



Topics and trends in Mountain Livestock Farming research: a text mining approach



A. Zuliani^{a,*}, B. Contiero^b, M.K. Schneider^c, G. Arsenos^d, A. Bernués^e, P. Dovc^f, M. Gauly^g, Ø. Holand^h, B. Martinⁱ, C. Morgan-Davies^j, W. Zollitsch^k, G. Cozzi^b

^a Department of Food, Agricultural, Environmental and Animal Science, University of Udine, Via Sondrio 2/A, 33100 Udine, Italy

^b Department of Animal Medicine, Production and Health, University of Padova, Viale dell'Università 2, Legnaro, 35020 Padova, Italy

^c Agroscope, Forage Production and Grassland Systems, Reckenholzstrasse 191, 8046 Zurich, Switzerland

^d Laboratory of Animal Husbandry, Department of Veterinary Medicine, School of Health Sciences, Aristotle University, 54124 Thessaloniki, Greece

^e Unidad de Producción y Sanidad Animal, Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Instituto Agroalimentario de Aragón-IA2 (CITA-Universidad de Zaragoza), Zaragoza, Spain

^f Biotechnical Faculty, University of Ljubljana, Groblje 3, 1230 Domžale, Slovenia

^g Livestock Production Systems, Faculty of Science and Technology, Free University of Bolzano, Universitätsplatz 5, 39100 Bolzano, Italy

^h Faculty of Biosciences, Norwegian University of Life Sciences, N-1434 Ås, Norway

ⁱ Université Clermont Auvergne, INRAE, VetAgro Sup, UMR Herbivores, F-63122 Saint-Genès-Champanelle, France

^j Scotland's Rural College (SRUC), South and West Faculty, Hill and Mountain Research Centre, Kirkton Farm, Criannlarich FK20 8RU, United Kingdom

^k Division of Livestock Sciences, Department of Sustainable Agricultural Systems, BOKU-University of Natural Resources and Life Sciences, Vienna, Austria

ARTICLE INFO

Article history:

Received 16 April 2020

Received in revised form 22 August 2020

Accepted 26 August 2020

Available online 18 December 2020

Keywords:

Animal production
Mountain agriculture
Research topics
Ruminants
Text mining

ABSTRACT

Pasture-based and small-scale livestock farming systems are the main source of livelihood in the mountain primary sector, ensuring socioeconomic sustainability and biodiversity in rural communities throughout Europe and beyond. Mountain livestock farming (MLF) has attracted substantial research efforts from a wide variety of scientific communities worldwide. In this study, the use of text mining and topic modelling analysis drew a detailed picture of the main research topics dealing with MLF and their trends over the last four decades. The final data corpus used for the analysis counted 2 679 documents, of which 92% were peer-reviewed scientific publications. The number of scientific outputs in MLF doubled every 10 years since 1980. Text mining found that milk, goat and sheep were the terms with the highest weighed frequency in the data corpus. Ten meaningful topics were identified by topic analysis: T1-Livestock management and vegetation dynamics; T2-Animal health and epidemiology; T3-Methodological studies on cattle; T4-Production system and sustainability; T5-Methodological studies; T6-Wildlife and conservation studies; T7-Reproduction and performance; T8-Dairy/meat production and quality; T9-Land use and its change and T10-Genetic/genomic studies. A hierarchical clustering analysis was performed to explore the interrelationships among topics, and three main clusters were identified: the first focused on sustainability, conservation and socioeconomic aspects (T4; T6 and T9), the second was related to food production and quality (T7 and T8) and the last one considered methodological studies on mountain flora and fauna (T1; T2; T3; T5 and T10). The 10 topics identified represent a useful and a starting source of information for further and more detailed analysis (e.g. systematic review) of specific research or geographical areas. A truly holistic and interdisciplinary research approach is needed to identify drivers of change and to understand current and future challenges faced by livestock farming in mountain areas.

© 2020 The Authors. Published by Elsevier Inc. on behalf of The Animal Consortium. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Implications

The present paper summarizes the main outcomes of a text mining and topic modelling analysis on the available scientific literature in the field of mountain livestock farming. By identifying the main explored

topics and their trends over the last 40 years, this work will serve as knowledge platform for future research actions and collaborations.

Introduction

Mountains occupy a quarter of the Earth's solid surface and host almost one billion people (Wymann von Dach et al., 2018). Human activities have been shaping mountain environments for millennia, leading to the transformation of natural ecosystems into agro-pastoral

* Corresponding author.

E-mail address: zuliani.anna.2@spes.uniud.it (A. Zuliani).

ecosystems (Faccioni et al., 2019). Mountain agro-pastoral ecosystems regulate water flows and quality, mitigate the consequences of natural hazards and support the delivery of material and non-material benefits such as food, biodiversity, attractive landscapes and mental well-being (Grêt-Regamey et al., 2012; Martín-López et al., 2019). It is increasingly recognized that low-input farming practices, such as pastoralism and livestock grazing, often in combination with grain and vegetable cultivation, maintain and enhance diversity of species, habitats and landscapes (O'Rourke et al., 2016). In mountain areas low temperatures, short growing seasons, limited sun exposure, topography and lack of organic soil further constrain the possibility to carry out many types of agricultural activities in an intensive and disruptive way (Briner et al., 2013). Indeed, these environments currently do not offer many profitable agronomic alternatives to forage production and the utilization of local grassland resources (Cozzi and Bizzotto, 2004).

Despite the above-mentioned constraints and the related marginal economic value of mountain primary production as compared to that in the lowlands (Wymann von Dach et al., 2013), the positive interdependence that exists between mountain agroecosystems and livestock farming has attracted research efforts from a wide variety of academic communities worldwide. However, this variety can also become a constraint when trying to truly acknowledge these research outcomes. This study aims at mapping mountain livestock farming (MLF) research by carrying out a deep analysis of topics and trends in the available scientific literature over the last 40 years, to inform future research actions and collaborations. Systematic literature reviews are useful tools for understanding the current state of an issue and for informing further studies (O'Connor and Sargeant, 2015). However, they involve the identification, appraisal and synthesis of all relevant studies investigating a defined theme and require significant resources and time. Text mining and topic modelling analysis represent suitable alternatives (Li et al., 2016) for reducing the burden associated with document screening as they produce, fully

unsupervised, a structured map of textual knowledge (Wang et al., 2016; Park and Kremer, 2017) by uncovering recurrent topics and latent themes in large sets of documents.

The present paper summarises the main outcomes of a text mining and topic modeling analysis on the available scientific literature carried out by the European Federation of Animal Production (EAAP) across commissions working group on MLF. By identifying the main explored topics and their trends, this work will update the scientific community on the developments of research on MLF in the different geographic areas of the planet, and it will serve as knowledge platform for future research actions and collaborations.

Material and methods

Identifying relevant papers

A literature search protocol was set up to identify the peer-reviewed papers dealing with mountain livestock farming using Scopus®, the abstracts and citation database of Elsevier©. The bibliographic search was developed starting from three terms (i.e. "mountain", "livestock" and "farm*"). Few additional terms were added according to the expertise of the authors in order to keep the search as broad as possible and capture the largest number of papers related to MLF (Fig. 1). The terms were searched for in article title, abstract and keywords fields. The search was performed in May 2019 and it considered all the published literature from the year 1980 to the year 2018.

Some descriptive statistics of the selected records were calculated to profile the scientific corpus, based on information retrieved from Scopus and SCImago database, which provide an overview of international research output and scientific influence, respectively. The information considered were publication year, publication source and its quartile within the scientific journal ranking in the year of publication. Geographic localisation of each record was set based on the affiliation of the corresponding author or first author.

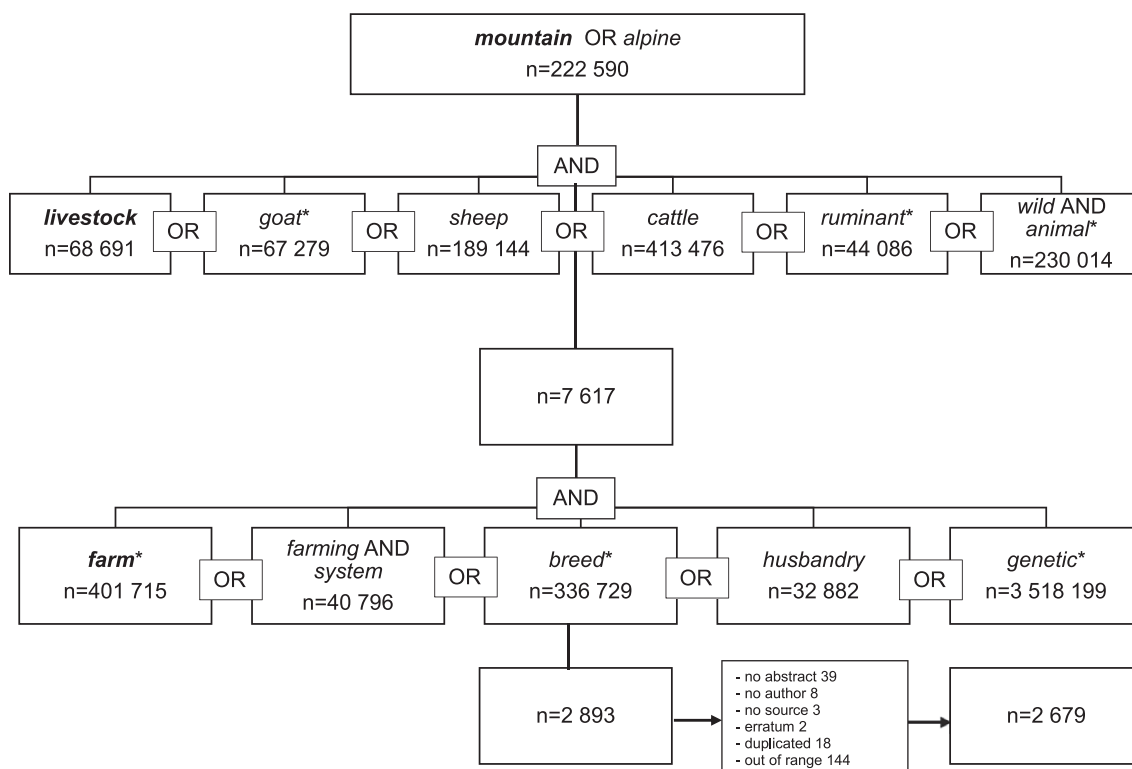


Fig. 1. Search string for document retrieval. The key words on which the string was developed (mountain, livestock and farm*) are in bold.

Table 1
The most probable 25 words defining the 10 topics emerged with Latent Dirichlet Allocation (LDA). Classification of topics according to hierarchical clustering is also shown.

Topic	Most probable words	No. of documents	Cluster
T4: Production system and sustainability	System product manag econom develop practice local agricultur land sustain tradit resourc crop ecology strategi research environment adapt improve food household social support income policy	363	1
T10: Genetic/genomic studies	Popul sheep genet gene divers sequenc genotyp allel result analysi variat studi region analys structur polymorph frequenc genom cluster microsatellite origin differenti found geography marker	351	3
T2: Animal health and epidemiology	Sampl infect studi collect test diseas detect found animal identifi sheep examin isol associ posit method risk strain preval caus presenc human report confirm virus	313	3
T1: Livestock management and vegetation dynamics	Graze pastur chang increase plant veget grassland soil climat summer forag year result area winter manag reduc impact site decreas stock meadow season grass density	303	3
T7: Reproduction and performance	Group signific season weight period rate male female time lamb increase reproduct treatment breed control effect perform bodi month year calv birth mate concentr adult	303	2
T8: Dairy/meat production and quality	Goat milk product dairi yield total higher protein lactat content feed concentr saanen compar diet chees acid composit produc qualiti average meat nubian intak cell	286	2
T9: Land use – land use change	Area region forest natur landscap activ land develop cultur human valley agriculture people villag water sourc country local tree century northern river zone import ecology	269	1
T6: Wildlife and conservation studies	Speci wild popul conserve rang habitat deer larg number domest distribut protect site ibex park wildlife success natur import bird captiv predat individu surviv nation	227	1
T5: Methodological studies	Effect data model estim studi select train size value factor correl individu number direct record genet mean paramet variabl measur analysi year influence breed method	163	3
T3: Methodological studies on cattle	Anim cattl differ studi type herd level result condit relat consid characterist present accord import wel investing organ small locat measure altitude rear lowland order	101	3

genetics/genomics. The least popular was related to methodological studies on cattle (T3), which merged 101 documents.

Considering the dendrogram produced by cluster analysis (Fig. 3), topics were sorted in three main clusters at the height (indicating the distance between clusters) of 0.98: cluster 1, grouped sustainability, conservation and socio-economic topics (i.e. T4, T6; T9); cluster 2, grouped topics related to livestock performance and food production and quality (i.e. T7 and T8) while cluster 3 grouped methodological studies on mountain flora and fauna (i.e. T1, T2, T3, T5 and T10). Cluster analysis grouped topics that had different trends throughout the study period. Clusters 1 and 3 were composed of topics that had both increasing and stable trends, while Cluster 2 encompassed only topics with steady trends.

The affiliation of the first or corresponding author was used to define the country in which the studies were carried out. Affiliates to European

research institutions were the most active on all topics, while the scientific contribution from other continents was more focused on specific topics (Fig. 4). More than 80% of publications allocated to T3-Methodological studies on cattle, were produced by European academics and only T6-Wildlife and conservation studies, saw less than 50% of the scientific production originating from European research communities. North American affiliates contributed to more than 20% of the scientific production in T2-Animal health and epidemiology, T5-Methodological studies and T6-Wildlife and conservation studies. Asian academics were more productive in T4-Production system and sustainability and T10-Genetic/genomics studies, publishing about 25% of the scientific production allocated to those topics. Affiliates to South American, African and Oceanian research institutes contributed with an average of 10% of publications across all topics.

Discussion

This paper aimed at mapping available scientific knowledge on MLF. Although this approach can be seen as methodological-orientated, such identification of the most explored research topics and their trends over the last 40 years represents a useful starting base of information to scientists for further studies. Once the main topics are identified within a large number of documents using our approach, a systematic review of the papers allocated to each topic could then be carried out for further in-depth analysis. Indeed, literature reviews are useful tools for understanding the current state of an issue and for informing further studies (O'Connor and Sargeant, 2015). When performing systematic literature reviews, scholars identify a specific research question to be answered through document reading and manual data extraction. However, when broader research questions are asked and larger document collections need to be screened, such as in this paper, text mining and topic modelling analysis, represent suitable methods for information retrieval (Wang et al., 2016; Park and Kremer, 2017).

Similar to global bibliometric trends (Bornmann and Mutz, 2015), the number of scientific outputs in MLF has doubled every 10 years since 1980. In addition to the widespread use of English for publication purposes even among non-native speakers, a reason for this sharp increase has been linked to the adoption of bibliometric indicators for the evaluation of scientific output (Fanelli and Larivière, 2016) and to the emergence of new academic communities in countries of the Global North and South, respectively. In fact, according to the Indicators Report

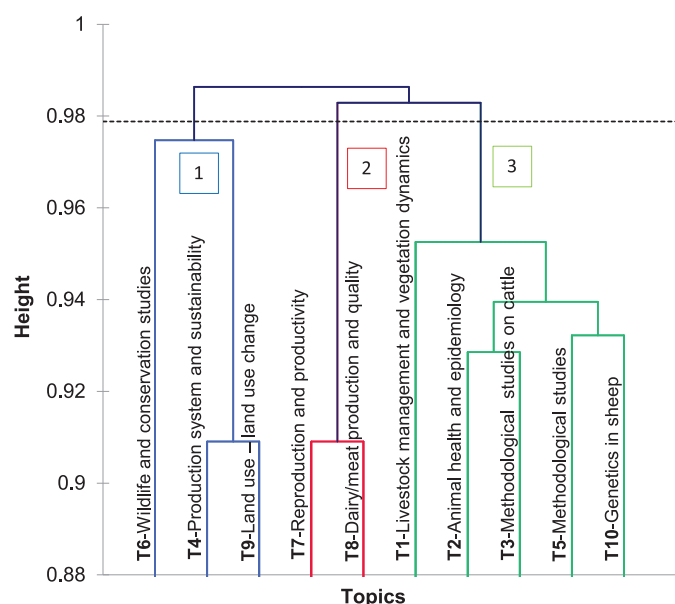


Fig. 3. Hierarchical clustering of Mountain Livestock Farming research topics (T). Branches belonging to cluster 1 are represented in blue, those belonging to cluster 2 are displayed in red while those of cluster 3 are shown in green.

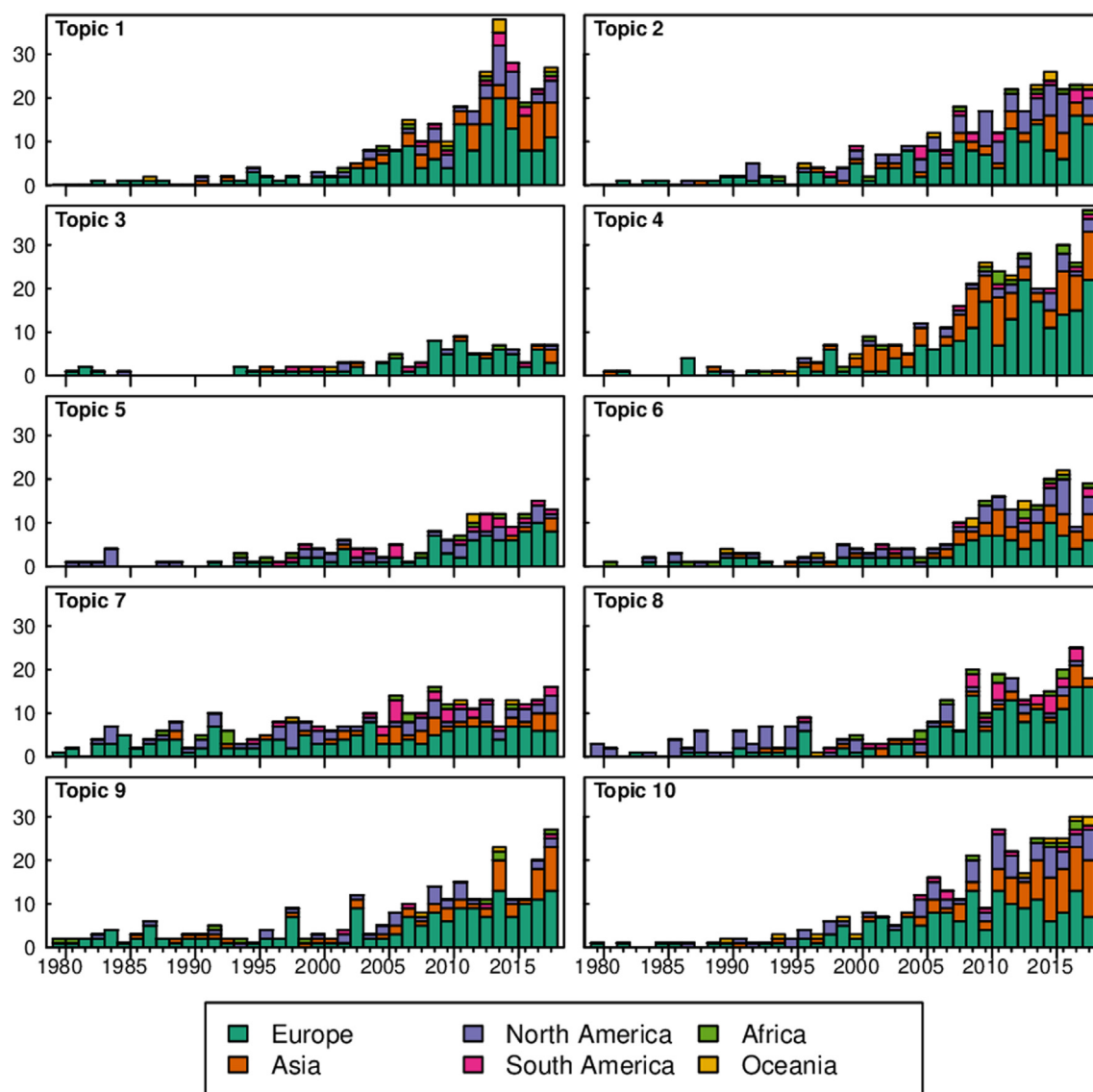


Fig. 4. Scientific production (n. of publications) by topic T (Table 1), continent and year (1980–2018).

of the National Science Foundation (National Science Foundation (NSF), 2018), in BRIC countries (i.e. Brazil, Russia, India and China), the number of scientific publications is growing at a fast pace due to the availability of research funds and to the increasing science and technology capacities. The increasing share of publications on T4 and T10 from Asian scholars is in line with this hypothesis (Yuan et al., 2018).

Pastoralism and extensive rearing of domestic herbivores are the most frequent farming systems found in harsh mountainous environments (Battaglini et al., 2014); therefore, it is not surprising that the text mining analysis identified milk, goat and sheep as terms with the highest weighed frequency in the data corpus. Additionally, the majority of the papers came from Europe, so the agricultural policies over that same period probably need to be considered, as research priorities and industry concerns would have been somehow aligned. For example, reproduction and performance (T7) and land use – land use change (T9) – attracted most of the European research interest between 1980 and 1990. Indeed, in 1980, the sheep and goat support regime within the EU Common Agricultural Policy (CAP) was implemented (Gordon et al., 1993), as well as the 1975 Less Favoured Areas directive (Brassley and Lobley, 2003) for challenged mountain areas in particular. This could explain the increasing number of studies relating to these

topics. In most recent years, a relevant number of studies have been related to emerging topics of agricultural and environmental sciences, such as production systems and sustainability (T4) and livestock management and vegetation dynamics (T1). This points at the essential role played by livestock farming systems in the management and protection of a fragile environment such as that of mountains (Grêt-Regamey et al., 2012). Agricultural and rural policies orientations as well as associated research priorities for these areas can also explain this emergence, especially at European level. Indeed, the CAP reform in 1992 (Gardner, 1996) and the Agenda 2000 reform saw the introduction of agri-environmental measures and rural development programmes support measures (Ackrill, 2000). In particular, the latest iteration of the CAP focused on protecting food quality, preserving the landscape and biodiversity and enabling vibrant rural areas (European Commission, 2018). Whilst this may not explain all the trends observed in these research topics, agricultural policies at European level would certainly have had an impact on the publications trend. For instance, in Europe, many of the animal-derived food from mountain regions are dairy products, often safeguarded and promoted by official EU or national quality schemes that protect their traditional production method and origin (Zuliani et al., 2018). New analytical technologies

(Coppa et al., 2012; Segato et al., 2019) such as those aiming at assessing product authenticity are often applied to European niche products, centering and almost limiting the scientific production on dairy/meat production and quality (T8) to Europe especially in the years that follow the launch of Agenda 2000 and food-quality schemes.

Additionally, in recent decades, the growing number of scientific articles has coincided with the establishment of more international collaborations (Bornmann and Mutz, 2015). In this study, the affiliation of the first or corresponding author was used to define the country in which studies were carried out. In the context of increased international collaborations, this assumption could have led to a geographic misrepresentation of MLF scientific output in favor of universities of the Global North, which may carry out international research projects also in countries of the Global South. Nonetheless, collaborations between academic communities of different geographical areas may have brought new perspective and tools to established research approaches and contributed to shifting the interest of MLF research from traditional topics such as reproduction and performance (T7) and methodological studies on cattle (T3) to innovative ones such as production system and sustainability (T4). Moreover, cluster analysis may suggest how different disciplines and expertise could work in an interdisciplinary manner to get a better picture of new challenges and opportunities for mountain livestock farming. For example, scientists focused on farming systems and sustainability (T4) could benefit from collaborations with experts working on land use – land use change (T9) and wildlife and conservation (T6). Similarly, traditional MLF studies (e.g. T3-Methodological studies on cattle) could use new approaches derived from genetic/genomic studies (T10) to optimize livestock robustness and resilience in a changing environment (Friggens et al., 2017). In addition, given the rich heritage of practices and traditions of mountain communities worldwide, the use of participatory approaches and the integration of traditional knowledge into scientific frameworks would further strengthen the impact and relevance of future research efforts on MLF.

Conclusion

The use of text mining and topic modelling analysis allowed to draw a unique and detailed picture of the main research topics dealing with MLF and their trends over the past 40 years. The 10 topics identified represent a starting source of information for further and more detailed analysis (e.g. systematic review) of specific research or geographical areas. Topics were grouped in three clusters: the first focused on sustainability, conservation and socioeconomic aspects, the second covered food production and quality, and the last one was related to methodological studies on mountain flora and fauna.

New collaborations between academic communities and disciplines, as well as changing policies orientations and research priorities, may have contributed to shift the interest of MLF research from traditional topics such as reproduction and performance and methodological studies on cattle to innovative ones such as production system and sustainability. Based on the knowledge platform set by this work, further integrations between research communities and disciplines may provide a sound interdisciplinary perspective for a deep understanding of current and future challenges faced by mountain livestock farming worldwide and help communicating its role for the conservation of mountain agroecosystems and the well-being of communities far beyond mountain areas.

Ethics approval

Not applicable.

Data and model availability statement

None of the data/models were deposited in an official repository but are available upon request.

Author ORCIDs

Zuliani A. ORCID 0000-0002-9193-8031.

Author contributions

Conceptualization: AZ, MKS, GA, AB, PD, MG, ØH, BM, CM-D, WZ, GC; Data curation, formal analysis, and methodology: BC; Visualization: BC, MKS; Writing – original draft: AZ, GC; Writing – review and editing: AZ, MKS, GA, AB, PD, MG, ØH, BM, CM-D, WZ, GC.

Declaration of interest

None.

Acknowledgements

None.

Financial support statement

This research received no specific grant from any funding agency, commercial, or not-for-profit section.

References

- Ackrill, R., 2000. *The Common Agricultural Policy*. Sheffield Academy Press, Sheffield, United Kingdom.
- Battagliani, L., Bovolenta, S., Gusmeroli, F., Salvador, S., Sturaro, E., 2014. Environmental sustainability of Alpine livestock farms. *Italian Journal of Animal Science* 13, 431–443.
- Bornmann, L., Mutz, R., 2015. Growth rates of modern science: a bibliometric analysis based on the number of publications and cited references. *Journal of the Association for Information Science and Technology* 66, 2215–2222.
- Brassley, P.W., Lobley, M., 2003. 20th edition the agricultural handbook. Blackwell Science Ltd, Oxford, United Kingdom.
- Briner, S., Huber, R., Bebi, P., Elkin, C., Schmatz, D.R., Grêt-Regamey, A., 2013. Trade-offs between ecosystem services in a mountain region. *Ecology and Society* 18, 35–48.
- Coppa, M., Martin, B., Agabriel, C., Chassaing, C., Sibra, C., Constant, I., Graulet, B., Andueza, D., 2012. Authentication of cow feeding and geographic origin on milk using visible and near-infrared spectroscopy. *Journal of Dairy Science* 95, 5544–5551.
- Cozzi, G., Bizzotto, M., 2004. Sustainability and environmental impact of the dairy production systems in mountain areas. *Acta Agriculturae Slovenica (suppl. 1)*, 21–28.
- European Commission, 2018. Future CAP. Retrieved on 3 August 2020, from: https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/future-cap_en#objectives.
- Faccioni, G., Sturaro, E., Ramanzin, M., Bernués, A., 2019. Socio-economic valuation of abandonment and intensification of Alpine agroecosystems and associated ecosystem services. *Land Use Policy* 81, 453–462.
- Fanelli, D., Larivière, V., 2016. Researchers' individual publication rate has not increased in a century. *PLoS ONE* 11, e0149504.
- Friggens, N.C., Blanc, F., Berry, D.P., Puillet, L., 2017. Review: deciphering animal robustness. A synthesis to facilitate its use in livestock breeding and management. *Animal* 11, 2237–2251.
- Gardner, B., 1996. *European agriculture. Policies, production and trade*. London, United Kingdom, Routledge.
- Gordon, D.V., Hobbs, J.E., Kerr, W.A., 1993. A test for price integration in the EC lamb market. *Journal of Agricultural Economics* 44, 126–134.
- Grêt-Regamey, A., Brunner, S.H., Kienast, F., 2012. Mountain ecosystem services: who cares? *Mountain Research and Development* 32 (suppl. 1), S23–S34.
- Grün, B., Hornik, K., 2011. Topicmodels: an R package for fitting topic models. *Journal of Statistical Software* 40, 1–30.
- Li, D., Wang, Z., Wang, L., Sohn, S., Shen, F., Murad, M.H., Liu, H., 2016. A text-mining framework for supporting systematic reviews. *American Journal of Information Management* 1, 1–9.
- Martín-López, B., Leister, I., Lorenzo Cruz, P., Palomo, I., Grêt-Regamey, A., Harrison, P.A., Lavorel, S., Locatelli, B., Luque, S., Walz, A., 2019. Nature's contributions to people in mountains: a review. *PLoS ONE* 14, e0217847.
- National Science Foundation (NSF), 2018. *Science and Engineering Indicators 2018*. NSF, Alexandria, VA, USA.
- O'Connor, A., Sargeant, J., 2015. Research synthesis in veterinary science: narrative reviews, systematic reviews and meta-analysis. *The Veterinary Journal* 206, 261–267.
- O'Rourke, E., Charbonneau, M., Poinot, Y., 2016. High nature value mountain farming systems in Europe: case studies from the Atlantic Pyrenees, France and the Kerry Uplands, Ireland. *Journal of Rural Studies* 46, 47–59.
- Park, K., Kremer, G.E.O., 2017. Text mining-based categorization and user perspective analysis of environmental sustainability indicators for manufacturing and service systems. *Ecological Indicators* 72, 803–820.

- R Core Team, 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing Retrieved on 18 July 2019, from. <http://www.R-project.org/>.
- Salton, G., Buckley, C., 1988. Term-weighting approaches in automatic text retrieval. *Information Processing and Management* 24, 513–523.
- Sebastiani, F., 2002. Machine learning in automated text categorization. *ACM Computing Surveys* 34, 1–47.
- Segato, S., Caligiani, A., Contiero, B., Galaverna, G., Bisutti, V., Cozzi, G., 2019. ¹H NMR metabolic profile to discriminate pasture based alpine Asiago PDO cheeses. *Animals* 9, 722.
- Wang, S.H., Ding, Y., Zhao, W., Huang, Y.H., Perkins, R., Zou, W., Chen, J.J., 2016. Text mining for identifying topics in the literatures about adolescent substance use and depression. *BMC Public Health* 16, 279–287.
- Wymann von Dach, S., Romeo, R., Vita, A., Wurzinger, M., Kohler, T., 2013. Mountain farming is family farming: a contribution from mountain areas to the International Year of Family Farming 2014. Food and Agriculture Organization of the United Nations, Rome, Italy.
- Wymann von Dach, S., Bracher, C., Peralvo, M., Perez, K., Adler, C., 2018. Leaving no one in mountains behind: localizing the SDGs for resilience of mountain people and ecosystems. Issue brief on sustainable mountain development. Centre for Development and Environment and Mountain Research Initiative. Bern Open Publisher, Bern, Switzerland.
- Yuan, L., Hao, Y., Li, M., Bao, C., Li, J., Wu, D., 2018. Who are the international research collaboration partners for China? A novel data perspective based on NSFC grants. *Scientometrics* 116, 401–422.
- Zuliani, A., Esbjerg, L., Grunert, K.G., Bovolenta, S., 2018. Animal welfare and mountain products from traditional dairy farms: how do consumers perceive complexity? *Animals* 8, 207.