Ref. Ares(2020)4007814 - 29/07/2020

H2020-SC5-01-2017

MED-GOLD

Turning climate-related information into added value for traditional **MED**iterranean **G**rape, **OL**ive and **D**urum wheat food systems

Deliverable 2.2

Report on the tool performance

This project has received funding from the European Union's Horizon 2020 research and innovation programme under gran agreement No 776467.

DOCUMENT STATUS SHEET

Deliverable Title	**Report of the tool performance**	
Brief Description	This report - "Report on the tool performance" – describes the three climate service tools developed for the olive/olive oil sector and the methodology followed for the calculation of the products that are available through the tools.	
WP number	WP2	Co-design of climate service for olive/olive oil
Lead Beneficiary	ec2ce	
Contributors	*ec2ce (Arjona R.), NOA (Giannakopoulos C., Varotsos K., Gratsea M.), ENEA (Luigi Ponti), GMV (Rivas F.), BSC (Raül Marcos, Chihchung Chou, Nube Gonzalez-Reviriego), BEETOBIT (Federica Stocchino, Federico Caboni), CREA (Adolfo Rosati, Damiano Marchionni), CASAS Global (Andrew Paul Gutierrez)*	
Creation Date	18/06/2020	
Version Number	1.2	
Version Date	27/07/2020	
Deliverable Due Date	31/07/2020	
Actual Delivery Date		
Nature of the Deliverable	R	*R - Report* *P - Prototype* *D - Demonstrator* *O - Other*
Dissemination Level/ Audience	PU	*PU - Public* *PP - Restricted to other programme participants, including the Commission services* *RE - Restricted to a group specified by the consortium, including the Commission services* *CO - Confidential, only for members of the consortium, including the Commission services*

REVISION HISTORY LOG

Version	Date	Created / Modified by	Pages	Comments
1.0	18-06-2020	NOA, EC2CE	20	Initial Draft
1.1	13-7-2020	NOA, ENEA, GMV, BSC, BEETOBIT	63	Final Draft

All partners involved in the production/implementation of the deliverable should comment and report (if needed) in the above table. The above table should support the decisions made for the specific deliverable in order to include the agreement of all involved parties for the final version of the document.

Finally, after the peer review process, the deliverable should be modified accordingly to the comments and the reflections to the comments should be reported in the above table.

Disclaimer

The information, documentation, tables and figures in this deliverable are written by the MED-GOLD project consortium under EC grant agreement 776467 and do not necessarily reflect the views of the European Commission. The European Commission is not liable for any use that may be made of the information contained herein

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

EXECUTIVE SUMMARY

This report - "Report on the tool performance" – describes the three climate service tools developed for the olive/olive oil sector and the methodology followed for the calculation of the products that are available through the tools. The tool applications were developed through continuous feedback and interaction with the end-users. More specifically, the applications are:

i) The Olivia platform, a web platform which includes a predictive pest management support system based on artificial intelligence

ii) The MED-GOLD Dashboard, a web-based application that allows the users to easily visualise, interact and even download climate data and indices referring to seasonal or longer time scales. The calculation of the indices and the essential climate values (ECVs), which are the main products of the Dashboard, are described in a different section of this report.

iii) The physiologically based demographic model (PBDM), an artificial intelligence (AI) based application, which captures the weather-driven biology of the interaction between olive and the olive fruit fly.

The tools will be further improved in accordance with the feedback provided by the end-users who evaluated them during the second workshop that was held in May 2020 [RD.2].

With this deliverable, the project has contributed to the achievement of the following objectives (DOA, PartB Table1.1):

No.	Objective	Yes
1	To co-design, co-develop, test, and assess the added value of proof-of-concept climate services for olive, grape, and durum wheat	X
2	To refine, validate, and upscale the three pilot services with the wider European and global user communities for olive, grape, and durum wheat	
3	To ensure replicability of MED-GOLD climate services in other crops/climates (e.g., coffee) and to establish links to policy making globally	
4	To implement a comprehensive communication and commercialization plan for MED-GOLD climate services to enhance market uptake	
5	To build better informed and connected end-user communities for the global olive oil, wine, and pasta food systems and related policy making	X

1. INTRODUCTION

MED-GOLD's objective for the olive/olive oil sector – which is one of the most important crops in the Mediterranean region – is to create innovative climate services in order to help to the adaptation of the agricultural management to climate change; climate information for the next days, seasonal or longer time scales will be available. The first version of the climate service tool was developed by the scientific team of the MED-GOLD in order to meet the specific needs of the olive/olive oil sector and was evaluated by the end-users [RD.1]. The second version, following this evaluation, comprises three different tools: i) the Olivia platform, a web platform which includes a predictive pest management support system based on artificial intelligence, ii) the MED-GOLD Dashboard, a web-based application that allows the users to easily visualise, interact and even download climate data and indices and iii) the physiologically based demographic model (PBDM), an artificial intelligence (AI) based application, which captures the weather-driven biology of the interaction between olive and the olive fruit fly. The methodology and the description of the tools and the methodology for the calculation of the indices and the essential climate values for the short and long-term, which are the main products of the Dashboard, are presented in this report.

1.1 PURPOSE

This report summarises the main features of the different tools that have been developed for the olive/olive oil sector in accordance to the end-users' needs. The tools presented in this report are the Olivia platform, the Dashboard and the PBDM model. The methodology followed for the calculation of the climate indices - which are the main output of the Dashboard – is also described. It should be mentioned that the methodology is different for seasonal forecasts and climate projections, thus presented separately.

1.2 SCOPE

The scope of this report is to provide essential information about the performance of the tools developed in the frame of the MED-GOLD project. The aim of the climate service tools is to enable the end-users to adapt their decision making strategies to the climate conditions and climate change. This report focuses on the olive/olive oil agricultural sector.

1.3 DEFINITIONS AND ACRONYMS

1.1.1. Definitions

Concepts and terms used in this document and needing a definition are included in the following table:

Table 1-1 Definitions

Concept / Term	Definition
Dashboard	Web-based application that was developed to apply climatic services to the agricultural sectors of MED-GOLD project

OLIVIA	ICT platform based on Artificial Intelligence developed by ec2ce that provides a predictive pest management and decision farming tool to improve productivity

1.1.2. ACRONYMS

Acronyms used in this document and needing a definition are included in the following table:

Table 1-2 Acronyms

Acronym	Definition
AoI	Area of Interest
CDS-C3S	Climate Data Store of the Copernicus Climate Change Service
DF	Decision Farming
ECMWF	European Centre for Medium-range Weather Forecasts
IFS	Integrated Forecast System
LST	Land Surface Temperature
MED-GOLD	Turning climate related information into added value for traditional Mediterranean Grape, OLive and Durum wheat food systems
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NEMO	Nucleus for European Modelling of the Ocean
NMDI	Normalized Multi- band Drought Index
ORAS5	Ocean analysis and Reanalysis ensemble
PBDM	Physiologically Based Demographic Model
RCM	Regional Climate Model
SIGPAC	(Acronym Spanish, Agriculture Parcels Geographic Information System)
TVX	Temperature Vegetation Index

2. REFERENCES

Reference Documents

The following documents, although not part of this document, amplify or clarify its contents. Reference documents are those not applicable and referenced within this document. They are referenced in this document in the form [RD.x]:

Table 2-1 Reference Documents

Ref.	Title	Code	Version	Date
[RD.1]	First feedback report from users on olive oil pilot service development	Deliverable 2.6 / D2.6		2019
[RD.2]	Second feedback report from users on olive oil pilot service development	Deliverable 2.7 / D2.7		2020
[RD.3]	Report assessing the quality of seasonal forecast information and climate projections and their appropriateness for use for climate services for each sector	Deliverable 1.4 / D1.4		2019
[RD.4]	Bartók B, Tobin I, Vautard R, Vrac M, Jin X, Levavasseur G, Denvil S, Dubus L, Parey S, Michelangeli P-A, Troccoli A, Saint-Drenan Y-M (2019) A climate projection dataset tailored for the European energy sector Climate Services, 16:100138, doi:https://doi.org/10.1016/j.cliser.2019.100138	Bartok et al., 2019		2019
[RD.5]	Moss RH, Edmonds JA, Hibbard KA, Manning MR, Rose SK, van Vuuren DP, Carter TR, Emori S, Kainuma M, Kram T, Meehl GA, Mitchell JFB, Nakicenovic N, Riahi K, Smith SJ, Stouffer RJ, Thomson AM, Weyant JP, Wilbanks TJ (2010) The next generation of scenarios for climate change research and assessment Nature 463:747-756 doi:10.1038/nature08823	Moss et al., 2010		2010
[RD.6]	van Vuuren DP, Edmonds J, Kainuma M, Riahi K, Thomson A, Hibbard K, Hurtt GC, Kram T, Krey V, Lamarque J-F, Masui T, Meinshausen M, Nakicenovic N, Smith SJ, Rose SK (2011) The representative concentration pathways: an overview Climatic Change 109:5 doi:10.1007/s10584-011-0148-z	van Vuuren et al., 2011		2011
[RD.7]	Iturbide M, Bedia J, Herrera S, Baño-Medina J, Fernández J, Frías MD, Manzanas R, San-Martín D, Cimadevilla E, Cofiño AS, Gutiérrez JM (2019) The R-based climate4R open framework for reproducible climate data access and post-processing Environmental Modelling & Software 111:42-54 doi:https://doi.org/10.1016/j.envsoft.2018.09.009	Iturbide et al., 2019		2019

[RD.8]	Casanueva A, Herrera S, Iturbide M, Lange S, Jury M, Dosio A, Maraun D, Gutiérrez JM Testing bias adjustment methods for regional climate change applications under observational uncertainty and resolution mismatch Atmospheric Science Letters n/a:e978 doi:10.1002/asl.978	Casanueva et al., 2020		2020
[RD.9]	Giannakopoulos C, Kostopoulou E, Varotsos KV, Tziotziou K, Plitharas A (2011) An integrated assessment of climate change impacts for Greece in the near future Regional Environmental Change 11:829-843 doi:10.1007/s10113-011-0219-8	Giannakopoulos et al., 2011		2011
[RD.10]	Varotsos KV, Giannakopoulos C, Tombrou M (2019) Ozone-temperature relationship during the 2003 and 2014 heatwaves in Europe Regional Environmental Change 19:1653-1665 doi:10.1007/s10113-019-01498-4	Varotsos et al., 2019		2019
[RD.11]	Tebaldi C, Arblaster JM, Knutti R (2011) Mapping model agreement on future climate projections Geophysical Research Letters 38 doi:10.1029/2011gl049863	Tebaldi et al., 2011		2011
[RD.12]	Payload-Mass Trends for Earth-Observation and Space-Exploration Satellites bulletin/bullet97/rast.pdf	Rast, M., Schwehm, G., & Attema, , 1999		1999
[RD.13]	A statistical analysis of the relationship between climatic factors and the Normalized Difference Vegetation Index in China. International Journal of Remote Sensing. https://www.researchgate.net/publication/232062766_A_statistical_analysis_of_the_relationship_between_climatic_factors_and_the_Normalized_Difference_Vegetation_Index_in_China	Song et al, 2011		2011
[RD.14]	Monthly precipitation in mm at 1 km resolution based on SM2RAIN-ASCAT 2007-2018, IMERGE, CHELSA Climate and WorldClim	Brocca et al. 2019		2019
[RD.15]	Biome-level relationships between vegetation indices and climate variables using time-series analysis of remotely-sensed data. https://www.tandfonline.com/doi/full/10.1080/15481603.2020.1733325	Phiri et al. 2020		2020

[RD.16]	Gutierrez, A.P., 1996. Applied population ecology: a supply-demand approach. John Wiley and Sons, New York, USA	Gutierrez et al., 1996		1996
[RD.17]	Gutierrez, A.P., Ponti, L., Cossu, Q.A., 2009. Effects of climate warming on olive and olive fly (Bactrocera oleae (Gmelin)) in California and Italy. Climatic Change 95, 195–217. https://doi.org/10.1007/s10584-008-9528-4	Gutierrez et al., 2009		2009
[RD.18]	Ponti, L., Cossu, Q.A., Gutierrez, A.P., 2009a. Climate warming effects on the Olea europaea–Bactrocera oleae system in Mediterranean islands: Sardinia as an example. Global Change Biology 15, 2874–2884. https://doi.org/10.1111/j.1365-2486.2009.01938.x	Ponti et al., 2009a		2009
[RD.19]	Toko, M., Neuenschwander, P., Yaninek, J.S., Ortega-Beltran, A., Fanou, A., Zinsou, V., Wydra, K.D., Hanna, R., Fotso, A., Douro-Kpindou, O., 2019. Identifying and managing plant health risks for key African crops: cassava, in: Neuenschwander, P., Tamò, M. (Eds.), Critical Issues in Plant Health: 50 Years of Research in African Agriculture, Burleigh Dodds Series in Agricultural Science. Burleigh Dodds Science Publishing, Cambridge, UK, p. doi:10.19103/AS.2018.0043.07. https://doi.org/10.19103/AS.2018.0043.07	Toko et al., 2019		2019
[RD.20]	IPCC, Intergovernmental Panel on Climate Change, 2014. Climate change 2014: Impacts, Adaptation, and Vulnerability. Part A: global and sectoral aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA	IPCC, 2014		2014
[RD.21]	IPCC, Intergovernmental Panel on Climate Change, 2007. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.	IPCC, 2007		2007
[RD.22]	Gutierrez, A.P., Ponti, L., Gilioli, G., 2010. Climate change effects on plant-pest-natural enemy interactions, in: Hillel, D., Rosenzweig, C. (Eds.), Handbook of Climate Change and Agroecosystems: Impacts, Adaptation, and Mitigation. Imperial College Press, London, UK, pp. 209–237. https://doi.org/10.1142/9781848166561_0012	Gutierrez et al., 2010		2010
[RD.23]	Ponti, L., Gutierrez, A.P., Ruti, P.M., Dell'Aquila, A., 2014. Fine-scale ecological and economic assessment of climate change on olive in the Mediterranean Basin reveals winners and losers. Proceedings of the National Academy of Sciences, USA 111, 5598–5603. https://doi.org/10.1073/pnas.1314437111	Ponti et al., 2014		2014
[RD.24]	Ponti, L., Gutierrez, A.P., Ruti, P.M., 2009b. The olive-Bactrocera oleae (Diptera Tephritidae) system in the Mediterranean Basin: a physiologically based analysis driven by the ERA-40 climate data. Notiziario sulla Protezione delle Piante – III Serie 1, 113–128	Ponti et al., 2009b		2009
[RD.25]	Gutierrez, A.P., 1992. The physiological basis of ratio-	Gutierrez, 1992		1992

	dependent predator-prey theory: the metabolic pool model as a paradigm. Ecology 73, 1552–1563. https://doi.org/10.2307/1940008			
[RD.26]	Gutierrez, A.P., Baumgärtner, J.U., 1984. Multitrophic level models of predator-prey energetics: I. Age-specific energetics models - pea aphid Acyrthosiphon pisum (Homoptera: Aphididae) as an example. Can. Entomol. 116, 924–932. https://doi.org/10.4039/Ent116923-7	Gutierrez and Baumgartner, 1984		1984
[RD.27]	Petrusewicz, K., MacFayden, A., 1970. Productivity of terresrtial animals: principles and methods, IBP Handbook. Blackwell, Oxford, UK	Petrusewicz and MacFayden, 1970		1970
[RD.28]	Gutierrez, A.P., Wang, Y., 1977. Applied population ecology: models for crop production and pest management, in: Norton, G.A., Holling, C.S. (Eds.), Pest Management: Proceedings of an International Conference, October 25-29, 1976, IIASA Proceedings Series. Pergamon Press, Oxford, UK, pp. 255–280.	Gutierrez and Wang, 1977		1977
[RD.29]	Gutierrez, A.P., Baumgärtner, J.U., 2007. Modeling the dynamics of tritrophic population interactions, in: Kogan, M., Jepson, P. (Eds.), Perspectives in Ecological Theory and Integrated Pest Management. Cambridge University Press, Cambridge, UK, pp. 301–360.	Gutierrez and Baumgartner, 2007		2007
[RD.30]	Gutierrez, A.P., Ponti, L., 2014. Assessing and managing the impact of climate change on invasive species: the PBDM approach, in: Ziska, L.H., Dukes, J.S. (Eds.), Invasive Species and Global Climate Change. CABI Publishing, Wallingford, UK, pp. 271–288. https://doi.org/10.1079/9781780641645.0271	Gutierrez and Ponti, 2014		2014
[RD.31]	Marcelis, L.F.M., Heuvelink, E., 2007. Concepts of modelling carbon allocation among plant organs, in: Vos, J., Marcelis, L.F.M., Visser, P.H.B. de, Struik, P.C., Evers, J.B. (Eds.), Functional-Structural Plant Modelling in Crop Production. Springer, The Netherlands, pp. 103–111	Marcelis and Heuvelink, 2007		2007
[RD.32]	Rodríguez, D., Cure, J.R., Cotes, J.M., Gutierrez, A.P., Cantor, F., 2011. A coffee agroecosystem model: I. Growth and development of the coffee plant. Ecological Modelling 222, 3626–3639. https://doi.org/10.1016/j.ecolmodel.2011.08.003	Rodriguez et al., 2011		2011
[RD.33]	Gutierrez, A.P., Ponti, L., 2013. Eradication of invasive species: why the biology matters. Environmental Entomology 42, 395–411. https://doi.org/10.1603/EN12018	Gutierrez and Ponti, 2013		2013
[RD.34]	Regev, U., Gutierrez, A.P., Schreiber, S.J., Zilberman, D., 1998. Biological and economic foundations of renewable resource exploitation. Ecological Economics 26, 227–242. https://doi.org/10.1016/S0921-8009(97)00103-1	Regev et al., 1998		1998
[RD.35]	Bongi, G., 2002. Freezing avoidance in olive tree (Olea europaea L.): from proxies to targets of action. Advances in Horticultural Science 16, 117–124	Bongi, 2002		2002

[RD.36]	Fiorino, P., 2003. Olea: Trattato di olivicoltura. Edagricole, Bologna.	Fiorino, 2003		2003
[RD.37]	Vitagliano, C., Sebastiani, L., 2002. Physiological and biochemical remarks on environmental stress in olive (Olea europaea L.). Acta Horticulturae 586, 435–440.	Vitagliano and Sebastiani, 2002		2002
[RD.38]	Katsoyannos, P., 1992. Olive pests and their control in the Near East. Food and Agriculture Organization of the United Nations (FAO) Plant Production and Protection Paper 115, Rome, Italy.	Katsoyannos, 1992		1992
[RD.39]	Daane, K.M., Johnson, M.W., 2010. Olive fruit fly: managing an ancient pest in modern times. Annual Review of Entomology 55, 151–169. https://doi.org/10.1146/annurev.ento.54.110807.090553	Daane and Johnson, 2010		2010
[RD.40]	Nardi, F., Carapelli, A., Dallai, R., Roderick, G.K., Frati, F., 2005. Population structure and colonization history of the olive fly, Bactrocera oleae (Diptera, Tephritidae). Molecular Ecology 14, 2729–2738. https://doi.org/10.1111/j.1365-294X.2005.02610.x	Nardi et al., 2005		2005
[RD.41]	de Wit CT, Goudriaan J (1978) Simulation of ecological processes. PUDOC, Wageningen, The Netherlands	de Wit, 1978		1978
[RD.42]	Gutierrez, A.P., Villacorta, A., Cure, J.R., Ellis, C.K., 1998. Tritrophic analysis of the coffee (Coffea arabica) - coffee berry borer [Hypothenemus hampei (Ferrari)] - parasitoid system. Anais da Sociedade Entomologica do Brasil 27, 357–385. https://doi.org/10.1590/S0301-80591998000300005	Gutierrez et al, 1998		1998
[RD.43]	Di Cola, G., Gilioli, G., Baumgärtner, J., 1999. Mathematical models for age-structured population dynamics, in: Huffaker, C.B., Gutierrez, A.P. (Eds.), Ecological Entomology. John Wiley and Sons, New York. USA, pp. 503–534.	Di Cola et al., 1999		1999
[RD.44]	Vansickle, J., 1977. Attrition in distributed delay models. IEEE T. Syst. Man Cyb. 7, 635–638. https://doi.org/10.1109/TSMC.1977.4309800	Vansickle 1977		1977
[RD.45]	Ponti, L., Gutierrez, A.P., Basso, B., Neteler, M., Ruti, P.M., Dell'Aquila, A., Iannetta, M., 2013. Olive agroecosystems in the Mediterranean Basin: multitrophic analysis of climate effects with process-based representation of soil water balance. Procedia Environmental Sciences 19, 122–131. https://doi.org/10.1016/j.proenv.2013.06.014	Ponti et al., 2013		2013
[RD.46]	Ponti, L., Gutierrez, A.P., Cure, J.R., Rodríguez, D., Caboni, F., Boggia, A., Neteler, M., 2019. Bioeconomic analogies as a unifying paradigm for modeling agricultural systems under global change in the context of geographic information systems. Geophysical Research Abstracts 21, EGU2019-13677.	Ponti et al., 2019		2019
[RD.47]	Rosati, A., Metcalf, S.G., Lampinen, B.D., 2004. A simple method to estimate photosynthetic radiation use	Rosati et al., 2004		2004

	efficiency of canopies. Annals of Botany 93, 567–574. https://doi.org/10.1093/aob/mch081			
[RD.48]	Huld, T., Müller, R., Gambardella, A., 2012. A new solar radiation database for estimating PV performance in Europe and Africa. Solar Energy 86, 1803–1815. https://doi.org/10.1016/j.solener.2012.03.006	Huld et al., 2012		2012
[RD.49]	Ruane, A.C., Goldberg, R., Chryssanthacopoulos, J., 2015. Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. Agricultural and Forest Meteorology 200, 233–248. https://doi.org/10.1016/j.agrformet.2014.09.016	Ruane et al., 2015		2015
[RD.50]	Fernandez, J.E., Moreno, F., 1999. Water use by the olive tree. Journal of Crop Production 2, 101–162.	Fernandez and Moreno, 1999		1999
[RD.51]	Sofo, A., Manfreda, S., Fiorentino, M., Dichio, B., Xiloyannis, C., 2008. The olive tree: a paradigm for drought tolerance in Mediterranean climates. Hydrology and Earth System Sciences 12, 293–301.	Sofo et al., 2008		2008
[RD.52]	GRASS Development Team, 2015. Geographic Resources Analysis Support System (GRASS) Software, Version 6.4. Open Source Geospatial Foundation. URL http://grass.osgeo.org.	GRASS		2015
[RD.53]	Report assessing the quality of European climate observations and their appropriateness for use in climate services for each sector	Deliverable 1.3 / D1.3		2018
[RD.54]	Simulations of the ENSO Hydroclimate Signals in the Pacific Northwest Columbia River Basin. Leung et al. (1999) Bull. Amer. Meteor. Soc. 80 (11): 2313-2330. doi.org/10.1175/1520-0477(1999)080	Leung et al., 1999		1999
[RD.55]	Statistical methods in the atmospheric sciences. Wilks, Daniel S. (2011) Academic press Vol. 100	Wilks, 2011		2011
[RD.56]	Forecast verification: a practitioner's guide in a atmospheric science. Jolliffe, Ian T., and David B. Stephenson, eds. (2012) John Wiley & Sons	Jolliffe and Stephenson, 2012		2012
[RD.57]	Three recomendations for evaluating climate predictions. Fricke et al. (2013) Meteorological Applications 20.2: 246-255. doi.org/10.1002/met.1409	Fricke et al., 2013		2013
[RD.58]	Fair scores for ensemble forecasts. Ferro, C. A. T. (2014) Quarterly Journal of the Royal Meteorological Society 140.683: 1917-1923. doi.org/10.1002/qj.2270	Ferro et al., 2014		2014
[RD.59]	Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V.: Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble, Geosci. Model Dev., 7, 1297–1333, https://doi.org/10.5194/gmd-7-1297-2014,	Kotlarki et al., 2014		2014

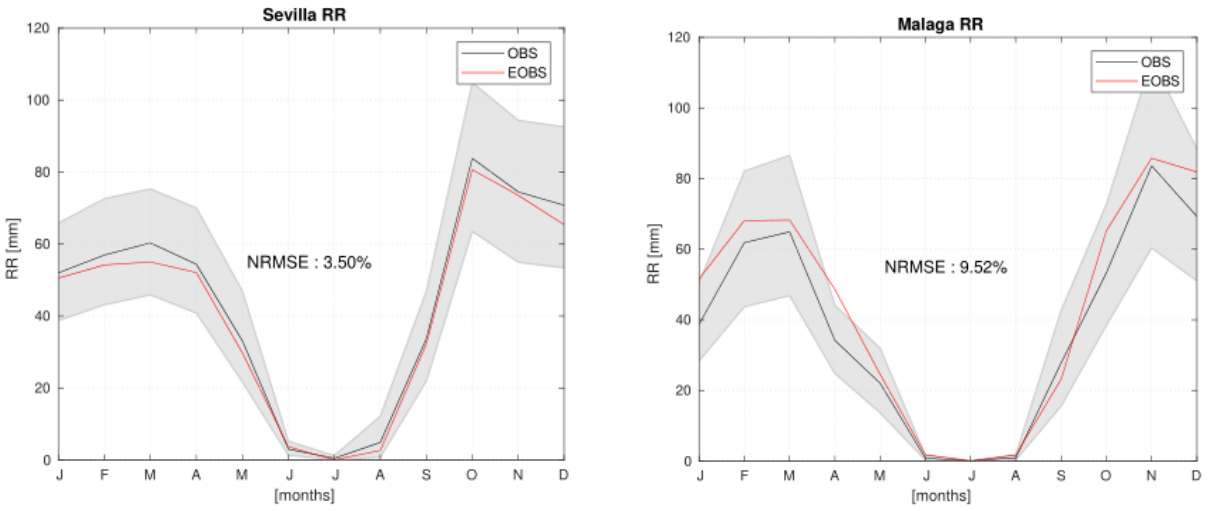
	2014.			
[RD.60]	C. Piani, J. O. Haerter and E. Coppola (2009) Statistical bias correction for daily precipitation in regional climate models over Europe, Theoretical and Applied Climatology, 99, 187-192	Piani et al., 2009		2009
[RD.61]	Bedia J, Baño-Medina J, Legasa M, Iturbide M, Manzanas R, Herrera S, Casanueva A, San-Martín D, Cofiño A, Gutiérrez J (2020). "Statistical downscaling with the downscaleR package (v3.1.0): contribution to the VALUE intercomparison experiment." Geoscientific Model Development. doi: 10.5194/gmd-13-1711-2020	Bedia et al., 2020		2020
[RD.62]	A. Michelangeli, M. Vrac, H. Loukos. "Probabilistic downscaling approaches: Application to wind cumulative distribution functions", Geophys. Res. Lett., doi:10.1029/2009GL038401, 2009	Michelangeli et al., 2009		2009
[RD.63]	Gutiérrez, JM, Maraun, D, Widmann, M, et al. An intercomparison of a large ensemble of statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross- validation experiment. Int. J. Climatol. 2019; 39: 3750– 3785. https://doi.org/10.1002/joc.5462	Gutierrez et al., 2019		2019
[RD.64]	Report on the Methodology followed to implement the wine pilot services	Deliverable 3.2/D3.2		2020

3. DATA AND SPATIAL DOMAINS

3.1 OBSERVATIONAL DATASETS EOBS

E-OBS is a dataset that provides the key variables of temperature and rainfall for Europe and is created by gridding, i.e., interpolation of irregularly-spaced surface weather stations from the ECA&D database onto a regular grid. In the early stages of the project, E-OBSv17 (horizontal resolution 0.22o) was evaluated against data from seven independent weather stations (the seven weather stations are not part of the network used to create the E-OBS dataset) over Andalucia for the years 2001-2017. The analysis revealed that E-OBS could be used for evaluating and bias adjusting the regional climate model (RCM) simulations [RD.53]. During the MED-GOLD progress a new E-OBS version was released, E-OBSv19 with a higher horizontal resolution (~0.1o) which was adapted for the needs of the WP2. Indicatively, some results from the evaluation of the E-OBS are shown in Figures 3-1 and 3-2.

Figure 3-1: Annual cycle of total monthly precipitation for the period 2001-2017. The black and red curves correspond to the weather stations and to the E-OBS dataset, respectively. The Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) are shown in the panels.



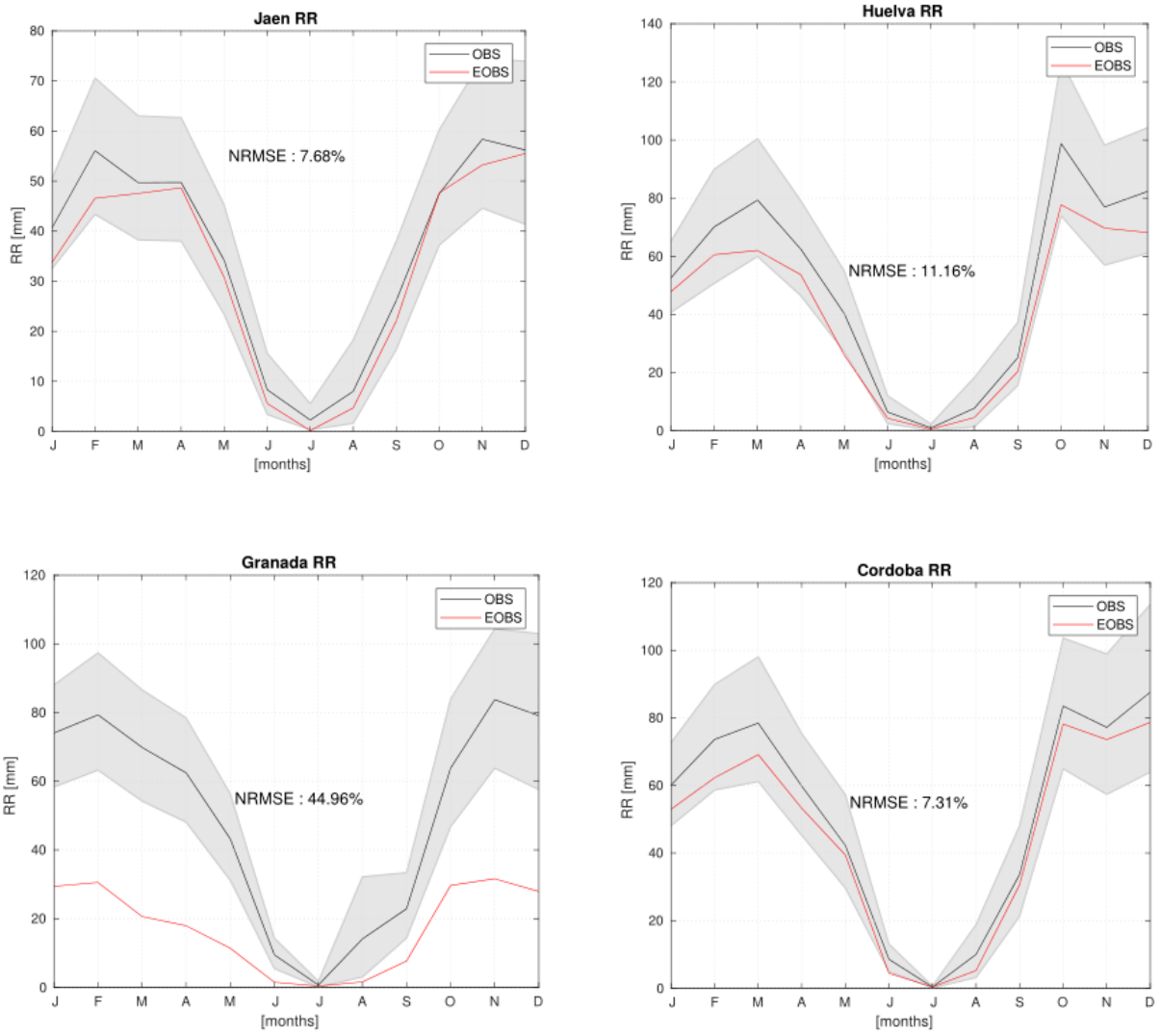
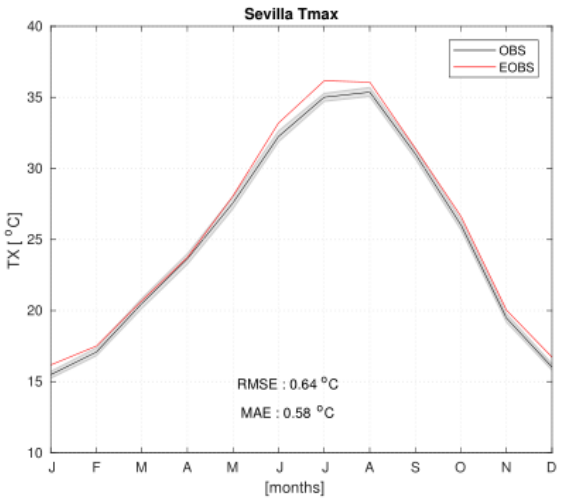
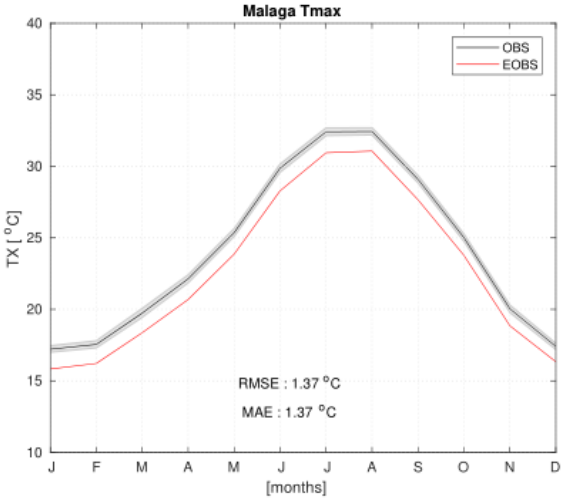
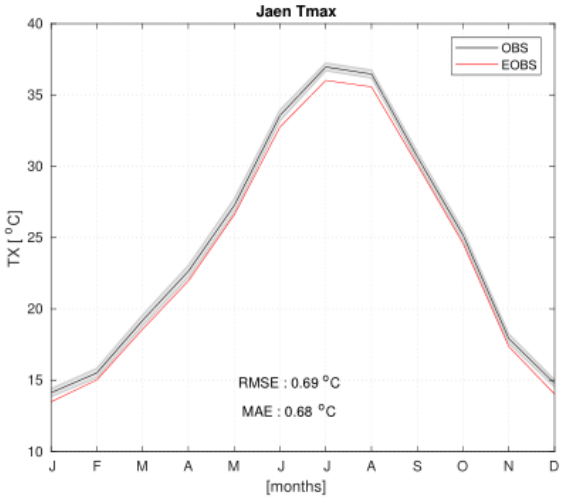
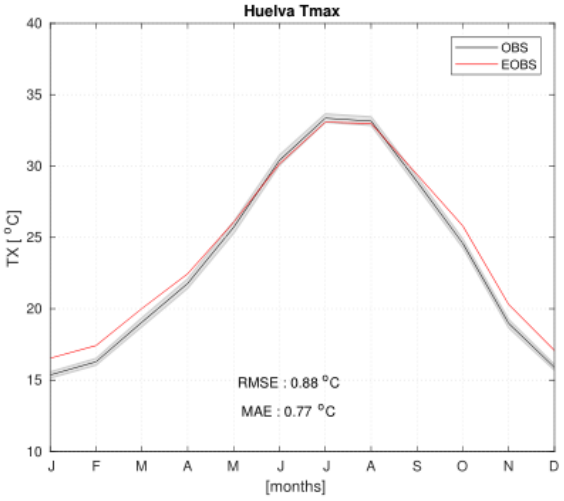


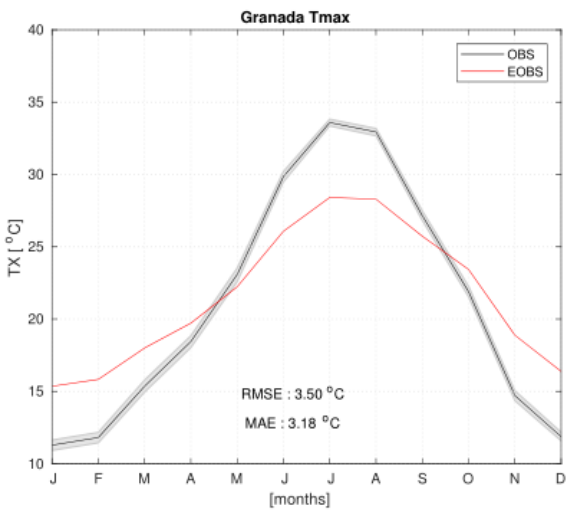
Figure 3-2: Annual cycle of mean monthly maximum temperature for the period 2001-2017. The black and red curves correspond to the weather stations and to the E-OBS dataset, respectively. The Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) are shown in the panels.

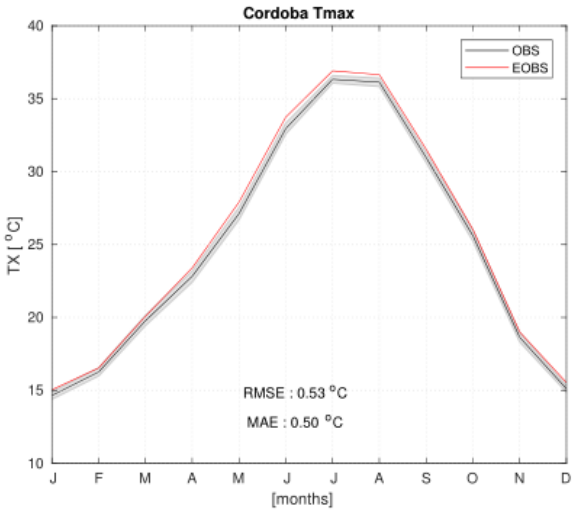

Sevilla Tmax
OBS
EOBS
TX [°C]
RMSE : 0.64 °C
MAE : 0.58 °C
[months]


Malaga Tmax
OBS
EOBS
TX [°C]
RMSE : 1.37 °C
MAE : 1.37 °C
[months]


Jaen Tmax
OBS
EOBS
TX [°C]
RMSE : 0.69 °C
MAE : 0.68 °C
[months]


Huelva Tmax
OBS
EOBS
TX [°C]
RMSE : 0.88 °C
MAE : 0.77 °C
[months]





3.2 SEASONAL DATASET ECMWF SEAS5

The ECMWF SEAS5 seasonal predictions dataset obtained from the Climate Data Store of the Copernicus Climate Change Service (CDS-C3S) is the fifth generation of seasonal prediction which replaces the former System 4 and uses the Integrated Forecast System (IFS) Cycle 43r1. There are 25 ensemble members for the hindcast period from 1993 to 2016.

SEAS5 includes enhancements in the land-surface initialization, atmospheric resolution and the ocean model when compared to S4. For instance, regarding the land-surface initialization, the former includes a new offline recalculation at the native atmospheric resolution with an improved precipitation forcing. The performance tests show a good degree of consistency between the initialization of SEAS5 re-forecast and the real-time predictions while the initialization is not perfect (i.e., the real-time assimilation is not identical as reanalysis). In addition, the SEAS5 uses the new version of ocean model NEMO (Nucleus for European Modelling of the Ocean) with upgraded ocean physics and resolution. Finally, the ocean and sea-ice initial conditions are provided by the new ocean analysis and reanalysis ensemble (ORAS5).

3.3 SATELLITE DATA

Earth observation remote sensing missions acquire information on the state of the planet without making physical contact with the target. Space-based remote sensing compliments in-situ measurements, taken at the target location. The remote-sensing observations made from orbit can be directly validated in the terrestrial environment that is being investigated. [RD.12] Regarding the sensors, the first remote sensing instruments were radiometers in the visible, infrared and microwave spectrum. Passive remote sensing was a precursor to active remote sensing given that the system is simpler (no transmission), the processing easier, and the requirements on power on the satellite are lower.

The evolution and impact of space-based remote sensing is accelerating as the technologies develop in the domain of electronics, ground processing, communications, and the inclusion of the private sector.

Earth observation missions provide a wealth of information on the planet. Forty years of space-based remote sensing has transformed our knowledge on the planet and how it is changing. Remote sensing has also an immeasurable effect on day to day applications.

Operational versus research missions.

Operational missions are launched to fulfil specific goals, to provide the users with information, images and data products they need. Examples of operational missions are the Copernicus Sentinel missions. Operational missions have strong requirements on reliability (low risk, redundancy) and longer lifetime to provide data continuity. Many operational missions span decades.

Research missions are intended for a one period study of a certain area. New technologies are tested in research missions, anything from guidance, navigation and control techniques, to calibration, data processing to non-space tested instrumentation certification.

In the framework of the MED-GOLD project, some satellite missions were identified as complementary to the validation (indirect way) tasks as well as extra information sources which could be useful to the users.

Considering the outputs defined to WP2 (although not exclusively), the Table 3-1 shows the missions used according to aims exposed in the previous paragraph. This information is currently available in the MED-GOLD ICT platform.

Table 3-1 Identified Earth Observation Datasets for WP2

WP2 output	EO dataset	Associated variable
Mean summer maximum temperature	ERA5-Land MODIS/Terra LANDSAT8	Temperature
Mean winter minimum temperature	ERA5-Land MODIS/Terra LANDSAT8	Temperature
Number of winter cold stress days	ERA5-Land MODIS/Terra LANDSAT8	Temperature
Number of annual and spring heat stress days	ERA5-Land MODIS/Terra LANDSAT8	Temperature
Number of summer heat stress days	ERA5-Land MODIS/Terra LANDSAT8	Temperature
Total annual, summer and winter precipitation	ERA5-Land SM2RAIN-ASCAT	Precipitation
Number of annual and winter dry days	ERA5-Land SM2RAIN-ASCAT	Precipitation

At this point it is important to highlight that all indirect measurements could accumulate errors (random or systematics) derived from several sources. Some factors such as clouds, sensor problems, transmission error, etc. can appear in the scenes affecting derived products like indices or mapping.

EO description.

SENTINEL 2

Copernicus is the European Union's Earth Observation programme aimed at providing remote sensing data services and Copernicus services information for the benefit of European citizens though products and applications, and to support policy and decision making for social, economic and environmental benefits.

Sentinel-2 is a multispectral operational imaging mission formed by two satellites. This mission compliments the SPOT and Landsat missions. The satellites are in the same orbital plane phased 180 degrees. The orbit is an SSO with a repeat cycle of 10 days. The mean altitude is 786 km and the LTDN is 10:30.

Sentinel-2 carries a MSI, Multi Spectral Instrument. The acquisition strategy consists in systematic push-broom acquisitions, plus lateral mode capability for emergency events monitoring. It samples 13 spectral bands: four bands at 10 m, six bands at 20 m and three bands at 60 m spatial resolution. The orbital swath width is 290 km.

MODIS

MODIS (or Moderate Resolution Imaging Spectroradiometer) is a key instrument aboard the Terra (originally known as EOS AM-1) and Aqua (originally known as EOS PM-1) satellites. Terra's orbit around the Earth is timed so that it passes from north to south across the equator in the morning, while Aqua passes south to north over the equator in the afternoon. Terra MODIS and Aqua MODIS are viewing the entire Earth's surface every 1 to 2 days, acquiring data in 36 spectral bands, or groups of wavelengths. Spatial Resolution: 250 m (bands 1-2), 500 m (bands 3-7), 1000 m (bands 8-36)

LANDSAT8

The Landsat 8 satellite orbits the Earth in a sun-synchronous, near-polar orbit, at an altitude of 705 km (438 mi), inclined at 98.2 degrees, and circles the Earth every 99 minutes. The satellite has a 16-day repeat cycle with an equatorial crossing time: 10:00 a.m. +/- 15 minutes. Landsat 8 acquires about 740 scenes a day on the Worldwide Reference System-2 (WRS-2) path/row system, with a swath overlap (or sidelap) varying from 7 percent at the Equator to a maximum of approximately 85 percent at extreme latitudes. The scene size is 185 km x 180 km (114 mi x 112 mi).

Spatial Resolution 30m except panchromatic Band 8, 15m

Thermal Infrared Sensor (TIRS) Two spectral bands: Band 10 and Band 11 100 m

ERA5 LAND.

Although ERA5 could be seen as model data, in this case it is included considering that several inputs used in the reanalysis comes from Satellite Observations.

ERA5-Land is a reanalysis dataset providing a consistent view of the evolution of land variables over several decades at an enhanced resolution compared to ERA5. ERA5-Land has been produced by replaying the land component of the ECMWF ERA5 climate reanalysis. Reanalysis combines model data with observations from across the world into a globally complete and consistent dataset using the laws of physics. Reanalysis produces data that goes several decades back in time, providing an accurate description of the climate of the past. Spatial Resolution 9km.

SM2RAIN-ASCAT

This dataset describes global daily satellite rainfall from ASCAT (real aperture radar.) soil moisture dataset [RD.14]. This Dataset was uploaded to MED-GOLD ICT platform as an ancillary dataset. Spatial Resolution 1Km.

Earth Observation in the context of MED-GOLD WP2

In the context of spatial coverage, The MED-GOLD WP2 OLIVE/OIL is focused on the Olive crops area, obtained from 2018 SIGPAC (Spanish, Agriculture Parcels Geographic Information System), previously filtered according to Olive use cover. The approximate extension of this region is 15700km^2 framed in the following geographic coordinates, N =38.73, S =36.16, W =-7.56, E =-1.63.

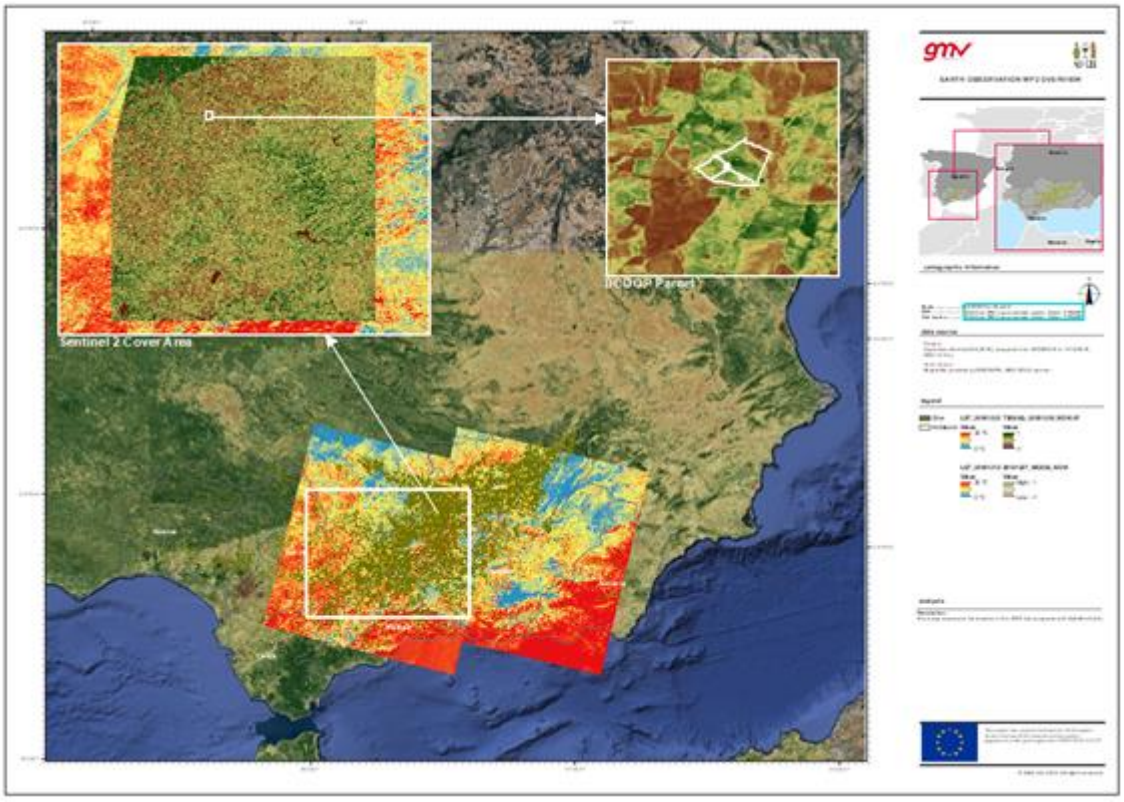
Despite the fact that the spatial covering of the satellite missions described above guarantees the covering of the Mediterranean Region, its usefulness depends on technical factors like Spectral covering (Bands) Covering Area (maximum extension covered) Spatial Resolution, temporal covering (when the mission started and frequency)

Sentinel2 (S2) covers an area of 12000km^2 , however the orbits described by the satellite are not enough to cover the total extension of the main AoI. This issue was solved by selecting the orbit which covers the maximum numbers of parcels of interest shared by the MED-GOLD partner DCOOP. Figure 3-3 shows the location of the AoI (Area of Interest) compared to S2 extension. The inset shows in detail a land parcel owned by DCOOP. The scene selected to calculate the Indexes corresponds to the orbit 30SUG.

The Landsat8 scene covers approximately 34200km^2 and the Area of Interest needs at least two scenes to be covered due its geographic location. The main inconvenience derived from this kind of situation is that each contiguous scene has a different acquisition day. So the analysis requires different treatment.

In the case of Modis Imagery, the total extension of the Area of interest is covered as well as by ERA5-Land. Figure 3-3 shows the area covered by Sentinel2, Landsat8 and MODIS imagery.

Figure 3- 3: EO imagery covered area.



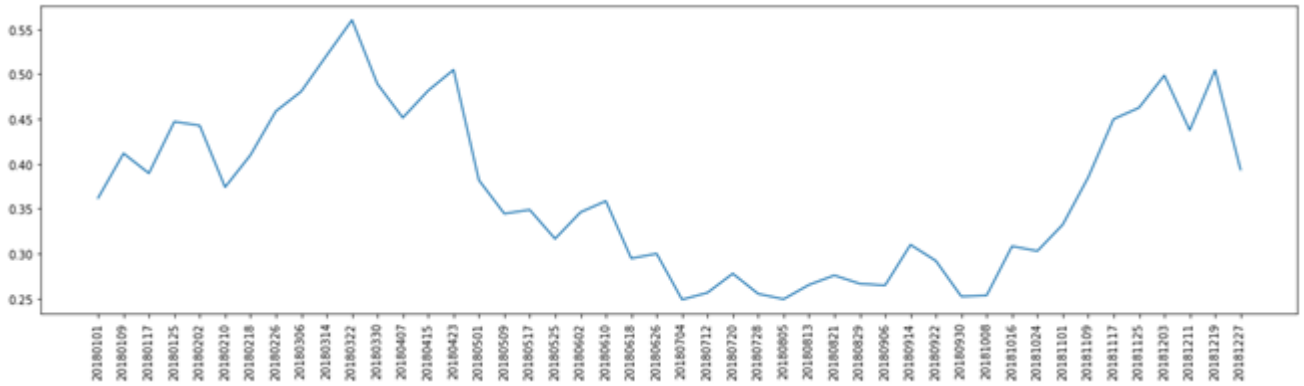
As indicated above, the use of MSI imagery in MED-GOLD is focused on the calculation of some indexes and measures as LST which could indirectly show climate variations through the comparison between EO results and climate behaviour. Figure 3-4 shows an example of the NDVI behaviour during 2018 on a specific location in the Olive region. Some studies suggest [RD.13] the relation between climatic factors and the Vegetation Index.

The calculation of different indexes depends on the factor associated with satellite specification. Table 3-2 shows the indexes calculated for the MED-GOLD project.

Table 3-2: Indices calculated for WP2

satellite	index	spatial resolution	time coverage
S2	NDVI	20m	2017
S2	NDWI	20m	2017
L8	NDVI	30m	2013
L8	NDWI	30m	2013
L8	LST	30m	2013
L8	TVX	30m	2013
MODIS	NDVI	460m	2000 to 2018
MODIS	NDWI	460m	2000 to 2018
MODIS	NMDI	460m	2000 to 2018

Figure 3- 4: NDVI (calculated from MODIS) variations during 2018



Geographical metadata

All subproducts (indexes) calculated in the MED-GOLD framework have a standardized geographic metadata based on ISO **TC 211. (**https://committee.iso.org/home/tc211). This metadata contains relevant information about the dataset and is available as *xml* file:

- fileIdentifier
- language
- contact
- dateStamp
- metadataStandardName
- metadataStandardVersion
- spatialRepresentationInfo
- referenceSystemInfo
- identificationInfo
- distributionInfo
- applicationSchemaInfo

EO Analysis

The EO imagery will be analysed as a way to verify results. As mentioned above, some of the currently operational Earth Observation system will be used to verify indirectly, when the physical conditions allow it, some results obtained by MED-GOLD. In the same way this information will be available to be shared as geospatial information linked to MED-GOLD through the dashboard if supported.

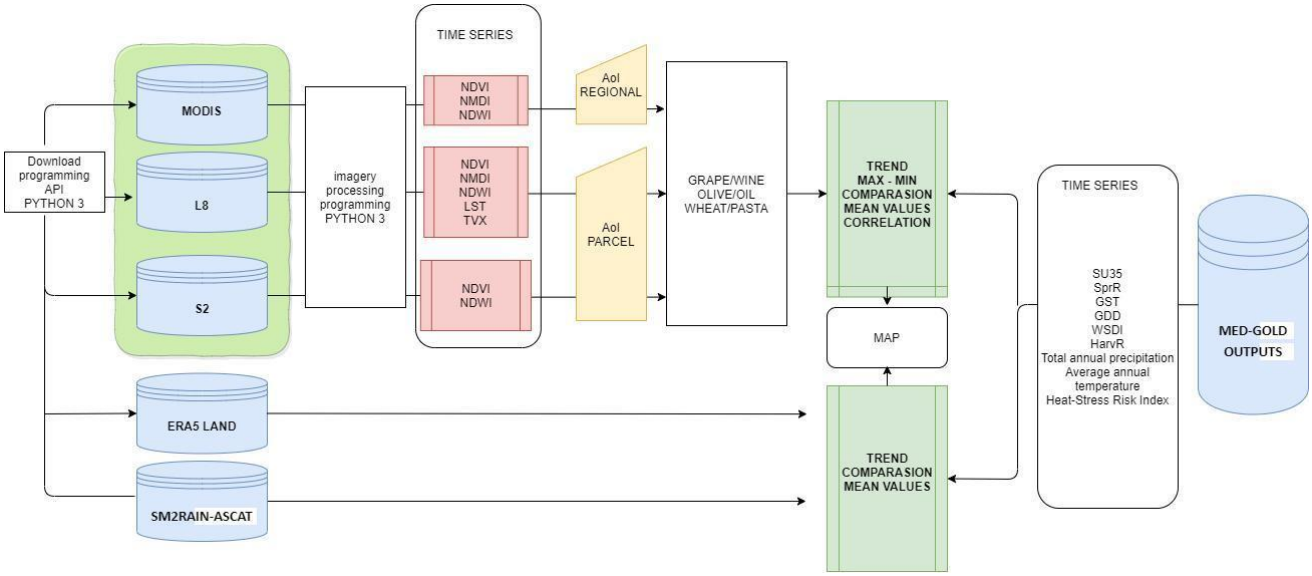
A total of 868 MODIS (2000-2018), 95 L8 (2014-2018), 31 (2017-2018) to WP2 has been currently processed and uploaded to MED-GOLD ICT platform focused on the second objective of these EO datasets.

The availability of the imagery depends on external providers, atmospheric conditions (cloud covering), sensor transmission, etc.

The imagery analysis in the context of MED-GOLD aims to verify some of the results obtained by the project. This process will be made through comparison based on the behaviour (trend, correlations, standard deviation) of the indexes curve compared to MED-GOLD results as well as statistical analysis.

Figure 3-5 shows a diagram to summarize the methodology that will be applied for validation purposes.

Figure 3-5: Methodology proposal to MED-GOLD validation using EO dataset

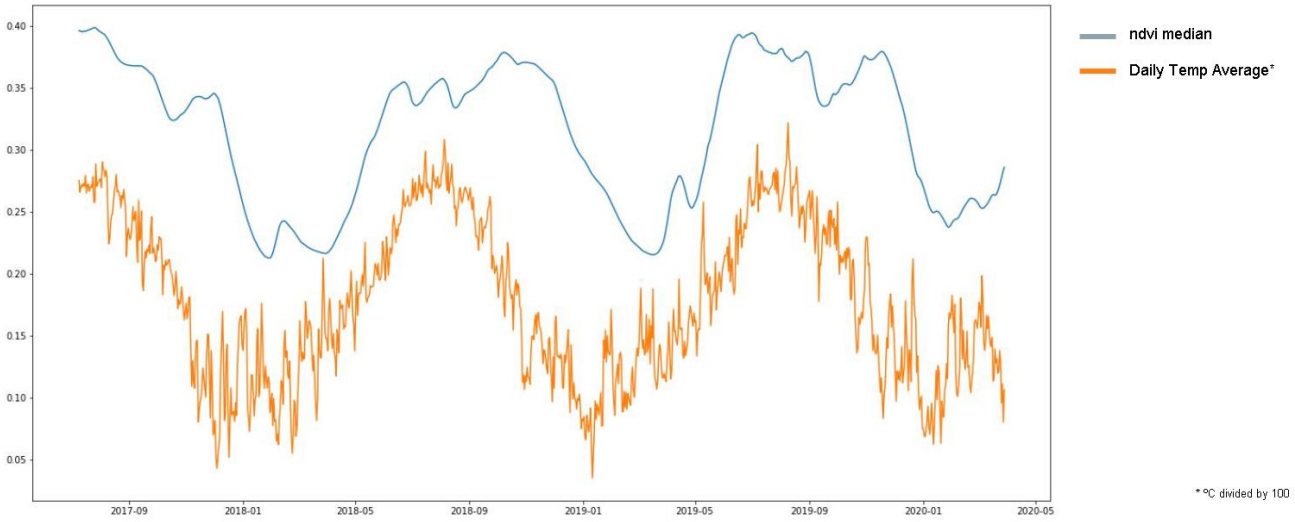


The first part of this proposal implied to develop python scripts to download dataset using API's, direct download or download under request approbation. Then another set of scripts were developed to calculate the indexes based on time series considering the spectral availability according to each sensor. Some of the spatial analysis libraries used were geopandas, gdal, shapely, pandas.

Considering the *AoI* (region or parcel) some aleatory areas inside the Olive region will be used to compare the results obtained by MED-GOLD. Depending on the imagery spatial resolution some parcels will be used to analyse the average value (derived from L8 and S2). These parcels correspond to Olive orchards provided by the MED-GOLD partner DCOOP. Indexes calculated from MODIS imagery will be used to compare largest extensions (maintaining homogeneity based on land cover) considering the spatial resolution of the MED-GOLD results.

This method is based on the assumption that different stages of the soil or vegetation are correlated with climate variations [RD.13][RD.15]. Figure 3-6 shows an example of the correlation between NDVI average (blue) and Temperature average (orange) of the orchard region in Spain (2017-2020).

Figure 3-6: NDVI average and Temperature average comparison.



The results of this validation applied to WP2 (olive/oil sector) will be described in the deliverable specified to this purpose.

3.4 Climate Projections Datasets

For the climate projections, daily maximum, daily minimum and daily mean temperatures are employed, as well as daily precipitation for Andalucia from a sub-ensemble of five RCMs (Table 3-3) from the EURO-CORDEX modelling experiment (http://www. euro-cordex.net). The model sub-ensemble used in this study is similar to the five member sub-ensemble identified and proposed in [RD.4]. The horizontal resolution of the models is 0.11° (~12km) and the simulated data of this project covers three periods: the 1971-2000 which is used as the reference period and the future periods 2031-2060 and 2071-2100 under two Representative Concentration Pathways (RCP) scenarios, the RCP4.5 and RCP8.5 [RD5-6].

Table 3-3 List of Regional Climate models used.

Institute	RCM	GCM
SMHI	RCA4	HadGEM2-ES CNRM-CERFACS-CNRM-CM5
IPSL-INERIS	WRF331F	IPSL-IPSL-CM5A-MR
KNMI	RACMO22E	ICHEC-EC-EARTH
MPI-CSC	REMO2009	MPI-M-MPI-ESM-LR

4. Climate Indices and Bioclimatic Indicators for the Olive/Olive Oil Sector

List of indices, identified and selected by the end-users as important for key decisions concerning the olive cultivation, have been calculated for the short (seasonal forecasts) and the long term (projections). The indices are presented in Table 4-1.

In addition, two indices (WINRR and SU38) have been used by DCOOP in order to identify a number of 'good' and 'bad' years in terms of olive yield and of infestation by the olive fly in the last 15 years. Based on the data of the report about the 'good' and 'bad' years, the return periods of the 'bad' years (the inverse probability of this event's occurrence in any given year) have been calculated for the near and distant future in Andalusia.

Table 4-1 List of selected indices for the olive/olive oil sector

Index	Description
SPRTX	Mean spring (Apr-May) maximum temperature
SPR32	Number of spring heat days, T>32 = count of days with Tmax above 32ºC during spring months (21 April - 21 June)
SU36	Number of summer heat days, T>36 = count of days with Tmax above 36ºC during summer months (21 June - 21 September)
SU40	Number of summer heat days, T>40 = count of days with Tmax above 40ºC during summer months (21 June - 21 September)
ANNRR	Total annual precipitation
WINRR	Total winter (Oct-May) precipitation
WINRRLT2	Number of dry days Oct-May, RR<2mm = cumulative number (count) of days with total precipitation below 2 mm
WINTN	Average winter (Nov-Jan) minimum temperature
WINFD	Number of consecutive frost days (Nov-Jan)
WINLT-7	Number of cold winter days cumulative number (count) of days with Tmin below -7ºC in November, December and January (temperature computed as SAT)
SPRTXHT28	Number of spring heat days = cumulative number (count) of days with Tmax above 28ºC, 21 April - 21 June (temperature computed as SAT)
SAT = 1.41 - 1.162 * V + 0.980 * T + 0.0124 * V ^ 2 + 0.0185 * (V * T) SAT is in degree Celsius (°C), V is wind speed (m/s) and T is air temperature (°C)	

4.1 SEASONAL FORECASTS

One of the components of the olive-oil climate service is based on seasonal predictions. BSC has developed this component of the climate service based on the seasonal predictions obtained from CDS-C3S, in particular, with the ECMWF SEAS5 dataset described in section 3 Data. The products obtained from this component of the climate service are based on seasonal predictions of essential climate variables (hereafter ECVs; i.e. mean temperature, maximum temperature, minimum temperature and precipitation) and bioclimatic indicators. The visualisation of these products is done through the MED-GOLD Dashboard.

The following section provides the description of the four climatic indices together with the equations as well as the methodology which was used to generate the forecast of the indices for the year 2020. The seasons considered for the indices below (i.e., spring and summer) are for the Northern Hemisphere.

Regarding bias correction, ERA5 is used as a reference to correct, after being upscaled, the seasonal forecast from ECMWF system 5. Please note that the start month chosen for most indices below for forecasts is February. However, due to the 7-month maximal forecast time, forecast data in September 2020 was not available and replaced with the climatology of observational data set (ERA5). Details of these configurations are given below for each index.

Calculations of the climatic indices

1. SPRTX – mean spring (April – May) maximum temperature

SPRTX is computed by averaging the daily temperature maximums during spring time, between 1 April and 31 May. As for bias correction, the period of observational data set used ranges between 1993 and 2019, the start month used for the forecast is February from 1993 to 2020. The equation for SPRTX is shown as follows

$$SPRTX = \frac{1}{n} \sum_{1\ April}^{30\ May} tasmax$$

where *tasmax* is daily temperature maximum and *n* is the total number of days.

There are similar basic concepts for the following three indices – SPR32, SU36 and SU40 –, such as the total count of heat stress days and a threshold for daily temperature maximum. The differences among the indices below are seasons and the threshold temperature considered. Therefore, a general equation is provided as below and in which the seasons (i.e., the inclusive days for each index) and the thresholds of temperature can be adapted accordingly. The general definition is the sum of days, during the period considered, when the percentile of daily temperature maximum forecast exceeds the percentile of observational temperature corresponding to the threshold. The formula is given below.

$$SPR_j \text{ or } SU_j = \sum_{initial\ day}^{final\ day} \left[P^{mod,i,j} > P_t^{obs,i,j} \right]$$

where $p^{mod,i,j}$ is the percentile corresponding to the ensemble-mean (excluding the *j*th year) modeled temperature at the *i*th day of the *j*th year. The $p_t^{obs,i,j}$ is the percentile of observational temperature corresponding to the threshold temperature *t* at the *i*th day of the *j*th year.

Instead of directly using the threshold temperature, the observational percentile corresponding to the threshold was used to examine the heat stress days in forecasts. This modification could avoid getting the unrealistic zeros for heat stress days in those areas where the temperature hardly exceeds the actual threshold. The observational data set used for the three indices below ranges between 1993 and 2019.

2. SPR32 – number of spring heat days (threshold: 32 °C)

SPR32 takes the period from 21 April to 21 June and the threshold of temperature is 32 °C. Forecasts with the start month of February were used for the years from 1993 to 2020. The equation is shown in *Eq 3*.

3. SU36 – number of summer heat days (threshold: 36 °C)
4. SU40 – number of summer heat days (threshold: 40 °C)

The inclusive days for both SU36 and SU40 ranges between 21 June and 21 September. For forecasts from 21 June to the end of August, start month of February was chosen from 1993-2020. For the rest of the days (i.e., first 21 days in September), the climatology of observation for the period from 1993 to 2019 was used. The equations are given below, and the threshold for indices is higher during summer time than spring: 36 °C and 40 °C for SU36 and SU40, respectively.

$$SU36_j = \sum_{21\ Jun}^{21\ Sep} \left[P^{mod,i,j} > P_{36}^{obs,i,j}\right]$$

$$SU40_j = \sum_{21\ Jun}^{21\ Sep} \left[P^{mod,i,j} > P_{40}^{obs,i,j}\right]$$

Seasonal forecast

Forecasts of the most-likely tercile category of the indices

Rather than the actual values of seasonal forecasts, the probabilistic forecasts of the most likely tercile category (i.e., below normal, normal and above normal with reference to climatology) are more useful for users to prepare labors/materials required for a few months ahead. This is because that users have been making decisions based on climatology together with their internal adjustments. As such, the probability of index forecasts which deviate from climatology will add values for users.

Bias-adjustment methods

The state-of-the-art forecast systems produce climate estimates with systematic errors due to model imperfections as well as other sources of uncertainty such as initial and boundary conditions. Thus, the bias in predictions needs to be adjusted before the provision of the forecast to the end-user. After testing different bias adjustment techniques, the simple bias correction approach has been chosen for implementing the workflow of the wine pilot service.

Simple bias correction generates an ensemble of the predicted data sets in which the mean and standard deviation are the same as the observations. Note that it assumes that the distributions of both the observational and predicted data sets well approximate the Gaussian (normal) distribution. At most of the times, this assumption is valid for monthly/seasonal mean data. This method has been widely used to correct temperature and precipitation [RD.54].

The simple bias correction can be formulated as shown in *Eq 6*,

$$y_{j,i} = (x_{i,j} - \overline{x})\frac{\sigma_{obs}}{\sigma_{mod}} + \overline{O}$$

where $y_{j,i}$ is the bias corrected forecast, computed by adjusting the anomalies with the ratio of standard deviation of observation to hindcast and adding the climatological observation (denoted as $\underline{O}$). The daily anomaly is calculated by subtracting the ensemble mean $\underline{x}$ of the hindcast data set from the daily value ($x_{i,j}$) for each member *i* and each year *j*. σ_{obs} (σ_{mod}) is the standard deviation of observation (hindcast).

This correction is done for each grid point separately, resulting in a new forecast ensemble that has the identical ensemble mean and standard deviation as the observation.

Verification methods

To evaluate the quality of seasonal predictions for both non bias-corrected and bias-corrected predictions, Pearson's correlation and Fair Ranked Probability Skill Score are applied as a verification metrics for all the indices. The methods are described as follows.

Correlation coefficient

The Pearson correlation coefficient [RD.55] between the ensemble-mean predicted and the observed data sets is used as a measure of the linear correspondence between the retrospective predictions and the reference. This can be defined as:

$$r_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})\,(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

where x_i and y_i are the observed and the ensemble-mean predicted values in each season over the *i* = 1, 2, … *n* years. The $\underline{x}$ and $\underline{y}$ are the average of the ensemble-mean predictions and the observations over the *n* years.

The correlation coefficient ranges between -1 and 1. If $r_{x,y}$ = 1 there is a perfect association between the ensemble-mean of the predictions and the observations. When $r_{x,y}$ = 0 indicates that there is no association between the ensemble-mean of the predictions and the reference dataset, which in turn, shows that the ensemble-mean of the predictions does not provide any added value relative to the retrospective climatology. A negative correlation coefficient indicates that the observed climatology should be used instead of the predictions. In addition to its predictability, only positive values of correlation coefficient indicate that the seasonal predictions are able to provide added values because of the need for the precise distribution of the data [RD.56].

Fair Ranked Probability Skill Score (FRPSS)

For the evaluation of the categorical events (e.g., below normal, normal and above normal here) obtained from probabilistic predictions, the Ranked Probability Score (RPS; [RD.55]) is one of the comprehensive verification metrics which has been frequently used in the context of seasonal predictions.

RPS is the sum of the squared distance between the cumulative probabilities of the n predictions-reference pairs (for the whole inter-annual time series) for the k equiprobable forecast categories (e.g., tercile when *k* = 3). The formula is given as follows:

$$RPS = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K}\left[\left(\sum_{j=1}^{k} y_{i,j}\right) - \left(\sum_{j=1}^{k} x_{i,j}\right)\right]^2$$

where $y_{i,j}$ and $x_{i,j}$ are the cumulative predicted and observed probabilities, respectively, assigned by the *i*th forecast (*i* = 1, … *n*) to the kth category (*i* = 1, … *k*). The $x_{i,j}$ = 1 indicates that the observation is in category *k*, and $x_{i,j}$ = 0 otherwise.

To better infer the predictive skill of forecast using *RPS*, it is often expressed as a skill score (i.e., Ranked Probability Skill Score, *RPSS*) which represents the added value of the prediction relative to the observations (usually referring to the climatology). The *RPSS* is given by:

$$RPSS = 1 - \frac{RPS}{RPS_{clim}}$$

As shown in the equation, RPSS ranges from $-\infty$ to one. The prediction is considered as unskillful when a negative RPSS returns. In other words, the climatology would be preferred in this case because the predictions do not provide additional value. In the case of a positive RPSS, the higher the RPSS, the better the predictions are than the climatology. RPSS = 1 corresponds to a 'perfect' prediction.

The RPSSs shown in this deliverable have been calculated for the verification of tercile (three equiprobable categories which are related with the two thresholds of the climatological distribution of the reference data set). The probabilities of forecast have been computed as the fraction of ensemble members (leave-one-out cross-validation) in the corresponding category.

Fair (Skill) Scores to ensemble predictions [RD.57], [RD.58] favors the prediction with its ensemble members being performed as if they had been sampled from the same distribution than the reference data set. The fair version of the RPSS (i.e., FRPSS) estimates the skill of prediction when an infinite ensemble size is used (a measure of potential skill).

4.2 CLIMATE PROJECTIONS

Regional Climate Models evaluation

The RCM simulations used in the project have an horizontal resolution of 0.11° (about 12 km), and are compared with the E-OBSv17 dataset (resolution 0.22°; approximately 25 km) for the period 1971-2000. The E-OBSv17 data have a lower resolution than the RCMs, therefore the RCM simulations were aggregated onto the 0.22° grid used by E-OBSv17. More specifically, the aggregation was achieved by calculating an area-weighted average of all grid cells of the RCM grid that overlapped with each of the E-OBS grid boxes [RD.59].

The evaluation analysis regarding the models performance against E-OBSv17 was performed for a selected number among the 27 ETCCDI (Expert Team on Climate Change Detection and Indices) core climate change indices for both the Mediterranean and the Andalucia with the analysis revealing considerable biases for the examined temperatures and precipitation indices respectively (RD.2). Therefore, we opted to perform bias adjustment to each model's output.

Testing bias adjustment methods

The following bias adjustments methods were considered as candidates to bias adjust the climate projections in MED-GOLD:

Variance Scaling of Temperature: the method corrects the mean and variance of temperature time series. It is only applicable for temperatures

Empirical Quantile Mapping (eqm): this is a very extended bias correction method which consists on calibrating the simulated Cumulative Distribution Function (CDF) by adding to the observed quantiles both the mean delta change and the individual delta changes in the corresponding quantiles. This method is applicable to any kind of variable.

Cumulative Distribution Function-transformation (CDFt): The method aims to link the cumulative distribution function (CDF) of a large-scale variable with the CDF of the same variable at a much smaller scale, and to downscale and correct CDFs from which local-scale data can be generated. When observations are of similar resolution as the climate model, CDft can be viewed as a bias-adjustment method. This method is applicable to any kind of variable.

Parametric Quantile Mapping (pqm): It is based on the initial assumption that both observed and simulated intensity distributions are well approximated by a given distribution therefore is a parametric q-q map that uses the theoretical instead of the empirical distribution. For instance, the gamma distribution is described in [RD.60] and is applicable to precipitation.

Generalized Quantile Mapping (gpqm) is also a parametric quantile mapping method but using two theoretical distributions, the gamma distribution and Generalized Pareto Distribution (GPD). By default, It applies a gamma distribution to values under the threshold given by the 95th percentile and a general Pareto distribution (GPD) to values above the threshold. The threshold above which the GPD is fitted is the 95th percentile of the observed and the predicted wet-day distribution, respectively. For variables other than precipitation, values below the 5th percentile of the observed and the predicted distributions are additionally fitted using GPD and the rest of the values of the distributions are fitted using a normal distribution.

All methods except CDFt are part of the climate4R statistical toolbox ([RD.7]; [RD.61]) whereas CDFt [RD.62] can be also found as a separate package in the R statistical computing project.

In order to assess and compare the performance of the different bias adjusting methods the cross validation framework was used. More specifically, cross validation tests whether the relationship established between the predictor (RCMs) and predictand (E-OBSv17) remains valid outside the training period [RD.63]. In the framework of Med-GOLD the available data (n = 30 years) were partitioned into k-non-overlapping "folds" or subsets, each containing n/k elements (k=5 1971:1976, 1977:1982, 1983:1988, 1989:1994,1995:2000). The bias adjusted methods were then calibrated and validated k times, considering in turn each of the folds as a test set and training the method with the remaining (k−1) sets. The resulting k-test series are typically joined

and validated together in a single series spanning the whole analysis period. The major findings of the comparison can be summarized as follows. 1) Regarding the annual cycle, all methods capture the observed annual cycle of the essential climate variables in Andalucía with no significant differences among the methods. 2) Among the methods better performance was found for the eqm method for threshold based indices such as the number of days with daily maximum temperatures higher that 35oC. Based on these findings NOA opted to use the eqm method for bias adjusting the RCM simulations in MED-GOLD. For more details on the eqm method and its implementation the reader is referred to the studies of [RD.7] and [RD.8].

During the MED-GOLD progress E-OBSv17 was updated to E-OBSv19 with the latest version having a higher horizontal resolution (0.11°; ~11 km). In order to integrate EOBSv19 in the WP2 workflow, the RCMs daily output were remapped onto the EOBSv19 grid using bilinear interpolation and the eqm bias adjustment method was then performed. Consequently the results were obtained for the selected bioclimatic indicators as well as for the ECVs.

For each one of the indices and for each grid point over Andalucia, the differences between each one of the future periods and the reference one are considered robust; the changes in at least three out of five models are found statistically significant and the change in the same models is of the same sign. The first criterion is examined by using the 95th percentile confidence intervals as derived by bootstrap [RD.9], [RD.10]. If only one of the criteria is met, the change at the specific grid point is not considered significant. This simple and transparent method, which was proposed by Tebaldi et al. [RD.11], summarizes multi-model projections and clearly separates lack of climate change signal from lack of model agreement by assessing the degree of consensus on the significance of the change as well as the sign of the change. The main idea is that if multiple models agree on a result, there is a higher confidence than if the result is based on a single model, or if models disagree on the result.

All the produced results have been uploaded to the MED-GOLD ICT platform.

Finally, apart from the daily bias adjusted temperature and precipitation data, daily data for relative humidity, wind speed and surface downwelling solar radiation were uploaded in the ICT platform. The data uploaded cover Andalucia for the three periods of interest and under the two RCPs and were used as input in the CASAS-PBDM modelling system.

5. OLIVIA PLATFORM

Olivia is a web platform, which is available to the farmers, and includes a predictive pest management support system based on artificial intelligence.

An invivo campaign in 2019 (from the end of June to the end of November) provided insights on pest evolution for the following 4 weeks based on climate, agronomic and geographical data. 380,000 ha in eight provinces in Andalucía contributed to the campaign; the tool manages efficiently and under sustainable conditions more than 191,000 farms, and it has been used by more than 220 technicians from Integrated Production organizations, that have an integrated pest management system in place.

The tool has been widely tested, the local government supports it and it is well known in the sector as the most advanced tool for pest control.

Decision Farming Tool:

The type of information that can be accessed via the tool are:

- Expected productivity (based on seasonal climate projections)
- Irrigation (a recommendation is provided month by month throughout the whole campaign)
- Meteorological data (past and seasonal)
- Access to a database with all the farm's information
- Information about agricultural practices

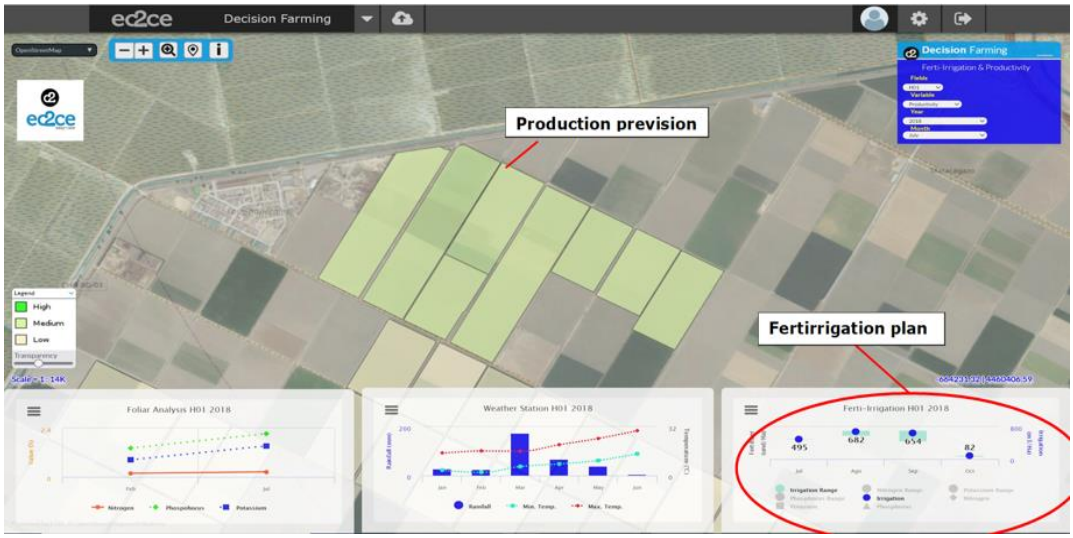
The different version of this tool was presented at the MED-GOLD workshops in April 2019 [RD.1] and May 2020 [RD.2] to receive feedback from technicians and users.

The tool was also tested during two campaigns in 2018 and 2019 in Portugal. It was tested in 300 ha in 2018 and in three areas of 1,800 ha in total in 2019, operating with DF tool. In a total of 92 ha, a comparison was made between operating a farm according to the tool recommendations and according to recommendations by a very experienced technician. In brief, the results for sectors A and B were +36% (not very homogeneous, the comparison was less realistic) and +6% productivity (homogeneous) following the tool recommendations, respectively. The results regarding consumption of water and fertilizers are summarised in Table 5-1. The tool interface was specific for this application (Fig. 5-1)

Table 5-1 Consumption of water and fertilisers

	N (UF)	P (UF)	K (UF)	Irrigation (mm)
DF tool vs Experienced	-30%	17%	21%	-28%

Figure 5-1: The decision farming tool interface



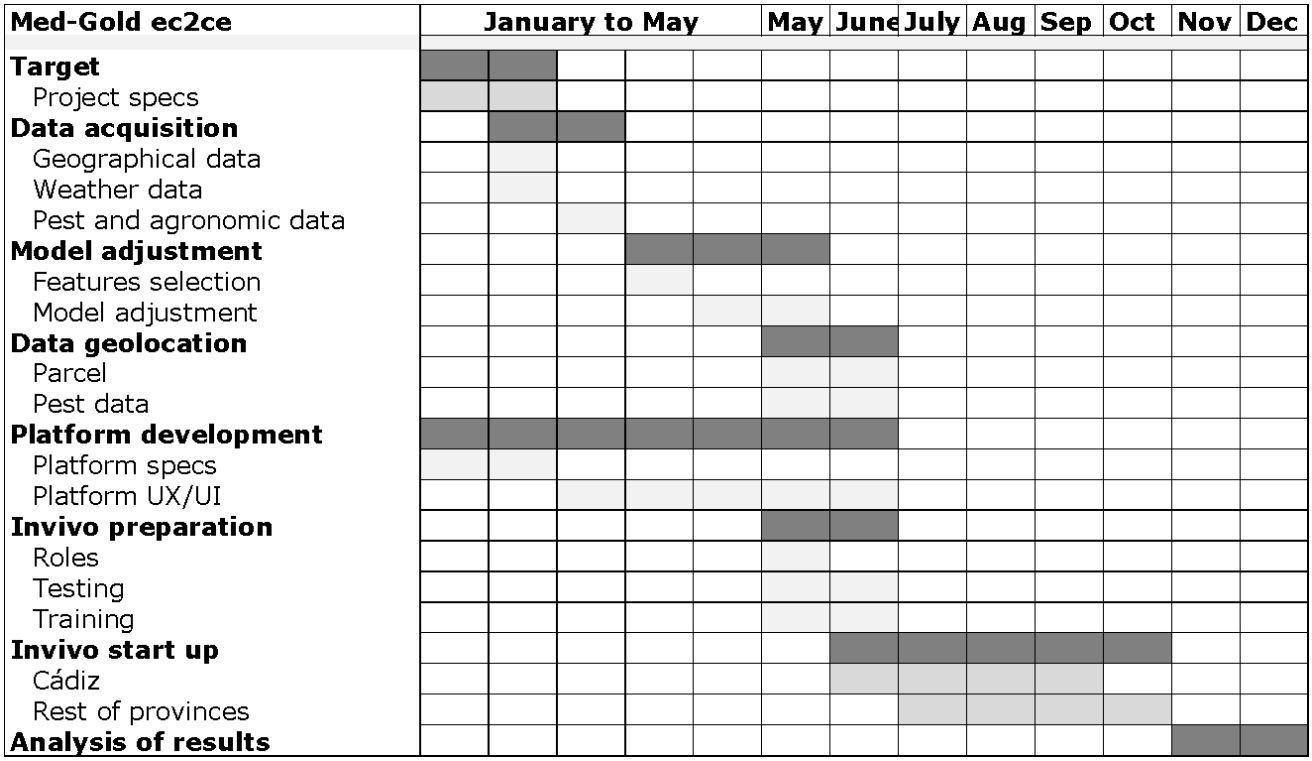
Overall, the use of the Decision Farming Tool resulted to:

- optimised productivity
- lower N and water consumption recommendations
- higher P and K consumption under recommendations

Olivia: Pest control tool

The tasks developed throughout the project are listed in Figure 5-2.

Figure 5-2: The tasks developed throughout the project



The data to adjust the models were collected from more than 400,000 ha in Andalucía using historical info from 2005. The target of the models is to depict the percentage of fruits affected by the fly. The data used include field scouting, weather information, geographical information and satellite images from Copernicus.

Once the modes were adjusted, they were put in vivo throughout the 2019 campaign. The models were developed to forecast up to four weeks the pest evolution (the users required just two). The results were provided as a classification with two different groups (Table 5-1):

- Group one: five intervals (0% / 1% / 3% / 5% / 10% / >10%)

- Group two: three intervals (0% / 1% / 5% / >5%), according to the standardized Integrated Pest Management protocols.

Some numbers of the project:

- In vivo in over 378,000 Ha
- More than 190,000 parcels in eight provinces (Fig. 5-3)
    - More than 7,930,000 predictions throughout the campaign
    - 223 technicians using the tool
    - 108 integrated production associations using the tool
        - Four training workshops (in different provinces) were celebrated before the campaign
        - 72% visiting the tool daily
        - 83% visiting the tool weekly
        - All of them providing the required input data weekly
        - All of them providing parcels information at the beginning of the campaign

Figure 5-3: The parcels of the 2019 campaign

Province	Surface (Ha)	Number of parcels
Cádiz	11,243	6,124
Córdoba	61,878	21,056
Granada	88,790	52,369
Huelva	733	97
Jaén	85,140	67,757
Málaga	26,175	5,032
Sevilla	105,395	39,427
Total	**379,354**	**191,862**

The results are summarised in Table 5-2.

Table 5-2: Accuracy of the predictive percentage of the fruits affected by the fly

Horizon (weeks)	Five intervals classification	Three intervals classification
1	76%	93%
2	71%	88%
3	66%	82%
4	63%	78%

The interface (Olivia) provides through a web platform (Fig. 5-4, 5-5) weather and climate information, satellite images and the expected pest evolution at parcel level for different time horizons and with different roles.

Figure 5-4: Olivia platform, Manager role vision

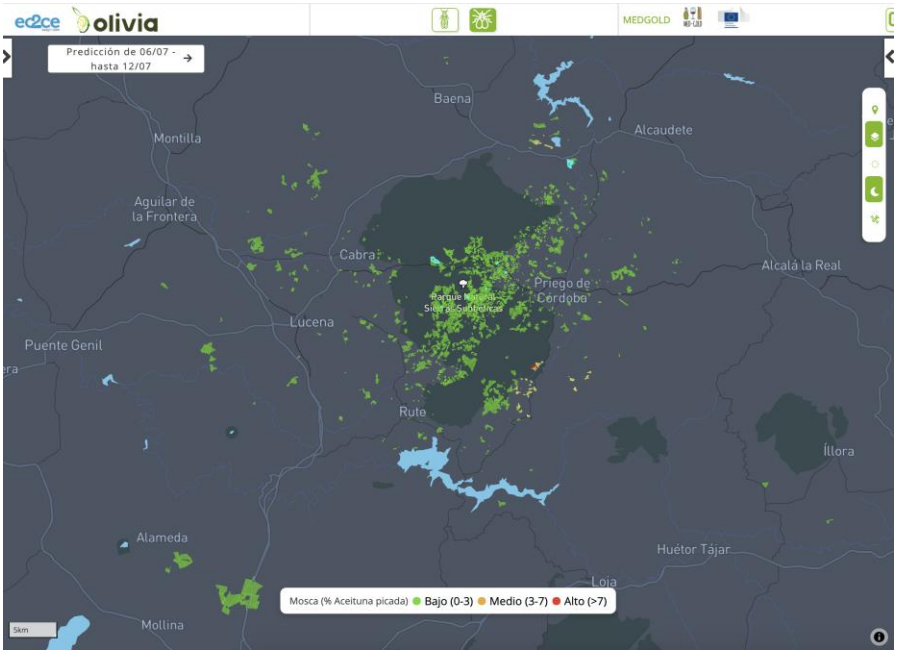
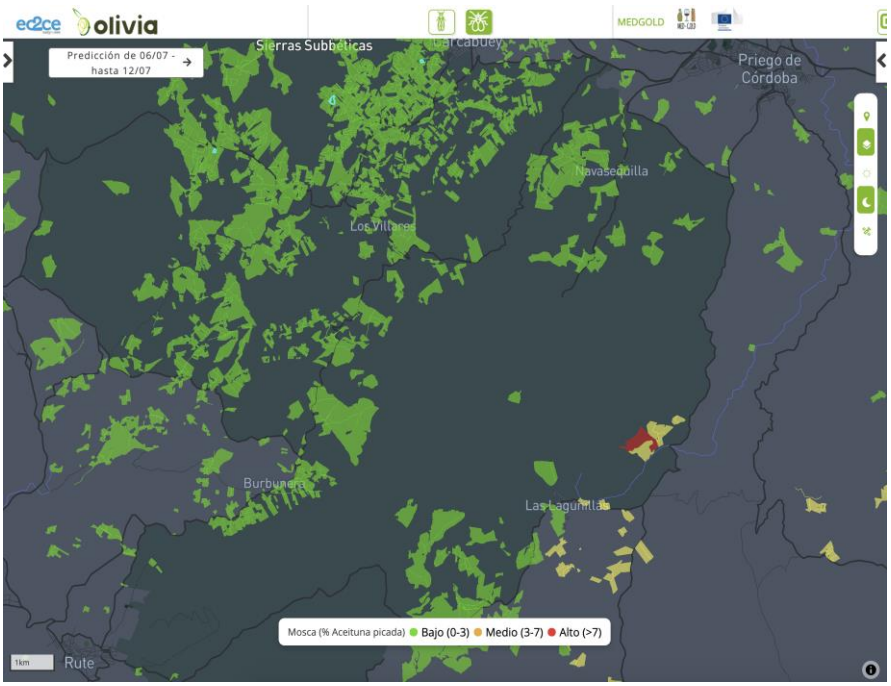


Figure 5-5: Olivia platform, Technician role vision



6. CASAS-PBDM FOR OLIVE AND OLIVE FLY

CASAS Global (Center for the Analysis of Sustainable Agricultural Systems, (http://www.casasglobal.org/) physiologically based demographic models (CASAS-PBDMs) [RD.16] are one of the key existing technology components of the MED-GOLD project. The CASAS-PBDMs application programming interface (API) is already part of the MED-GOLD ICT platform (https://platform.med-gold.eu/) as the *pbdm* workflow.

The PBDM used in WP2 explicitly captures the weather-driven biology of the interaction between the olive tree and the olive fruit fly *Bactrocera oleae* [RD.17]. The PBDM for olive and olive fly predicts the geographical distribution and relative abundance of the two species across time and space, independently of the observed species distributions, using extant and climate change weather scenarios as drivers for the system [RD.18].

The added value of PBDMs generally accrues mostly in terms of regional recommendations for crop management as opposed to precise prediction at field level (see e.g., the Olivia platform in this document). This is because PBDMs provide an assessment of the olive/olive fly system at the regional level that is independent of space and time, and hence provides insight on how to best allocate limited resources for agroecosystem management. This kind of insight would be impossible logistically and economically to obtain otherwise [RD.19].

How the biological component, which is itself a system (i.e., a food web), interacts with the climate component of the Earth system and how to assess this interaction is an open scientific question that the last two Assessment Reports of the Intergovernmental Panel on Climate Change (IPCC) pointed out clearly [RD.20, RD.21]. By bringing a realistic multitrophic biological layer in a climate service context using PBDMs, the MED-GOLD project aimed to address this scientific question, in an effort to complement the more widely-used index-based proxy approach to assessing agroecosystem complexity under climate change [RD.22].

This section outlines the methodology used for the PDBM analysis of olive and olive fly in the MED-GOLD case study on olive and olive oil.

6.1 MODELING THE DYNAMICS OF OLIVE AND OLIVE FLY

A weather-driven PBDM has been used to simulate the phenology, growth, and population dynamics of olive and olive fly at different geographic locations and scales [RD.23], [RD.24], [RD.18]. This model is being used in the MED-GOLD project to analyze this plant-pest system in Andalusia, Spain.

The PBDM approach identifies common processes across trophic levels and imbeds them in models having the same functional (resource acquisition model) and numerical (birth-death rates) response models [RD.25], [RD.16], [RD.26]. The model captures the biology and physiology of organisms in a general way describing the acquisition (i.e., the *supply*, S) and allocation of resources in priority order to egestion, conversion costs, respiration (i.e., the Q10 rule in poikilotherms) and growth and reproduction, using site-specific weather to drive the population dynamics. The physiology of assimilation falls under the rubric of the metabolic pool [RD.27]. The biology of biomass acquisition (i.e. the *supply*, S) is captured using a ratio-dependent functional response model where the sum of maximal genetic *demand (D*) is the major parameter [RD.26]. The $0 \leq S/D < 1$ ratio measures the extent to which assimilation demands are met and turns out to be always less than unity due to imperfect consumer search. The S/D ratio is used in the model to scale maximal vital rates of species [RD.26], [RD.28].

The models driven by weather have been used to simulate the dynamics of species in a broad variety of systems [RD.29],[RD.30]. The development of PBDMs for plants is well established in the literature [RD.31] with an excellent detailed example for coffee reported by Rodríguez et al. [RD.32] that under the MED-GOLD umbrella extends PBDM-based climate services to coffee in Colombia (WP6) with the real potential of its extension globally. These plant models provide a realistic base for linkages to herbivore and higher trophic

levels (e.g., [RD.33]. This is the structure used to model olive and olive fly dynamics. The model for olive growth and development is a plant canopy model with subunit populations of leaves, stem, root, and healthy and attacked fruit that allows capturing the bottom-up effects on the dynamics of olive fly. The mathematics of the model are reported in Gutierrez et al. (2009). The underlying model has also been used as a foundation for developing economic theory [RD.34] (see Fig. 6-1).

Olive is a long-lived species known for its tolerance to drought, with a distribution limited mostly by frost and high temperature, and to a lesser extent by soil water and other factors [RD.35], [RD.36], [RD.37]. Temperature influences nearly all aspects of olive's biology (see [RD.17]) and in the model, this is captured by a concave scalar function of temperature that represents the normalized net of the photosynthetic and respiration rates that also define the optimum and upper and lower thermal thresholds for development. The olive model predicts flowering phenology that is controlled by vernalization, age-structured growth and yield, and fruit mortality due to temperature and fly attack.

Olive fly is endemic in the Mediterranean Basin and the Middle East [RD.38], and is the major pest in most commercial olive-growing regions of the world [RD.39], [RD.40]. Its biology is closely linked to olive fruit age and availability, and as in olive, the effect of temperature on olive fly's vital rates is captured by a concave scalar function [RD.17].

6.2 IMPROVING THE OLIVE YIELD COMPONENT

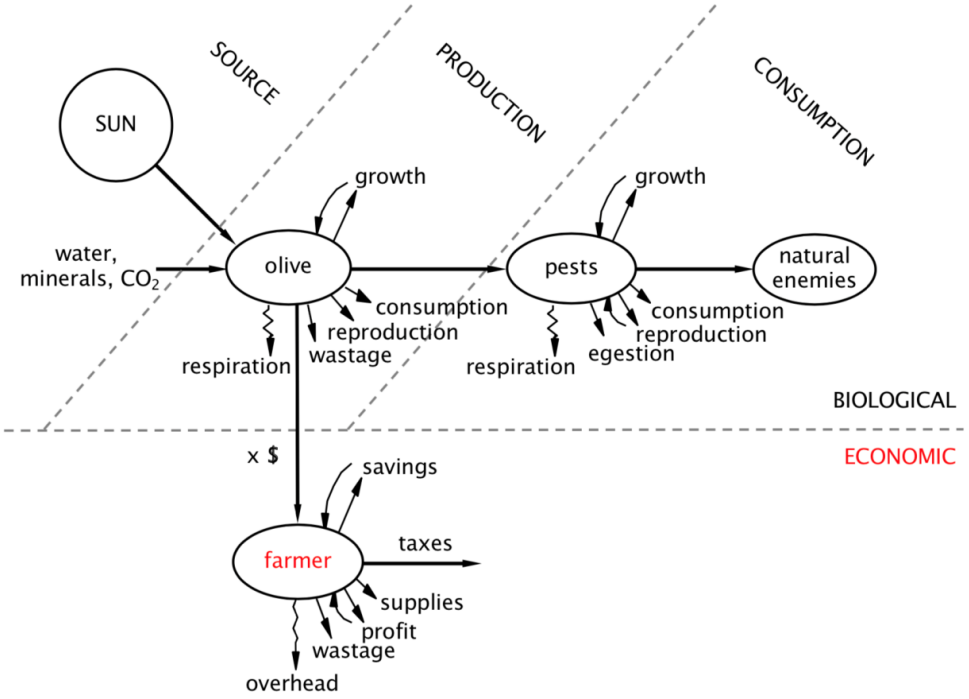
For a general description of how olive biology is modeled in the PBDM of olive, please see [RD.17] where the model is used to examine the effects of observed weather and three climate warming scenarios on the distribution and abundance of olive and olive fly in Arizona-California and Italy.

In general, the photosynthesis model common to all of the **plant** PBDMs is based on pioneering work by de Wit and Goudriaan [RD.41]. The total leaf mass per unit plant is converted to leaf area index (LAI) and the incident light in the area occupied by the plant is converted to g photosynthate per plant (P) – the Metabolic Pool (MP, see below). The actual P is computed as the proportion of the light intercepted (potential P, see the potAss variable in the olive code below) using Beer's Law that is concave saturating at about LAI = 4.5 (i.e., the demand driven Gutierrez-Baumgärtner functional response model; Gutierrez and Baumgärtner, 1984). The allocation of P then follows the allocation rules that also had Dutch influence but not in an age-structured context (i.e., the *demographic* part of PBDMs). Probably, only the coffee model developed by Gutierrez et al. [RD.42] and the newer version by Rodríguez et al. [RD.32] calculates light capture by branch level because berry initiation is a function of branch growth.

6.2.1 Overview

Common processes across trophic levels allow the same population dynamics and functional response models to be used in the PBDM to model the number and mass dynamics and interactions of olive and olive fly [RD.43], [RD.44] (Fig. 6-1).

Figure 6-1: Conceptual linkages of per capita metabolic pool (MP) resource acquisition and allocation across trophic levels including economics of human firm harvesting renewable resources (Gutierrez, 1996).



The core idea is that each species must acquire and allocate resources (Fig. 6-1; the metabolic pool; *MP*). This idea is captured by the Eqn. 1 below for all trophic levels (e.g., *i, i+1,…, i+n*) as simple mass (energy) flow *demand-driven* models of biomass dynamics per unit consumer in trophic levels M_i and M_{i+1} that determine the growth and reproductive rates [RD.26]. The notation could be simplified by aggregating terms, but we keep the biological detail.

$$\frac{dM_i}{dt} = \dot{M}_i = \lambda_i\{(\beta D_i h(u_i) - v_i - \mu_i)M_i\} - D_{i+1}h(u_{i+1})M_{i+1}$$

$$\frac{dM_{i+1}}{dt} = \dot{M}_{i+1} = \lambda_{i+1}\{(\beta D_{i+1} h(u_{i+1}) - v_{i+1} - \mu_{i+1})M_{i+1}\} - D_{i+2}h(u_{i+2})M_{i+2} \quad (1)$$

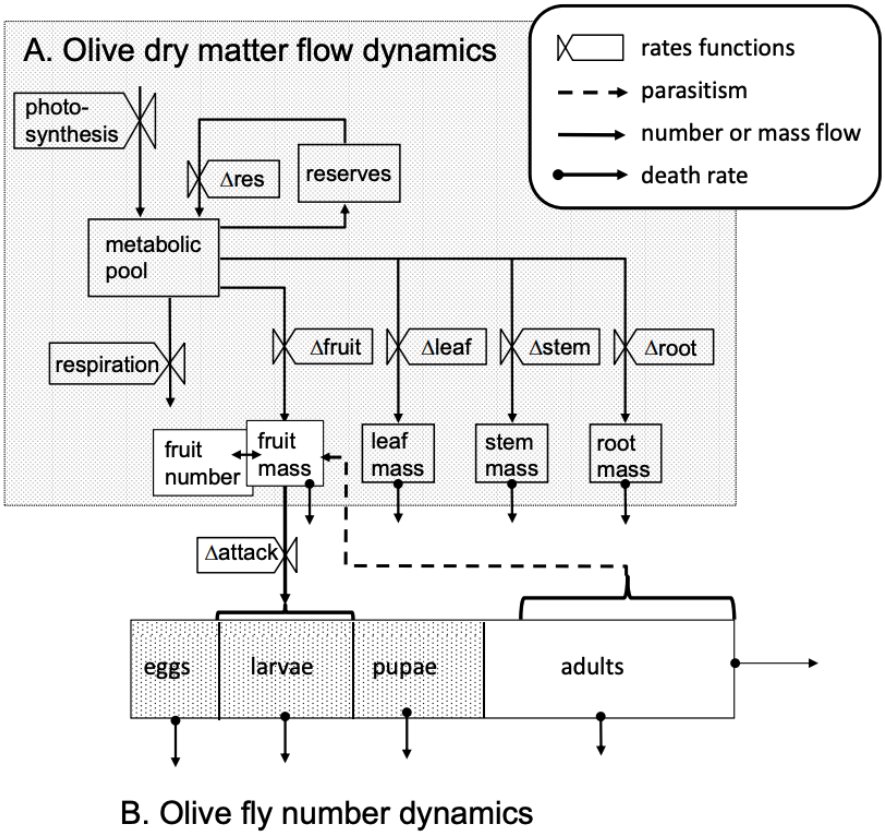
The concave ratio-dependent demand-driven functional response model ($0 \le h(u_i) \le 1$) estimates the proportion of the maximum resource *demand* (D_iM_i) acquired (*supply*; $S_i = D_i h(u_i)$) by trophic level *i* from trophic level *i-1* (see supplemental materials). The parameter $0 < \lambda_i < 1$ is the efficiency of converting the assimilated biomass to consumer biomass, $0 < \beta_i < 1$ is the correction for egestion, v_i is respiration rate that in poikilotherms is related to body size (*M*). Furthermore, the top trophic level in Eqn. 1 obeys the same relation, except that the right-most term for "*predation*" is absent, but as in all levels, an intrinsic net death rates occurs

(μ_i). The growth rates of individuals and the population are governed by the ratio of total resource acquired (supply, *S(t)*) to total demand at any time t (*D(t)*) (i.e. 0≤ *S/D=h(u)* <1), and provides timely reciprocal feedback between resource and consumer dynamics [RD.16]. Because search is imperfect (i.e. in $h(u_i)$), *S/D* <1 is the rule and measures the departure of the population growth rate from the maximum. Furthermore, the organisms may seek multiple resources (i.e. *j*=1, ..., *J*) using analogous acquisition functions (say *h(u,ij)* that modify growth and reproduction as follows $\prod_{j=1}^{J} h(u_{i,j}) \times h(u_i)$. In the olive model, the effects of water were implemented in previous analysis [RD.45] and could be implemented here too if required.

Structured population dynamics. In the field at any point in time, the olive tree has a population of leaves, some of which may be young and some older. A histogram of the frequency in each age class of leaves will vary over time. Similarly, stem, root and fruit also have age structured populations, and the key to modeling all of them (including leaves) is the development of population models for each subunit population. Specifically, the model for the olive plant simulates the structured population dynamics of nine functional populations (n = 1…9): the dynamics of olive leaf mass and numbers {sub models n = 1, 2}, stem plus shoot mass {n = 3}, root mass{4} and fruit mass and number {5, 6} (Fig. 6-2).

To include the olive fly, requires the development of three additional population models: one egg-larvae-pupal stages of olive fly in fruit {7}, and models for reproductive and dormant free-living adults {8, 9} (Fig. 6-2). More details are found in [RD.17]. In our model, the time step is a day that in physiological time varies with temperature, and may differ for each species.

Figure 6-2: Multitrophic biology of the olive/olive fly system. (A) Dry matter flow in olive and to olive fly, and (B) dynamics of olive fly number (Ponti et al., 2014).



Functional response model. All organisms are consumers (i.e., predators in a general sense), and all search for resources; for example, olive leaves search for light and the roots search the soil for water and nutrients, and adult olive flies seek olive fruit to deposit their eggs. The functional response model (Eqn. 2) in ecology determines the success of organisms in acquiring resources. To estimate the photosynthesis rate, we use the demand-driven concave functional response model (Eqn. 2) [RD.26].

$$S(u) = Dh(u) = D\left[1 - exp\left(\frac{-\alpha R}{DN}\right)\right] \quad (2)$$

S(u) is the per capita resource acquired by consumers of population *N* in the face of intra-specific competition (i.e. the exponent) for resource *R*, *D* is the per capita demand rate, and α is the search rate. In olive, $\alpha(N) = 1 - \exp(-sN)$ is Beer's Law of plant physiology, *N* is the density of leaf area (or roots) each with per capita (unit) search rate *s*. This makes Eq. 1 a type III functional response model.

In olive, with a known set of biological state variables (mass and age structure of plant subunits) and known temperature, light, water and nutrients, the quantity of photosynthate produced *S* = *S*(*u*) can be predicted

using Eqn. 2. This production is allocated first to wastage *(1−β)*, then respiration (i.e., $v = Q_{10}$), and after correction for conversion efficiency (λ) to reproduction and/or growth plus reserves (GR; Gutierrez and Baumgärtner, 1984).

$$GR = (S\beta - Q_{10})\lambda \quad (3)$$

We note that *S* depends on *D* in Eqb. 2 that can be estimated under conditions of non-limiting resource by solving Eqn. 3 and assuming $D \approx S_{max}$.

$$D \approx S_{max} = \left(\frac{GR_{max}(t)}{\lambda} + Q_{10}\right) / \beta \quad (4)$$

D is the sum of all plant subunit demands that may vary with age, stage, sex, size, temperature and other factors, and these and consumer preferences may be included in Eqn. 2. Dividing both sides of Eqn. 2 by *D* yields the consumers supply-demand ratio

$$0 \leq \phi_{\frac{S}{D}} = \frac{S}{D} = h(u) < 1 \quad (5)$$

$\phi_{\frac{S}{D}}$ is used to scale per capita growth and fecundity from the maximum rate under optimal conditions (e.g., $GR = \varphi \times GR_{max}$). The allocation is made to the subunits as the proportion they contribute to the total demand [RD.26]. In addition, if $O(t)$ is the number of fruit susceptible to shedding ($\mu_O(T)$), then at any time *t*, the number of surviving fruit equals

$$O(t+1) = \phi_{\frac{S}{D}}\big(1 - \mu_O(T)\big)O(t) \quad (6)$$

The reproduction in olive fly is modeled in a similar manner, except that the fly adult female seeks individual fruit to lay its eggs in [RD.17].

6.2.2 Implementation

The source code for the olive/olive fly PBDM is Borland Pascal code that is embedded in a larger code base of about ten thousand lines of code (without comments) including PBDMs for 40 different species of plants, herbivores, parasitoids, predators, and pathogens (a subset of the species modeled using PBDMs) that were published as PBDM analyses implemented in a GIS context [RD.30]. Hence, a variety of similar models coded in Pascal exist that draw on the same code base that is currently undergoing rewrite using an object-oriented programming paradigm in a more modern language for release as open source [RD.46]. Like the rest of the

PBDM code base developed in the last three decades, the Pascal code for olive/olive fly PBDM is currently not licensed nor it is deposited in a code repository, and is managed by the nonprofit scientific consortium CASAS Global (http://www.casasglobal.org/) and its CEO Prof. Andrew Paul Gutierrez who is a member of the MED-GOLD EAC. The PBDM algorithms as well as key innovative code such as the Pascal subroutine for distributed maturation times with and without attrition, were published in detail in Gutierrez (1996).

In the PBDM for olive, photosynthesis is the supply of photosynthates that the plant uses to meet its genetic demand by allocating it to conversion cost, respiration, egestion, reproduction, and growth.

A good summary of the meta-physiology underlying the olive plant model is found in Wermelinger et al. (1991) and the type III functional response model used (see Gutierrez, 1996, p. 81) is found in Gutierrez and Baumgärtner (1984).

The general idea is to use RUE (radiation use efficiency) as modeled by Rosati et al. [RD.47] in place of the constant daily solar radiation from weather data that in the olive PBDM is used to obtain potential assimilation A response curve for photosynthesis vs. photosynthetically active radiation (PAR) was developed under the MED-GOLD project for the olive leaf and was used to predict daily photosynthesis and then daily RUE, using the modeling approach of Rosati et al. (2004) and Rosati and Dejong (2003) based on an hourly solar radiation time series for the period 2007-2016 for Jaen, Spain obtained from the Photovoltaic Geographical Information System (PVGIS, https://ec.europa.eu/jrc/en/pvgis) [RD.48] under the assumption that 45% of global radiation is PAR. Daily RUE was plotted against daily PAR for all days and for two years, as two years is the period considered by Mariscal et al. (2000) in their field experiment on RUE in olive.

The PAR response curve was normalized with respect to average RUE across the same period. This results in a coefficient that varies from about zero on the darkest days to about 1.4 on medium PAR intensity days (best in terms of RUE) and then drops back to just below one for brighter days. This coefficient could be multiplied by the 3.85 constant in the olive PBDM model when computing potAss and will vary as a function of daily radiation.

An even more precise approach would be to calculate the actual RUE on that day not only from the global PAR but also from its hourly variability. In this case, the transition to the PBDM model would take more time, but it would still be feasible. However, it remains to be verified whether this further improvement would actually be worthwhile because over- and under-estimates would probably offset each other in the long run. Hourly solar radiation data would also have to be processed and integrated into the current weather dataset, which is itself a nontrivial task. Another approach would be to integrate the RUE values on a daily basis, and to develop a sub model function to correct observed daily solar radiation. Additionally, the modeling approach considered in Rosati et al. (2004) and Rosati and Dejong (2003) works best at daily time steps.

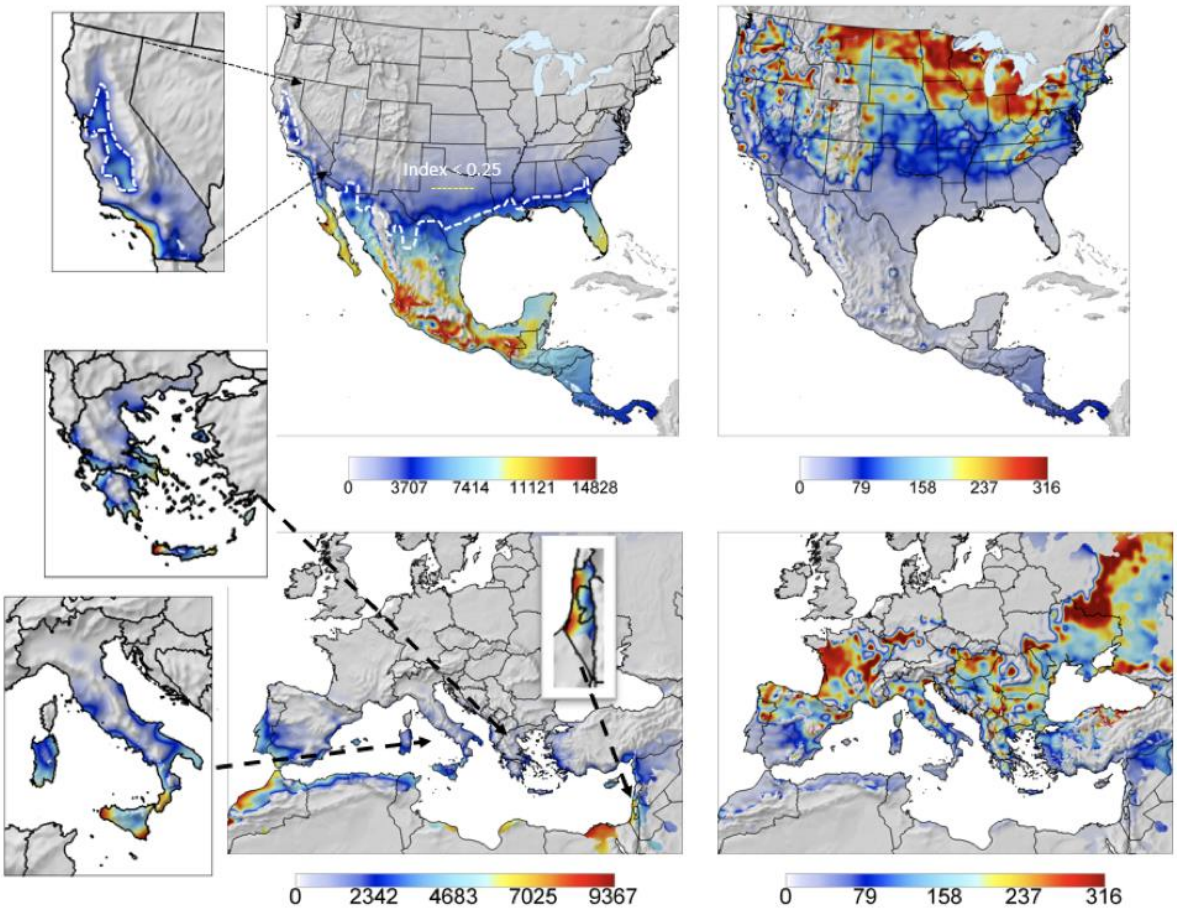
An extensive literature review has been carried out to confirm the radiation use efficiency value used in the PBDM and to analyze further possibilities of improving the model's ability to predict the effects of climate change on olive physiology/productivity. Literature review also included finding representative values of light interception by olive orchard, in order to give the possibility to adjust the model to different actual light interception values for different types of orchards in different locations, also based on tree management practices. Not having found all the values required, field work was carried out, and unpublished data were used to calculate light interception values for different olive orchards. The data has recently been submitted for publication to the international peer-reviewed journal Agriculture and Forest Meteorology.

6.3 IMPROVING THE OLIVE FLY COMPONENT

The effects of relative humidity on adult oviposition is being implemented. Furthermore, the models for olive fly and Mediterranean fruit fly (*Ceratitis capitata*) are being used as templates for development of models for oriental fruit fly (*Bactrocera dorsalis*), melon fly (*Bactrocera cucurbitae*), and Mexican fruit fly (*Anastrepha*

ludens). GIS maps for North and Central America and the Mediterranean Basin are under development to assess their geographic range and potential invasiveness. This work can easily be extended to explore the invasive potential of still other fruit flies and other invasive pests (*Tuta absoluta*, Ponti et al. under submission). These are but a few of the spinoffs of PBDM development. Below is a sample of the mapping for Mediterranean fruit fly with the favorability index < 0.25 in North America indicating the upper limits of potential favorability (Fig. 6.3). These maps were developed using PBDMs and the GIS technology described below.

Figure 6-3: Prospective geographic distribution and relative abundance of Mediterranean fruit fly (*Ceratitis capitata*) as affected by temperature and relative humidity for North and Central America and the Mediterranean Basin, obtained using the same PBDM methods used for olive fly in the MED-GOLD project.



6.4 THE CLIMATE DATASET

Daily weather from the AgMERRA climate dataset was used to run the PBDM of olive and olive fly because it includes the six climate variables required to run PBDMs: maximum and minimum temperature, precipitation, solar radiation, relative humidity, and wind. AgMERRA is a climate forcing dataset created as an element of the Agricultural Model Intercomparison and Improvement Project (AgMIP, https://agmip.org/) to provide consistent, daily time series over the 1980-2010 period with global coverage of the climate variables required for agricultural models [RD.49]. The AgMIP project is one of the initiatives with which MED-GOLD officially interacts (see D6.7), and the MED-GOLD External Advisory Committee (EAC) includes Prof. Bruno Basso (Michigan State University, USA) who is one of the research leaders at AgMIP.

Daily weather station data provided by Junta de Andalucia and obtained through the IFAPA (Instituto Andaluz de Investigación y Formación Agraria, Pesquera, Alimentaria y de la Producción Ecológica) Web portal (https://www.juntadeandalucia.es/agriculturaypesca/ifapa/ria/) were also used. The data include the full set of climate variables required for running PBDMs. Specifically, daily data for eight weather stations for a period covering 2000-2018 were used. IFAPA promotes the free use and reuse of the data available on its Portal over which it holds intellectual property rights. These data are available under a CC BY 4.0 Creative Commons Attribution 4.0 license (https://creativecommons.org/licenses/by/4.0/deed.en).

Most of the climate data processed and used in the project (observations, forecasts, and projections) only include temperature and precipitation (see D1.3 and D1.4). However, models that simulate biological processes in crop systems, require a larger set of variables that AgMERRA exemplifies as an international standard climate forcing dataset for agricultural modeling (see details here https://data.giss.nasa.gov/impacts/agmipcf/).

The standard climate scenarios produced in MED-GOLD also do not include the full set of climate variables required to run PBDMs (see D1.4) (e.g., solar radiation required for running the olive plant model is missing). As a consequence, an additional and slower climate data processing chain was implemented for PBDMs. The climate change scenario data complete with the six required variables were available in early January 2020, and will be used to produce the final release of the MED-GOLD prototype.

Seasonal climate forecast data available under the project also miss the full set of climate variables required to run PBDMs (see D1.4). Further, seasonal climate forecasts have not been previously used to run PBDMs, and hence this is an open and nontrivial research question that will be addressed as such, and not in a climate service prototyping context. For example, seasonal forecast data are generally provided and used on a monthly time scale, including in MED-GOLD, but PBDMs require daily climate data as input. This and other issues are being discussed at project level.

6.5 SIMULATION RUNS

Although many aspects of the daily age-mass structured dynamics of olive and olive fly are computed by the model, only Julian bloom dates, season yield, cumulative season long olive fly pupae, and the percent of fruit attacked are used in the PBDM analysis. Model runs of olive alone and others with the fly in the absence of pest control were carried out as a proxy for measuring the fly's potential damage. Soil moisture is assumed non-limiting given the ability of olive to resist prolonged drought [RD.50], [RD.51] but the effect of soil moisture could be implemented as needed given appropriate soil data [RD.45]. Using weather data as input, model runs were carried out via batch processing across years and locations, and the geo-referenced output data were written to an output text file for GIS processing and mapping. The same initial conditions for olive and olive fly were used at all locations, and the model was run continuously for each of the simulation periods (e.g., 1 January 1980 to 31 December 2010 for AgMERRA). The first year of simulation was used to allow the model dynamics to equilibrate to local weather, and hence the first year data was not used to compute means, coefficients of variation, and other summary statistics.

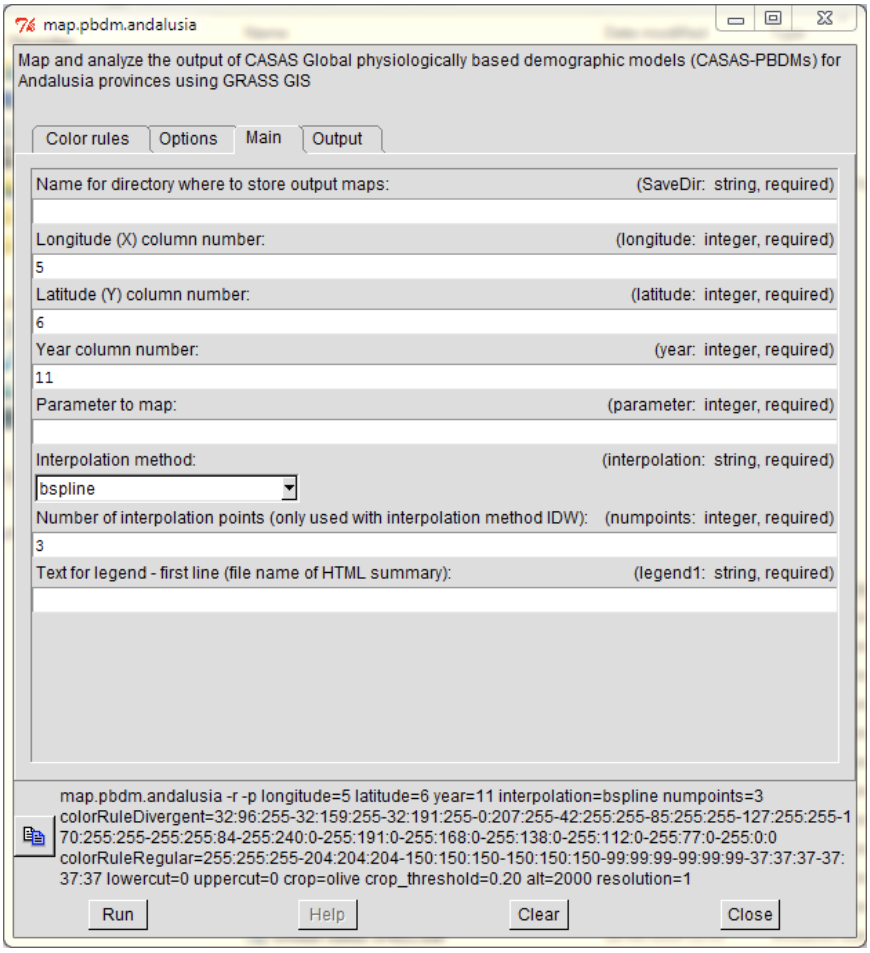
6.6 GIS AND MARGINAL ANALYSIS

All GIS datasets used in the analysis are available open access, and most of them were sourced from the public domain repository *Natural Earth* (https://www.naturalearthdata.com/). The open source GIS software GRASS (GRASS Development Team, 2015) (see http://grass.osgeo.org/) [RD.52] was used to map output data from the PBDM. Inverse distance weighting or bicubic spline interpolation was used to map PBDM output as a continuous raster surface, and hence the spatial patterns reflect not only the site specific effects of weather on the biology of the species but also the resolution and arrangement of the weather grid. The digital elevation

model used is the NOAA "Global Land One-km Base Elevation" (GLOBE) (www.ngdc.noaa.gov/mgg/topo/globe.html). Simulation data across years and locations were analyzed using linear multivariate regression implemented in R using function *lm* (R Core Team, 2020) retaining only independent variables having slopes significantly greater than zero (t-values, $p < 0.05$).

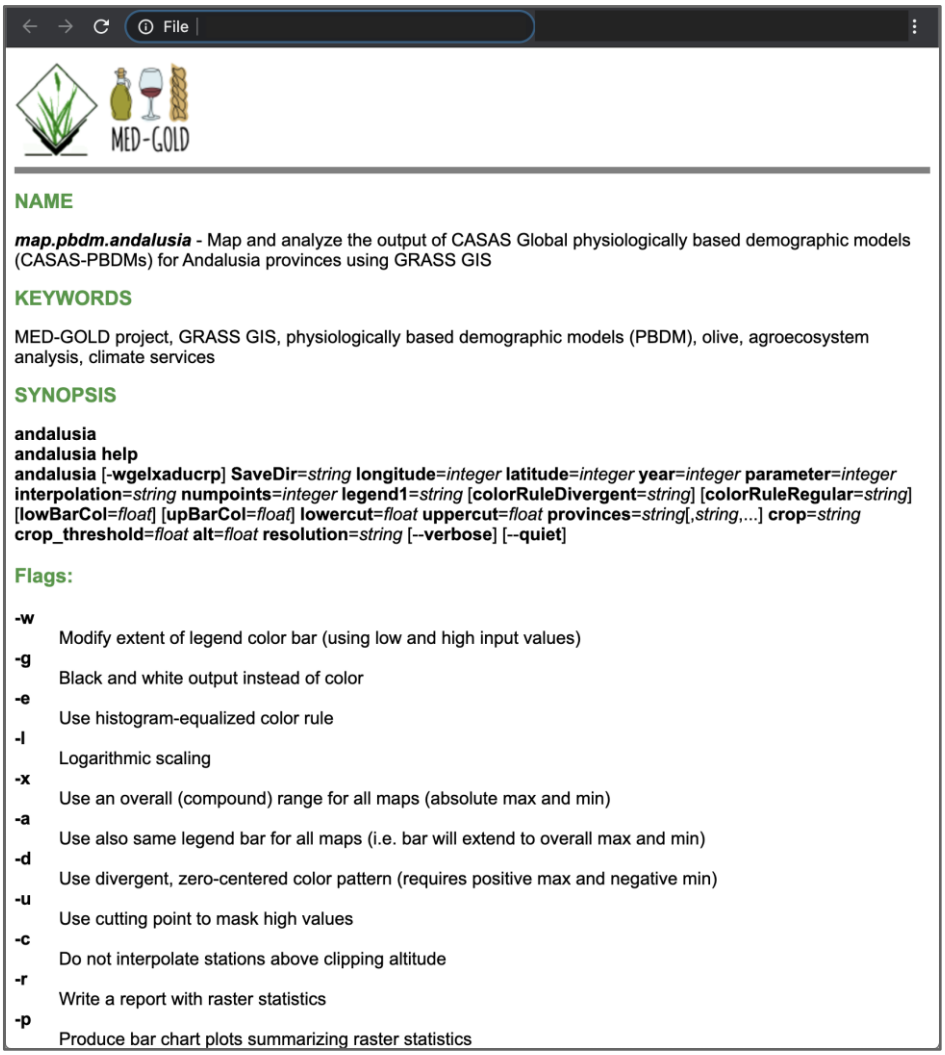
The GRASS-based GIS for mapping and analysis of PBDM model output in Andalusia was developed under MED-GOLD. It includes custom software called *map.pbdm.andalusia* written in the Bash shell scripting language that uses a wide variety of GRASS modules to map the output of PBDMs and to generate an HTML summary where maps are available, including summary statistics. Both the GRASS geographic database and script will be released as open source under the project. The script includes a graphical user interface (GUI) for easier interaction with its functionality (Fig. 6-4)

Figure 6-4: Graphical user interface of the map.pbdm.andalusia GRASS GIS script.



Documentation will also be made available and is under developing for the script (see screenshot in Fig. 6-5).

Figure 6-5: HTML documentation for the *map.pbdm.andalusia* GRASS GIS script as displayed in a web browser.



File

MED-GOLD

NAME

map.pbdm.andalusia - Map and analyze the output of CASAS Global physiologically based demographic models (CASAS-PBDMs) for Andalusia provinces using GRASS GIS

KEYWORDS

MED-GOLD project, GRASS GIS, physiologically based demographic models (PBDM), olive, agroecosystem analysis, climate services

SYNOPSIS

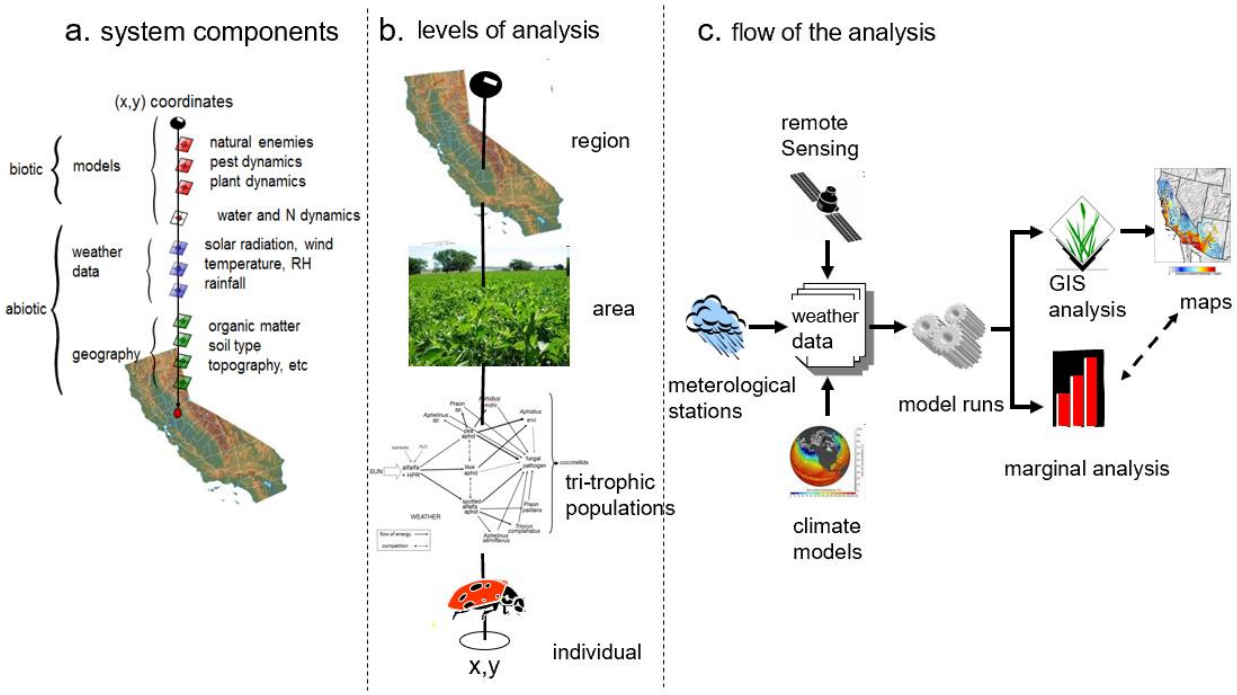
andalusia
andalusia help
andalusia [**-wgelxaducrp**] **SaveDir**=*string* **longitude**=*integer* **latitude**=*integer* **year**=*integer* **parameter**=*integer* **interpolation**=*string* **numpoints**=*integer* **legend1**=*string* [**colorRuleDivergent**=*string*] [**colorRuleRegular**=*string*] [**lowBarCol**=*float*] [**upBarCol**=*float*] **lowercut**=*float* **uppercut**=*float* **provinces**=*string*[,*string*,...] **crop**=*string* **crop_threshold**=*float* **alt**=*float* **resolution**=*string* [**--verbose**] [**--quiet**]

Flags:

-w
Modify extent of legend color bar (using low and high input values)
-g
Black and white output instead of color
-e
Use histogram-equalized color rule
-l
Logarithmic scaling
-x
Use an overall (compound) range for all maps (absolute max and min)
-a
Use also same legend bar for all maps (i.e. bar will extend to overall max and min)
-d
Use divergent, zero-centered color pattern (requires positive max and negative min)
-u
Use cutting point to mask high values
-c
Do not interpolate stations above clipping altitude
-r
Write a report with raster statistics
-p
Produce bar chart plots summarizing raster statistics

The components, levels, and flow of the PBDM analysis in a GIS context are summarized and illustrated in Figure 6-6.

Figure 6-6: Components, levels, and flow of the PBDM analysis in a GIS context as an outcome for the MED-GOLD climate service prototype (Gutierrez et al., 2010).



The PBDM analysis in a GIS context as illustrated in Figure 6-5 was integrated into the MED-GOLD ICT platform.

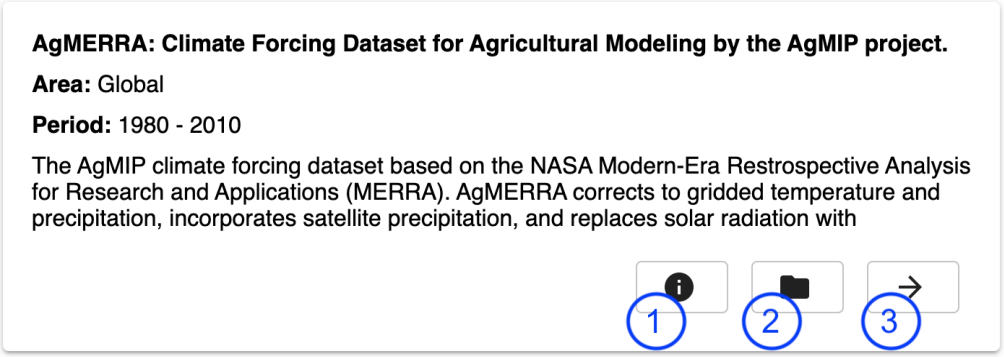
6.7 Integration into the MED-GOLD ICT platform

The idea underlying integration of the PBDM analysis into the MED-GOLD ICT platform is to show how the platform can support the operation of heritage software tools such as the Borland Pascal executables for PBDMs that only run on Windows operating systems, by connecting them to modern sources of climate data while at the same time providing the PBDM functionality as a Web service using an API, independent of which operating systems the client computer is running. This comes with the added advantage of scalability that is important when working towards the provision of climate services, as the cloud computing MED-GOLD ICT infrastructure is designed to scale efficiently with increasing computational loads and user requests.

Integration of the PBDM for olive and olive fly into the ICT platform requires development of several components. First, an adapter is required that reads and converts climate data files from their native NetCDF or GRIB binary format to the CSV text format that is used as input by PBDMs. This also involves taking care of unit conversion if needed, and of computing daily values as appropriate if the climate variables are provided on a different time step in the binary source file. Second, the PBDM is a Windows executable and hence needs to run simulations in a Windows environment using a virtual machine that is provided by the platform. Third, access to all this functionality needs to be made available in a general way using a representational state transfer (REST) API. A fourth step is to develop a software wrapper (i.e., a container) allowing PDBMs to use ICT platform functionality directly, without the need of a Windows virtual machine. Another step is the addition of a scheduler to run the PBDMs periodically unattended. Finally, PBDM output could be managed directly by the ICT platform itself, for example by mapping the CSV output files using the MED-GOLD dashboard rather than using the custom external GRASS-based GIS software (see section 6.6).

The PBDM API is accessible and was tested successfully on the MED-GOLD platform. For example, one of the datasets stored on the platform and listed on its Web interface (https://platform.med-gold.eu/) is AgMERRA, for which (see snapshot in Fig. 6-7 below), you can get information on the dataset (1), an interface to download the raw data (2), and a programmatic way to run CASAS-PBDMs (CASAS Global physiologically based demographic models) for olive and olive fly using AgMERRA as input, by assembling an appropriate application programming interface (API) call to the *pbdm* workflow (3).

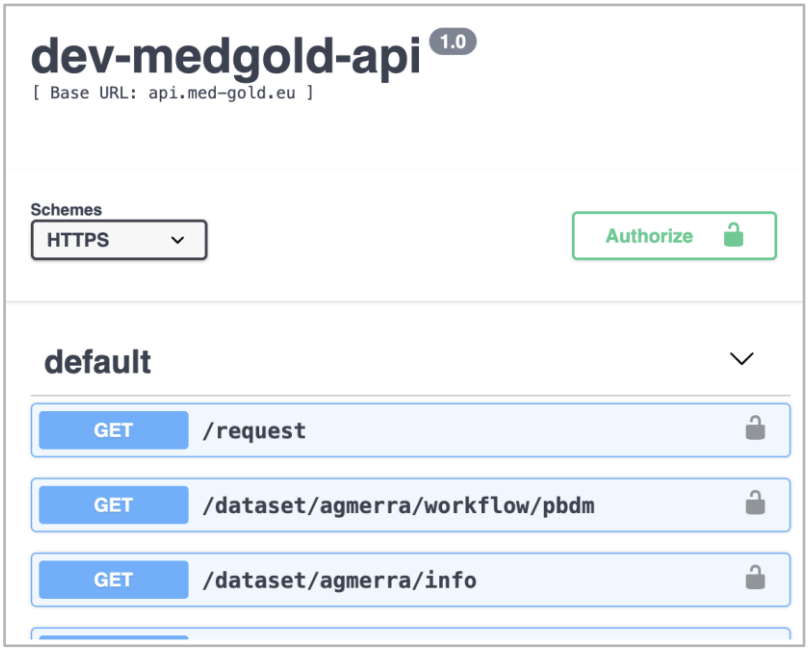
Figure 6-7: AgMERRA element of the MED-GOLD platform website (https://platform.med-gold.eu/) providing a short description of the dataset as well as buttons to access: (1) more detailed information about the dataset; (2) an interface to download the raw data; and (3) API documentation in OpenAPI format illustrating a programmatic way to run the PBDM for olive and olive fly using AgMERRA as input, by assembling an appropriate application programming interface (API) call to the pbdm workflow.



Clicking on the arrow button (3) (see Fig. 6-6) exposes the *pbdm* API functionality (see snapshot in Fig. 6-8) that can be used in WP2 for the olive/olive oil case study and that will be extended to the PBDM for coffee as part of WP6. After obtaining an authorization token, the API functionality can be accessed via the API server (https://api.med-gold.eu/).

Figure 6-8: Graphical interface of the MED-GOLD API for dataset AgMERRA on the MED-GOLD platform website (https://platform.med-gold.eu/) listing available API calls and related access to API documentation in OpenAPI

format, including the *pbdm* workflow that enables running the PBDM for olive and olive fly, for coffee under WP6, and prospectively for any PBDM.

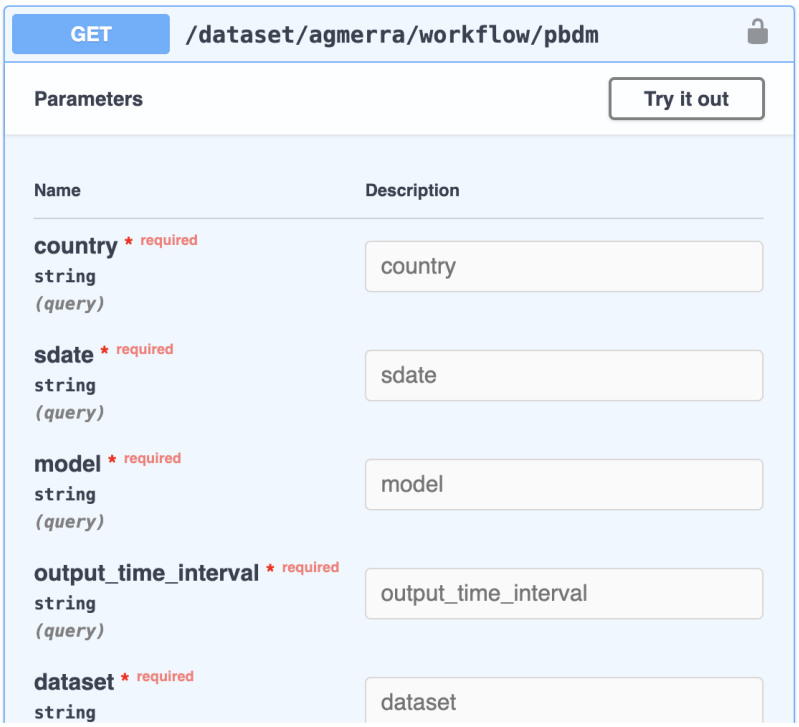


The *pbdm* API workflow, based on the AgMERRA dataset, creates text-based data files by merging source per-variable NetCDF files into columnar daily files, for given geographical locations, that are used to feed the olive model application. The workflow also takes care of uploading the output of the model to the data storage.

The documentation for the *pbdm* workflow is provided in standard OpenAPI format, is available on the web ICT platform, and includes all available endpoints (including the *pbdm* one), along with their input/output parameters and the structure of the output in JSON format. It is also possible to invoke APIs directly in the browser by clicking the "Try it out" and then the "Execute" button. An appropriate security token is injected directly by the platform. Figure 6-9 is a snapshot of the OpenAPI documentation for the *pbdm* workflow.

Figure 6-9: Graphical interface of the MED-GOLD API for dataset AgMERRA on the MED-GOLD platform website (https://platform.med-gold.eu/) showing the API documentation for the *pbdm* workflow in OpenAPI format. Note that

when the "olive" string value is assigned to the *model* parameter, the PBDM for olive and olive fly is run, while the "coffee" string will be required for running the PBDM for coffee.



APIs endpoints including the PBDM one are protected with a typical credentials/token security schema. Each user (and application) of the ICT platform is provided with a username and a password. When an user/application wants to call the *pbdm* API workflow, they must then first obtain a token (i.e., a unique identifier of an application/user requesting access to your service) via a dedicated API call:

https://api.med-gold.eu/security/token?username={username}&password={password}

The token, which expires after a set period of time, will be used as an Authorization header to all subsequent API calls:

```
curl -H "Authorization: {token}" https://api.med-gold.eu/...
```

Users/applications are responsible to request a new token when it expires.

Each API call to the *pbdm* workflow generates a request that is stored in a queue system and executed by the ICT platform in due time. The caller will get as a response a request ID which, in turn, will be passed to a dedicated API to check the status of the request. When the request is completed the response will also include an URL to the data storage containing the workflow's output.

For example, the following API call runs the olive CASAS-PBDM model for Spain and Portugal for the period 2008 to 2010 using AgMERRA climate data (line breaks are for presentation purposes only).

```
curl -H
"Authorization: SOME-VERY-LONG-API-KEY" 'https://api.med-gold.eu/
dataset/agmerra/workflow/pbdm
?country=ESP-POR
&sdate=2008/01/01
```

```
&edate=2010/12/31
&model=olive
&dataset=agmerra
&output_time_interval=365'
```

The PBDM model workflow's API implemented in the ICT platform is described below in ANNEX A.

7. DASHBOARD

MED-GOLD Dashboard is a web-based application meant as an easy to use end-to-end environment for MED-GOLD users and stakeholders. The Dashboard was meant as a horizontal tool to be utilized by all three sectors. As of now, indices particularly relevant to the olive/oil sector are not yet available on the Dashboard, and are expected to be included at a later time as a result of user interactions [RD.2]. In the meantime, historical climate data, seasonal forecasts and long term projections are available on the Dashboard for the entire Iberian peninsula and can be already used by the sector's stakeholders.

A more complete technical description of MED-GOLD Dashboard can be found in [RD.64], under section 4.6; a complete API reference can be found as the ANNEX A of the same document.

Figure 7-1: Utilization of ICT Platform's services (in orange) by PBDM REST APIs and the Dashboard.

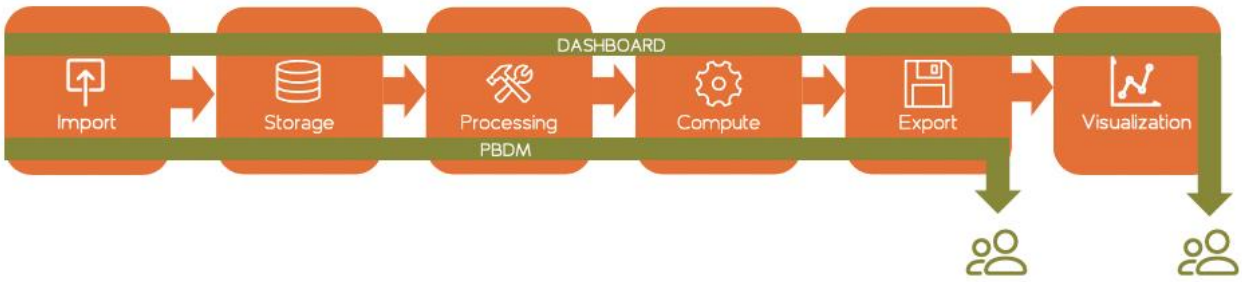
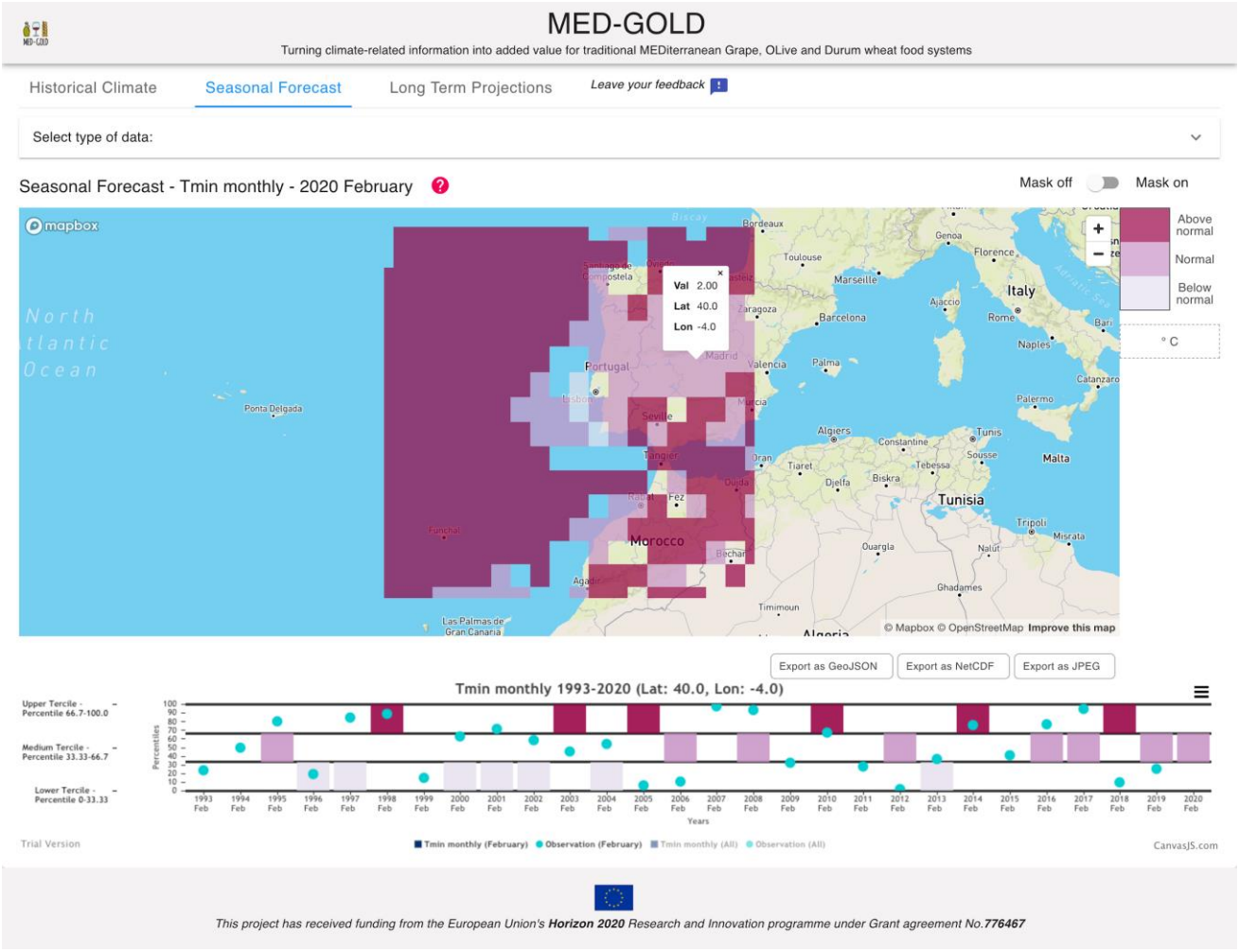


Figure 7-2: Seasonal Forecast of Tmin (monthly value) for the Iberian peninsula as displayed in the Dashboard



*This project has received funding from the European Union's **Horizon 2020** Research and Innovation programme under Grant agreement No.**776467***

8. CONCLUSIONS

This report presents the methodology followed by the scientific partners of the MED-GOLD project (EC2CE, ENEA, NOA, BSC, BEETOBIT, GMV) for the development of the climate service tools for the olive/olive oil sector. The three tools (Olivia platform, PBDM and Dashboard) were designed, developed and improved in accordance to the needs and the continuous feedback of the end-users (DCOOP) [RD.1], who are mostly interested in the Olivia platform.

However, the tools reported in this deliverable will be further improved following the results of the second end-users' feedback survey [RD.2], which took place in May 2020. Better graphs and few more indices relevant to the olive/olive oil sector are expected to be included in the Dashboard, while a more efficient identification of the farms is expected for the Olivia platform.

ANNEX A. PBDM API DOCUMENTATION

3.1 GET /dataset/agmerra/workflow/pbdm

PBDM workflow elaboration

PBDM workflow elaboration
The endpoint returns a link of the file based on information passed

REQUEST

QUERY PARAMETERS

NAME	TYPE	DESCRIPTION
*dataset	enum ALLOWED: agmerra	name of dataset
*model	enum ALLOWED: pbdm	name of model
*country	enum ALLOWED: ESP-POR	required location
*sdate	string	starting date
*edate	string	end date
*output_time_interval	enum ALLOWED: 365	time interval

RESPONSE

STATUS CODE - 200: Link of .zip file which contains OliveSummary.txt, OliveDaily.txt, GisFilesList.txt and a files for each year in interval required

STATUS CODE - 401: Unhautorized

RESPONSE MODEL - application/json

```
{
  message  string  unhautorized error message
}
```

STATUS CODE - 404: Not found

RESPONSE MODEL - application/json

```
{
  message  string  not found error message
}
```

3.2 GET /request

Horta workflow elaboration

Horta workflow elaboration
The endpoint returns a link of the file based on information passed

REQUEST

QUERY PARAMETERS

NAME	TYPE	DESCRIPTION
*id	string	id of request return by api workflow elaboration

RESPONSE

STATUS CODE - 200: returns id of elaboration

STATUS CODE - 401: Unhautorized

RESPONSE MODEL - application/json

```
{
  message  string  unhautorized error message
}
```

STATUS CODE - 404: Not found

RESPONSE MODEL - application/json

```
{
  message  string  not found error message
}
```

END OF DOCUMENT