A spatialised information system to support decisions regarding grazing management in mountainous and Mediterranean rangelands

Urcel Kalenga, Bassira Djibo, Lucile Sautot, Maguelonne Teisseire, Magali Jouven

A spatialised information system to support decisions regarding grazing management in mountainous and Mediterranean rangelands

Urcel KALENGA[a,b], Bassira DJIBO[b], Lucile SAUTOT[a], Maguelonne TEISSEIRE[a], Magali JOUVEN[b]

[a] *Univ Montpellier, AgroParisTech, CIRAD, CNRS, INRAE, TETIS, 500, rue Jean-François Breton, Montpellier, 34093, France*
[b] *Univ Montpellier, Institut Agro / Montpellier SupAgro, CIRAD, INRAE, SELMET, 2 Place Pierre Viala, Montpellier, 34060, France*

Abstract

Agro-sylvo-pastoral systems are common around the Mediterranean Basin, where they provide a variety of goods and services to the local populations. Their sustainability relies on efficient grazing management, especially in Mediterranean rangelands. The diversity of pastoral resources, combined with the variety of grazing management techniques and farming objectives, has slowed down the development of digital tools to assist grazing management in these conditions. However, digital technologies could serve agro-sylvo-pastoral farms by improving the efficiency of grazing management and reducing the difficulty of the associated work. In this objective, and to take into account the variety of situations, we suggest developing an information system based on a variety of (contextualised) complementary data. For

southern France, the following data can be combined: Sentinel 2 images, Registre Parcellaire Graphique (RPG), OSO land use and cover, RGE Alti altimetry data, pastoral technical references, herd GPS location and feedback from farmers. However, designing and implementing such an information system requires overcoming methodological and technical limitations concerning the integration of heterogeneous structured and unstructured data and the definition of meaningful ways to combine them to produce relevant insights for decision-making. In this article, we describe an approach to produce a spatialised information system aimed at providing farmers with relevant information to support decisions regarding grazing management in mountainous and Mediterranean rangelands. The information system was codesigned with stakeholders, including farmers, to best match their needs and to facilitate its integration into farm management. The various stakeholders were involved in choosing the types of data to be associated and, defining the functions and the conceptual model of the information system. We propose to structure the information system as a spatial data lake, designed to integrate and associate the identified heterogeneous data, and to produce decision-making insights for grazing management in mountainous and Mediterranean rangelands.

Keywords: Agro-sylvo-pastoral system, Heterogeneous data, Spatial data lake, Decision-making

1. Introduction and motivation

Agro-sylvo-pastoral systems are widely represented in Mediterranean areas of southern Europe (Nori, 2019). To feed their herds, these systems rely on the diverse and variable natural forage resources of Mediterranean rangelands, woodlands and mountain pastures. These diverse and variable resources enable agro-sylvo-pastoral systems to play a fundamental role in the socioeconomic development of rural populations (Bernués et al., 2014) and the preservation of natural resources (Hadjigeorgiou et al., 2005).

The forage resources of rangelands are highly variable between seasons and years in terms of quantity and quality due to the effect of climatic factors (Dumont et al., 2015; Dono et al., 2016). In addition to this temporal variability, the wide range of altitudes, the diversity of geological substrates and geomorphology favour the great diversity of pastoral resources. This variability forces farmers to continually adapt grazing management. Short-term adaptations may include changes in the daily grazing circuit (for shepherding), in the timing and distribution of paddock grazing, temporary reduction of animal performance or supplementation. Long-term adaptations may include mobility to new grazing areas, changes in the seasonal use of pastoral resources, water development, fence development or shrub clearing. In addition, farmers face other constraints, such as the concurrent use of space with other activities (e.g., hunting, tourism) and the presence of wildlife, which can significantly impact the behaviour and production of animals due to predation (Vincent, 2010; Meuret et al., 2018).

Under these conditions, grazing management is at the heart of several issues, especially economic viability (di Virgilio et al., 2018), adaptation to climate change and variability (Dono et al., 2016), maintenance and sustainability of natural environments (Bailey et al., 2019), adaptation to the European Common Agricultural Policy, and the fight against predation (Meuret et al., 2017; Linnell and Cretois, 2018). To graze while remaining compliant with these challenges, a complete understanding of the grazed ecosystem and the actors involved in its management is needed.

The grazing ecosystem can be modelled through the use of various sources of information available from the public sector and at the farm level. High or very high public satellite images such as Sentinel 2 and SPOT 6/7 images associated with various reference data can be used to model the grazing environment (Shaqura and Lasseur, 2019; Castro et al., 2020); while in regard to farm herds, GPS location can be used to understand the use of rangelands by herds (Akasbi et al., 2012; Feldt and Schlecht, 2016; Handcock et al., 2009). However, this information taken individually is often of low precision for decision-making (Bahlo et al., 2019). Their combination can help to produce new indicators (Capalbo et al., 2017), which can allow farmers to adapt grazing management to the local context in the short and long term.

Initiatives have been carried out in different projects (e.g., E-Pasto[1] and Clochète[2]) to integrate digital technology (GPS and accelerometers) into grazing management in the mountains and Mediterranean areas of France. Experiments on different farms have also been conducted for different applications: (i) to locate and track animal routes in free ranging (Buerkert and Schlecht, 2009; Akasbi et al., 2012); (ii) to assess the spatiotemporal differences in grazing patterns (Feldt and Schlecht, 2016); (iii) to monitor animal behaviour and environmental interactions (Handcock et al., 2009); and (iv) to quantify the time spent by animals while grazing in different areas (Rutter et al., 1997). The results of these experiments show that the analysis of these kinds of data can produce relevant information for grazing management.

[1] https://www.epasto.fr/

[2] https://idele.fr/detail-article/projet-clochete-caracterisation-du%2Dcomportement-et-localisation-des-ovins-et-caprins-grace-a-des%2Dtechnologies-embarquees

Currently, embedded GPS devices are commonly marketed for hunting dogs and for cattle. Naturally, an increasing number of farmers are equipped with at least one device to locate and track the movements of their herds. Other sources of information can be used to provide more insight, such as digital terrain models and meteorological and climatic data. Moreover, the adoption of smartphones by farmers (Michels et al., 2020), is an opportunity for the provision of additional information (Pongnumkul et al., 2015; Mendes et al., 2020), which can bring more precision regarding the grazing environment and herd behaviour (e.g., forage quantity and quality, microclimate, animal physiology and feeding).

These public and farm data mentioned above are heterogeneous in terms of their nature (spatial, nonspatial, structured, or unstructured), format and resolution. Thus, the challenge remains to find ways to meaningfully combine these heterogeneous data and resolve questions relating to data integration, analytics, ownership and quality (Bahlo et al., 2019). This challenge cannot easily be managed by traditional Information Systems such as relational databases and data warehouses designed specifically for structured data. Data lakes currently being adopted, now represent a solution to handle these kinds of data (Ravat and Zhao, 2019; Suriarachchi and Plale, 2016;

Khine and Wang, 2018; Quix and Hai, 2019). We assume that the same approaches can also be applied in the context of agro-sylvo-pastoral livestock management to integrate heterogeneous farming data and provide users with relevant indicators for decision-making.

In this paper, we propose an approach to implement a spatial data lake adapted for the integration and analysis of heterogeneous data to produce new indicators for decision support in the context of agro-sylvo-pastoral farming. To do this, we codesigned the information system in the context of the P@stor-All project involving 7 farms, 2 experimental sites, 2 research units and 1 technical institute in French Mediterranean areas.

2. Characteristics of agro-sylvo-pastoral systems in French Mediterranean areas

2.1. Rangelands and agro-sylvo-pastoral farming in French Mediterranean areas

French Mediterranean areas represented by the Occitania, PACA and Corsica regions have large and dense rangeland areas. The estimations made by (Agreste, 2015) in 2014 on the basis of analyses of land use and cover show the importance of rangelands in these areas. In 2018, the spatial distribution of rangelands resulting from the RGP[3] showed their preponderance in these three French Mediterranean regions (Nozieres et al., 2021). The availability and spatial distribution of these rangelands promote the existence of agro-sylvo-pastoral farming, characterised by the persistence of traditional management practices in these regions (Bernués et al., 2011).

Agro-sylvo-pastoral farming systems in French Mediterranean areas vary depending on two principal factors: the type of production and the land resources and management. The farms mainly breed sheep, with goats additionally and, to a lesser extent, cattle and horses. The major production is sheep meat and goat and sheep milk (Bataille et al., 2016).Wool, cheese and cattle meat are produced to a lesser extent. Equine production is intended for slaughter or leisure. Regarding land resources and management, the importance of rangeland resources in feeding herds, herd mobility and grazing management are the factors of variability. Rangeland resources used by herds vary considerably depending on each type of farm. Farming can also be sedentary, characterised by the movement of animals close to the farm, or mobile. In the latter case, it implies a seasonal shift to a greater or lesser distance (Nozieres et al., 2021). Grazing management also varies,

[3] https://geoservices.ign.fr/rpg

from rotational or continuous grazing in fenced paddocks to shepherding and free grazing. Different combinations of these various factors thus create a diversity of agro-sylvo-pastoral systems and land use. This diversity of agro-sylvo-pastoral farming systems is also found more widely in southern Europe (de Rancourt et al., 2006).

The diversity of agro-sylvo-pastoral systems raises a number of issues, particularly for farmers and research and extension services. The functioning of agro-sylvo-pastoral systems is different from conventional systems, and their recognition remains difficult. In addition, the difficulty of accurately characterising the areas used by the herds can make it difficult for farmers to access certain subsidies, including the European CAP. However, it is difficult to study and produce knowledge on this varied form of livestock farming. There are many particular cases of grazing that require studies on a large number of farms to produce generic results. Furthermore, this diversity creates a need to collect a great deal of information in the field to consolidate the tools that make it possible to document and diagnose the use of rangelands by herds.

2.2. A favourable context for the emergence of an information system to manage rangeland utilisation

Several initiatives have been carried out to improve pastoral practices in French Mediterranean areas, ranging from technical aspects to the testing of technological tools to improve rangeland use.

In France, various services are involved in the technical support of farms, including the development of references and tools to assist grazing management: the chambres d'agriculture [4], technical institutes, associations of actors interested in grazing and pastoralism, and pastoral services. Since the pastoral law of 1972 (Charbonier, 1972), various pastoral support structures have over the years produced technical references adapted to rangelands in the form of manuals and more recently in digital format. These references provided by the different services can be mobilised and fed into an information system (IS) to improve pastoral use.

The increasing availability of free satellite data with better spatial, spectral and temporal resolution, such as Sentinel 2 data (10 m in the visible and near infrared and 20 m and 60 m spatial resolution; 13 spectral bands and a 5-day revisit period), and associated processing algorithms, favours the

[4] https://chambres-agriculture.fr/

characterisation of land use at large territorial scales (Inglada et al., 2017). Other types of spatial reference data can be associated with it, such as the Registre Parcellaire Graphique (RPG), the Forestry Database (BD-Forêt[5]) and the Altimetric Database(BD-Alti[6]) produced by the the National Institute of Geographic and Forestry Information (IGN) to better characterise rangelands. Current work (Inglada et al., 2017) was able to capitalise on the complementarity of these data to map the land use and cover of all French metropolitan and overseas territories with better accuracy (Kappa of 0.88 for 2019). Taken individually, these data produce less accurate information for improved decision-making for farmers. However, combining them in the best possible way increases the level of precision of the new information produced to facilitate grazing management (Bahlo et al., 2019).

Over the last decade, there has been a growing interest among pastoral farmers for digital tools, especially for geolocation. Some farmers use hiking tools (hiking GPS or smartphone GPS) for their own use. Others actively participate in seminars and discussions on digital tools for breeding and grazing. In general, their interest stems from a desire to lighten the drudgery of

[5] https://geoservices.ign.fr/bdforet
[6] https://geoservices.ign.fr/bdalti

their work of guarding, monitoring and documenting the activities of the herd. However, they are concerned about data ownership issues and do not want the digital transition to lead to increasing demands for justification of the spatial distribution of grazing herds. In addition, farmers are seeking exchanges with their peers (Berrier and Girard, 2021), hence the development of networks for sharing experiences. There are also experimental pastoral farms, which are increasingly equipped with digital tools to study grazing. They constitute a potential source of information to analyse more precisely the use of spaces by herds.

Through these opportunities, diverse and heterogeneous available data from a variety of sources could be gathered and combined to improve the understanding of the grazed ecosystem in pastoral conditions, both locally (for farmers) and generally (for research).

3. What are the functionalities for an IS in regard to rangeland utilisation?

The characteristics of the agro-sylvo-pastoral system described earlier suggest that specific and continuous information is needed for farmers to best adapt to the many constraints that they have to face for grazing management. The documentation of a variety of pastoral systems would also enable research to draw more general knowledge regarding rangeland utilisation in Mediterranean environments, depending on the environmental and management conditions.

3.1. Associating farmers to the design of the IS and the identification of its functionalities

To best respond to the different needs concerning grazing management, the P@stor-All project involving 7 farmers and 2 experimental farms (Table 1), researchers and a technical institute was established. The aim was to bring together the stakeholders in a process of codesigning the IS by alternating moments of individual exchange with moments of collective discussion. The quota of farmers participating in this process was limited to a small number of highly motivated individuals, whose different farms represent the diversity according to the characteristics described in Section 2.

The codesign process to define the functionalities of the IS included 3 steps (Figure 1): (i) individual semidirective interviews with farmers to determine their needs in the form of questions, (ii) collective discussions on farmer needs together with the stakeholders, and (iii) final validation by farmers and selection of relevant questions to address by the IS.

Table 1: Characteristics of the farming systems surveyed involved in the P@stor-All projetct. Le Merle and La Fage are experimental farms of the project.

Farms	Production	Nb. mothers	Grazed area	Complementation
1	Meat sheep and goat	300	Rangelands + vineyard	Any
2	Dairy goat	40	Rangelands	Crau hay
3	Meat sheep	140	Rangelands + vineyard	Any
4	Meat sheep	280	Rangelands + grassland	Any
5	Dairy sheep	550	Rangelands + culture	Hay + grain
6	Meat sheep and goat	123	Rangelands + grassland	Hay + barley
7	Dairy goat	-	Rangelands + grassland	Any
Le Merle	Meat sheep	1670	Rangelands + grassland	Any
La Fage	Meat sheep	300	Rangelands	Forage + concentrate

First, individual farmer questions were collected exhaustively (51 questions, step 1), and then the relevant questions, those related to the use of rangelands by herds of herbivores, were selected by all project stakeholders (38 questions, step 2) during an exchange session. At this stage, the exchanges also made it possible to group the questions into 3 themes according to the type of use to which they relate: (i) benchmarks for adjusting pasture management, ii) methods to achieve a given objective and iii) better understanding the ecology of the agro-sylvo-pastoral ecosystem. Finally, a ranking was made in order to assign a priority score to each question. Three scores were defined: i) priority question, ii) interesting question, and iii) not frequent question. The questions retained were those with the highest number of votes (8 to 9 by stakeholders) as "priority question". 9 questions have been retained (see table 2) to constitute the list of farmers' needs. The table A.1 in Appendix A shows the list of 38 questions with their priority ratings.

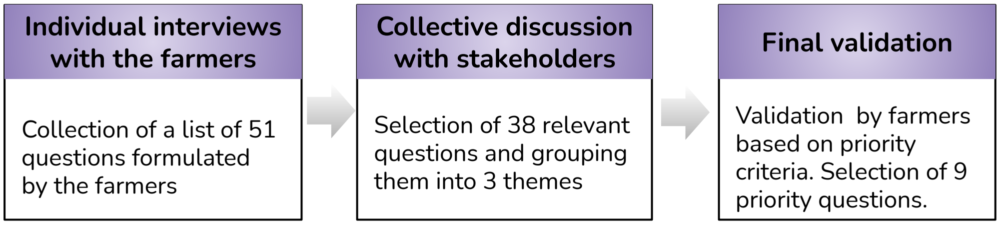

Figure 1: The three stages of defining the priority needs of farmers.

3.2. What are the functionalities of the P@stor-All IS?

The three stages of defining the needs of the farmers allowed the selection of the nine priority questions that constitute the functionalities of the IS. Table 2 presents the list of these questions grouped into themes. The majority of the functionalities of the IS concerns questions related to benchmarks for adjusting pasture management (5 questions), followed by those related to methods to achieve a given objective (3 questions), then by a functionality related to the understanding of the ecology of the agro-sylvo-pastoral system, which includes 1 question.

The IS is able to incorporate a range of information that can be used to provide farmers with new information to meet their needs formulated as a question in Table 2. However, the possibility of answering the questions depends on both the availability of adequate data and the appropriate processing methods to produce new information.

Table 2: Farmers' needs expressed in the form of questions, considered as a basis for the IS functionalities.

Questions
Function 1: benchmarks for adjusting pasture management
What is the grazing area used at the moment?
What route does my herd take each day?
How much time does my herd spend on grazing?
What is the area covered by the herd?
How do my animals occupy the space? (weekly, monthly, seasonally, or yearly)
Function 2: methods to achieve a given objective
How can encroached areas be enhanced?
How can pastoral resources be developed to meet the needs of my herd?
How can animals be stimulated to move into unexplored areas?
Function 3: better understanding the ecology of the agro-sylvo-pastoral ecosystem
How does the behaviour of the animals change according to the season?

4. Representing the grazed ecosystem: main factors determining the spatial utilisation of rangelands

To address the needs of farmers through the analysis of heterogeneous pastoral data, it is first necessary to have a model of the grazing ecosystem and to understand the use of space by animals. This will then contribute to the structuring and definition of the cross-analyses of the information system.

The combination of the different factors of the grazing ecosystem determines the grazing behaviour of the herds. This behaviour refers to a set of spatiotemporal activities that animals perform or develop to ensure their feeding during grazing (Dumont et al., 2001; Ginane et al., 2008; Zampaligré and Schlecht, 2018); and includes the time devoted to searching for, selecting, prehending and consuming forage, usually on a daily basis (Allen et al., 2011). Thus, movement, space occupation, food choices and feed intake of animals may vary according to the effects of these factors. Knowledge of these factors will allow a better understanding of the use of space by grazing animals and thus allow us to model the entire agro-sylvo-pastoral ecosystem.

4.1. Codesigning the grazing ecosystem with farmers

The codesign of the grazed ecosystem was carried out through a process of identifying key factors of the feeding behaviour of the herds. Three stages were carried out for this purpose (Figure 2): i) first, an analysis of the scientific literature made it possible to globally identify the factors of grazing behaviour, ii) then a validation followed by a complementation of these factors was carried out through individual surveys of farmers, and iii) a final analysis made it possible to compile the list of factors specifically influencing the behaviour of herds while grazing.

The literature used corresponds to different conditions from the context of our study (different environments, sometimes different types of animals, and under experimental conditions) because of the limited availability of scientific studies in the mountainous and Mediterranean areas of southern Europe on the topic of grazing behaviour. We integrated the vision of the farmers involved in the project to move closer to the context of our study (both ge-

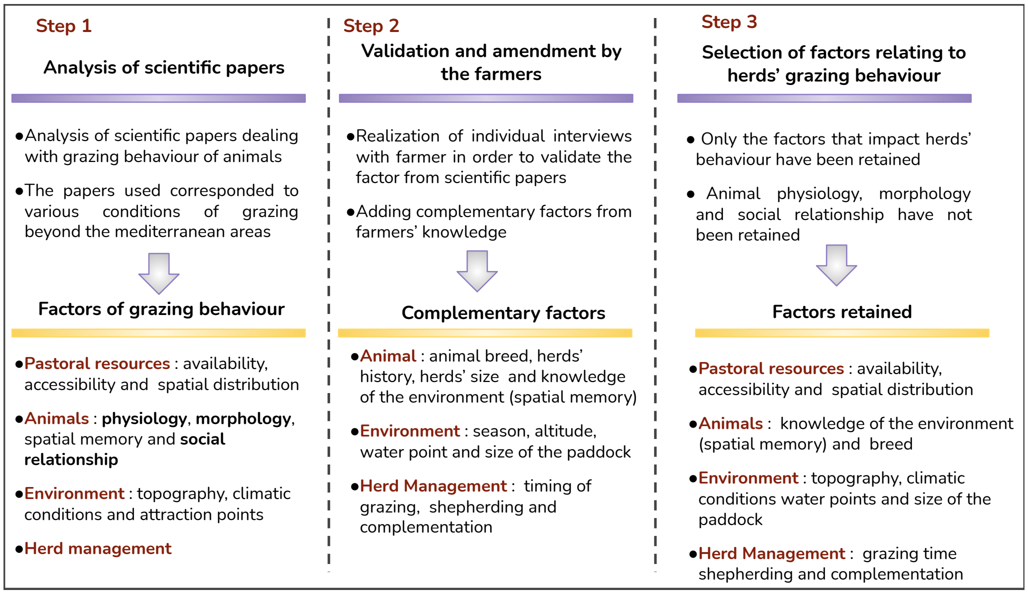

Figure 2: Different steps conducted to identify the key factors influencing the grazing behaviour of herds on pasture. Factors in bold were not retained at the end of the process.

ographic and in terms of uses). These processes have allowed us to validate, complete and establish a list of factors that enhances the scientific knowledge addressed while resituating it in a context (in terms of vocabulary and representation) that is customary for farmers in the European Mediterranean environment.

4.2. The agro-sylvo-pastoral ecosystem model used in the P@astor-All information system

The analysis of the scientific literature enabled the identification of various factors grouped into 4 categories: those concerning pastoral resources, those concerning the animals of the herd, those concerning the grazing environment and the grazing management method (Figure 2). The interviews with farmers showed that certain factors identified in the literature were not sufficiently consistent with their vision of the system and the parameters used to manage it. Their integration in the codesign process enabled the scientific factors to be validated and completed in the context of livestock farming in European Mediterranean areas.

Most factors in the literature retained for the conceptual model of the agro-sylvo-pastoral ecosystem were selected based on the number of times they were validated by farmers. The factors resulting from this process are shown in Table A.2 in the Appendix. The factors validated by all farmers were pastoral resource abundance, quality and diversity, presence of the farmer and heat. These factors correspond to pastoral resource availability, management practices and weather conditions in scientific knowledge. Some factors concerning herds (herd size, knowledge of the environment, animal breed, and animal species) were mainly derived from the point of view of the farmers. Although the farmers did not validate the impact of accessibility, we also maintained this factor because shrub encroachment is a characteristic of rangelands in mountainous and Mediterranean environments, and its impact on the feeding and spatial behaviour of herds is well known.

The unification of validated scientific factors and complementary factors coming from the farmers made it possible to constitute the final list of key factors of the feeding and spatial behaviour of the herds on pasture in an adapted context in mountainous and Mediterranean rangeland conditions. These factors were then used to model the agro-sylvo-pastoral ecosystem and to represent their different links without addressing their interactions.

We present in Figure 3 the conceptual model of the agro-sylvo-pastoral ecosystem under Mediterranean rangeland conditions. The scale of our study concerns the herd. Thus, only the factors that act directly on the grazing and spatial behaviour of herds were taken into account. Those that had effects on the individual (animal) were not retained. This was the case for animal morphology, physiology and social relationships.

The factors presented in the unified model can be modelled thanks to available data from various sources. The factors concerning pastoral resources and the environment can be represented using various available spatial data (satellite and GIS data) (Shaqura and Lasseur, 2019; Castro et al., 2020). Herd location data can also be used to represent animal factors (Buerkert and Schlecht, 2009; Handcock et al., 2009; Akasbi et al., 2012). Last, information

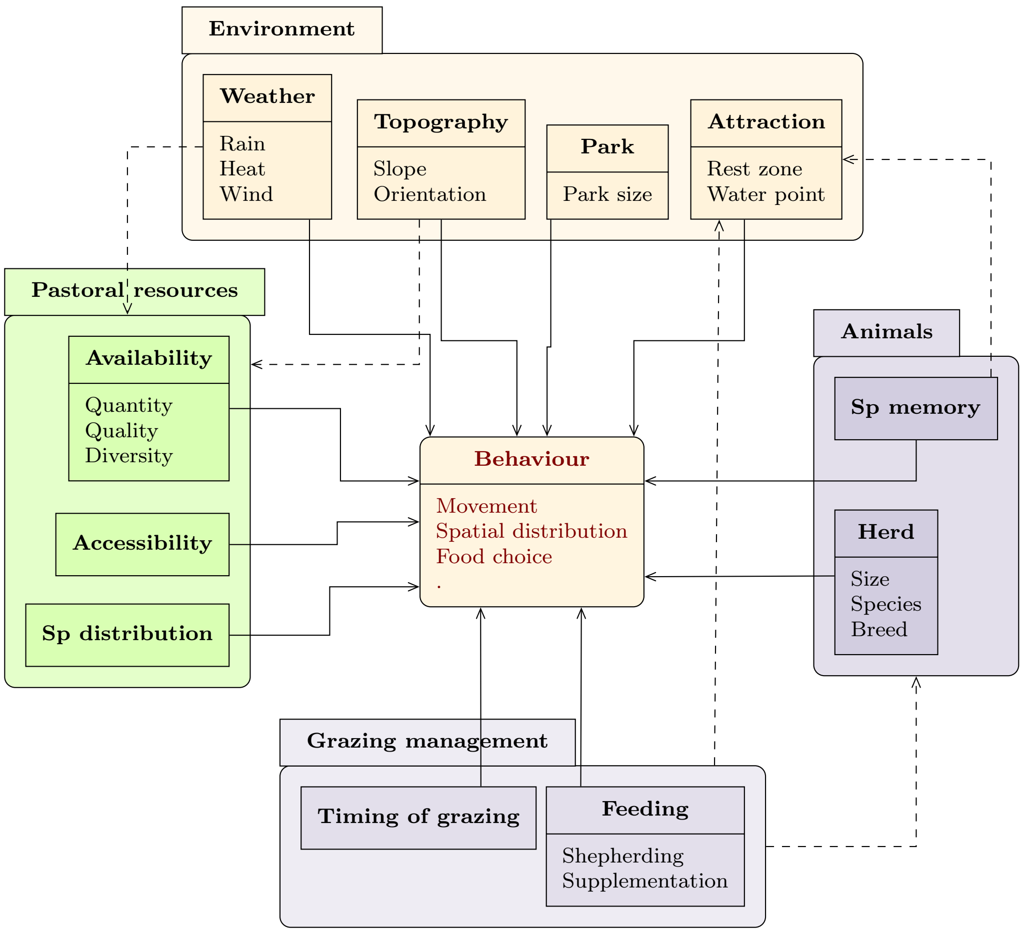

Figure 3: The conceptual model of the feeding and spatial behaviour of domestic herbivores on Mediterranean rangelands. the solid line represents the influence of the factor on herd behaviour and the dashed line represents the link between the factors.

provided by farmers at the level of agro-sylvo-pastoral farms allows us to understand the factors related to grazing management. The combination of this information allows us to achieve more insight to better understand the use of rangelands by herds to facilitate grazing management.

In the following section, we identify these data and present their relevance to provide insight for agro-sylvo-pastoral systems.

5. Selecting data to document the model of the grazed ecosystem

Various sources of data can be used to characterise the factors of feeding and spatial behaviour of the herds. To achieve the objectives of this study, we rely on the availability of acceptable accurate data individually or in combination with other data. The unavailability of data or the lack of accuracy of some of the available data can be overcome by the complementary information coming from the feedback of the farmers. However, with the evolution of increasingly accurate sensors and open data, the possibility of obtaining more accurate data reduces the implications to the farmers for the provision of complementary information.

The various data identified can be categorised into two types according to their sources: (i) external data, which are data from sources other than the farm site, and (ii) farm data that can be collected at the level of each agro-sylvo-pastoral farm.

5.1. External source data

The external data concern the Sentinel 2 satellite data, Registre Parcellaire Graphique (RPG), OSO land use and cover, IGN altimetry data, and technical references on grazing. They are mostly provided by public sources. Except for the technical references, all external data are provided under Etalab 2.0 Open Licence[7]. This licence grants a nonexclusive and free right to reuse information for commercial purposes or not for an unlimited period. The characteristics of the external data are presented in the following table 3. For interested readers, further information about these data are provided in the appendix B.

5.2. On farm data

On-farm data mainly concern herd location data using GPS and feedback from farmers. They are an important part of the data source of this study because they characterise the scale of analysis of our work, which is considered as the management unit, corresponding to the area grazed by the herds

[7] https://www.etalab.gouv.fr/licence-ouverte-open-licence

Table 3: Summary of characteristics of the external data that can be used to model rangelands utilization by herds

External data	Caracteristics
Sentinel 2	Description : Satellite image in TIFF format Source : Theia Availability : 5 days for Level 1C and 2A, 1 month for Level 3A Spectral domain : 13 bands of 10m, 20 m and 60 m of spatial resolution Wide swath : 290 km Contribution : Land use and cover
RPG	Description : database of agricultural parcels declared by farmers, provided in SHP or GeoPackage format Source : French Service and Payment Agency Availability : yearly Contribution : Identification of partoral classes for mapping rangelands
OSO land use and cover	Description : Land use and cover with 23 classes and a minimum mapping unit of 20 m Source : CNES Availability : yearly Contribution : description and mapping of rangeland
MNT RGE ALTI	Description : Relief model in the form of a regular rectangular grid Source : IGN Spatial resolution : 1 m to 5 m in X and Y Availabity : on update Contribution : Modelling of the terrain relief
Technical references	Description : documents describing technical practices in pastoral farming Source : technical institutes Availability : at each release Contribution : grazing management

during a given period.

5.2.1. Herd GPS location

GPS data from herds are collected in different ways by different actors. Some farmers, experimental farms, and researchers have started using GPS to analyse the behaviour of their herds. These data are in different forms depending on the variety of GPS equipment used and can be accessed in different ways. They can be used to analyse the movement of herds and their use of rangeland resources. Several studies have shown the potential of GPS data to model the grazing behaviour of herds. The grazing pattern (Buerkert and Schlecht, 2009; Akasbi et al., 2012; Feldt and Schlecht, 2016; McGavin et al., 2018), herds behaviours and interaction with grazing areas (Putfarken et al., 2008; Handcock et al., 2009) and grazing time (Rutter et al., 1997), are types of applications that can be realised from herd GPS location data. Different analytical methods can be applied to this information to produce new valuable information to model the use of pastoral areas by herds in combination with other sources of information.

5.2.2. Feedback from farmers

The level of precision provided by the data mentioned above may have limitations in describing agro-sylvo-pastoral systems in more detail. For instance, the 10 m spatial resolution of the Sentinel data does not allow us to identify the grazing resources available at the farm level. To provide more details on agro-sylvo-pastoral systems, feedback from farmers can be of great interest. In addition to the details of the information, the feedback from farmers also provides information that the other available data cannot provide regarding the grazing management and behaviour of the animals observed in the field.

The acquisition of farmer data can be achieved via collaborative smartphone applications to collect information on pastoral resources, soil, climate, animal intake, or any other information concerning grazing management. There exist several applications for making these acquisitions in an adapted way, allowing their analysis in an IS.

5.3. Connecting data and the modelling of the grazed ecosystem

The available data presented can allow us to identify basic analyses that can be carried out to characterise each factor of grazing behaviour. In Table 4, we present the different factors of the grazing behaviour of the herds and the available data identified that can be used to characterise them. The characterisation of the factors can be possible through a combination of different types of information of the data. The retained components of herd behaviour can also be modelled through a different combination at a second analytical level. This can allow us to produce additional information regarding the interactions between animals, pastoral resources, the environment and grazing management.

Not all data allow characterisation of the factors of animal grazing behaviour (see empty boxes in the table). However, other cross-references between the different types of information are likely to be identified a posteriori to improve the characterisation of certain factors. Farmers can also contribute much more precise information to complement other data and improve the characterisation of factors.

5.4. Issues related to the variability and heterogeneity of the identified data

The data presented are very varied in terms of their sources and their nature. First, the data presented come from a variety of public and farm sources. Depending on the source, the data may differ in terms of their nature or the level of accuracy they possess. Herd GPS location data may come from different sources and present variability in terms of the level of accuracy

Table 4: The factors of feeding and spatial behaviour of herds, their components and the available data allowing to characterise them.

Factors	**External/Public data**				**On farm data**	
	Sentinel 2	**RPG**	**OSO**	**MNT RGE**	**Herd location**	**Feedback**
Pastoral resources						
Availability	X	X	X			X
Accessibility	X	X	X	X		X
Space distribution	X	X	X			
Environment						
Meteorology						X
Topography				X		
Paddock						X
Attraction point						X
Animal/Herds						
Sp. memory					X	
Herd characteristics						X
Grazing management						
Timing of grazing					X	X
Feeding						X
Components of feeding behaviour						
Movement				X	X	
Space occupation	X	X	X	X	X	
Food choice	X	X	X	X	X	X

and frequency of acquisition. The same is true for all other types of data. Second, according to their nature, we categorise two types of data: spatial and nonspatial. Within these two types of data, the structure (structured and unstructured), format (ASCII Grid, PDF, SHP, TIFF, TXT, or HTML) and resolution can be variable. We distinguish from this list: (i) the structured data that mainly concern herd location (GPX or CSV formats), RPG (SHP formats), OSO land use and cover (SHP format) and feedback from farmers (form formats), and (ii) the unstructured data corresponding to satellite data (TIFF format) and technical references (PDF format). Other considerations can be made for the weight of these data, which also varies according to their format. The Sentinel 2 images, RPG, OSO land cover and MNT RGE ALTI are heavy by nature. They need sufficient storage to ensure their integration for analysis. These data may also have different levels of accuracy. This is the case for the Sentinel 2 data and MNT RGE ALTI, which have 10 m (visible and near-infrared) and 5 m spatial resolutions, respectively. Spatial vector data (OSO and RPG) and on-farm data can also vary in terms of accuracy.

This great variability of the identified data in terms of their sources, format, structure, weight and resolutions constitutes a challenge for the implementation of the information system intended to integrate them. The information system intended to integrate these types of data must take into account all these heterogeneous constraints. In addition to this challenge in regard to the heterogeneity of the data, the IS must be able to manage the increasing volume of data and propose adequate processing methods for extracting and crossing information from these data. Another important consideration for the information system is the fact that it needs to meet the principles of findability, accessibility, interoperability, and reusability (FAIR) of the data project.

6. The P@stor-All information system

6.1. Objectives and uses of the information system

The main goal of the IS is to provide both farmers (private and experimental farms) and researchers with a framework that allows them to gather their data and at the same time guarantees their access as well as the results of their combined analyses. Thus, this IS targets two categories of use: agro-sylvo-pastoral farming and research. On the one hand, farmers and experimental farms, which are the direct beneficiaries of this information system, contribute by providing the necessary data they collect on their respective farms. These are mainly herd location data and feedback. When stored and analysed together with external data, the resulting information contributes to the management of grazing at the level of individual farms. On the other hand, access to the data by researchers enables them to carry out different types of analysis on a wider scale (territorial or regional) than the agro-sylvo-pastoral farms, outside the IS or internally, depending on the processing possibilities offered. These cross-sectional analyses associated with different pastoral situations contribute to the development of scientific knowledge on rangeland and grazing management.

6.2. Conceptual data model

Different criteria were taken into account for the design of the conceptual data model of the IS: the functionalities of the IS defined with the stakeholders, availability of data and associated methods that can be used to produce new information, and consideration of the technical constraints of the IS. The functionalities of the IS are described in Table 2, and the data available for modelling the grazed ecosystem are presented in Section 5. As for the technical constraints:

- The IS must support the integration and combination of the heterogeneous identified data.
- Data standardisation must be guaranteed for better collaborative management of stakeholders.
- The information system must be scalable to manage the increase in data volume and the application of new functionalities.

The figure 4 presents the conceptual data model we propose. The different components of the agro-sylvo-pastoral system are described by the different associations between the classes corresponding to the data, as presented in Table 4. Information about the different factors can be obtained as a result of analysis applied to the interacting data in the IS. Different levels of analysis can be realised to produce new insight for decision-making. Other information resulting from the analysis between the herds and grazing area can be historicised to follow its dynamics in space and time. This new information will then contribute to other types of analysis involving the interaction between herd behaviour, pastoral resources, the environment and management practices.

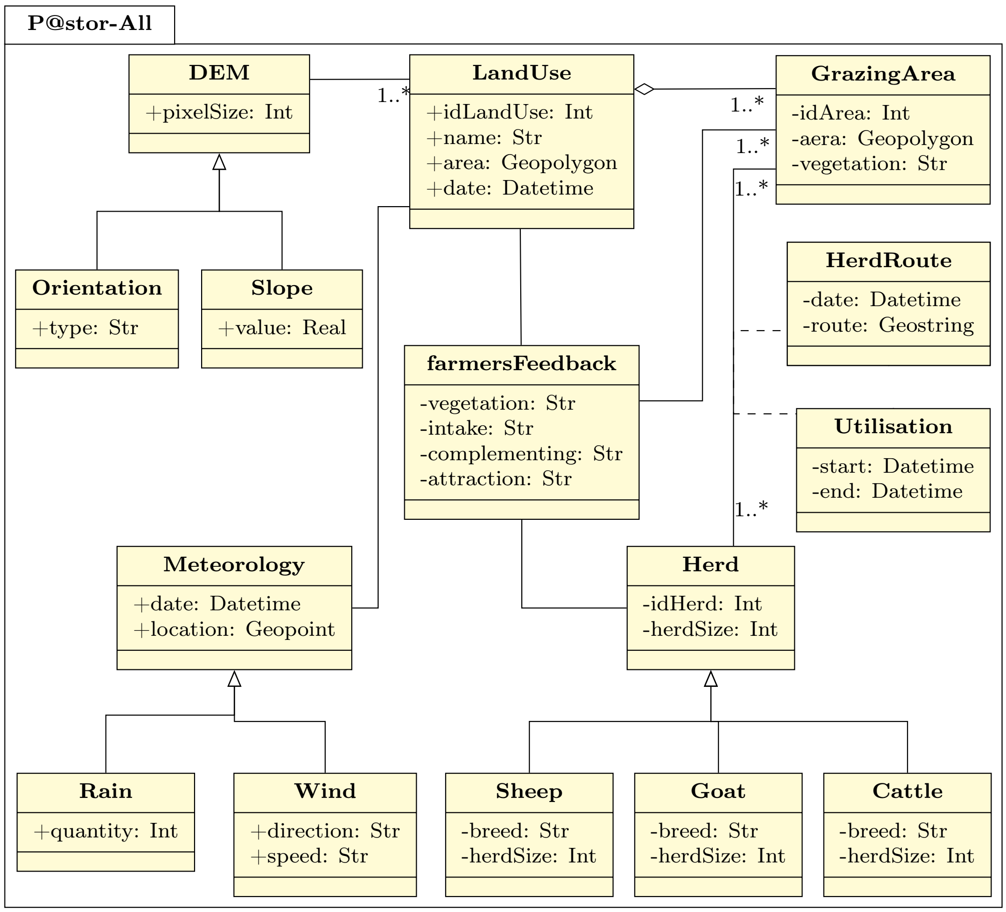

Figure 4: Conceptual data model of the spatial Information System for agro-sylvo-pastoral system in Mountain and Mediterranean European regions. This model describes the different relationships between the data allowing to model the use of rangelands by herds. A given environment characterised by one or more land uses (LandUse) contains one or more grazing areas (GrazingArea), visited by one or more groups of animals (Herd). These first two entities are characterised by the influence of two spatial phenomena that impact their use by herds: the relief (DEM) and the weather (Meteorology). The use in time and space of grazing areas (GrazingArea) by herds (Herd) allows the production of new spatio-temporal information: grazing trajectories (HerdRoute) and the use of the pastoral resources of a grazing area (Utilisation), which can help to understand how animals vary the use of rangelands. Finally, feedback from farmers (FarmersFeedback) will be provided on both the environment and areas used, as well as on animal management.

6.3. Information system infrastructure

The IS infrastructure must allow for the integration of the heterogeneous identified data. These types of data cannot be easily managed in a data warehouse ideally designed to manage structured data. We opted for the data lake to enable the integration of heterogeneous data. Data lakes allow the ingesting, storing and processing of heterogeneous raw data from various sources; provide access to different users for various analyses; and govern data to ensure data quality, data security and data life-cycles (Ravat and Zhao, 2019). These advantages presented by the data lake allow both the management of the identified pastoral data and offer several possibilities of analysis and a better accessibility of the information by the different users.

The data lake implementation we opted for is Apache Hadoop solution based based on two essential components: HDFS (Hadoop Distributed File Stystem) and MapReduce. HDFS (Honnutagi, 2014) ensures the distributed and parallelised storage of data when they are ingested in the data lake. The structure of the HDFS offers several advantages, including (i) the easy ingestion and retrieval of data from their metadata and (ii) the parallelisation of storage and processing, which brings more performance for data security and significantly improves the processing cost. Apart from the storage system, the Hadoop HDFS is flexible in regard to connecting with other processing engines. This allows the creation of an ecosystem that provides a full range of tools for data storage, analysis and exploration for grazing management. These various technical advantages offered by Hadoop HDFS-based data lakes better meet the constraints imposed by the IS in the present study.

7. Architecture and data management in the P@stor-All information system

In a data lake, access to the data is essential to ensure better management. This is accomplished through queries at the metadata level. Apache Hadoop does not have all the functionality needed to capture metadata for the wide variety of data that exists. Faced with the variety of data we identified for agro-sylvo-pastoral farming purposes, it is necessary to define suitable means to extract metadata in an objective manner. To overcome this constraint related to metadata extraction, various studies have proposed complementary solutions (Hai et al., 2016; Quix and Hai, 2019; Sawadogo et al., 2019). However, these studies did not propose an implementation capable of handling the variety of heterogeneous data (especially spatial) we identified. The implementation made by Kafando et al. (2020), fulfils these functionalities.

The metadata management system implemented in the data lake is based on GeoNetwork[8] which uses the ISO 19115 conceptual model implementation for spatial data. Geonetwork allows storage of the metadata, together with the HDFS path of the data lake, and facilitates data retrieval. This implementation is suitable for the integration of agro-sylvo-pastoral farming data in the spatial data lake.

7.1. The P@stor-All information system architecture

The data lake architecture we propose for pastoralism data integration and analysis has 5 main components: data collection, data preprocessing, data ingestion and storage, data processing and analysis, and visualisation. Figure 5 presents the overall architecture of the spatial data lake for the integration of heterogeneous pastoralism data and the cross-analysis of information for the production of new insight for decision-making.

The different processes in these components are further presented in the following points.

[8] https://geonetwork-opensource.org/

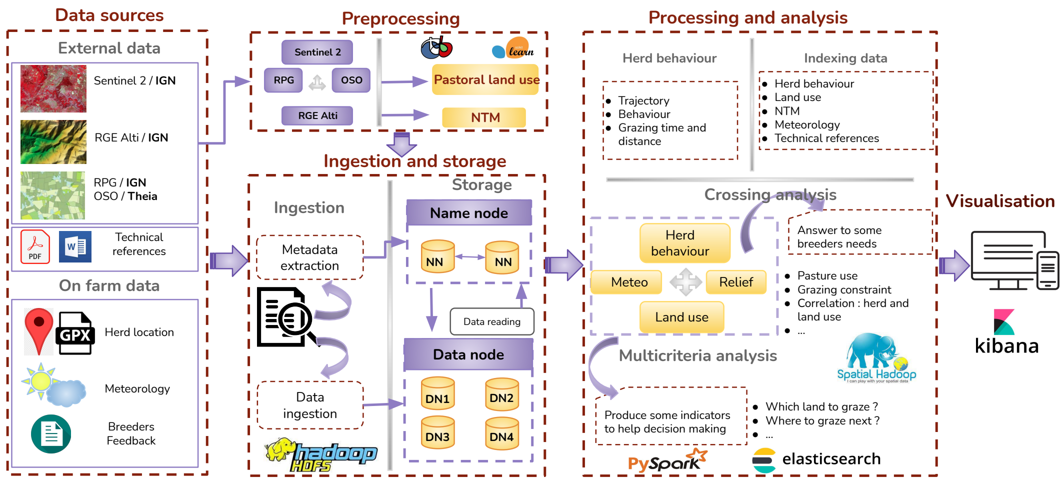

Figure 5: General architecture of the data lake for the integration of heterogeneous data from pastoralism in Mediterranean areas

7.1.1. Data source

We identified several data sources in Section 5. The frequency of acquiring external data is different depending on the dynamics of the type of information to be drawn. Considering that the relief is not affected by the change, the MNT RGE ALTI is acquired only once without any need for updating. Otherwise, to describe the land use that is susceptible to change, the data should be collected regularly. This means that Sentinel 2 data should be available at least monthly to cover large areas, and RPG and OSO land use and cover should be updated annually. Technical references can be acquired only once, with the possibility of being completed in case of new availability. On-farm data are acquired continuously and transmitted at well-defined periods for their integration into the IS. Herd GPS location data are acquired at different time intervals depending on whether they are from a private farm or an experimental area. Indeed, given their interest in research, experimental farms may be able to plan more frequent location point acquisitions than private farms. This difference in acquiring frequency is thus able to provide some precision in the description of herd behaviour. For feedback from farmers, a wide variety of information can also be collected from applications.

7.1.2. Data preprocessing

The goal of the processing carried out in this component concerns the production of land use and cover of pastoral interest from Sentinel 2 data, the RPG and the OSO land use and cover. These preprocessing steps are not a part of the various analyses within the data lake. The data are processed externally, and then the land use and cover produced are injected into the data lake thereafter.

The principle of this preprocessing is to improve the low accuracy of the pastoral interest classes present in the OSO land use and cover. From the Sentinel 2 data, the objects corresponding to the pastoral classes in the OSO land use and cover are reclassified using the RPG as a reference information source for training and validation of the classification machine learning model. We thus manage to produce improved OSO land use and cover, offering better precision for the classes of pastoral interest, which are permanent meadows, lawns, moorlands and grazed forests.

Preprocessing also concerns the DTMs in producing different relief variables, namely, slope and aspect. It is also possible to consider, if necessary, preprocessing for other types of data to facilitate their integration into the data lake. However, the aim remains to use the raw data produced by their sources.

Following training and validation of the classification algorithm, the processing chain is used for the annual update of land use and cover.

7.1.3. Data ingestion and storage

In this phase, we define two levels of operations: ingestion and storage of the data. The process is based on the data lake implementation by Kafando et al. (2020). First, the data undergo the ingestion process, where their preliminary analysis is performed to differentiate them (e.g., compared with their format) and extract their metadata. The metadata make it possible to identify and locate the data to which they refer from the GeoNetwork search engine. In the storage process, the metadata are stored in the Namenode, and the data are stored in a distributed manner in Datanodes thanks to HDFS. As input data into the data lake, we list the improved land use and cover from preprocessing, meteorological data, technical references, herd location data and farmer feedback.

7.1.4. Data processing and analysing

This component allows the IS to perform different kinds of queries and analysis to produce decision-making support indicators. Once the data are stored, the retrieval becomes easier through the links defined between data and metadata due to GeoNetwork. Thus, the data can be used according to different needs. Various analysis will be possible within this information system.

In the following table 2 we describe some of the analyses that can be carried out in relation to the needs of farmers. To meet these needs, several analysis techniques can be applied. For the first category of questions, "benchmarks for adjusting grazing behaviour", the following functionalities can be included: spatial queries (allowing to locate the places visited by the herds as well as some areas of interest), trajectory analysis (allowing to know the routes travelled and the distances made by the herds), point pattern analysis (allowing to know the distribution of the location points in a grazed environment), point clustering (allowing to discover the different phases of the herds' activities), detection of the patterns of moving objects (that can help to discover the differences or similarities in the exploration of the space by the herds). The second category of needs, "methods to achieve a certain goal", will mainly require text mining methods. They will help to discover new knowledge from farmers' experiences and technical references related to grazing. Finally, the third category of needs will require medium- and long-term analyses to assess variations in the use of rangelands by herds at different periods of the year. The combinations of the different information produced for the first category of needs will be carried out to produce new knowledge.

Beyond the needs of farmers, researchers are led to conduct various studies on a wider scale. Where the IS provides analytical results for farmer decision-making, research finds information to base new hypotheses and to verify them using other types of analysis. This new research knowledge can then be fed into the IS.

In addition to these analyses corresponding to the needs expressed in table 2, multicriteria analyses could be carried out with the aim of providing new indicators to facilitate decision-making for grazing. They provide a set of solutions or choices that farmers are led to select according to their grazing objectives. These solutions or choices can concern the identification of nearby pastoral areas, the avoidance of recently grazed pastoral areas, and the identification of optimal grazing routes (taking into account the topography of the environment).

7.1.5. Data visualisation

The information produced within the data lake is accessible to users via different visualisation interfaces depending on different access levels. Three levels of access are established: one dedicated to agro-sylvo-pastoral farms, one to researchers, and one to all types of users. The interface specific for agro-sylvo-pastoral farms allows each farmer or experimental site to access the data describing his or her farm. All information related to Table 2 in the data lake is used to characterise each individual farm. No farmer has access to data specific to a farm other than his or her own. The interface for researchers gives them access to information and tools that enable them to carry out large-scale analyses of livestock systems. They are able to combine the different types of information from the IS to produce new knowledge about pastoralism. Finally, the interface, which is dedicated to all users, concerns farmers, researchers and other users interested in the application.

Table 5: An overview of some of the analyses within the IS that can help to address farmers' needs

Farmers' needs	Kinds of analysis to integrate in the IS
Function 1: benchmarks for adjusting pasture management	
What is the grazing area used at the moment (1)? What route does my herd take each day (2)?	Data : land use and cover and daily GPS location data of animals Analysis : - Spatial queries allowing to identify in which land use and cover animals are located (1) - Transformation of GPS data in trajectories with trajectory data mining functions (2) - Application of spatial intersection functions between trajectories and land use to produce a summary of the areas explored each day (2)
How much time does my herd spend for grazing (3)? What is the area covered by the herd (4)? How do my animals occupy the space (5)?	Data : daily GPS location data of animals, meteo, land use, DEM Analysis : - Point pattern analysis: Spielman (2017) - Identification of approximate bouding area of grazing (4) - Analysis of the distribution of location points in the surveyed area : clustering or dispersion ? (5) - Trajectory clustering of Moving object (animal locations) analysis allowing to find out how differences in the use of rangelands (5): Chen et al. (2014) - Discovery of moving pattern : animals who move together for a certain timestamp (5): Zheng (2015) Cross analysis : - Correlation between point pattern and slope, land use and cover and, meteo parameters to discover the effect of factors on the use of rangelands (5)
Function 2: methods to achieve a given objective	
How can encroached areas be enhanced (6)? How can pastoral resources be developed to meet the needs of my herd (7)? How can animals be stimulated to move into unexplored areas (8)?	Data : technical references and feedbacks from farmers Analysis : Keyword search of indexed data (6, 7 and 8) - Application of text mining techniques to extract knowledge on pastoral management (6, 7, and 8)
Function 3: better understand the ecology of the agro-sylvo-pastoral ecosystem	
How does the behaviour of animals change according to the season (9)?	Data : GPS location data, land use and cover, DEM and meteo Analysis : - Application of the analytical methods described above to data collected over long periods of time in order to analyse daily and seasonal variations in animal behaviour and the factors that determine them (9)

It will provide access to all types of information except those collected at the individual farm level (GPS data or herd data). People who want to start agro-sylvo-pastoral farming can be able to find the necessary information on how to conduct it from this interface. It can also be used as a didactic reference to support the teaching of grazing practices.

8. Discussion

8.1. Relevance to addressing agro-sylvo-pastoral farming system issues

The integration of massive data in an information system for agro-sylvo-pastoral management is not currently as effective as in most agricultural applications. Data from different sources (public and farm levels) are often only partially used to contribute to livestock management. To take maximum advantage of the wide variety of data, their combination is necessary (Capalbo et al., 2017; Bahlo et al., 2019). However, the wide variety of agro-sylvo-pastoral systems in terms of environments, management modes and data make it difficult to implement a decision support tool. In this study, we proposed a spatial data lake for the integration of heterogeneous data to produce indicators and useful knowledge for short- and long-term decision-making in agro-sylvo-pastoral farming.

The originality of this study lies in considering the pastoral context of the southern Mediterranean regions to define IS functionalities; identifying and contextualising the key factors for modelling the agro-sylvo-pastoral ecosystem; designing an adapted and operational IS for the integration, cross analysis of data, and retrieval of information for decision support; and setting up a collaborative framework for farmers and between farmers and researchers.

Thanks to the cataloguing system offered by GeoNetWork, the IS data we present benefit from a standardisation that facilitates their access by different users. The latter can use it to find useful information according to their specific needs. They will find answers to the questions mentioned above in Table 2, and much more to discover new indicators to optimise the use of pastures. Indeed, the added value of combining data lies in the methods used to produce new information (Bahlo et al., 2019). According to (Teucher et al., 2014), the information system must have functionalities that provide precise information on the location of phenomena and recommendations on the best agricultural practices. The different levels of processing that we propose will allow users to become aware of the state of the farms and to provide them with proposals and references according to the needs and situation of each farmer.

In the medium and long term, the increase in example cases in the IS will open a way for new applications to enrich the knowledge regarding pastoral management. Apart from analyses based on short periods of time, research will find uses at this stage by widening its temporal scale of study. Thus, various applications can be considered, for example, the influence of grazing on the spatial distribution of vegetation, influence of climatic factors on the dynamics of vegetation, and medium- and long-term changes in animal behaviour. The spatial IS can also be applied at the territorial level. In (Capalbo et al., 2017), the authors showed that ISs can also be beneficial for management at levels above the farm level. This could allow us to analyse the state and dynamics of vegetation and to plan interventions at the territorial level.

8.2. Functional considerations

The proposed spatial data lake meets most of the constraints mentioned, including data heterogeneity and standardisation, correspondence with user needs, and scaling. Beyond these constraints, the IS presents other advantages facilitating the management of metadata, as presented in (Sawadogo et al., 2019)and summarised by (Kafando et al., 2020): semantic enrichment, data indexing, link generation and conservation, data polymorphism, data versioning and user tracking. These functionalities present in the IS based on ISO 19115 conceptual model implementation favour interoperability, security and confidentiality of pastoral data.

However, the practical applications of data integration in spatial ISs raise some practical considerations, especially in relation to the integration of spatial land cover and satellite imagery data. The processes of ingesting and storing data in the HDFS are carried out by replicating the sequences of blocks corresponding to each file in the Datanodes available for this purpose (Honnutagi, 2014). The size of each block is 128 mb by default but can be configured by the operator. For 500 mb data, for instance, 4 blocks including 3 of 128 mb and 1 of 116 mb, can be replicated as many times in the Datanodes. This process allows better management of less heavy data, such as technical references, GPS data, meteorological data and farmer feedback. However, ingesting and storing massive spatial data such as land cover and raster data requires a specific approach that takes into account the topology of factors. Hadoop does not handle better the topology of large vector data when ingesting and storing. Application of appropriate techniques are necessary. In Yao et al. (2017), the authors proposed an implementation based on a spatial partitioning matrix allowing a coherent replication of spatially close factors. Such an approach can be experimented with for the integration of land use data.

The data lake is dynamic, depending on the data cycle and its feeding over time. The increase in data volume raises the problem of scaling of the IS. In the production phase, it is important to evaluate the storage aspect of heterogeneous pastoral data by considering their dynamics. In essence, data lakes are designed to better manage the problem of increasing the volume of data (Fang, 2015; Khine and Wang, 2018).

8.3. Applicability in other European Mediterranean areas

Several aspects favoured the realisation of this study in French mountainous and Mediterranean areas. The presence of services and farmers involved in pastoralism, the recent work carried out within a framework of different projects that tested different technological tools to improve pastoral conditions, and the availability of various public data that can contribute to the modelling of the agro-sylvo-pastoral ecosystem.

As long as the conditions concerning data availability and farmer interest are met in other areas, there are sufficient similarities between EU agro-sylvo-pastoral farms and common issues (de Rancourt et al., 2006) to extend this method. Outside of the EU, the questions and needs of farmers might be slightly different since the CAP is a strong structuring factor in the EU.

If the IS is to be extended to other geographic areas, new specific databases would need to be built, since certain types of similar data (MNT RGE, OSO land use and cover, RPG and technical references) are specific to each country. Data of the same nature as the one used in this work can also be obtained or generated in these different countries. In the case of the existence of these types of data, the different approaches developed in this article may be applicable.

9. Conclusion and perspectives

In this study, we describe the codesign process of a spatial information system for grazing management in agro-sylvo-pastoral systems in French mountainous and Mediterranean regions. To address the operational concerns of grazing, we define with the stakeholders the main functionalities of the IS. We then mapped out the grazed ecosystem model that allowed us to identify the heterogeneous pastoralism data. The information system is a data lake based on the HDFS, offering functionalities for the integration of heterogeneous data and the analysis and visualisation of information to support decision-making in pastoral farming.

In the context of variability and heterogeneity of agro-sylvo-pastoral systems, this information system will provide information adapted to each farm thanks to the data that they provide. This will provide farmers with an effective tool to help them make decisions regarding the actions to be taken. This information system will also provide a source of information for research to carry out studies on larger scales and to document knowledge about these variable and heterogeneous systems. Furthermore, this IS could lead to an exchange network between pastoralists, with the aim of improving knowledge and practices. The design approach presented in this study can also be applied to other mountainous and Mediterranean regions in Europe to build an IS adapted to the context of each region.

The codesign process are not only limited to the applications defined in this study. Throughout the project, regular interactions between stakeholders made it possible to improve the established applications and to propose others to be integrated into the IS. These applications notably concern the cross-referencing and visualisation of information. They offer users relevant information adapted to their needs.

The next study will focus on the implementation of the spatial data lake presented here to provide concrete results in the practical situation of livestock farming in French Mediterranean areas.

Appendices

Appendix A : tables

Table A.1: Breeders' needs expressed in the form of questions, considered as a basis for the IS functionalities

N°	Questions
	Benchmarks for adjusting pasture management
1	Do my animals have enough forage to eat ? *
2	What is the instantaneous stocking rate on my rangeland? (no. of animals/ha) *
3	What is the seasonal or annual stocking rate on my rangeland? (no. of animals/ha * days) *
4	What type of pastoral resource is available in my paddock? **
5	What is the average daily distance travelled by the herd? **
6	What is the grazing area used at the moment? ***
7	What route does my herd take each day? ***
8	How much time does my herd spend for grazing? ***
9	What is the area covered by the herd? ***
10	How do my animals occupy the space? (weekly, monthly, seasonally, yearly) ***
	Methods to achieve a given objective
11	How to identify the vegetation? *
12	How to prevent fires on my rangelands? *
13	Where to place fences or enclosures? *
14	How to bring water on the rangelands? *
15	How can animals refusals be reduced? *
16	How to adapt the grazing circuit to the topography of the rangelands? *
17	How can several grazing methods be managed (guarding, free ranging, park)? *
18	How can animals be forced to consume woody plants? *
19	Would it be useful to split up a paddock? *
20	How can animals be forced to eat unpalatable species? *
21	How can digital tools be integrated to save time and accuracy? *
22	How can encroached areas be reopened? **
23	What information should be passed on to a breeder who takes over the herd? **
24	How to quantify the available pastoral resource? **
25	How to control encroached areas? **
26	How to adapt the grazing circuit to the weather of the day? **
27	How to change the attractiveness of an area for a herd? **
28	How can encroached areas be enhanced? ***
29	How to develop the pastoral resource to meet the needs of my herd? ***
30	How can animals be stimulated to move into unexplored areas? ***
	Better understand the ecology of the pastoral ecosystem
31	What are the differences between species in the use of rangeland? *
32	What are the differences between breeds of the same species in terms of rangeland utilisation? *
33	How does animal behaviour change according to the breeder? *
34	I'm setting up: How can I breed my herd in the open air and valorise the rangelands? *
35	What is the impact of climate change on rangelands? **
36	What is the influence of the type of pastoral resource on dairy production? **
37	How does the animal behaviour change depending on the grazing management? **
38	How does the behaviour of animals change according to the season? ***

Table A.2: Factors from the literature review validated by the breeders and those from the breeders' feedback only, shown in *italics*.

Factors	**Number of votes out of 8**
Pastoral resources	
Availability	8
Accessibility	3
Spatial distribution	7
Animals / herds	
Physiology	7
Morphology	1
Relationship between animals	7
Spatial memory : *Knowledge of the environment*	2
Herd size	2
Managment practices	
Guarding	8
Complementing	5
Grazing moment	7
Environment	
Topography : slope	5
Meteo : Ran	5
Meteo : Wind	4
Meteo : Warm	8
Attractions: water point	6
Attractions: resting area	7
Park size	3

Appendix B : description of external data that can be integrated into the information system for grazing management

Appendix B.1: Sentinel 2 satellite data from Theia

Sentinel 2 data[9] comprise 13 spectral bands at 10, 20 and 60 m spatial resolutions depending on the spectral band. A wide swath width of 290 km makes it possible to capture at once a large portion of the Earth's surface with a revisit period of 10 days (by 1 satellite) or 5 complementary days (for the two satellites, Sentinel 2A and Sentinel 2B). Its spatial resolution and its spectral richness are an asset for obtaining better discrimination of the different pastoral vegetation classes (see, Castro et al., 2020), which can bring, in association with other information (e.g., spectral indices of very high spatial resolution data), more detail for analyses at the level of agro-sylvo-pastoral exploitation (e.g., ligneous vegetation density). These data can be analysed together with other references, such as the RPG, which can provide better information for training and validation in the supervised classification process.

[9] https://www.theia-land.fr/

Appendix B.2: Registre Parcellaire Graphique (RPG)

RPG is a geographic information system managed by the French Service and Payment Agency (ASP), allowing the identification of agricultural parcels declared by farmers. It provides detailed information on land use and land structures. Several rangelands (lawns, open moorlands, and closed moorlands) and grazed forests are represented. In addition to identifying pastoral classes, the RPG can be useful in the land use classification process in combination with satellite data, as shown in Shaqura and Lasseur (2019). It is an important source of accurate and large-scale training data.

Appendix B.3: OSO land use and cover

The OSO land use and cover is processed and provided by the National Center for Spatial Studies (CNES) in collaboration with specialised research laboratories[10]. These data have a spatial resolution of 10 m in raster mode and a minimum mapping unit of 20 m in vector mode with 23 classes. The 23 classes are divided into 5 groups: artificial areas, annual crops, perennial crops, forests, natural vegetation and other nonvegetated natural areas.

The iota2 processing chain[11] that made it possible to obtain land use

[10] https://www.theia-land.fr/la-carte-doccupation-des-sols-millesime-2019%2Dfrance-entiere/

[11] https://www.theia-land.fr/product/iota-2/

data from Sentinel 2 data (Inglada et al., 2017) allowed us to obtain better outlines of objects. However, the accuracy of the rangeland classes (permanent grassland, lawn and moorlands) in OSO land use and cover is low. To improve this accuracy, the processing of OSO vector data can be performed for an oriented object classification process. To do this, the OSO data can be combined with the Sentinel 2 data and the RPG is utilised as reference data. The annual production of OSO data is a good opportunity for an annual update of land use for agro-sylvo-pastoral applications.

Appendix B.4: IGN altimetry data: the MNT RGE ALTI

The MNT RGE ALTI altimetry data compose a relief model in the form of a regular rectangular grid, also called the “altitude matrix”, where the altitude of each of the nodes of this grid corresponds to the nominal terrain altitude at the point defined by the grid node. The spatial resolution of the grids ranges from 1 m to 5 m in X and Y. The information drawn from these data allows the relief of the rangelands to be modelled with better accuracy. Interaction with other information, such as land use and cover, can allow us to characterise the shape of the land within an agro-sylvo-pastoral farm. These data might be collected only once for the areas to be studied, given that the topography of these environments can be considered constant.

Appendix B.5: Technical references

These are documents describing technical practices in the field of pastoralism. They are mainly collected from different agro-sylvo-pastoral farming institutions and associations. They can be integrated into the information system to serve as a reference source of information for various users, mainly farmers and shepherds, on agro-sylvo-pastoral practices.

Acknowledgements

This work was supported by the French National Research Agency under the Investments for the Future Program, referred as ANR-16-CONV-0004.

We thank the 7 farmers, the 2 experimental sites, the researchers and the technical institute (IDELE) that contributed to the realisation of this paper within the framework of the P@stor-All project.

We also thank Daria Renault, who contributed, following her internship, to the formalisation of the feedback from the farmers, which helped to define the functionalities of the information system.

References

Agreste, 2015. L'utilisation du territoire en 2014. Teruti-Lucas. Agreste Chiffres et Données Agriculture , 105.

Akasbi, Z., Oldeland, J., Dengler, J., Finckh, M., 2012. Analysis of GPS trajectories to assess goat grazing pattern and intensity in Southern Morocco. The Rangeland Journal 34, 415.

Allen, V., Batello, C., Berretta, E., Hodgson, J., Kothmann, M., Li, X., McIvor, J., Milne, J., Morris, C., Peeters, A., Sanderson, M., The Forage and Grazing Terminology Committee, 2011. An international terminology for grazing lands and grazing animals. Grass and Forage Science 66, 2–28.

Bahlo, C., Dahlhaus, P., Thompson, H., Trotter, M., 2019. The role of interoperable data standards in precision livestock farming in extensive livestock systems: A review. Computers and Electronics in Agriculture 156, 459–466.

Bailey, D.W., Mosley, J.C., Estell, R.E., Cibils, A.F., Horney, M., Hendrickson, J.R., Walker, J.W., Launchbaugh, K.L., Burritt, E.A., 2019. Synthesis Paper: Targeted Livestock Grazing: Prescription for Healthy Rangelands. Rangeland Ecology & Management 72, 865–877.

Bataille, J.F., Launay, F., Aussibal, G., Moulin, C.H., Nozieres, M.O., Lasseur, J., Tchakerian, E., 2016. Des scénarios de futurs possibles pour des systèmes d'élevage pastoraux méditerranéens. Innovations Agronomiques 49, 309–325.

Bernués, A., Rodríguez-Ortega, T., Ripoll-Bosch, R., Alfnes, F., 2014. Socio-Cultural and Economic Valuation of Ecosystem Services Provided by Mediterranean Mountain Agroecosystems. PLoS ONE 9, e102479.

Bernués, A., Ruiz, R., Olaizola, A., Villalba, D., Casasús, I., 2011. Sustainability of pasture-based livestock farming systems in the European Mediterranean context: Synergies and trade-offs. Livestock Science 139, 44–57.

Berrier, C., Girard, N., 2021. The construction of the meaning of work for livestock farmers engaged in the agroecological transition, in: 2nd international symposium of word of agriculture, Clermont-Ferrand, France. p. 11.

Buerkert, A., Schlecht, E., 2009. Performance of three GPS collars to monitor goats' grazing itineraries on mountain pastures. Computers and Electronics in Agriculture 65, 85–92.

Capalbo, S.M., Antle, J.M., Seavert, C., 2017. Next generation data systems and knowledge products to support agricultural producers and science-based policy decision making. Agricultural Systems 155, 191–199.

Castro, M., Ameray, A., Castro, J.P., 2020. A new approach to quantify grazing pressure under mediterranean pastoral systems using GIS and remote sensing. International Journal of Remote Sensing 41, 5371–5387.

Charbonier, Q. (Ed.), 1972. La loi pastorale. La Cardère, France.

Chen, W., Ji, M., Wang, J., 2014. T-DBSCAN: A Spatiotemporal Density

Clustering for GPS Trajectory Segmentation. International Journal of Online and Biomedical Engineering (iJOE) 10, 19.

Dono, G., Cortignani, R., Dell'Unto, D., Deligios, P., Doro, L., Lacetera, N., Mula, L., Pasqui, M., Quaresima, S., Vitali, A., Roggero, P.P., 2016. Winners and losers from climate change in agriculture: Insights from a case study in the Mediterranean basin. Agricultural Systems 147, 65–75.

Dumont, B., Andueza, D., Niderkorn, V., Lüscher, A., Porqueddu, C., Picon-Cochard, C., 2015. A meta-analysis of climate change effects on forage quality in grasslands: specificities of mountain and Mediterranean areas. Grass and Forage Science 70, 239–254.

Dumont, B., Meuret, M., Boissy, A., Petit, M., 2001. Le pâturage vu par l'animal : mécanismes comportementaux et applications en élevage. Fourrages 166, 2013–238.

Fang, H., 2015. Managing Data Lakes in Big Data Era, in: The 5th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems, Shenyang, China. p. 5.

Feldt, T., Schlecht, E., 2016. Analysis of GPS trajectories to assess spatiotemporal differences in grazing patterns and land use preferences of domestic livestock in southwestern Madagascar. Pastoralism 6, 5.

Ginane, C., Dumont, B., Baumont, R., Praches, S., Farrugia, A., 2008. Comprendre le comportement alimentaire des herbivores au pâturage : intérêts pour l'élevage et l'environnement, in: Rencontres autour des Recherches sur les Ruminants, Paris. pp. 315–322.

Hadjigeorgiou, I., Osoro, K., Fragoso de Almeida, J., Molle, G., 2005. Southern European grazing lands: Production, environmental and landscape management aspects. Livestock Production Science 96, 51–59.

Hai, R., Geisler, S., Quix, C., 2016. Constance: An Intelligent Data Lake System, in: Proceedings of the 2016 International Conference on Management of Data - SIGMOD '16, ACM Press. pp. 2097–2100.

Handcock, R., Swain, D., Bishop-Hurley, G., Patison, K., Wark, T., Valencia, P., Corke, P., O'Neill, C., 2009. Monitoring Animal Behaviour and Environmental Interactions Using Wireless Sensor Networks, GPS Collars and Satellite Remote Sensing. Sensors 9, 3586–3603.

Honnutagi, P.S., 2014. The Hadoop distributed file system. International Journal of Computer Science and Information Technologies 5, 6238–6243.

Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., Rodes, I., 2017. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. Remote Sensing 9, 95.

Kafando, R., Decoupes, R., Sautot, L., Teisseire, M., 2020. Spatial Data Lake for Smart Cities: From Design to Implementation. AGILE: GIScience Series 1, 1–15.

Khine, P.P., Wang, Z.S., 2018. Data lake: a new ideology in big data era. ITM Web of Conferences 17, 03025.

Linnell, J.D.C., Cretois, B., 2018. Research for AGRI Committee - The revival of wolves and other large predators and its impact on farmers and their livelihood in rural regions of Europe. European Parliament. Policy Department for Structural and Cohesion Policies, Brussels.

McGavin, S.L., Bishop-Hurley, G.J., Charmley, E., Greenwood, P.L., Callaghan, M.J., 2018. Effect of GPS sample interval and paddock size on estimates of distance travelled by grazing cattle in rangeland, Australia. The Rangeland Journal 40, 55–64.

Mendes, J., Pinho, T.M., Neves dos Santos, F., Sousa, J.J., Peres, E., Boaventura-Cunha, J., Cunha, M., Morais, R., 2020. Smartphone Applications Targeting Precision Agriculture Practices—A Systematic Review. Agronomy 10, 855.

Meuret, M., Garde, L., Moulin, C.H., Nozières-Petit, M.O., Vincent, M., 2018. Élevage et loups en France : historique, bilan et pistes de solution. INRA Productions Animales 30, 465–478.

Meuret, M., Laurent, G., Moulin, C.H., Nozieres-Petit, M.O., Vincent, M., 2017. L'élevage de plein air confronté aux loups protégés et adaptables Bilan sur 25 ans, in: Duclos, J.C., Fabre, P., Garde, L. (Eds.), Elevage pastoral, espaces protégés et paysages: en Provence-Alpes-Côte d'Azur, Cardère, Maison de la Transhumance, CERPAM & ARPE-PACA, Aix-en-Provence, France. pp. 100–106.

Michels, M., Fecke, W., Feil, J.H., Musshoff, O., Pigisch, J., Krone, S., 2020. Smartphone adoption and use in agriculture: empirical evidence from Germany. Precision Agriculture 21, 403–425.

Nori, M., 2019. Herding through Uncertainties – Regional Perspectives. Exploring the interfaces between pastoralists and uncertainty. Results from a literature review. Robert Schuman Centre for Advanced Studies Research. Paper No. RSCAS 2019/68 , 84.

Nozieres, M., Etienne, L., Launay, F., Moulin, C., 2021. Les grands traits de l'élevage pastoral aujourd'hui en France. Revue de l'association francophone pour les prairies et les fourrages , 3–11.

Pongnumkul, S., Chaovalit, P., Surasvadi, N., 2015. Applications of Smartphone-Based Sensors in Agriculture: A Systematic Review of Research. Journal of Sensors 2015, 1–18.

Putfarken, D., Dengler, J., Lehmann, S., Härdtle, W., 2008. Site use of grazing cattle and sheep in a large-scale pasture landscape: A GPS/GIS assessment. Applied Animal Behaviour Science 111, 54–67.

Quix, C., Hai, R., 2019. Data Lake, in: Sakr, S., Zomaya, A. (Eds.), Encyclopedia of Big Data Technologies. Springer International Publishing, Cham, pp. 1–8.

de Rancourt, M., Fois, N., Lavín, M., Tchakérian, E., Vallerand, F., 2006. Mediterranean sheep and goats production: An uncertain future. Small Ruminant Research 62, 167–179.

Ravat, F., Zhao, Y., 2019. Data Lakes: Trends and Perspectives, in: Hartmann, S., Küng, J., Chakravarthy, S., Anderst-Kotsis, G., Tjoa, A.M., Khalil, I. (Eds.), Database and Expert Systems Applications. Springer International Publishing, Cham. volume 11706, pp. 304–313.

Rutter, S., Beresford, N., Roberts, G., 1997. Use of GPS to identify the grazing areas of hill sheep. Computers and Electronics in Agriculture 17, 177–188.

Sawadogo, P.N., Scholly, E., Favre, C., Ferey, E., Loudcher, S., Darmont, J., 2019. Metadata Systems for Data Lakes: Models and Features, in: 1st International Workshop on BI and Big Data Applications, Springer, Bled, Slovenia. pp. 440–451.

Shaqura, I., Lasseur, J., 2019. Characterizing extensively grazed areas between forests and pastureland to better assess the resilience of Mediterranean livestock activities. Lessons learned from a study on the scale of the Provence Alpes Côte d'Azur region (France)., in: FAO-CIHEAM Network on Sheep and Goats and Mediterranean Pastures "Efficiency and resilience of forage resources and small ruminant production to cope with global challenges in Mediterranean areas", Meknes, Moroco. p. 2.

Spielman, S.E., 2017. Point Pattern Analysis, in: International Encyclopedia of Geography. John Wiley & Sons, Ltd, pp. 1–9. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118786352.wbieg0849, doi:https://doi.org/10.1002/9781118786352.wbieg0849. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118786352.wbieg0849.

Suriarachchi, I., Plale, B., 2016. Crossing analytics systems: A case for integrated provenance in data lakes, in: 2016 IEEE 12th International Conference on e-Science (e-Science), IEEE, Baltimore, MD, USA. pp. 349–354.

Teucher, M., Hornetz, B., Jätzold, R., Mairura, Z., 2014. Assessing the users' need for a spatial decision support system of smallholder farming in kenya , 11.

Vincent, M., 2010. Les pratiques des bergers dans les Alpes bouleversées par le retour de loups protégés, in: Un savoir-faire de bergers.. meuret m.(coord.) ed.. Dijon & Versailles. Educagri & Quæ, pp. 221–243.

di Virgilio, A., Morales, J.M., Lambertucci, S.A., Shepard, E.L., Wilson, R.P., 2018. Multi-dimensional Precision Livestock Farming: a potential toolbox for sustainable rangeland management. PeerJ 6, e4867.

Yao, X., Mokbel, M.F., Alarabi, L., Eldawy, A., Yang, J., Yun, W., Li, L., Ye, S., Zhu, D., 2017. Spatial coding-based approach for partitioning big spatial data in Hadoop. Computers & Geosciences 106, 60–67.

Zampaligré, N., Schlecht, E., 2018. Livestock foraging behaviour on different land use classes along the semi-arid to sub-humid agro-ecological gradient in West Africa. Environment, Development and Sustainability 20, 731–748.

Zheng, Y., 2015. Trajectory Data Mining: An Overview. ACM Transactions on Intelligent Systems and Technology 6, 1–41. URL: https://dl.acm.org/doi/10.1145/2743025, doi:10.1145/2743025.