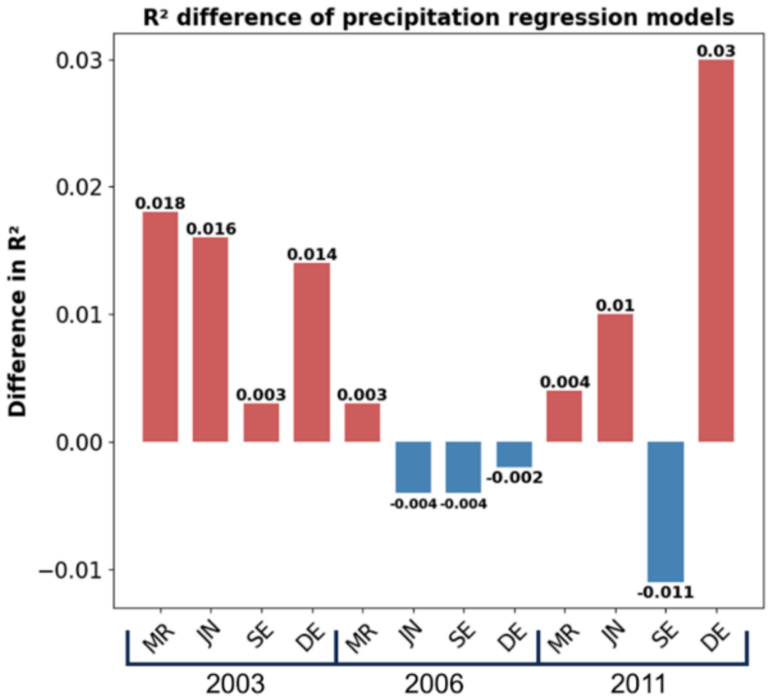
MODEL	OFFICIAL adjusted R^2	EXTENDED adjusted R^2
MR 2003	0.596	0.615
JN 2003	0.706	0.624
SE 2003	0.726	0.670
DE 2003	0.585	0.529
MR 2006	0.587	0.574
JN 2006	0.753	0.778
SE 2006	0.175	0.137
DE 2006	0.263	0.256
MR 2011	0.261	0.302
JN 2011	0.411	0.393
SE 2011	0.478	0.447
DE 2011	0.321	0.316

MR: March **JN**: June
SE: September **DE**: December

Figure 7: Adjusted R^2 difference of mean maximum temperature MLR models between the official models and the extended models (left graphic) and adjusted R^2 values of each model (right table).

5.1.2 Precipitation data

The results of the MLR analysis using the 100 % of the meteorological stations as model input showed that the extended dataset obtained lower adjusted R^2 values than the official dataset (Figure 8). Although the adjusted R^2 values of the model using the extended dataset were higher in September 2011 and in most months of 2006, this improvement consists of much smaller increments than the decrease in the adjusted R^2 values that occurs using this extended dataset in the remaining months.



MODEL	OFFICIAL adjusted R^2	EXTENDED adjusted R^2
MR 2003	0.309	0.291
JN 2003	0.264	0.248
SE 2003	0.064	0.061
DE 2003	0.456	0.442
MR 2006	0.445	0.442
JN 2006	0.222	0.226
SE 2006	0.015	0.019
DE 2006	0.301	0.303
MR 2011	0.186	0.182
JN 2011	0.117	0.107
SE 2011	0.028	0.039
DE 2011	0.125	0.095

MR: March **JN**: June
SE: September **DE**: December

Figure 8: Adjusted R^2 difference of precipitation MLR models between the official models and the extended models (left graphic) and adjusted R^2 values of each model (right table).

5.2 Predictor Selection

When calculating the variances of the weather stations in each dataset for the independent variables (Table 2), it was observed that for temperature most of the extended datasets have a higher variance. For precipitation, only two variables have a higher variance in the extended dataset, but the differences between the datasets are not as large as for temperature.

Table 2: Variance in the independent variables of the official and the extended dataset for temperature (T) and precipitation (P).

MODEL	Official Variance (T)	Extended Variance (T)	Official Variance (P)	Extended Variance (P)
Altitude (m)	105 035	113 703	112 841	113 515
Slope (º)	21.28	25.92	26.95	26.64
Aspect (º)	11 828	12 915	11 696	11 775
Distance to the Atlantic (km)	2 350	2 580	2 270	2 250
Distance to the Mediterranean (km)	432	368	412	406
Solar Radiation (10 $kJ*m^{-2}*dia^{-1}$)	11 297	9 756	15 026	14 891

5.2.1 Multiple Linear Regression with temperature data

When examining the geographical variables incorporated into the *MLR* models, several key observations can be made. Altitude was identified as the most important variable, as it was included in 100 % of both the official models and the extended models. Distance to the Mediterranean Sea was included in 89 % of the official models and 83 % of the extended models, highlighting its significant influence. Similarly, the distance to the Atlantic Ocean was considered important, appearing in 67 % of the official models and 75 % of the extended models. Aspect, representing the direction a slope faces, was found to be included in 64 % of the official models and 69 % of the extended models. Slope, indicating the steepness of the terrain, was included in 50 % of the official models and 53 % of the extended models, suggesting its moderate importance. Interestingly, solar radiation was deemed insignificant in the majority of the models, appearing in only 25 % of the official models and 22 % of the extended models.

In the *MLR+I* run, with only 75 % of the AEMET stations, altitude emerged as the most important variable, being included in 100 % of both the official and the extended models. Distance to the Mediterranean Sea was included in 81 % of the official models and 86 % of the extended models. Distance to the Atlantic Ocean appeared in 61 % of the official models and 81 % of the extended models. Aspect was found to be included in 53 % of the official models and 67 % of the extended models. Slope and solar radiation were found to be the least significant independent variables, with slope being included in 33 % of the official models and 58 % of the extended models and solar radiation being involved in 47 % of the official models and 33 % of the extended models.

In general, it can be noticed that the introduction of auxiliary networks to the regression models leads to a slight increase in the inclusion of predictor variables. The most substantial distinction becomes evident when modeling monthly mean temperatures. For further details see Appendix B.

5.2.2 Multiple Linear Regression with precipitation data

Looking at the independent variables included into the regression models, thus with 100 % of the AEMET stations, three of the geographical variables stood out from the rest, altitude, distance to the Mediterranean Sea and distance to the Atlantic Ocean being the most important. Altitude and distance to the Mediterranean Sea were included in 83 % of the official models and 100 % of the extended models, while distance to the Atlantic Ocean was included in 92 % of the official and the extended models. Slope showed a moderate importance, being incorporated in 100 % of the official models but only in 58 % of the extended models. Aspect and radiation were clearly the less important ones, with a similar performance. Aspect was included in only 58 % of the official models and in 67 % of the extended models. Meanwhile solar radiation followed the opposite pattern, being incorporated in 67 % of the official models and in 58 % of the extended models.

In general, for the *MLR* run, the addition of the auxiliary networks did not show a significant improvement in terms of the geographical variables included in the regression model. In certain months, for some excluded variables such as altitude, the auxiliary networks allowed them to be integrated into the regression model, but this contribution is canceled out by the opposite effect that the extended dataset had on other variables such as slope, where the use of the auxiliary networks meant that slope was less included in the regression.

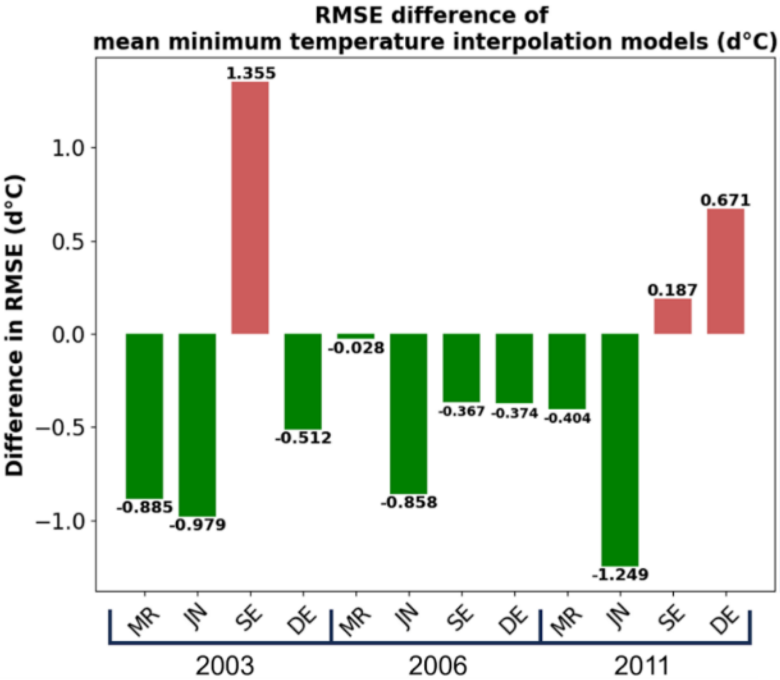
For the *MLR+I* run, therefore without 25 % of the AEMET stations, distance to the Atlantic Ocean and distance to the Mediterranean Sea were the most important independent variables, being present in the same months as in the previous *MLR* run. On the other hand, the altitude variable, although still quite important, was only included in 75 % of the official models, but still 100 % of the extended models. Slope and aspect showed major changes in their behavior, with aspect being more present in the models, included in 67 % of the official models and in 83 % of the extended models and slope being more frequently excluded from the regression models, being in 67 % of the official and the extended models. Solar radiation was again the less important independent variable, included in only 50 % of the official models and 58 % of the extended models.

In this case, fewer variables were included in the models when using the official dataset compared to the previous *MLR* run. However, using the extended dataset allowed more independent geographical variables to be included in the regression models. For further details see Appendix B.

5.3 RMSE of the interpolation

5.3.1 Temperature data

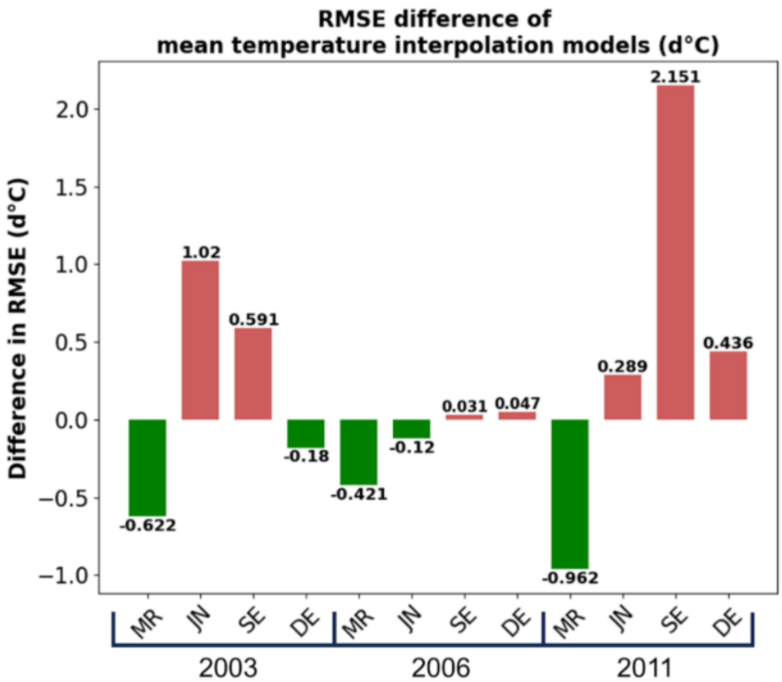
Regarding the results of the *MLR+I* models, the inclusion of the auxiliary meteorological networks improved 19 of the 36 temperature models generated, see Figures 9 to 11. Particularly, in the case of mean minimum temperature, the addition of the complementary networks led to an improvement in nine out of the twelve months processed, with a lower RMSE ranging between 0.028 and 1.249 d°C. For both mean and maximum temperature models, five out of twelve months obtained a lower RMSE by adding meteorological auxiliary networks, with March 2011 being the most improved month (-0.962 d°C RMSE). Regarding the seasonal distribution of the results, March models improved eight out of nine times, June models improved five out of nine times, December models improved five out of nine times, and only one model for September (2006 mean minimum) obtained a lower RMSE when adding auxiliary meteorological networks. Looking at the processed years, the 2003 temperature models achieved a lower RMSE five out of twelve times, models for 2006 eight out of twelve times, and the error of the 2011 models decreased six out of twelve times.



MODEL	OFFICIAL RMSE (d°C)	EXTENDED RMSE (d°C)
MR 2003	15.48	14.59
JN 2003	19.06	18.09
SE 2003	15.87	17.22
DE 2003	13.31	12.80
MR 2006	17.94	17.91
JN 2006	11.15	10.30
SE 2006	20.57	20.20
DE 2006	17.88	17.50
MR 2011	16.85	16.44
JN 2011	18.44	17.20
SE 2011	18.48	18.66
DE 2011	14.76	15.43

MR: March **JN**: June
SE: September **DE**: December

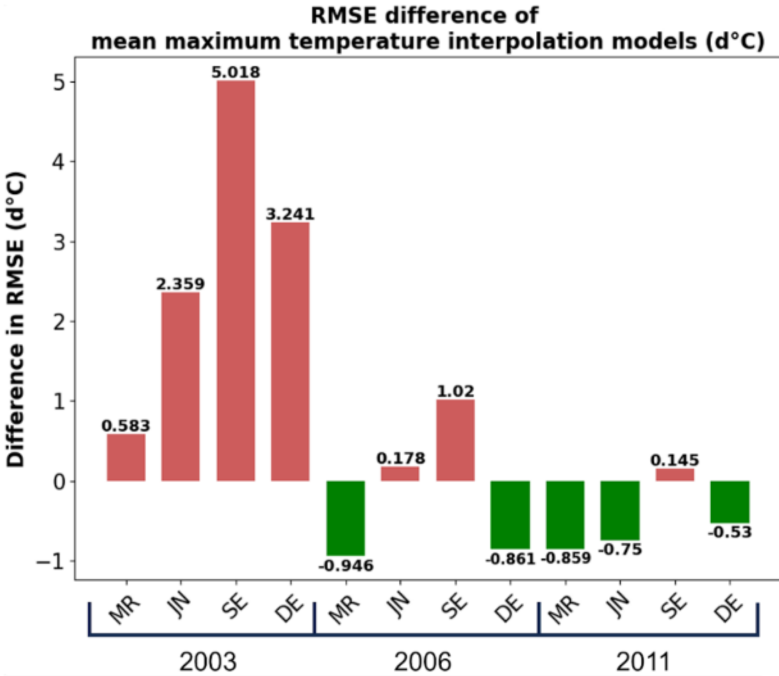
Figure 9: RMSE difference of mean minimum temperature MLR+I models between the official and the extended models (left graphic) and RMSE values of each model (right table).



MODEL	OFFICIAL RMSE (d°C)	EXTENDED RMSE (d°C)
MR 2003	10.15	9.53
JN 2003	11.92	12.94
SE 2003	9.21	9.81
DE 2003	8.63	8.45
MR 2006	14.69	14.27
JN 2006	8.53	8.41
SE 2006	15.51	15.54
DE 2006	14.01	14.06
MR 2011	12.59	11.63
JN 2011	13.04	13.33
SE 2011	13.85	16.00
DE 2011	8.45	8.88

MR: March **JN**: June
SE: September **DE**: December

Figure 10: RMSE difference of mean temperature MLR+I models between the official and the extended models (left graphic) and RMSE values of each model (right table).



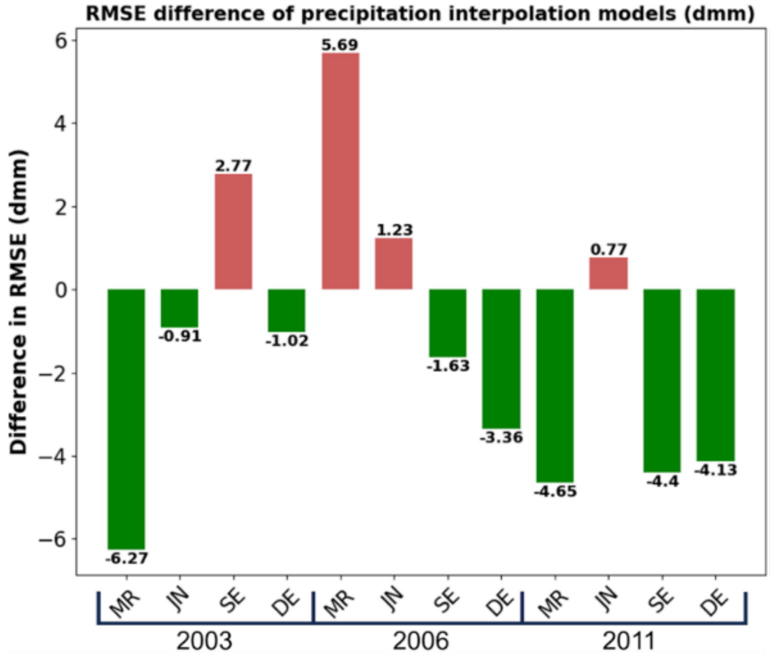
MODEL	OFFICIAL RMSE (d°C)	EXTENDED RMSE (d°C)
MR 2003	10.58	11.16
JN 2003	12.25	14.61
SE 2003	9.70	14.72
DE 2003	11.23	14.47
MR 2006	15.96	15.01
JN 2006	11.47	11.65
SE 2006	18.84	19.86
DE 2006	18.78	17.92
MR 2011	16.15	15.29
JN 2011	18.28	17.53
SE 2011	17.06	17.21
DE 2011	11.42	10.89

MR: March **JN**: June
SE: September **DE**: December

Figure 11: RMSE difference of mean maximum temperature MLR+I models between the official and the extended models (left graphic) and RMSE values of each model (right table).

5.3.2 Precipitation data

The results of the *MLR+I* models showed that the RMSE values also varied depending on the dataset used (Figure 12), with a predominance of lower RMSE values when using the extended dataset. For four months the RMSE values were lower when using the official dataset, but for the remaining eight months the addition of the auxiliary networks reduced the RMSE error associated with the resulting maps, with a RMSE reduction between -0.91 dmm and -6.27 dmm.



MODEL	OFFICIAL RMSE (dmm)	EXTENDED RMSE (dmm)
MR 2003	161.74	155.47
JN 2003	68.17	67.26
SE 2003	120.24	123.01
DE 2003	235.52	234.50
MR 2006	206.56	212.25
JN 2006	143.26	144.49
SE 2006	158.07	156.44
DE 2006	102.12	98.76
MR 2011	232.54	227.89
JN 2011	107.43	108.20
SE 2011	83.38	78.98
DE 2011	59.59	55.46

MR: March **JN**: June
SE: September **DE**: December

Figure 12: RMSE difference of precipitation MLR+I models between the official models and the extended models (left graphic) and RMSE values of each model (right table).

6. Discussion

For precipitation and in most of the temperature cases, the official model achieved higher adjusted R^2 values than the extended models. However, there were slight improvements in the adjusted R^2 when modeling mean minimum temperature and mean maximum temperature by including the auxiliary stations. In general, the auxiliary networks do not contribute positively to explain the distribution of the dependent variable.

When considering the RMSE, the results were more varied. The extended models generally had a slightly lower RMSE score for precipitation and mean minimum temperature, and in almost half of the cases for mean and mean maximum temperature. The auxiliary networks provide more ground point locations that contribute to the spatial interpolation, which is expected to reduce the uncertainty of the maps.

Regarding the selection of predictor variables, the addition of the auxiliary networks generally allowed slightly more independent variables to be included in the regression models. It seems that the addition of the auxiliary networks provided further information or variability within these geographical variables that was not captured by the official data alone. This assertion is supported by the variance calculations for station values in each geographical variable, which showed a consistent pattern where the inclusion of independent variables was contingent upon the variance. In general, when the inclusion of auxiliary data allowed for the incorporation of more independent variables into the multiple linear regression models, it was observed that the variance of the extended dataset in these independent variables was higher compared to the official dataset (Table 2). Conversely, when the inclusion of auxiliary data resulted in a reduction in the number of independent variables included in the model, it was usually observed that the variance of the geographical variables that were now excluded from the model was higher or similar in the official dataset. The auxiliary data contributes to a wider range of measurement scenarios that may capture nuances or fluctuations in the variables that the official data might miss. As a result, the model is better able to explain the variance in the dependent variable, leading to the inclusion of more independent variables.

7. Conclusions and future work

The complexity of climate spatiotemporal variability makes it difficult for meteorological networks, limited to specific ground station locations, to capture climate variability patterns. To assess this data availability problem, we have worked with the inclusion of auxiliary networks of meteorological stations in the AEMET network for climate mapping. We aimed to study their contributions to the variance explained by regression models, predictor selection and map uncertainty, to assess whether these auxiliary networks provide crucial information or add more noise to the models.

For all variables, the inclusion of the auxiliary networks had a light negative effect on the models' performance, resulting in lower adjusted R^2 values, especially for mean minimum temperature and mean temperature. However, these networks capture more variability and allow for the inclusion of more independent variables in the regression models. In terms of map uncertainty, the results were more varied. For mean minimum temperature and precipitation, the RMSE values decreased in the vast majority of months with the extended dataset, but for mean and maximum temperature, the error of the maps decreased only in a small number of months and to a lesser extent than in the months in which it increased.

Considering these results and their magnitude, which consist in a maximum adjusted R^2 decrease of 0.082 for temperature models and 0.03 for precipitation models, and a maximum RMSE improvement of 1.24 d°C for temperature models and 6.27 dmm for precipitation models when using the extended dataset, the key takeaway message is that the incorporation of these auxiliary networks does not yield substantial improvements to consider their implementation in the official network for climate mapping, especially when considering the amount of additional work of filtering, data harmonization and transformation they require. Indeed, the addition of auxiliary networks provides insights for an extension of the AEMET network in this region: more variability on topographic features and a light reduction in map uncertainty.

However, it is important to acknowledge the limitations of this study. Expanding the dataset with a larger sample size can provide a richer perspective in the metrics behavior. Additionally, focusing on geographically challenging regions like mountainous areas, where the lack of meteorological stations can have a serious impact on capturing climate variability patterns, the incorporation of the auxiliar networks can have significant effects in climate mapping.

For future studies, it would be interesting to further explore the spatialization of map uncertainty to better identify conflicting areas and gain additional insight at a regional scale. It may also be important to conduct a similar study in different regions, where the availability of auxiliary networks is high compared to official data. These areas may provide a clearer perspective on the impact of auxiliary networks on climate models. In addition, given the potential shown by automatic learning algorithms in regression processes (de Burgh-Day & Leeuwenburg, 2023), these algorithms could be used within the methodology to compare the results with those obtained from conventional methods.

6. Acknowledgements

The authors would like to express their gratitude to AEMET-Spain for the information provided (data, metadata, and technical reports) <https://www.aemet.es/es/sede_electronica>, as well as the institutions that maintain the complementary meteorological networks, for providing the meteorological data that have made this study possible.

This work was supported by European Union's Horizon 2020 research and innovation programme under I-CISK project, Grant Agreement 101037293, and it was partially funded by the Catalan Government under grant no. SGR2021 00554.

References

Agencia Estatal de Meteorología. (n.d.). *Valores Extremos: Andalucía*. Retrieved July 10, 2023, from https://www.aemet.es/es/serviciosclimaticos/datosclimatologicos/efemerides_extremos?w=0&k=and&datos=img

Barry, R. G., Chorley, R.J. (1968): Atmosphere, Weather and Climate. *The Geographical Journal*, Vol. 134, pp. 591-592. https://doi.org/10.2307/1796429

Confederación Hidrográfica del Guadalquivir. (n.d.). *Datos Históricos*. Retrieved May 25, 2023, from https://www.chguadalquivir.es/saih/

Confederación Hidrográfica del Guadalquivir. (n.d.). *La Demarcación Hidrográfica del Guadalquivir*. https://www.chguadalquivir.es/la-demarcacion-hidrografica-del-guadalquivir

Confederación Hidrográfica del Guadiana. (n.d.). *Ámbito Geográfico*. https://www.chguadiana.es/cuenca-hidrografica/medio-natural/biodiversidad

Consejo Superior de Investigaciones Científicas. (n.d.). *SPEI Global Drought Monitor.* Retrieved July 9, 2023, from https://spei.csic.es/map/maps.html#months=1#month=5#year=2023

Daly, C., Gibson, W.P., Taylor, G.H., Johnson, G.L., Pasteris, P. (2002). A knowledge-based approach to the statistical mapping of climate. *Climate Research*, Vol. 22, pp. 99-113. https://doi.org/10.3354/cr022099

Daly, C., Neilson, R.P., Phillips, D.L. (1993). A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *Journal Of Applied Meteorology*, Vol. 33, pp. 140-158. https://doi.org/10.1175/1520-0450(1994)033<0140:astmfm>2.0.co;2

de Burgh-Day, C. O., & Leeuwenburg, T. (2023). Machine learning for numerical weather and climate modelling: A Review. Geoscientific Model Development, 16(22), 6433–6477. https://doi.org/10.5194/gmd-16-6433-2023.

García de Pedraza, L. (1976). Clima de cuencas hidrográficas: Aplicación para el Ebro y Guadalquivir. *Calendario MeteorofenolóGico. 1977*, p.177-189. https://doi.org/http://hdl.handle.net/20.500.11765/8498

Harel, O. (2009). The estimation of R^2 and adjusted R^2 in incomplete data sets using multiple imputation. *Journal of Applied Statistics*, *36*(10), 1109–1118. https://doi.org/10.1080/02664760802553000

I. E. S. Los Pedroches. (n.d.). *Datos meteorológicos*. Retrieved July 11, 2023, from https://ieslospedroches.com/

Instituto Andaluz de Investigación y Formación Agraria, Pesquera, Alimentaria y de la Producción Ecológica. (n.d.). *Estaciones Meteorológicas*. Retrieved July 11, 2023, from https://www.juntadeandalucia.es/agriculturaypesca/ifapa/web/

Intergovernmental Panel on Climate Change. (2018). Summary for Policymakers. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. [Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA, pp. 3-24. https://doi.org/10.1017/9781009157940.001.

Junta de Andalucía. (2023). *Subsistema de Información de Climatología Ambiental.* Retrieved July 11, 2023, from https://www.juntadeandalucia.es/medioambiente/servtc5/WebClima/

Marquínez, J., Lastra, J., & García, P. (2003). Estimation models for precipitation in mountainous regions: the use of GIS and multivariate analysis. *Journal of Hydrology*, 270(1–2), 1–11. https://doi.org/10.1016/s0022-1694(02)00110-5

Mitasova, H., Mitas, L. (1993). Interpolation by regularized spline with tension: I. Theory and implementation. *Mathematical Geology*, Vol. 25, pp. 641-656. https://doi.org/10.1007/bf00893171

Ministerio de Agricultura, P. y A. (n.d.). *Anuario de Estadística*. Retrieved July 9, 2023, from https://www.mapa.gob.es/es/estadistica/temas/publicaciones/anuario-de estadistica/default.aspx

Moral, F. J. (2010). Comparison of different geostatistical approaches to map climate variables: application to precipitation. *International Journal of Climatology*, *30*(4), 620–631. https://doi.org/10.1002/joc.1913

Ninyerola, M., Pons, X., & Roure, J. M. (2005). *Atlas Climático Digital de la Península Ibérica.* http://opengis.uab.es/wms/iberia/index.htm

Ninyerola, M., Pons, X., Roure, J.M. (2007): Objective air temperature mapping for the Iberian Peninsula using spatial interpolation and GIS. *International Journal of Climatology*, Vol. 27, pp. 1231 – 1242. https://doi.org/10.1002/joc.1462

Ninyerola, M., Pons, X., & Roure, J. M. (2007). Monthly precipitation mapping of the Iberian Peninsula using spatial interpolation tools implemented in a Geographic Information System. *Theoretical and Applied Climatology*, *89*(3–4), 195–209. https://doi.org/10.1007/s00704-006-0264-2

Pesquer, L., Masó, J., & Pons, X. (2007). Integración SIG de regresión multivariante, interpolación de residuos y validación para la generación de rásters continuos de variables meteorológicas. *Revista de Teledetección* , *28*, 69–76. ISSN: 1133-0953

Plan Nacional de Ortofotografía Aérea. (n.d.). *Productos de descarga PNOA*. Retrieved May 5, 2023, from https://pnoa.ign.es/pnoa-imagen/productos-a-descarga

Pons, X. (2004). MiraMon. Sistema d'Informació Geogràfica i software de Teledetecció. *Centre de Recerca Ecològica i Aplicacions Forestals*. Bellaterra. ISBN: 84-931323-4-9

Schermerhorn, V. P. (1967). Relations between topography and annual precipitation in western Oregon and Washington. *Water Resources Research*, *3*(3), 707–711. https://doi.org/10.1029/WR003i003p00707

Shrestha, A.K., Thapa, A., Gautam, H. (2019). Solar Radiation, Air Temperature, Relative Humidity, and Dew Point Study: Damak, Jhapa, Nepal. *International Journal of Photoenergy*, Vol. 2019, pp. 1-7. https://doi.org/10.1155/2019/8369231

Sistema de Información Agraria de Murcia. (2023). *Efemérides de Estaciones*. Retrieved July 15, 2023, from http://siam.imida.es/apex/f?p=101:1:7718480372670535

Sistema de Información Agroclimática para el Regadio. (2023). *Consulta de datos*. Retrieved July 15, 2023, from https://eportal.mapa.gob.es/websiar/

Tullot, I. F. (2000). *Climatología de españa y portugal* (Vol 76). Universidad de Salamanca.

Appendix A: Independent Variables raster maps.

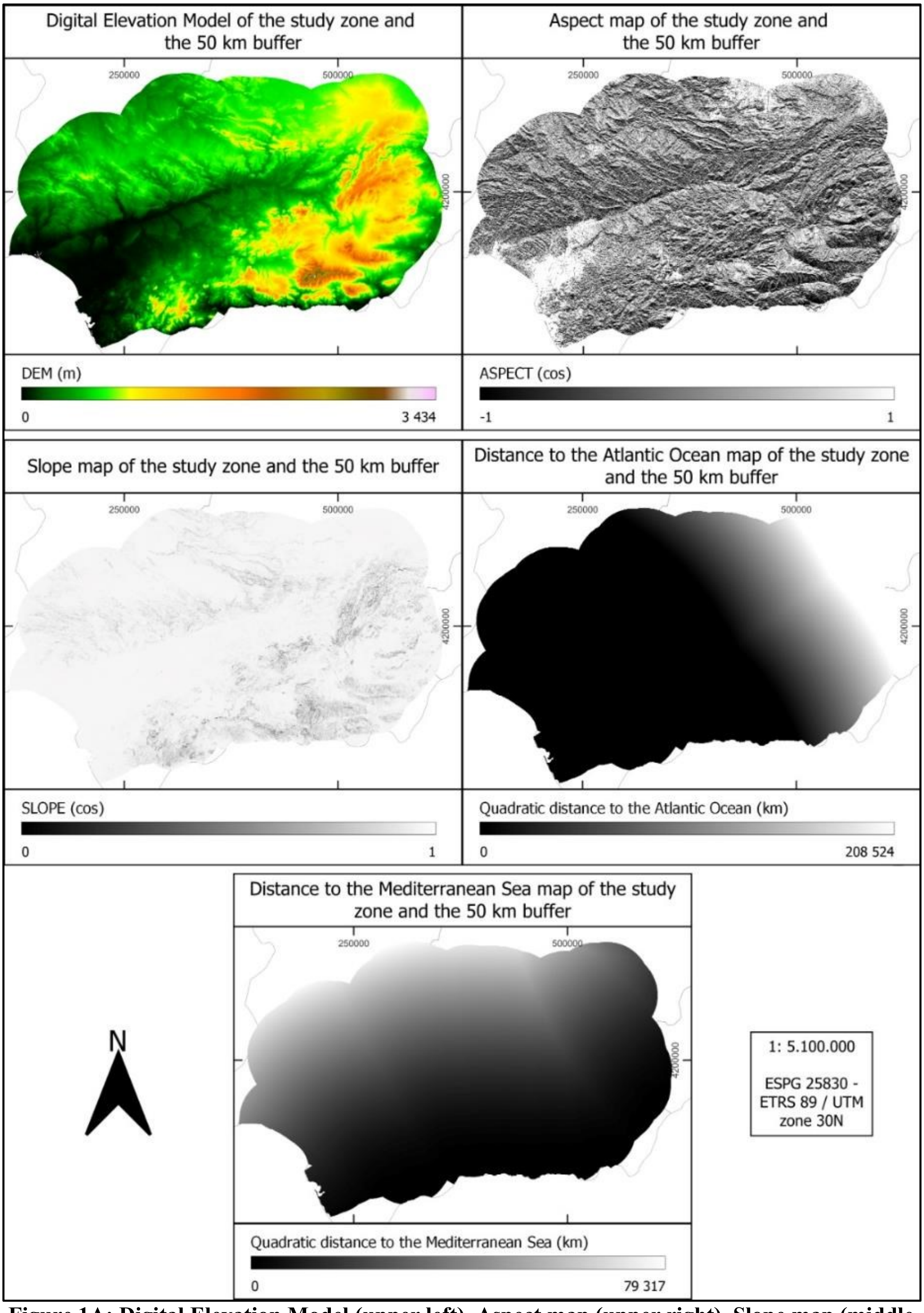


Figure 1A: Digital Elevation Model (upper left), Aspect map (upper right), Slope map (middle left), Distance to the Atlantic Ocean (middle right) and Distance to the Mediterranean Sea (bottom center) of the study zone and the 50 km buffer.

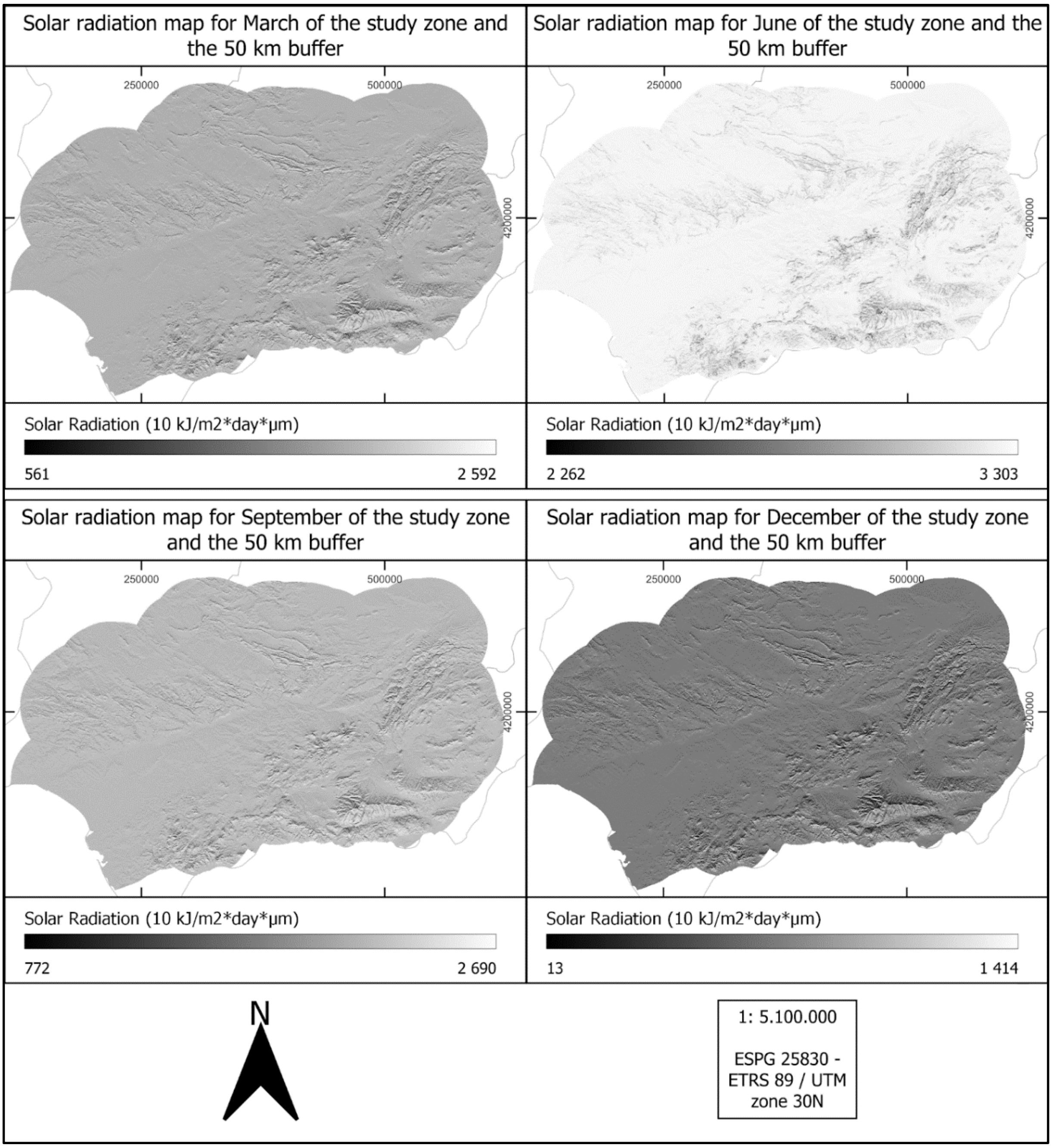


Figure 2A: Solar radiation map for March (upper left), Solar radiation map for June (upper right), Solar radiation map for September (bottom left) and Solar radiation map for December (bottom right) of the study zone and the 50 km buffer.

Appendix B: Predictor Selection of the Independent Variables.

Table 1B: Geographical variables that were included or excluded (X) for the mean minimum temperature MLR official models using 100 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Official Regression Mean Minimum	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003		X				
JUNE 2003		X		X		
SEPTEMBER 2003				X	X	
DECEMBER 2003				X		
MARCH 2006		X	X	X		
JUNE 2006		X	X	X		X
SEPTEMBER 2006			X			X
DECEMBER 2006		X	X			
MARCH 2011						
JUNE 2011		X			X	
SEPTEMBER 2011		X			X	
DECEMBER 2011		X				

Table 2B: Geographical variables that were included or excluded (X) for the mean temperature MLR official models using 100 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Official Regression Mean	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003		X	X			X
JUNE 2003		X	X	X		
SEPTEMBER 2003		X			X	
DECEMBER 2003		X				
MARCH 2006				X		
JUNE 2006		X	X	X		
SEPTEMBER 2006				X		
DECEMBER 2006		X				
MARCH 2011		X		X		
JUNE 2011		X	X		X	
SEPTEMBER 2011		X	X	X	X	X
DECEMBER 2011		X				

Table 3B: Geographical variables that were included or excluded (X) for the mean maximum temperature MLR official models using 100 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Official Regression Mean Maximum	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003		X	X	X	X	
JUNE 2003		X	X	X		
SEPTEMBER 2003		X	X	X		
DECEMBER 2003		X		X		
MARCH 2006				X		
JUNE 2006			X	X		
SEPTEMBER 2006			X	X	X	
DECEMBER 2006		X				
MARCH 2011						
JUNE 2011		X			X	
SEPTEMBER 2011		X			X	
DECEMBER 2011		X				

Table 4B: Geographical variables that were included or excluded (X) for the mean minimum temperature MLR extended models using 100 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Extended Regression Mean Minimum	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003		X				
JUNE 2003		X				X
SEPTEMBER 2003				X		
DECEMBER 2003		X		X		
MARCH 2006		X	X	X		
JUNE 2006		X		X	X	X
SEPTEMBER 2006		X	X		X	X
DECEMBER 2006		X	X			
MARCH 2011						
JUNE 2011		X			X	
SEPTEMBER 2011					X	
DECEMBER 2011		X				

Table 5B: Geographical variables that were included or excluded (X) for the mean temperature MLR extended models using 100 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Extended Regression Mean	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003		X	X			X
JUNE 2003		X		X		
SEPTEMBER 2003		X				
DECEMBER 2003		X				
MARCH 2006						
JUNE 2006		X	X	X		
SEPTEMBER 2006				X		
DECEMBER 2006						
MARCH 2011		X				
JUNE 2011		X	X			
SEPTEMBER 2011		X	X	X		
DECEMBER 2011		X			X	X

Table 6B: Geographical variables that were included or excluded (X) for the mean temperature MLR extended models using 100 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Extended Regression Mean Maximum	Independent Variables					
	ALT	RAD	AS P	SLO	DIST ATL	DIST MED
MARCH 2003		X		X		
JUNE 2003		X	X	X		
SEPTEMBER 2003		X	X	X		
DECEMBER 2003		X		X		
MARCH 2006				X		
JUNE 2006		X	X	X		
SEPTEMBER 2006				X		
DECEMBER 2006		X	X	X	X	
MARCH 2011				X	X	X
JUNE 2011		X				
SEPTEMBER 2011		X			X	
DECEMBER 2011		X			X	

Table 7B: Geographical variables that were included or excluded (X) for the mean minimum temperature MLR+I official models using 75 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Official MLR+I Mean Minimum	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003		X		X		
JUNE 2003			X	X		
SEPTEMBER 2003				X	X	
DECEMBER 2003				X		
MARCH 2006		X	X	X		
JUNE 2006			X	X	X	X
SEPTEMBER 2006		X	X	X		X
DECEMBER 2006		X				
MARCH 2011				X		
JUNE 2011			X		X	
SEPTEMBER 2011			X		X	
DECEMBER 2011		X				

Table 8B: Geographical variables that were included or excluded (X) for the mean temperature MLR+I official models using 75 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Official MLR+I Mean	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003		X		X		X
JUNE 2003		X	X			X
SEPTEMBER 2003				X	X	X
DECEMBER 2003				X		
MARCH 2006				X		
JUNE 2006		X	X	X		
SEPTEMBER 2006				X	X	
DECEMBER 2006		X				
MARCH 2011		X	X	X		
JUNE 2011		X	X		X	X
SEPTEMBER 2011		X	X	X	X	
DECEMBER 2011		X				

Table 9B: Geographical variables that were included or excluded (X) for the mean maximum temperature MLR+I official models using 75 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Official MLR+I Mean Maximum	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003		X	X	X	X	
JUNE 2003			X			
SEPTEMBER 2003			X	X	X	
DECEMBER 2003		X	X	X		
MARCH 2006				X		
JUNE 2006			X	X		
SEPTEMBER 2006				X	X	
DECEMBER 2006		X	X	X	X	
MARCH 2011				X		X
JUNE 2011		X			X	
SEPTEMBER 2011		X			X	
DECEMBER 2011		X				

Table 10B: Geographical variables that were included or excluded (X) for the mean minimum temperature MLR+I extended models using 75 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Extended MLR+I Mean Minimum	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003		X				
JUNE 2003		X				
SEPTEMBER 2003						
DECEMBER 2003				X		
MARCH 2006		X	X			
JUNE 2006		X			X	X
SEPTEMBER 2006		X	X		X	
DECEMBER 2006		X				
MARCH 2011		X				
JUNE 2011		X			X	
SEPTEMBER 2011			X			
DECEMBER 2011		X				

Table 11B: Geographical variables that were included or excluded (X) for the mean temperature MLR extended models using 75 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Extended MLR+I Mean	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003		X	X			X
JUNE 2003		X	X	X		
SEPTEMBER 2003				X	X	X
DECEMBER 2003		X				
MARCH 2006				X		
JUNE 2006		X	X	X		
SEPTEMBER 2006				X		
DECEMBER 2006		X				
MARCH 2011		X				
JUNE 2011		X	X			
SEPTEMBER 2011		X	X	X	X	X
DECEMBER 2011		X				

Table 12B: Geographical variables that were included or excluded (X) for the mean maximum temperature MLR extended models using 75 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Extended MLR+I Mean Maximum	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003		X		X		
JUNE 2003			X			
SEPTEMBER 2003			X	X		
DECEMBER 2003		X		X		
MARCH 2006				X		
JUNE 2006			X	X		
SEPTEMBER 2006				X		
DECEMBER 2006		X	X	X	X	
MARCH 2011				X	X	X
JUNE 2011		X				
SEPTEMBER 2011		X				
DECEMBER 2011		X				

Table 13B: Geographical variables that were included or excluded (X) for precipitation MLR official models using 100 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Official MLR Precipitation Models	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003						
JUNE 2003		X				
SEPTEMBER 2003			X			
DECEMBER 2003		X	X			
MARCH 2006						
JUNE 2006						
SEPTEMBER 2006		X				
DECEMBER 2006		X				X
MARCH 2011	X					
JUNE 2011			X			
SEPTEMBER 2011	X		X		X	
DECEMBER 2011			X		X	

Table 14B: Geographical variables that were included or excluded (X) for precipitation MLR extended models using 100 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Extended MLR precipitation Models	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003						
JUNE 2003		X		X		
SEPTEMBER 2003			X			
DECEMBER 2003		X				
MARCH 2006						
JUNE 2006						
SEPTEMBER 2006		X		X		
DECEMBER 2006		X	X			X
MARCH 2011		X				
JUNE 2011			X			
SEPTEMBER 2011				X		
DECEMBER 2011			X	X		

Table 15B: Geographical variables that were included or excluded (X) for precipitation MLR+I official models using 75 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Official MLR+I precipitation Models	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003						
JUNE 2003		X				
SEPTEMBER 2003	X	X				
DECEMBER 2003		X	X			
MARCH 2006						
JUNE 2006		X		X		
SEPTEMBER 2006		X		X		
DECEMBER 2006						X
MARCH 2011	X	X	X			
JUNE 2011			X			
SEPTEMBER 2011	X		X	X	X	
DECEMBER 2011				X	X	

Table 16B: Geographical variables that were included or excluded (X) for precipitation MLR+I extended models using 75 % of AEMET data (ALT = Altitude; RAD = Solar Radiation; ASP = Aspect; SLO = Slope; DIST ATL = Distance to the Atlantic Ocean; DIST MED = Distance to the Mediterranean Sea).

Extended MLR+I precipitationModels	Independent Variables					
	ALT	RAD	ASP	SLO	DIST ATL	DIST MED
MARCH 2003						
JUNE 2003		X				
SEPTEMBER 2003	X	X				
DECEMBER 2003		X				
MARCH 2006						
JUNE 2006		X		X		
SEPTEMBER 2006		X		X		
DECEMBER 2006						X
MARCH 2011			X			
JUNE 2011						
SEPTEMBER 2011				X		
DECEMBER 2011			X	X		

Appendix C: Regression maps

Precipitation

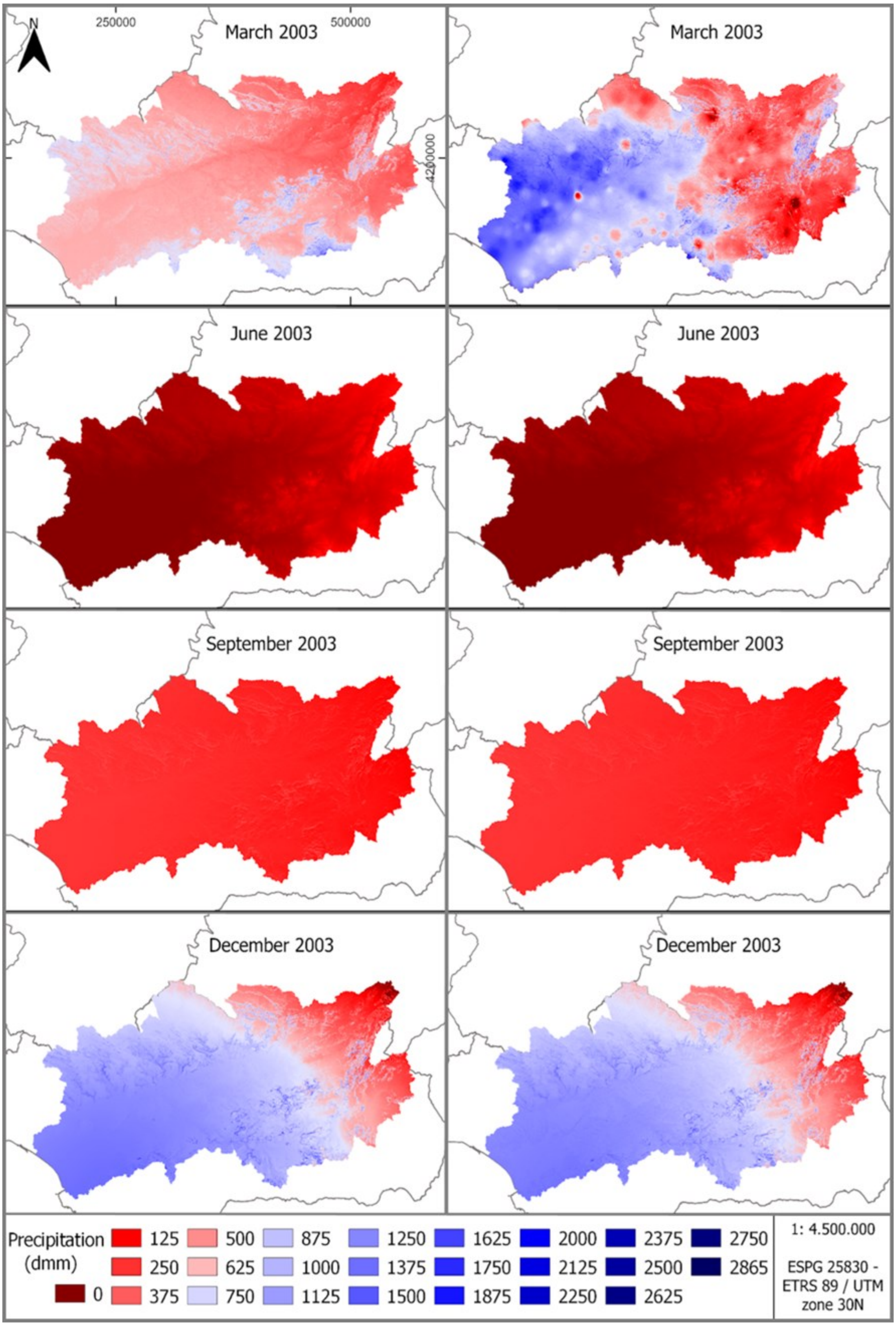


Figure 1C: 2003 MLR precipitation maps using the official dataset (left) and the extended dataset (right).

Precipitation

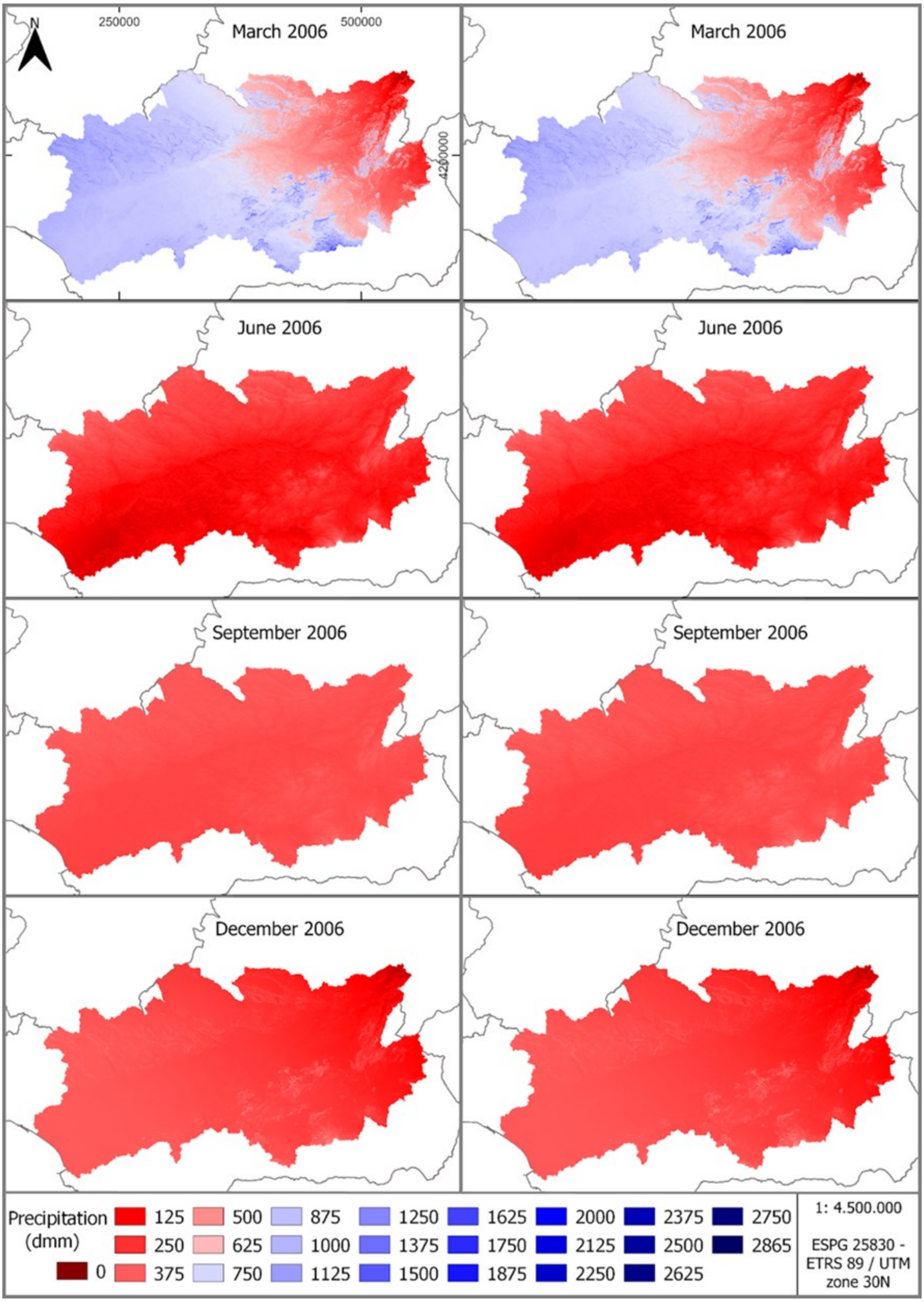


Figure 2C: 2006 MLR precipitation maps using the official dataset (left) and the extended dataset (right).

Precipitation

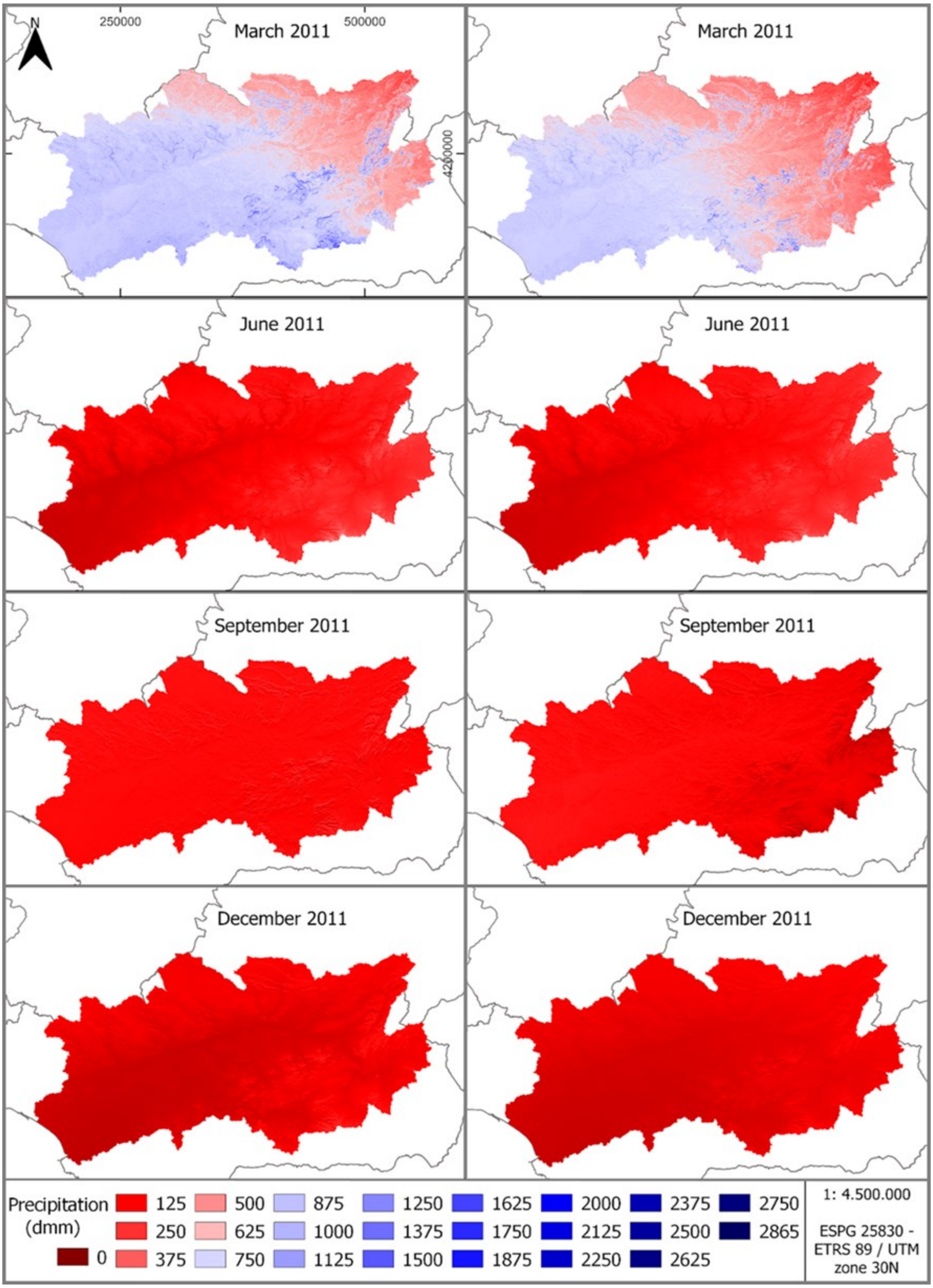


Figure 3C: 2011 MLR precipitation maps using the official dataset (left) and the extended dataset (right).

Mean Minimum Temperature

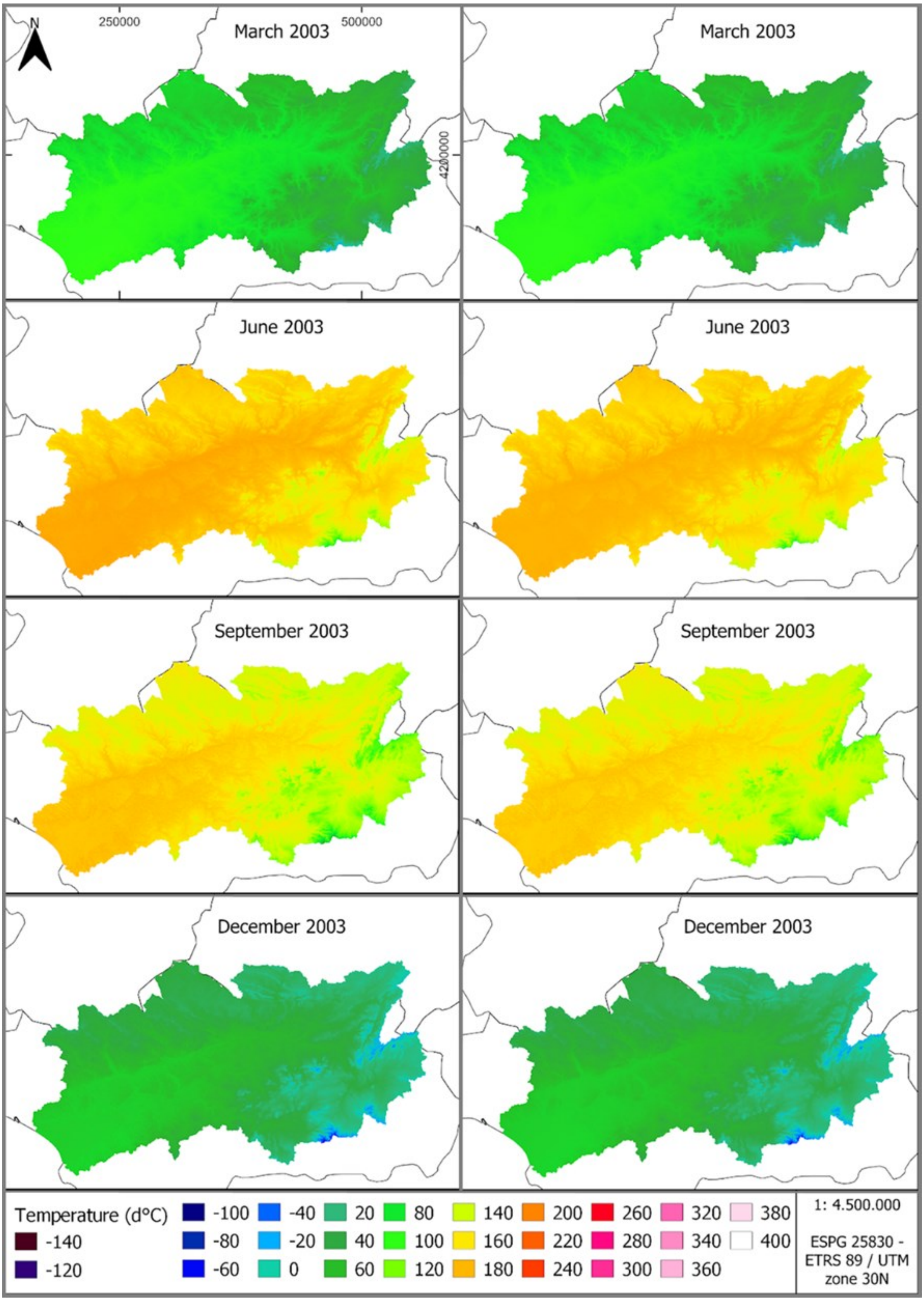


Figure 4C: 2003 MLR mean minimum temperature maps using the official dataset (left) and the extended dataset (right).

Mean Minimum Temperature

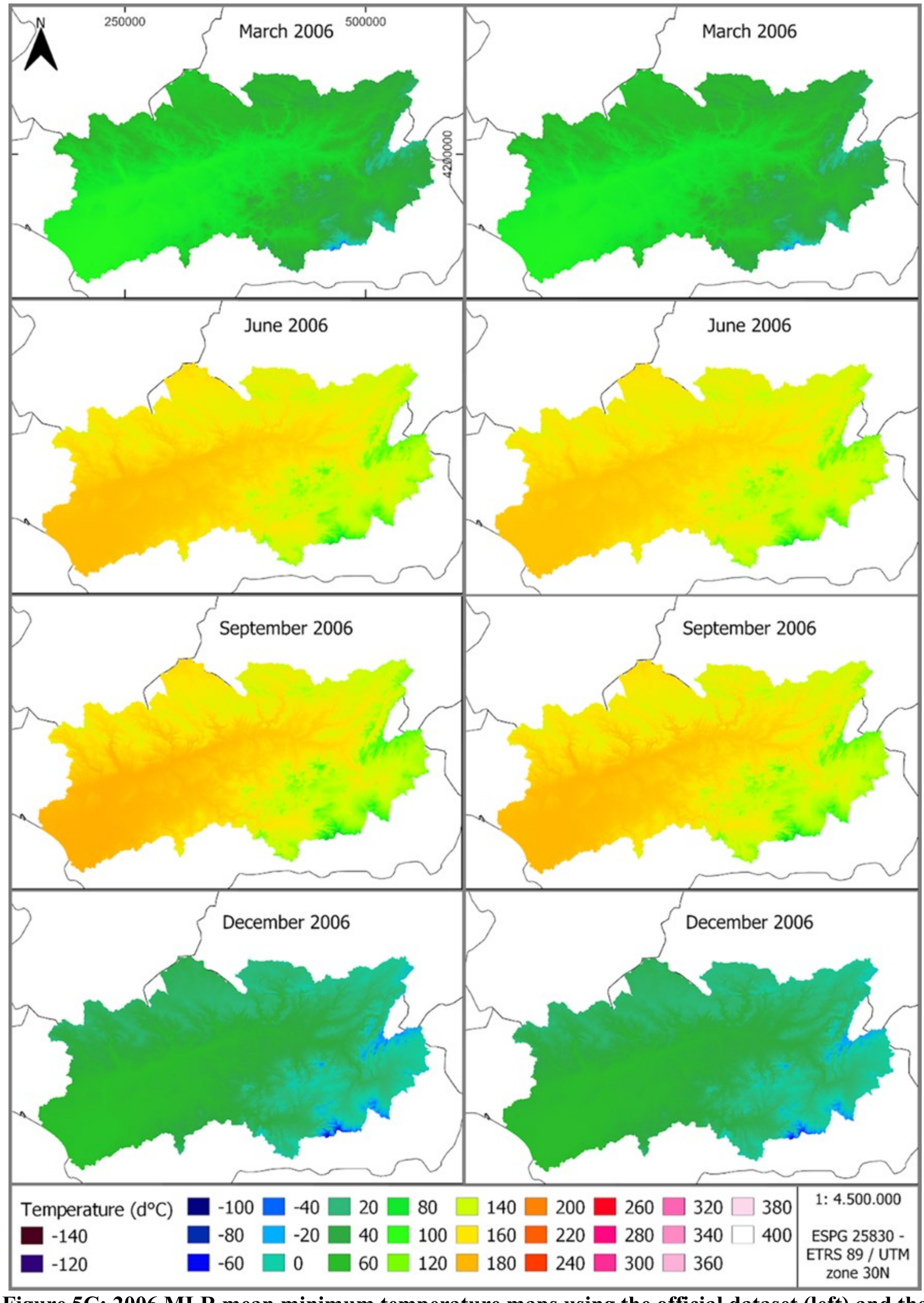


Figure 5C: 2006 MLR mean minimum temperature maps using the official dataset (left) and the extended dataset (right).

Mean Minimum Temperature

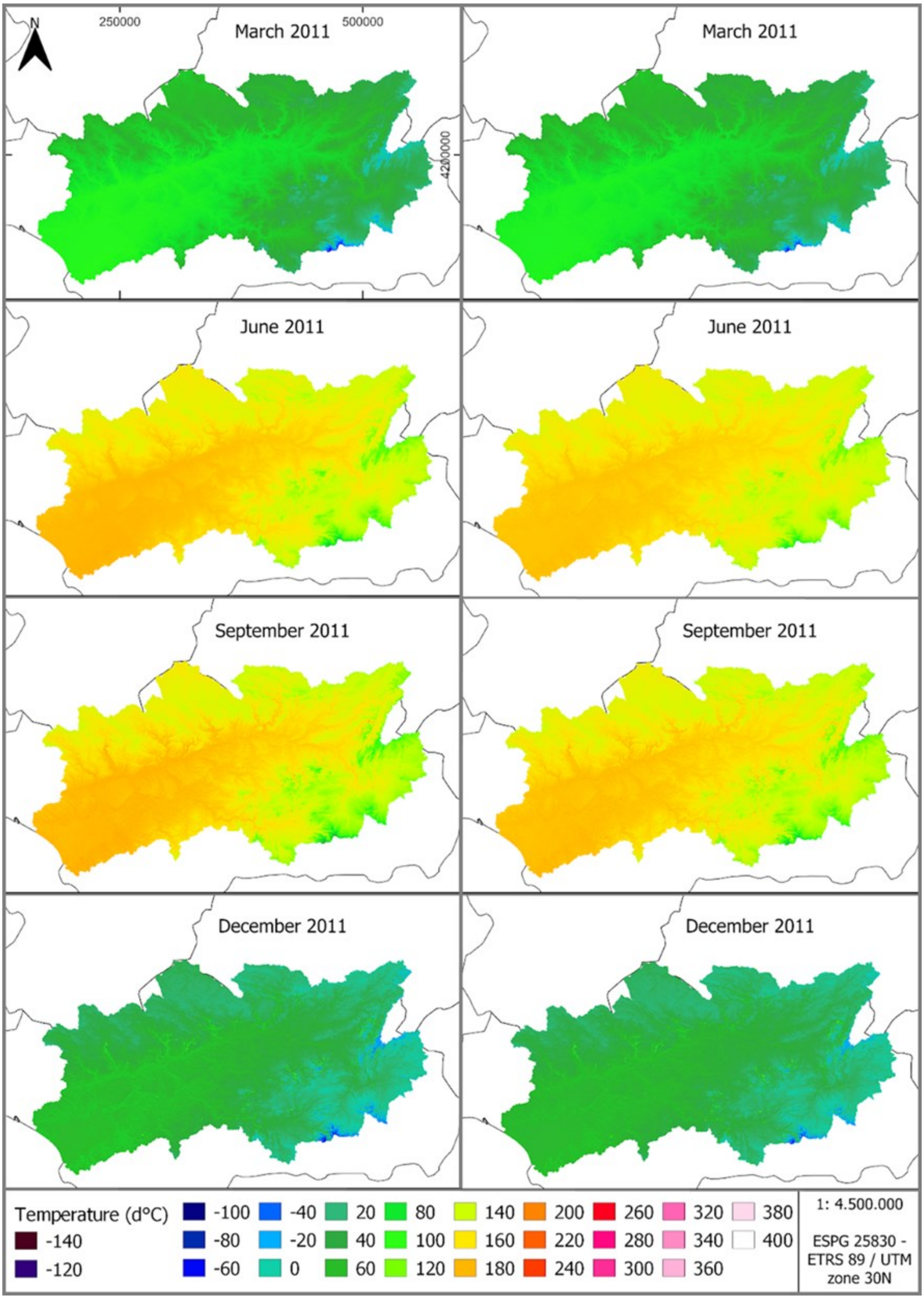


Figure 6C: 2011 MLR mean minimum temperature maps using the official dataset (left) and the extended dataset (right).

Mean Temperature

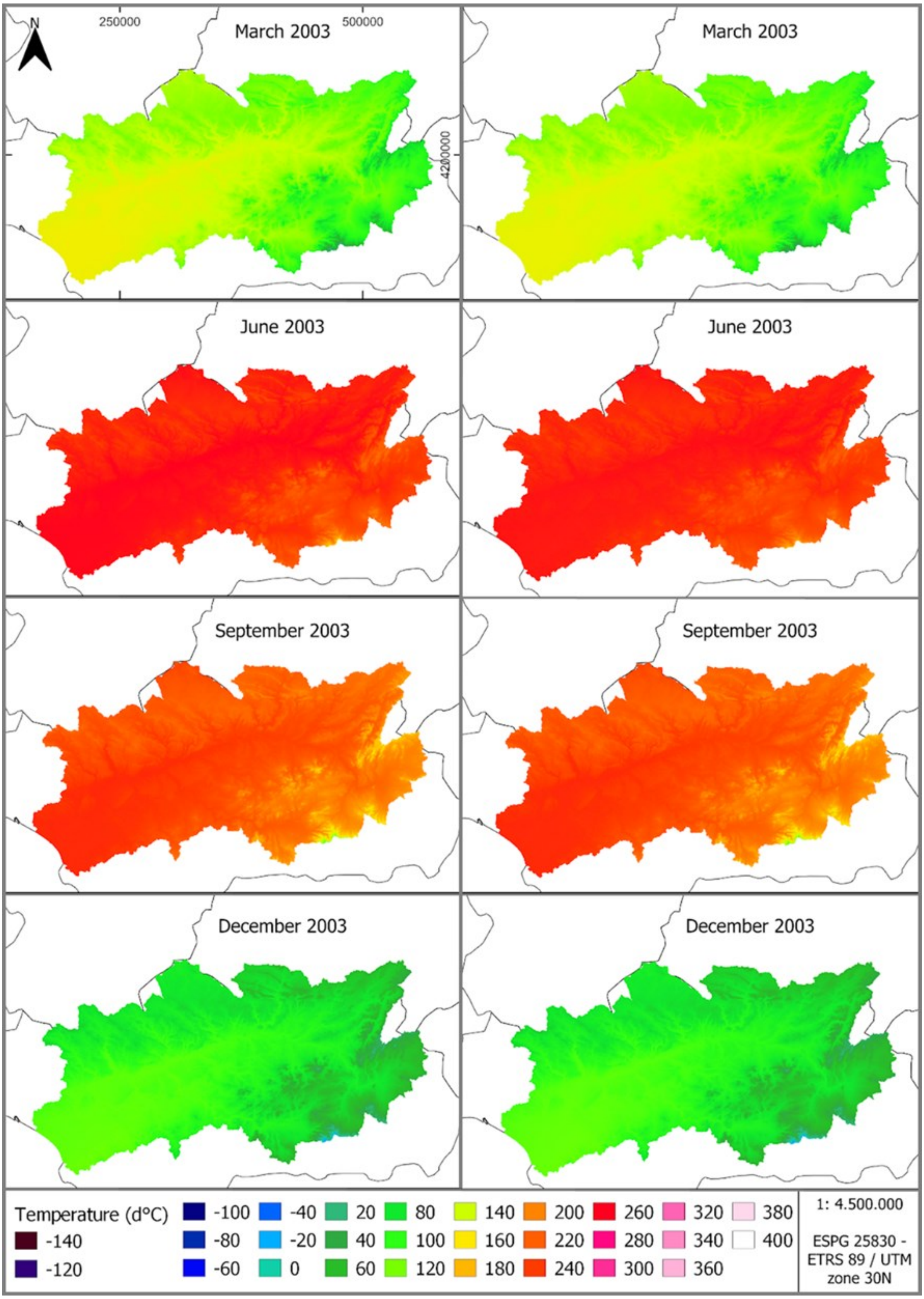


Figure 7C: 2003 MLR mean temperature maps using the official dataset (left) and the extended dataset (right).

Mean Temperature

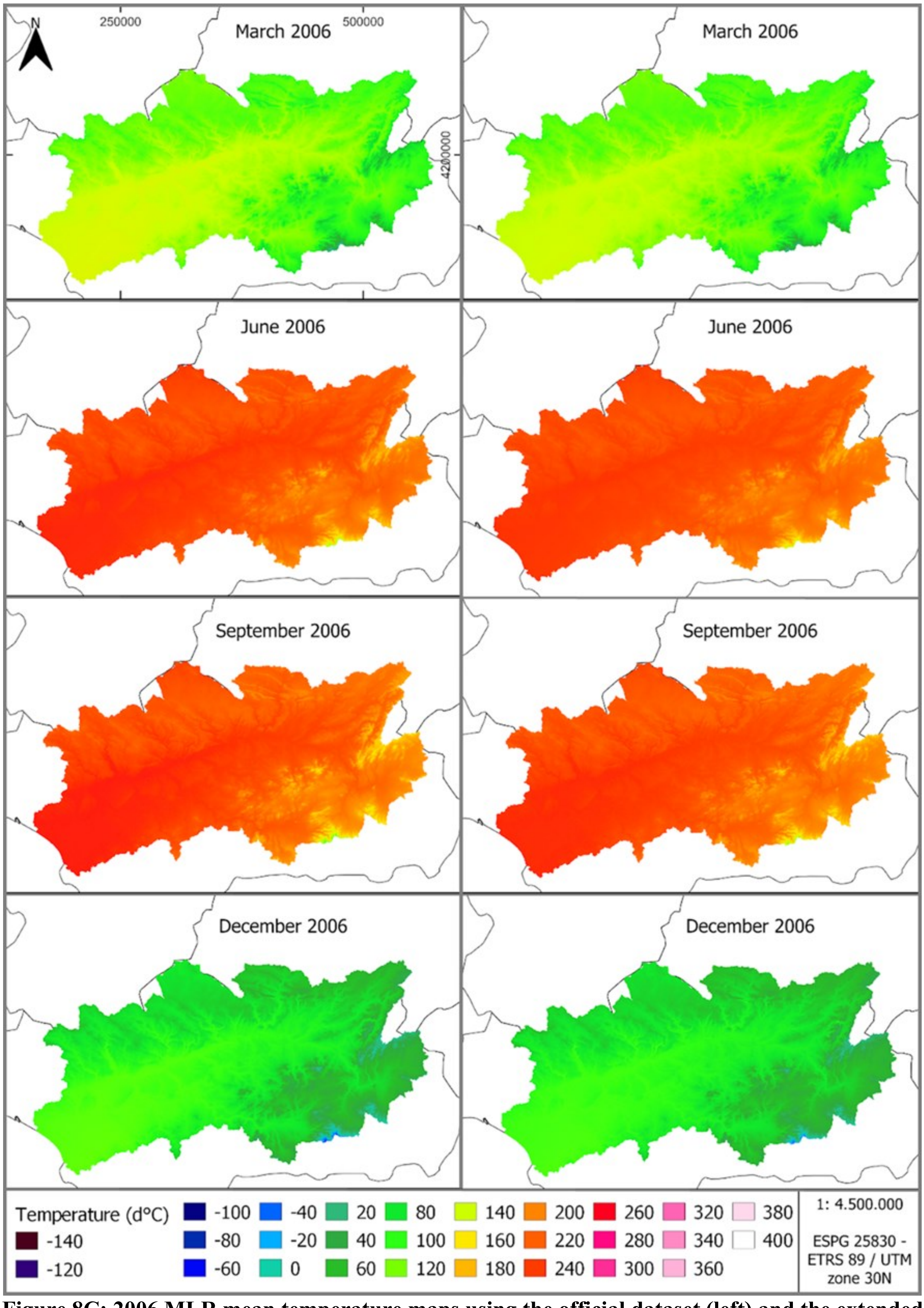


Figure 8C: 2006 MLR mean temperature maps using the official dataset (left) and the extended dataset (right).

Mean Temperature

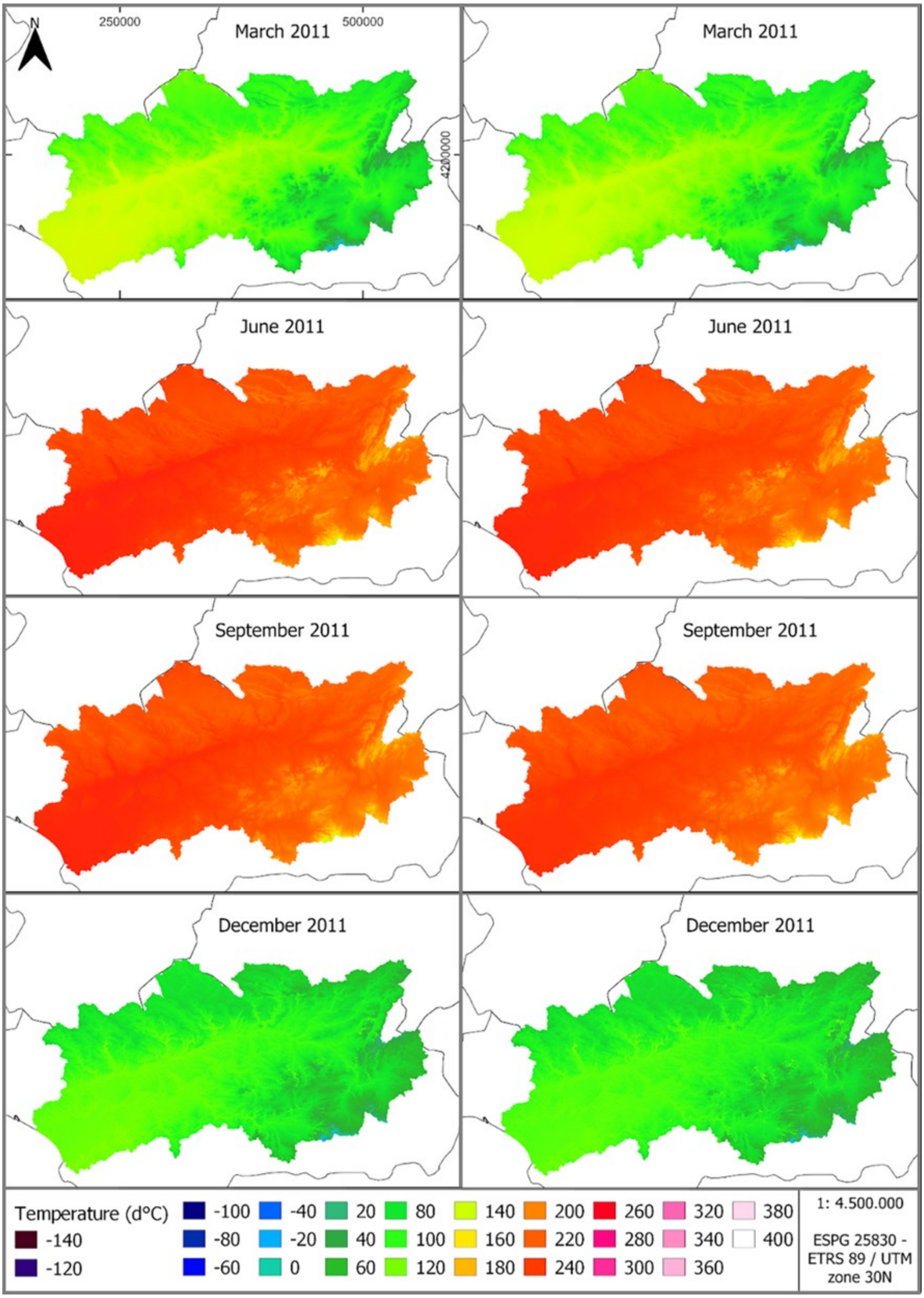


Figure 9C: 2011 MLR mean temperature maps using the official dataset (left) and the extended dataset (right).

Mean Maximum Temperature

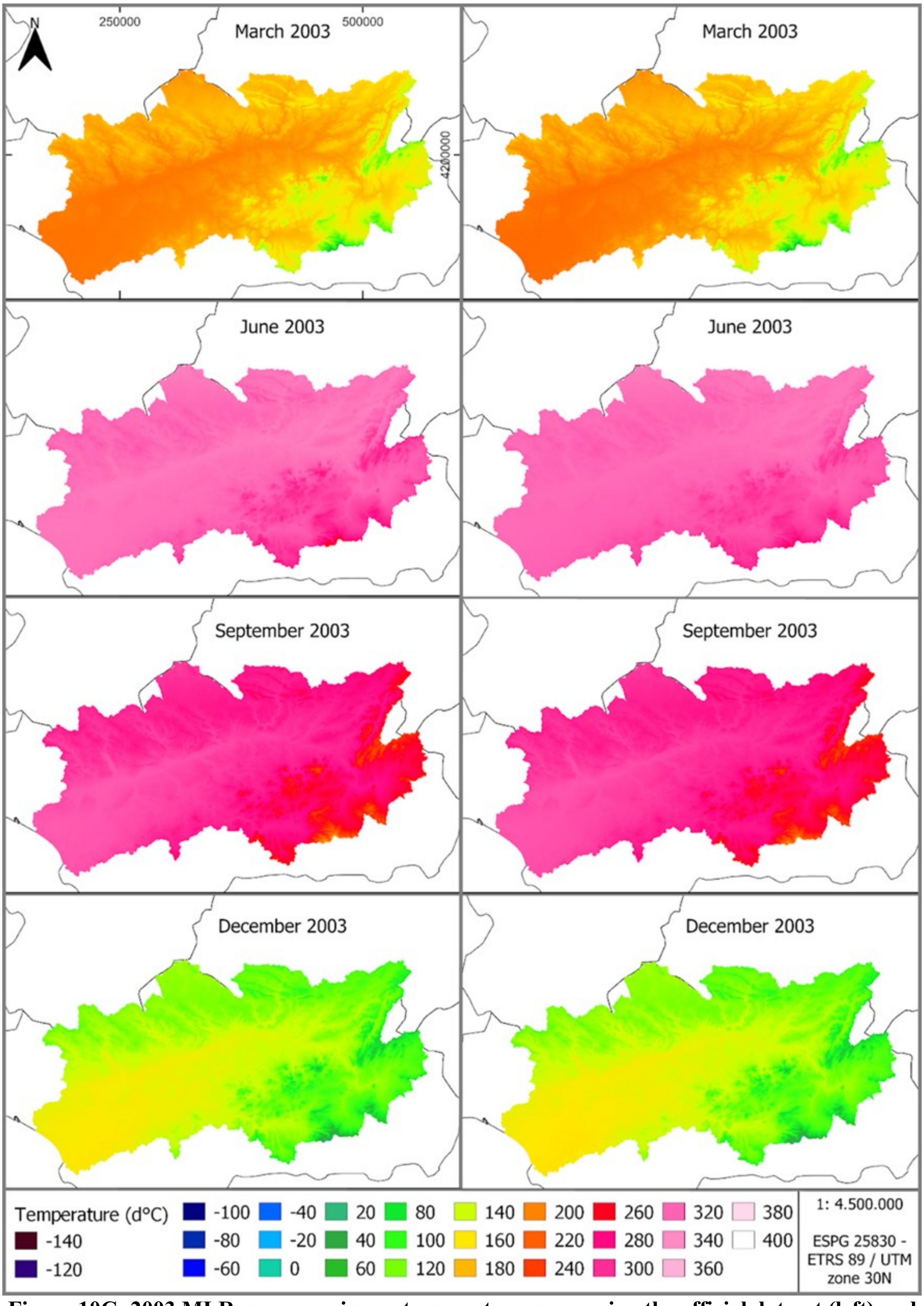


Figure 10C: 2003 MLR mean maximum temperature maps using the official dataset (left) and the extended dataset (right).

Mean Maximum Temperature

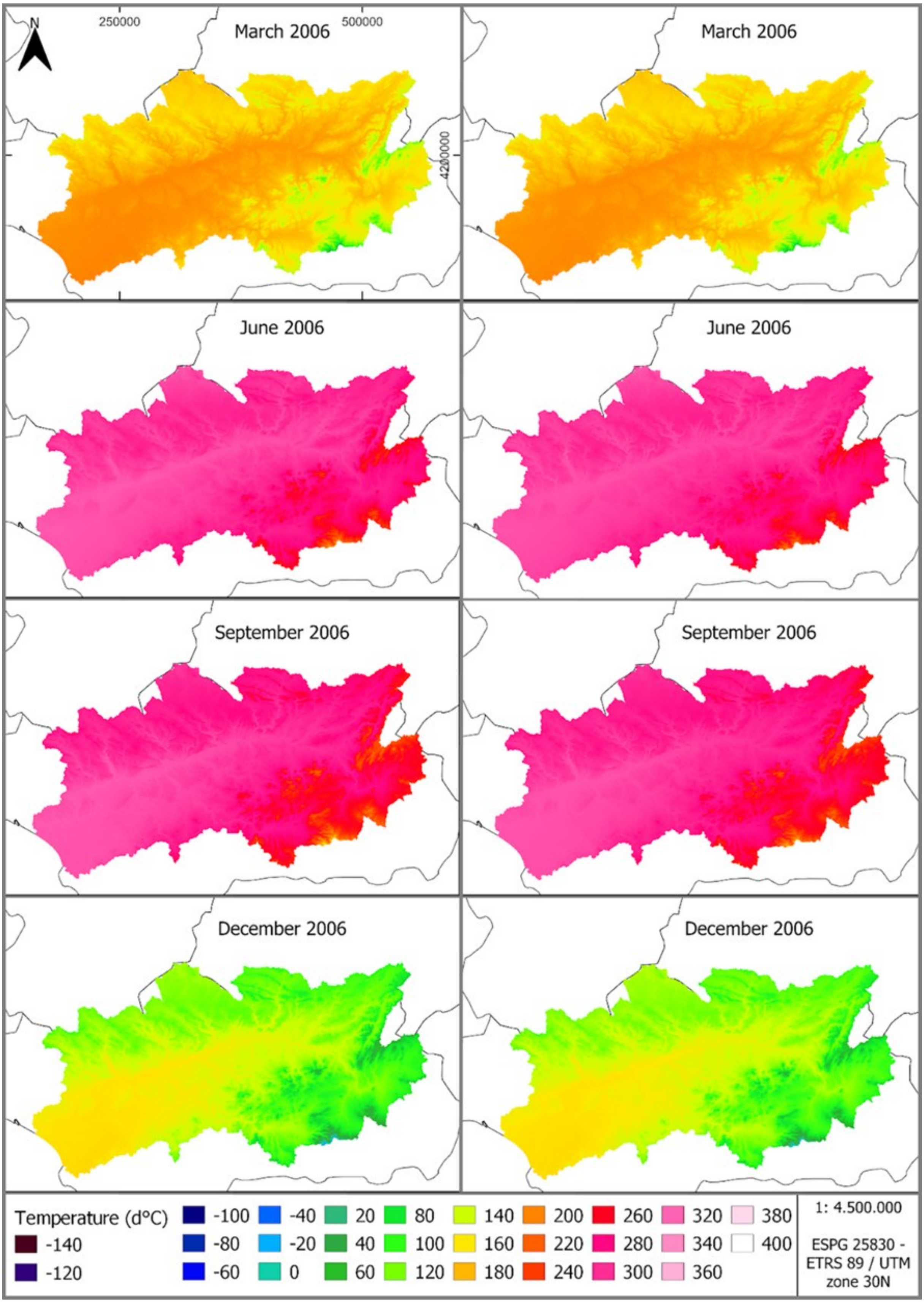


Figure 11C: 2006 MLR mean maximum temperature maps using the official dataset (left) and the extended dataset (right).

Mean Maximum Temperature

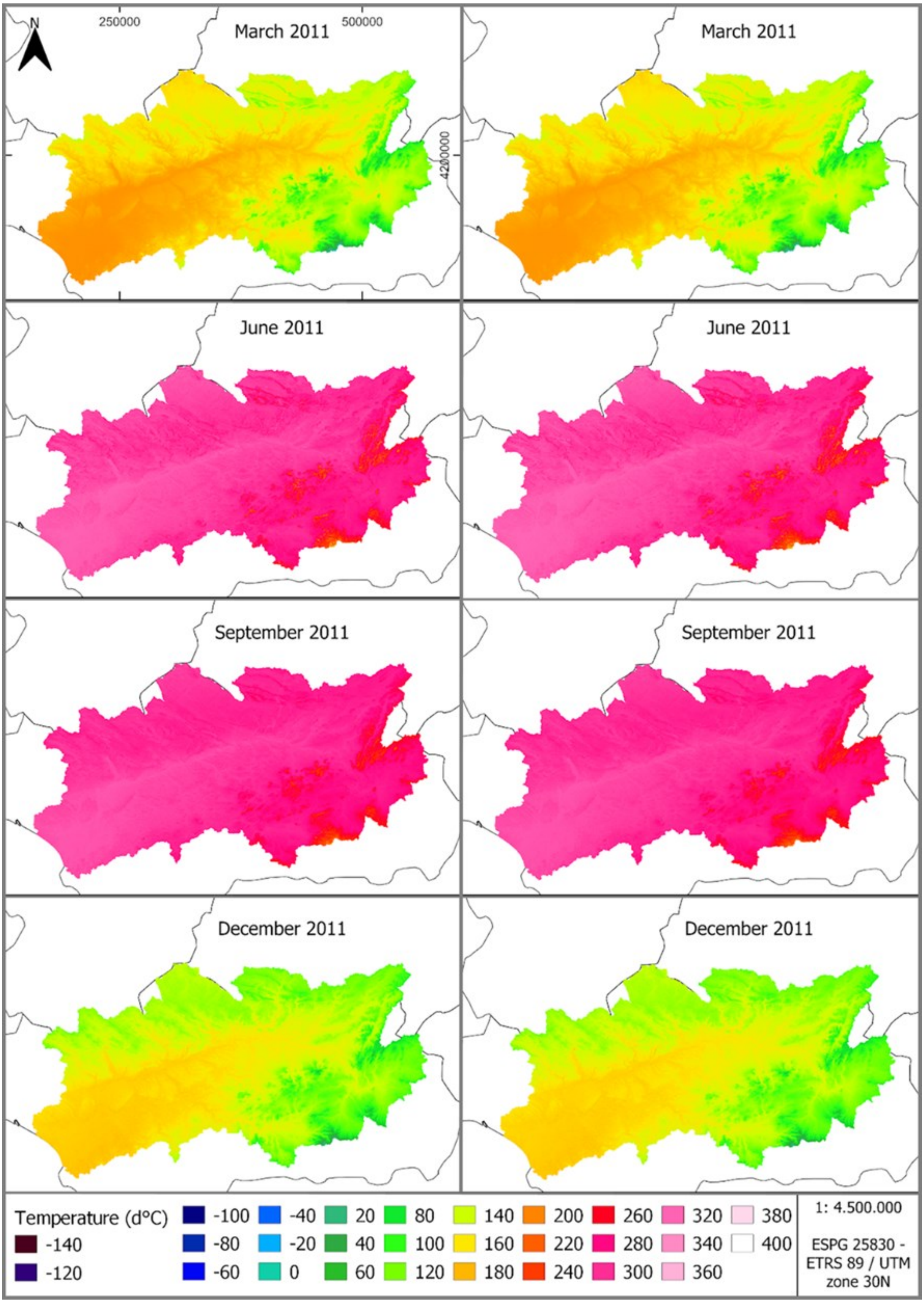


Figure 12C: 2011 MLR mean maximum temperature maps using the official dataset (left) and the extended dataset (right).

Appendix D: Regression and Interpolation maps.

Precipitation

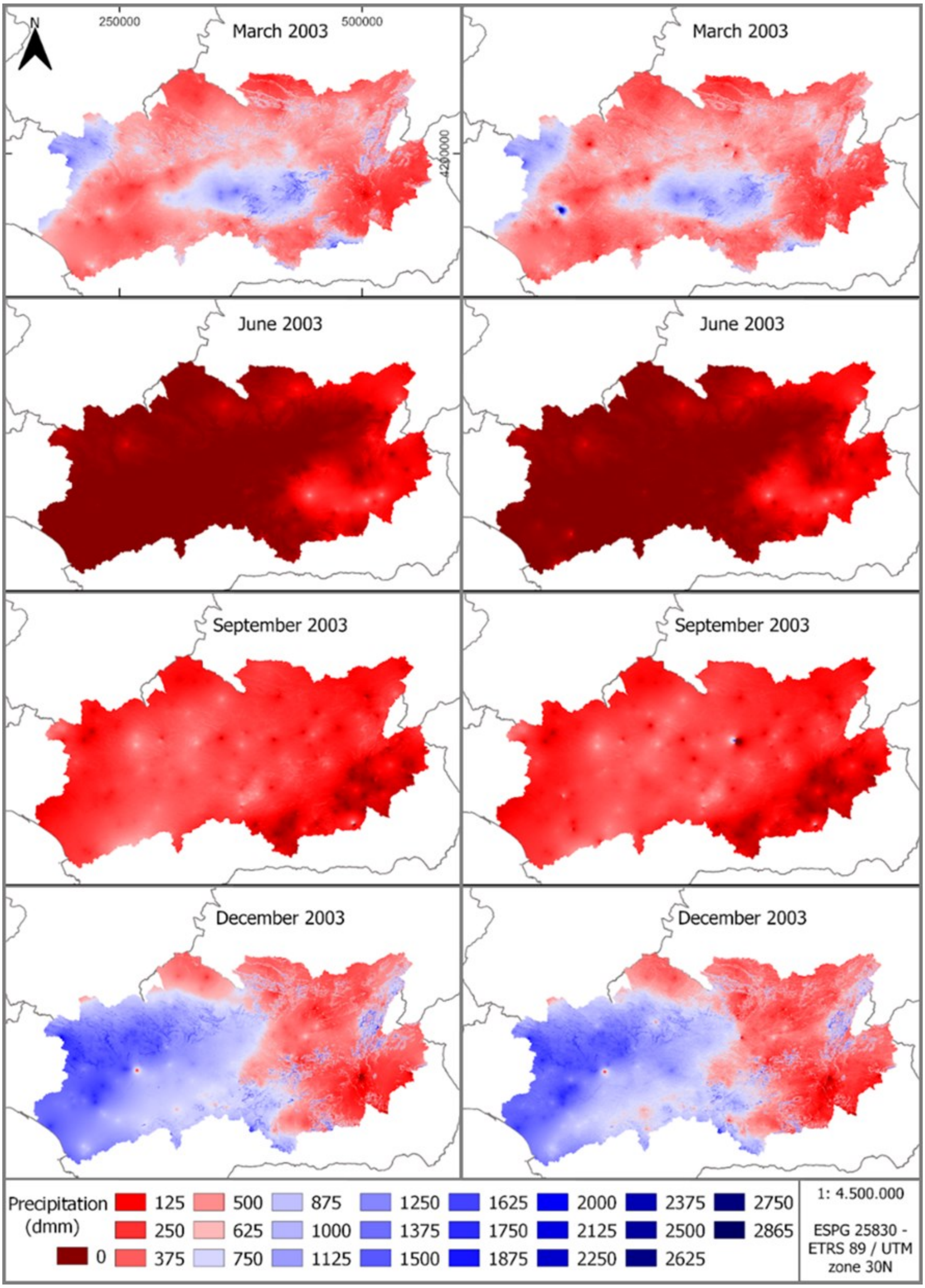


Figure 1D: 2003 MLR+I precipitation maps using the official dataset (left) and the extended dataset (right).

Precipitation

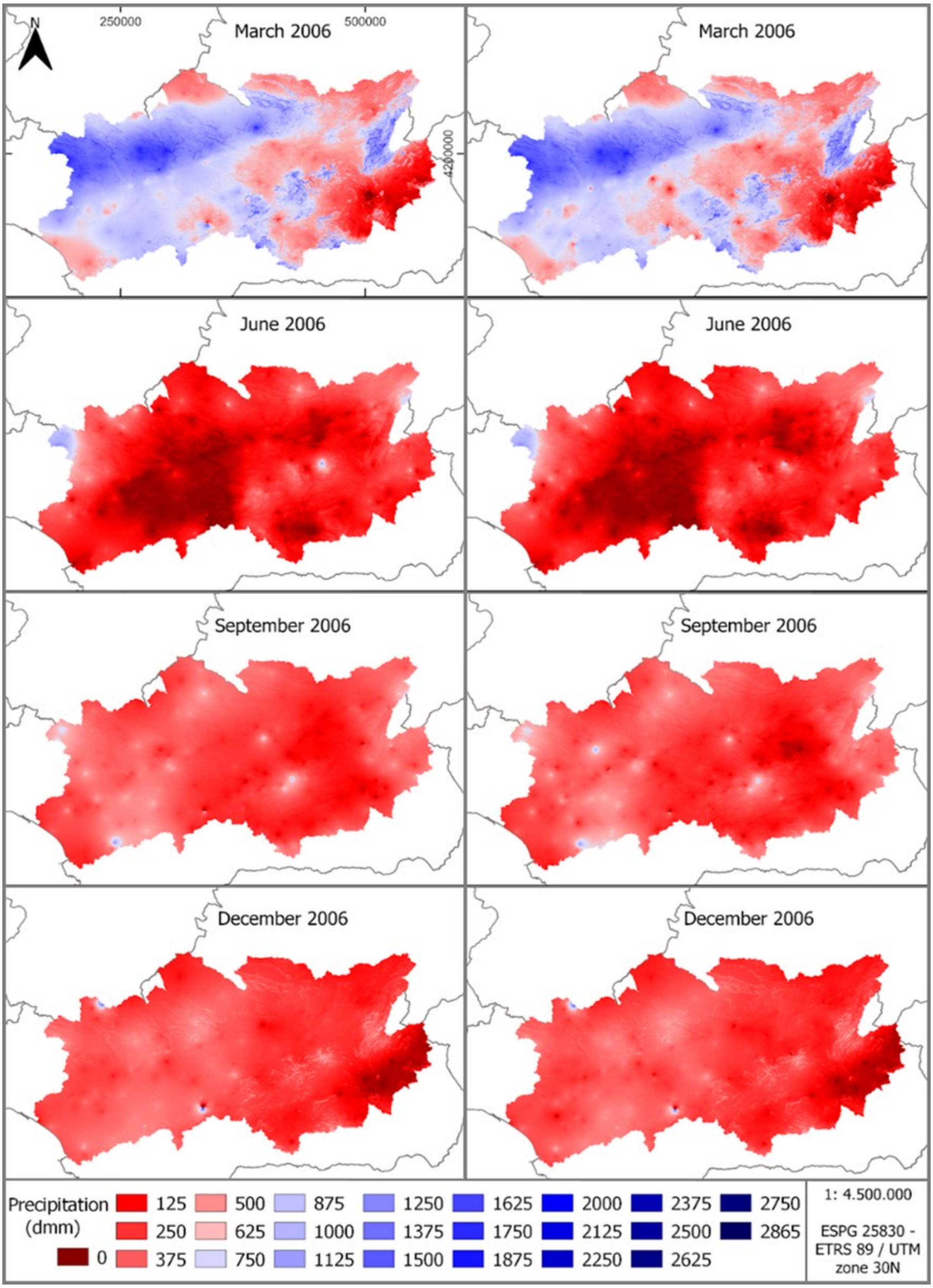


Figure 2D: 2006 MLR+I precipitation maps using the official dataset (left) and the extended dataset (right).

Precipitation

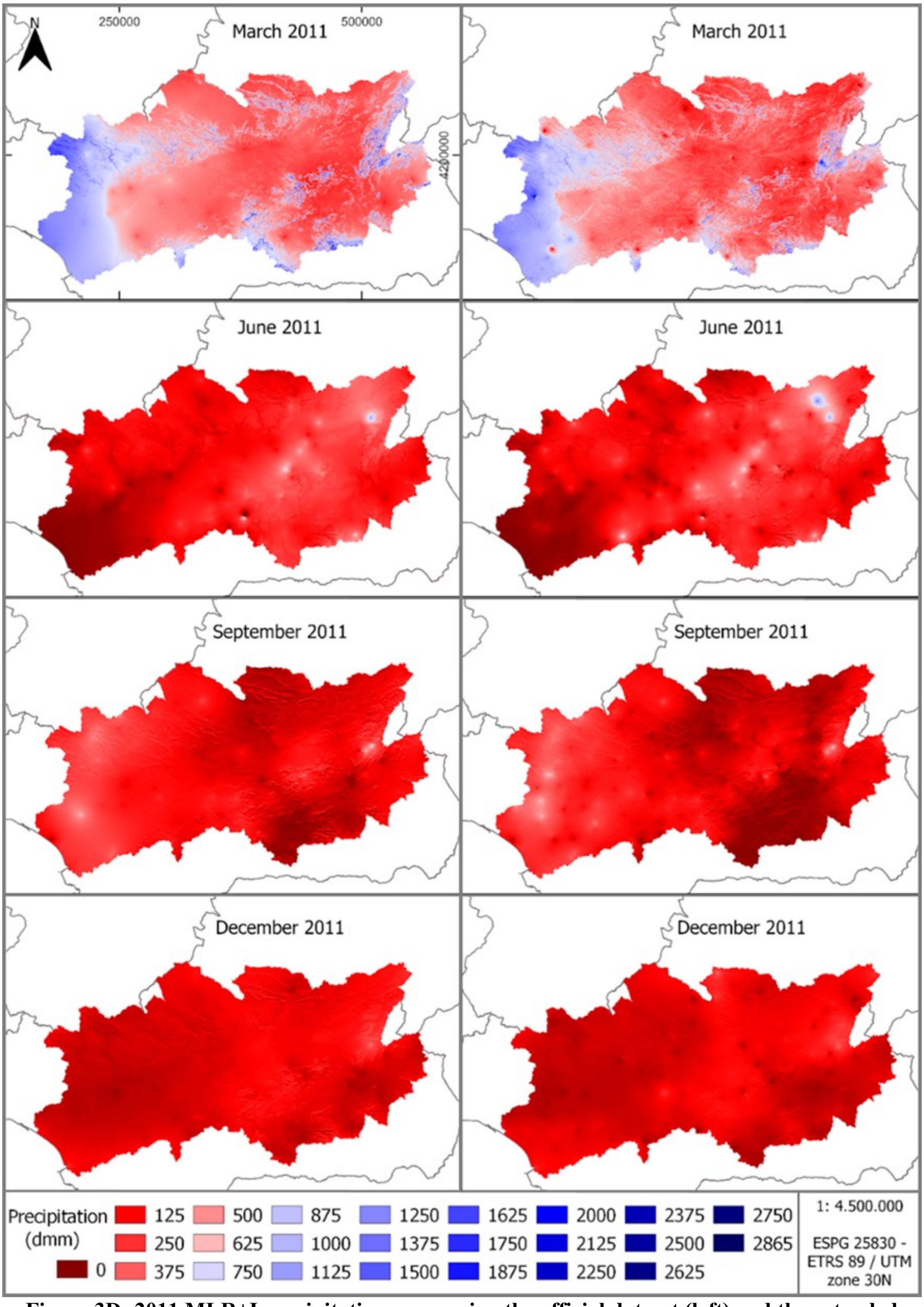


Figure 3D: 2011 MLR+I precipitation maps using the official dataset (left) and the extended dataset (right).

Mean Minimum Temperature

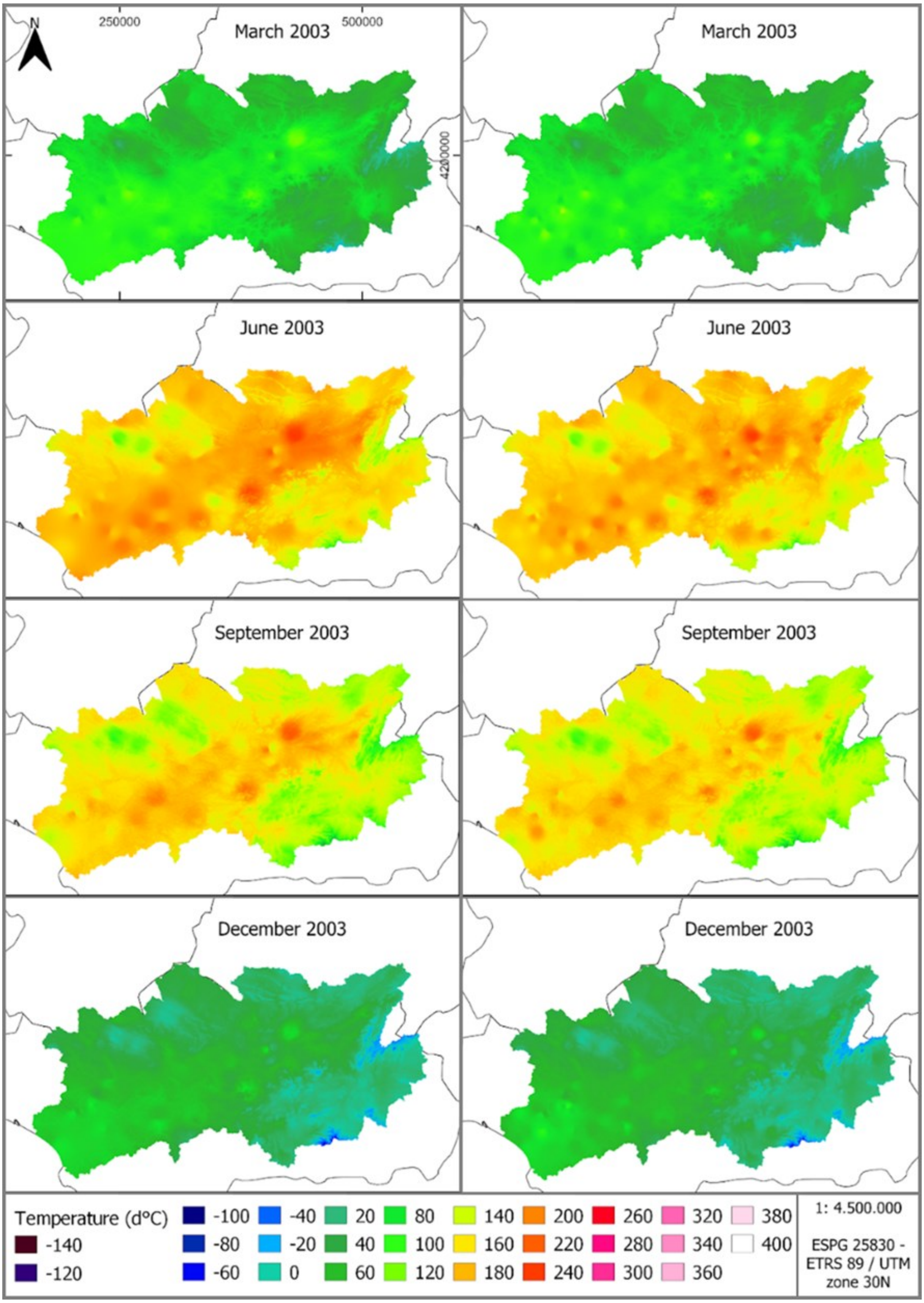


Figure 4D: 2003 MLR+I mean minimum temperature maps using the official dataset (left) and the extended dataset (right).

Mean Minimum Temperature

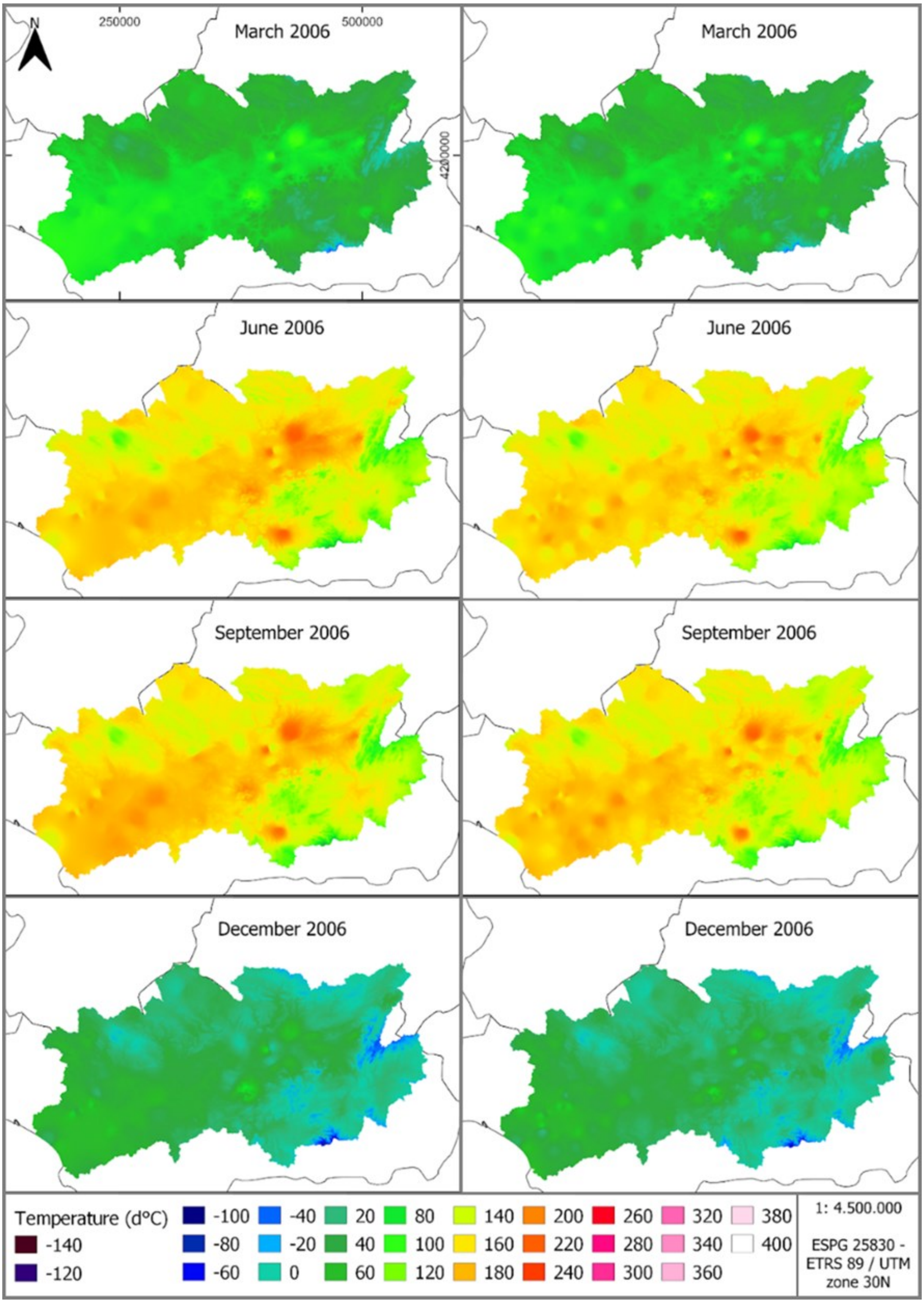


Figure 5D: 2006 MLR+I mean minimum temperature maps using the official dataset (left) and the extended dataset (right).

Mean Minimum Temperature

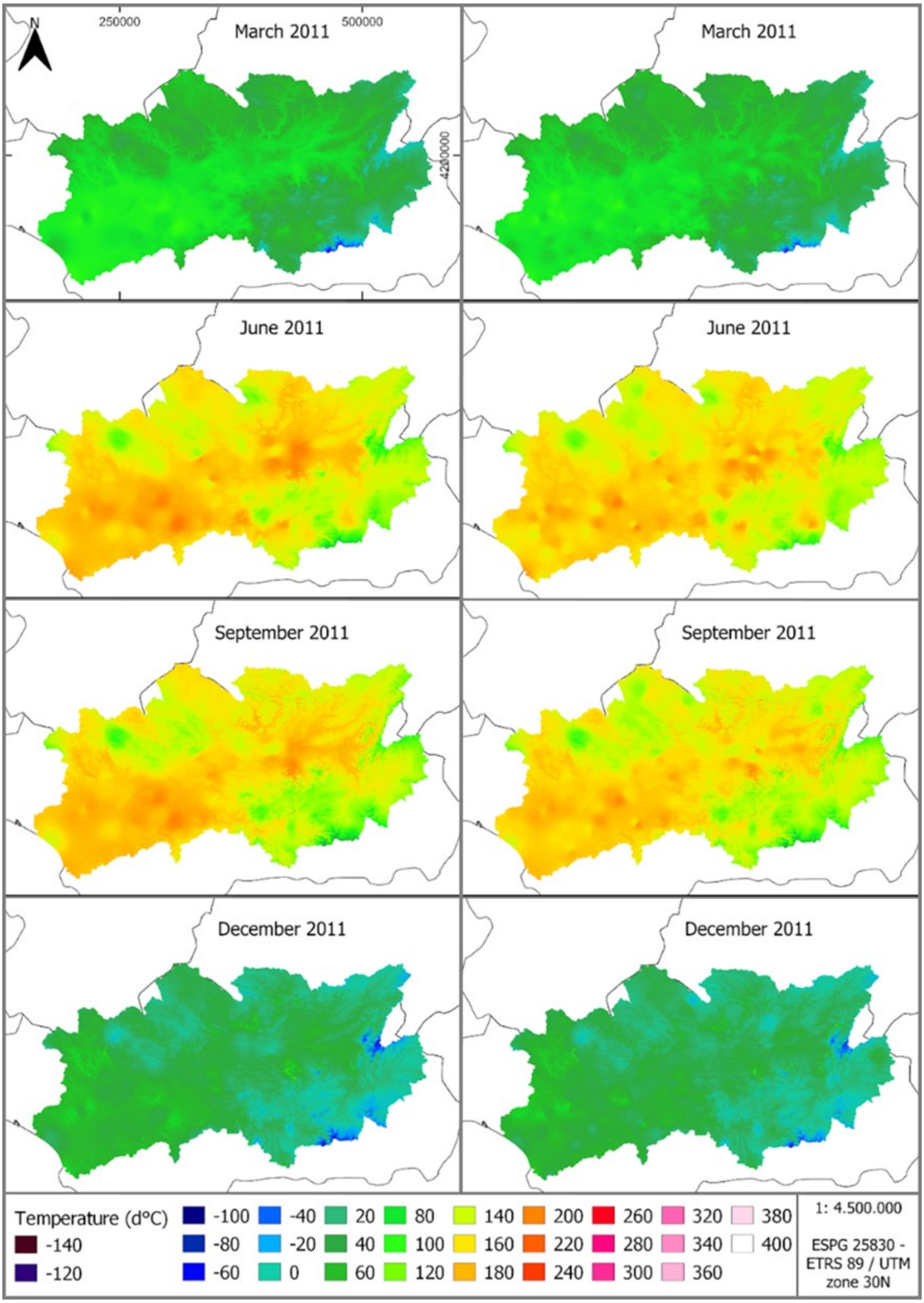


Figure 6D: 2011 MLR+I mean minimum temperature maps using the official dataset (left) and the extended dataset (right).

Mean Temperature

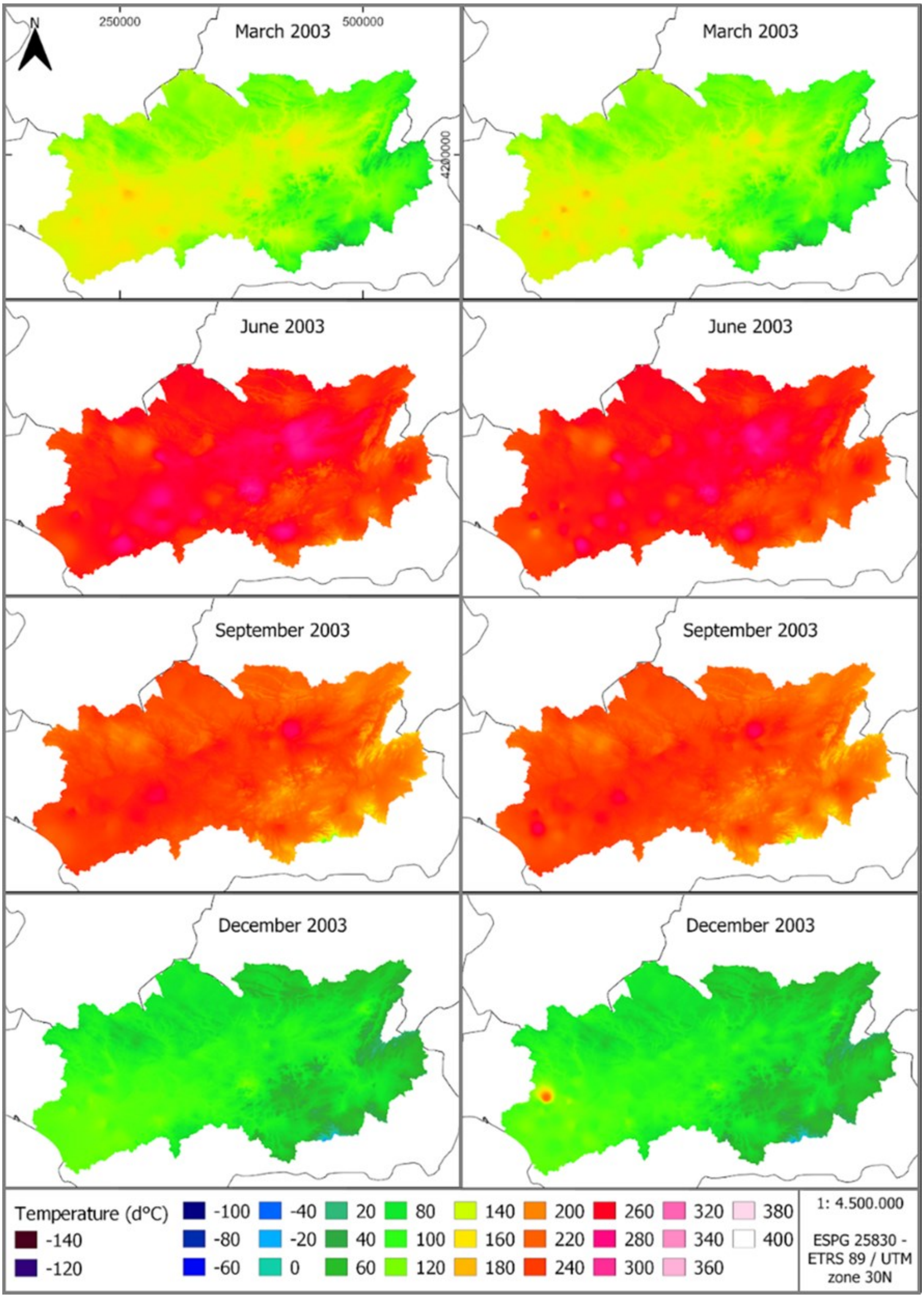


Figure 7D: 2003 MLR+I mean temperature maps using the (left) and the extended dataset (right).

Mean Temperature

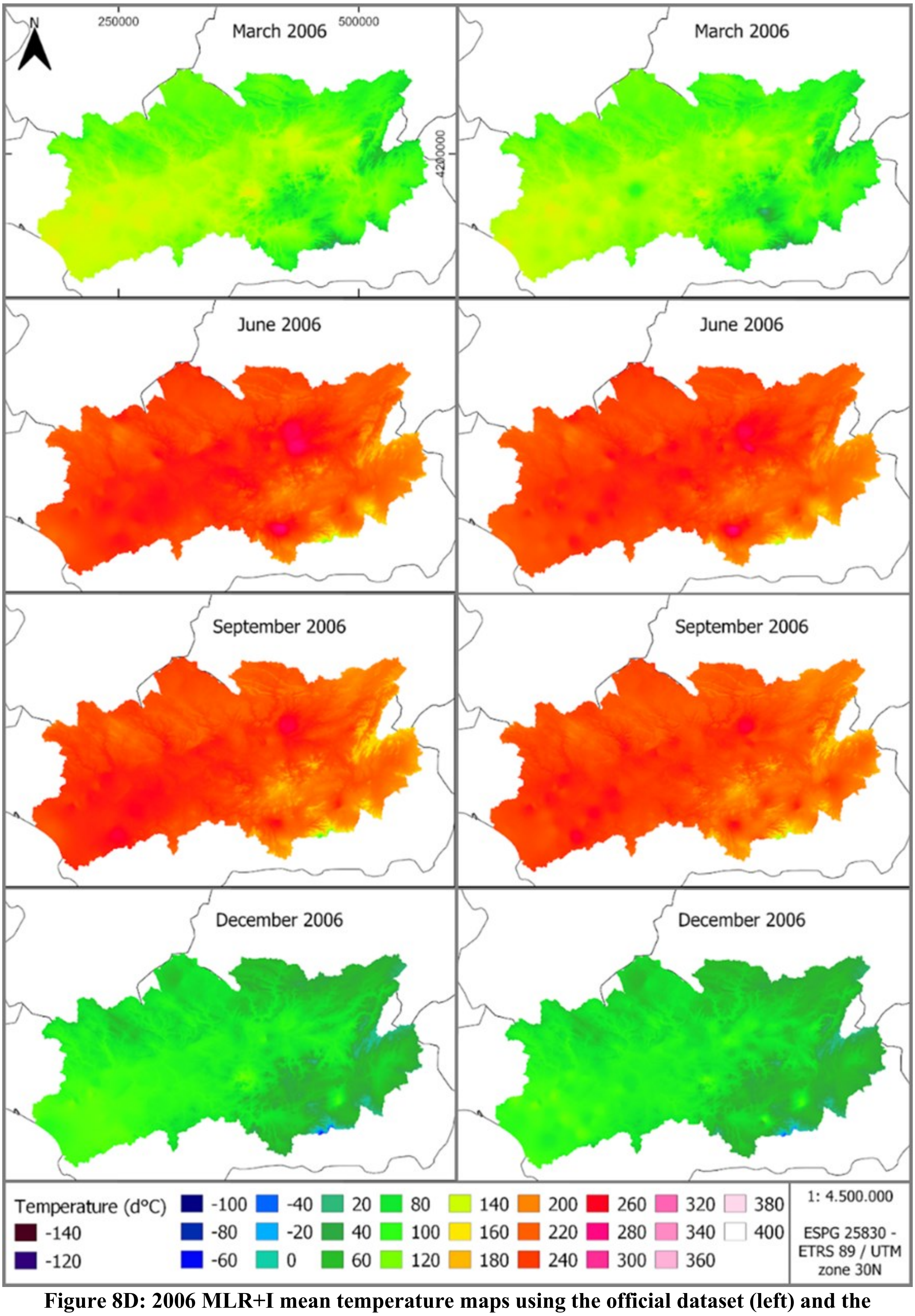


Figure 8D: 2006 MLR+I mean temperature maps using the official dataset (left) and the extended dataset (right).

Mean Temperature

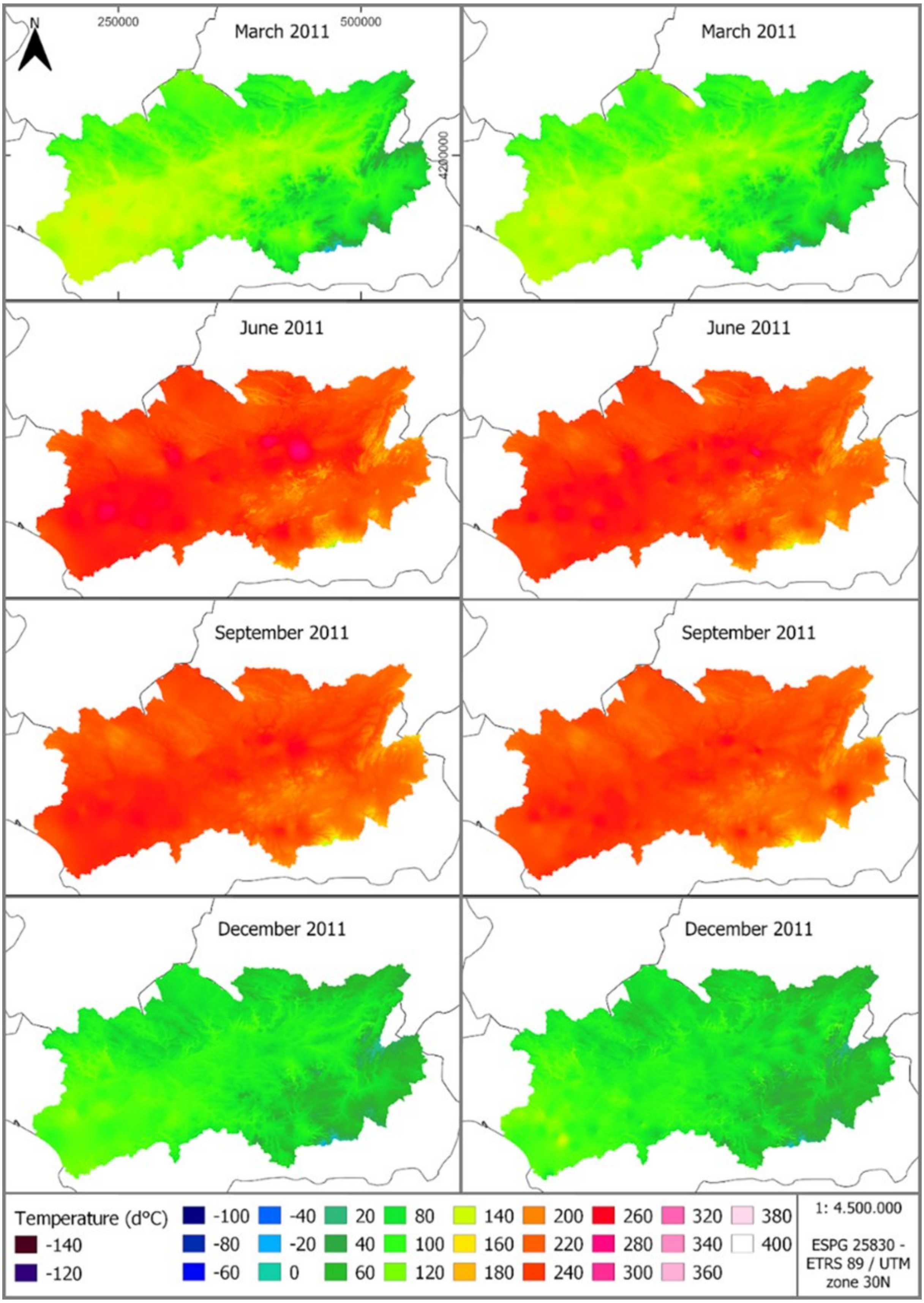


Figure 9D: 2011 MLR+I mean temperature maps using the official dataset (left) and the extended dataset (right).

Mean Maximum Temperature

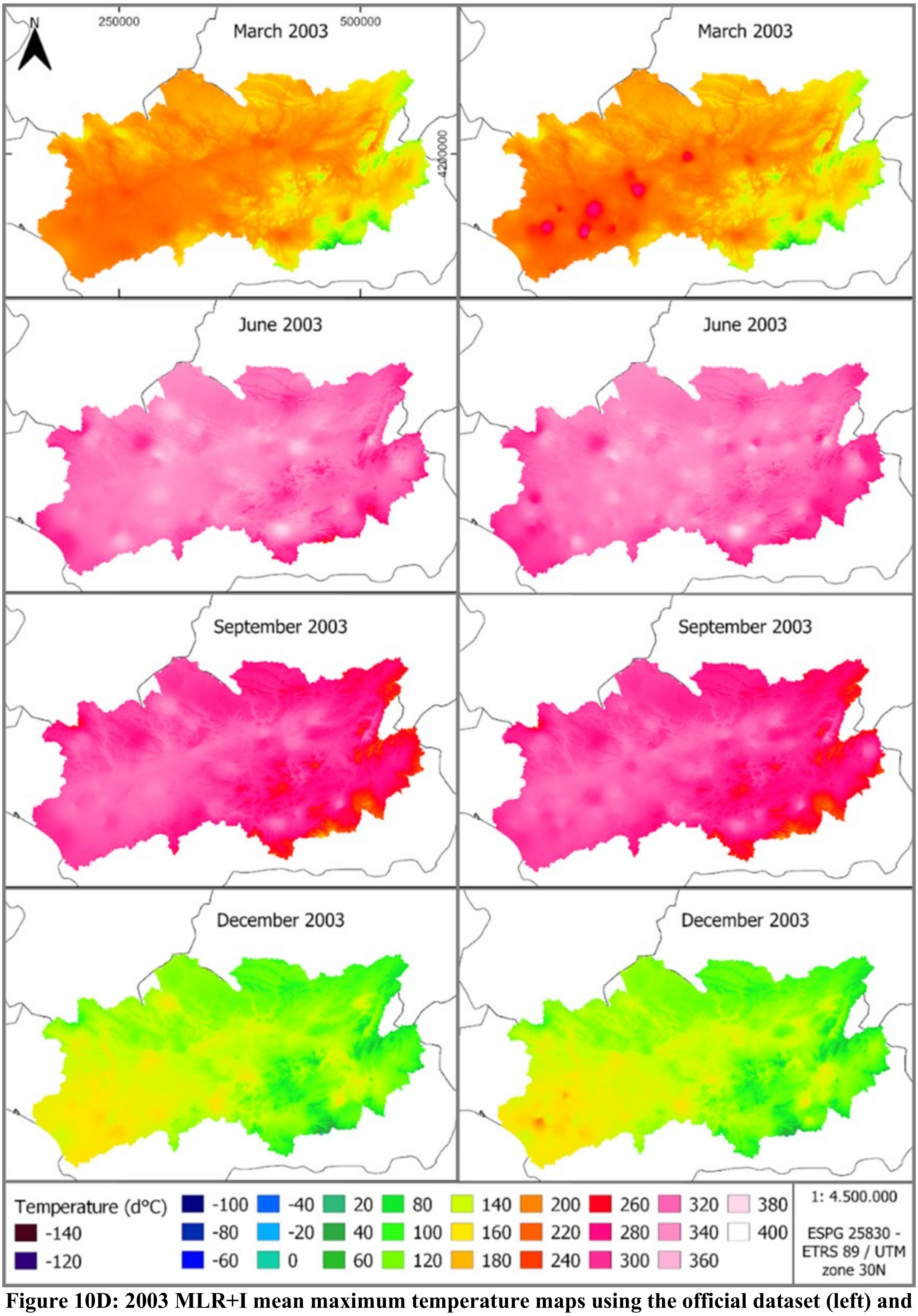


Figure 10D: 2003 MLR+I mean maximum temperature maps using the official dataset (left) and the extended dataset (right).

Mean Maximum Temperature

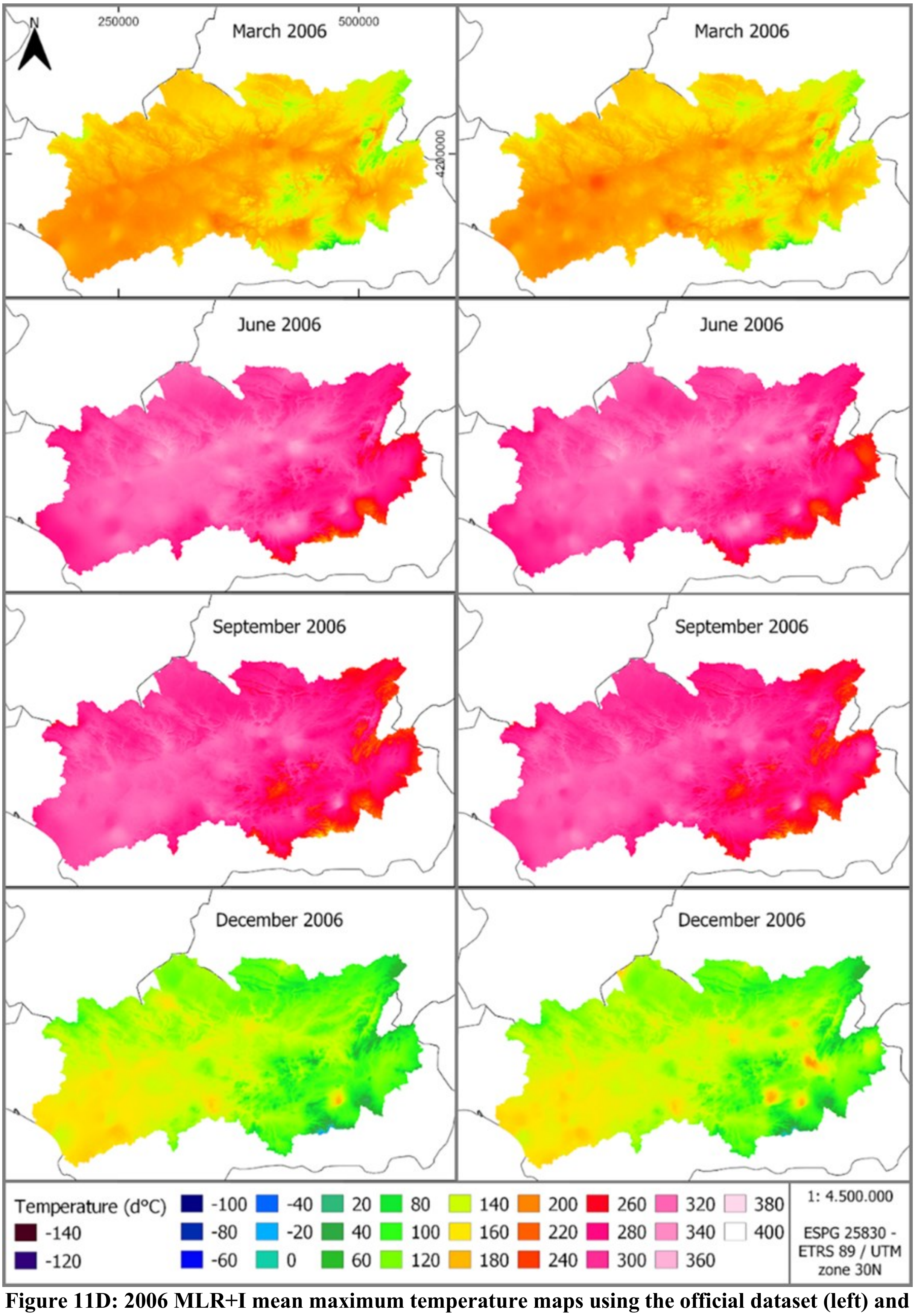


Figure 11D: 2006 MLR+I mean maximum temperature maps using the official dataset (left) and the extended dataset (right).

Mean Maximum Temperature

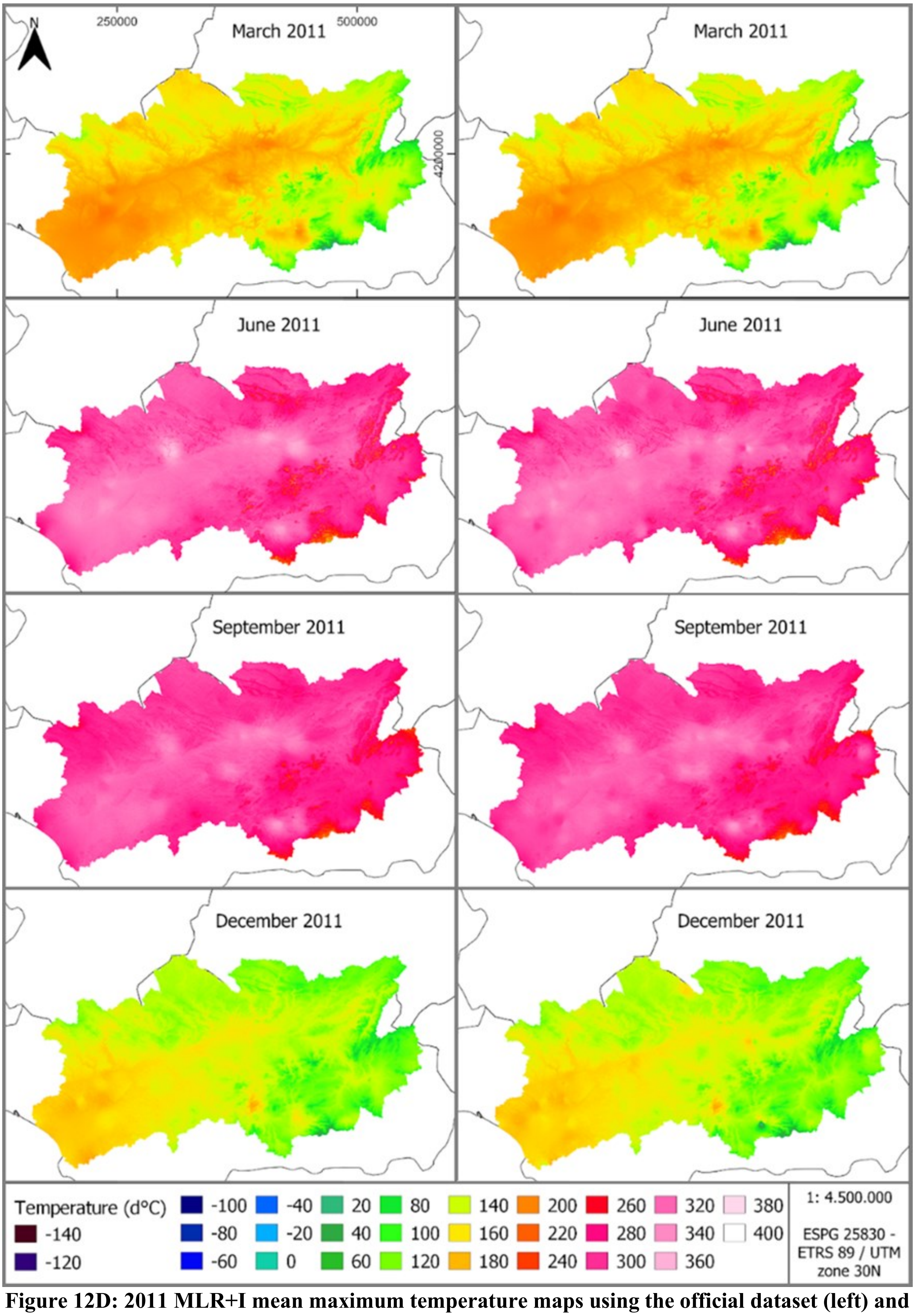


Figure 12D: 2011 MLR+I mean maximum temperature maps using the official dataset (left) and the extended dataset (right).