Introduction to Precision Agriculture

written by Marta Alexy PhD

Budapest

2024

1 Contents

1. Introduction

I am sure everyone has direct contact with agricultural products several times daily. Regardless of what you do or your occupation, whether you are an engineer, a teacher, or a computer scientist, every human consumes more or less animal—and plant-based food every day. I think I am, therefore, right in thinking that this subject affects everyone closely. Moreover, if you are also curious about how what you eat ends up on your table, this book is well worth reading.

Initially, the exhausting search for plant and animal food that humans can eat, and then the conscious production of it, is the most essential task of humankind. The science of agriculture, which began more than 10,000 years ago and has evolved, albeit very slowly, since then, now faces many challenges. Humans have brought about some of these challenges by exploiting the natural environment. Agricultural activity, initially extensive and then intensifying from the 20th century onwards, has upset the balance of the environment, causing the damaging processes we are witnessing today. However, there was a particular reason for this intensive agricultural activity, damaging the natural environment: the world's population is growing, and developing countries' living standards and purchasing power are increasing. In contrast, natural resources are limited in number, size, and availability for resource-efficient and profitable food production. Given the growing need to provide good quality food for more and more people, there is no alternative but to optimize resources.

One of the specificities of agricultural activity is that the 'product' to be produced is a living plant and animal organism with different biological characteristics over time and space. One of the most influential discoveries has been genetics, which has helped humans to understand how they can influence the quality and quantity of agricultural products that benefit them through targeted breeding. This knowledge was complemented by the use of animal power to replace human muscle power, then by the use of mechanical power through mechanization and automation, and, from the 21st century onwards, by the science of information technology, which has complemented the capacity of the human brain to facilitate cognitive work. Through thousands of years of non-conscious and then conscious breeding, humans have selected the 15 most valuable plants and five animal species, taken them out of their natural environment, and placed them in conditions they can more or less control. These are the so-called intensive, large-scale agricultural production systems. The using of the relatively young science of information technology, especially data science, in intensive agrifood systems is precision agriculture, which this book provides insight into.

Although the benefits of using information technology (IT) are beyond doubt, its widespread practical application in food production and its acceptance by farmers is more complex than one might first think. Agriculture is a traditional sector in which thousands of years of knowledge and experience have accumulated. Farmers' wisdom, practical experience, and passion for the land and livestock largely determine the extent to which they 'let go' of control. Although international trends clearly show a concentration of agricultural production on increasing profitability in human capital, the sector is very human-oriented, and it takes work to absorb new and innovative solutions. In contrast, information technology is rapidly changing and reacting quickly to environmental challenges. The most significant added value of IT to agricultural production is the vast amount of digital data that can be collected at all stages of production using IoT devices (as opposed to the traditional manual production data that has been collected so far). Digital data can also be used for economic decisions that farmers have yet to collect or have collected with much less efficiency. At the same time, the framework for acquiring digital data needs to be developed, which includes the architecture for transmitting, storing, and analyzing large amounts of data. This goes far beyond the statistical calculations used so far. The analytical methodology of Big Data databases, not only because of their size but also because of the heterogeneous sources of raw data in different forms, has taken traditional agricultural data processing to a new level.

Therefore, it is essential to learn about data science methods and methodologies that can help the actors in the agricultural sector to meet today's ecological, economic, and social expectations, achieving resource-efficient and sustainable food production.

Due to the course's interdisciplinary nature, precision agriculture is presented in **three main chapters**: the **first describes the specificities of large-scale crop and livestock farming**. In the **second** chapter, I present the areas of data science that are most relevant for agriculture: **the importance and relevant areas of artificial intelligence and, within this, machine learning**. In the **third** chapter, I will present **some practical applications of digitalization in crop and livestock production without claiming to be exhaustive**. Through a detailed description of our PLF research in large-scale poultry production, I will show the process of taking an idea from implementation to product development.

2. Some critical definitions of terms

Agriculture is the practice of farming and cultivating plants, animals, and other natural resources for food, fiber, and other products. It involves various activities such as planting, harvesting, raising livestock, and managing land for agricultural purposes. Agriculture is crucial in providing food security, supporting rural economies, and sustaining the environment.

Crop production refers to cultivating and growing crops for human consumption, livestock feed, and other purposes. It involves preparing the soil, planting seeds or seedlings, providing proper nutrients and water, controlling pests and diseases, and harvesting the crops at the right time. Crop production is a vital component of agriculture and plays a significant role in ensuring food security and meeting the demands of a growing population. Various crops such as grains, fruits, vegetables, and cash crops are grown through crop production methods.

Phenophases: crops go through several distinct growth stages, known as phenophases, during their development. These phenophases vary depending on the crop type, but some standard stages include 1. germination, 2. vegetative growth, 3. reproductive growth, and 4. ripening.

Livestock production refers to a farming system where many animals are raised in a confined space or controlled environment to maximize production efficiency. In this system, animals are typically housed in large buildings or feedlots and are provided with feed, water, and other necessities to promote rapid growth and high productivity. Intensive livestock production systems are characterized by: 1. High stocking densities: Animals are kept near each other to maximize space utilization and production efficiency; 2. Controlled environment: Facilities are designed to regulate temperature, lighting, ventilation, and other environmental factors to optimize animal growth and health; 3. Feedlot feeding: Animals are fed high-energy diets to promote rapid weight gain and efficient meat production; 4. Use of technology: Intensive livestock production often involves technology and automation to monitor and manage animal health, feed intake, and production parameters. While intensive livestock production can increase productivity and profitability, it raises concerns about animal welfare, environmental impact, and food safety. Sustainable practices and regulations are essential to ensure the well-being of animals, minimize environmental degradation, and produce safe and high-quality animal products.

Ruminants refer to animals with a unique digestive system called rumen, allowing them to digest fibrous plant materials efficiently. Ruminants have a specialized stomach with four compartments: the rumen, reticulum, omasum, and abomasum. This digestive system enables ruminant animals to ferment and break down complex carbohydrates in plant materials through microbial action, converting them into nutrients that can be absorbed and utilized by the animal. Common examples of ruminant livestock include cattle, sheep, and goats. These animals are herbivores and feed on grasses, forages, and other plant materials. Ruminant livestock play a significant role in agriculture and food production by providing meat, milk, wool, and other products for human consumption.

Monogastric refers to animals with a simple single-chambered stomach, unlike ruminant livestock with a multi-chambered stomach. Monogastric animals have a digestive system that consists of one stomach compartment where food is digested and nutrients are absorbed. Common examples of monogastric livestock include pigs, poultry (such as chickens and turkeys), and horses. These animals are omnivores or herbivores and have different dietary requirements than ruminant animals. Monogastric livestock are raised for meat, eggs, milk, and other products for human consumption.

Precision agriculture (PA), also known as precision farming or precision ag, is an approach to agriculture that uses technology and data analysis to optimize crop production and management practices. It involves using tools such as GPS, sensors, drones, satellite imagery, and data analytics to monitor and manage various aspects of farming operations with high accuracy and efficiency.

Precision livestock farming (PLF) is a technology-driven approach to livestock production that uses sensors, data analytics, and automated systems to monitor and manage individuals or groups of animals more precisely and efficiently.

The *Internet of Things (IoT)* refers to a network of interconnected devices, objects, and systems that can communicate, exchange data, and perform tasks without human intervention. In agriculture, IoT technology connects various sensors, devices, and equipment to collect and share real-time data, enabling farmers to monitor and manage their operations more efficiently.

Data science is a multidisciplinary field involving scientific methods, algorithms, and systems to extract knowledge and insights from structured and unstructured data.

Artificial intelligence refers to the simulation of human intelligence processes by machines, particularly computer systems. AI technologies enable machines to perform tasks that typically require human intelligence, such as learning, reasoning, problem-solving, perception, and decision-making.

Machine learning is a subset of artificial intelligence that involves the development of algorithms and models that enable computers to learn from and make predictions or decisions based on data. In machine learning, computers are trained to recognize patterns, relationships, and trends in data without being explicitly programmed to perform specific tasks.

3. Characteristics of food production and today's challenges

Food production is concerned with producing food of plant and animal origin that humans can consume, either raw or processed. Plant foods mainly give us carbohydrates, and animal foods provide protein.

The production of agricultural products is an economic activity, and therefore, the primary objective of the farmer is to maximize income. This is determined by the characteristics of the plant and the livestock species and the parameters of the farming environment. The 'output' provided by the plant and the animal is reflected in the phenotype, i.e., the appearance, from which the farmer derives a profit. The phenotype is determined by the interaction of genotype (the performance of a plant or animal as recorded in its genetic makeup) and environment. By default, the genotype is considered 100% in the production process of large-scale commercial farming (which can be surpassed by bio- and gene technologies, not discussed in this paper). However, the genotype includes the natural, species-specific characteristics and behaviors of farm species that largely determine the breeding technology and management required. The farmer must ensure the optimal environmental factors throughout the year to achieve the phenotype, i.e., the highest possible profit.

On this basis, agricultural production can be divided into two sectors: crop production (which includes the production of fodder crops important for large-scale livestock farming) and livestock production. *Figure 1* further subdivides these two sectors.

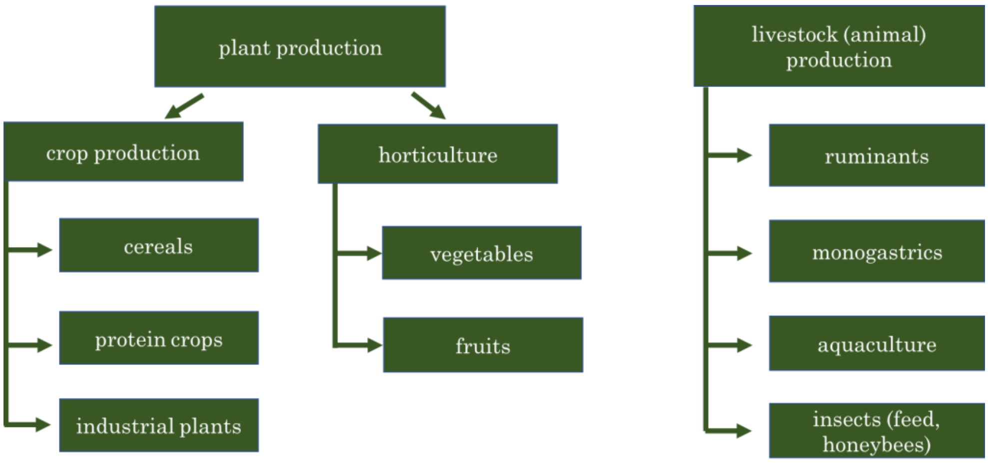


Figure 1. Structure of agricultural production system

Figure 1 shows the complexity of food production essential for human consumption. In addition to crops grown on arable land, crop production includes the horticultural subsector (this tutorial focuses on crop and livestock production). The large-scale livestock subsector comprises the leading commercial livestock: ruminants, monogastric, aquaculture, and insects.

The following is a brief description of each subgroup's main characteristics from the perspective of precision agriculture.

4. Crop production

Only 10-11% of the Earth's land area is suitable for arable farming. This amounts to roughly 1400 million ha. In addition, grassland, which covers approximately 3,400 million ha, is also valuable for agriculture. To meet the world food challenge, all that seems to be needed is to decide at what cost to expand the area under cultivation. Deforestation of forests, including tropical rainforests (e.g., Indonesia, Brazil), could be considered, but this would lead to severe environmental damage in many respects. Trees use carbon dioxide for photosynthesis, thus increasing their mass and emitting oxygen as part of this process. If we cut down the forests, we stop removing carbon dioxide from the atmosphere, which is a natural process that contributes to mitigating the greenhouse effect, i.e., increasing global warming; by cutting down forests, the carbon dioxide already sequestered in the trees is released into the atmosphere, further increasing the greenhouse effect; the Earth's forests are the most important source of oxygen in the atmosphere, and if these sources disappear, the atmosphere's ability to renew oxygen will be severely weakened. Another way of increasing the area under cultivation is to clear the heathlands and grasslands. However, intensive crop production is mostly limited by drought and is only possible in the rain-fed regions of Latin America and Africa. Elsewhere, only if newly cultivated areas are irrigated. However, this is very capital-intensive, so countries at constant risk of starvation do not have this option. Foreign investors are generally not attracted to the opportunity, as the return on investment is prolonged, and the risk of success is high. Partly for similar reasons, countries facing starvation do not use other means of increasing yields (e.g., mechanization, use of chemicals). National programs to increase arable land have been announced and launched with limited results. Few countries (e.g., the United States, Argentina, Australia, Ukraine, and Russia) still have limited arable land for cultivation. The biggest challenge for crop production today and in the future is how to produce increasing quantities of products for an accelerating human population on minimally expanding arable land. The difficulties are compounded by the fact that, on the one hand, optimum conditions for crop production are limited on a large part of arable land, and, on the other hand, only crop production technologies that produce safe and high-quality products and are environmentally and economically sustainable can be used. Due to the rapid growth of the world's population and the constant amount of arable land globally, the arable land per capita is gradually decreasing. This means that less and less land (0.42 ha/ha in 1961, 0.20 ha/ha in 2010, and some projections suggest that only 0.06 ha/ha in 2050) is available for crop production under increasing quantitative and qualitative constraints. In particular, countries with a large population relative to their land area will be in serious trouble (e.g., Pakistan, Nigeria, Indonesia, Ethiopia, and Bangladesh). In addition to the declining specific arable area, the further problem is that in areas used for arable crops, adverse phenomena and processes (e.g., processes, drought, nutrient stress, shallow soil cover, etc.) reduce soil fertility and the effectiveness of agrotechnical interventions. Agriculture is also the cause of two global environmental problems (soil nitrogen pollution and loss of water quality and biodiversity). In addition to soil scarcity, water scarcity is one of the main constraints to increasing food production. Some experts believe the key to ending hunger is reorganizing the world's water management. They say that increasing water tariffs and rational water use will begin. The balance of the world's water use and water resources indicates a huge deficit - a deficit that is largely invisible, recent in historical terms, and growing fast. Half of the world's population lives in countries where groundwater levels are sinking while groundwater resources are depleted. With around 70% of the world's water consumption used for irrigation, water scarcity can quickly lead to food shortages. Increasingly severe water scarcity could be significantly alleviated by sprinkler irrigation, which is much more efficient than traditional flood, furrow, or spray irrigation. Drip irrigation, a method developed in Israel, may not be economical for cereal production but can reduce water consumption by up to 70% for producing special-quality fruit and vegetables. In addition to increasing water use, water pollution is significant due to using pesticides and fertilizers, especially in fruit and vegetable production. Using chemicals and pesticides harms the environment, but even with them, high yields will not be sustainable in the long term. Increasing nitrogen loads on soil, water, and the atmosphere have serious long-term environmental impacts. One example is the eutrophication of water bodies, lakes, and canals through increased fertilizer use. Chemical use and intensive production technologies also contribute to biodiversity loss.

As world food demand has risen sharply, millions of farmers have drilled irrigation wells to increase crop yields. Because there were generally no government restrictions, too many wells were drilled. The result is that in some 20 countries, including China, India, and the United States, which produce half of the world's grain, groundwater levels have recently been falling, and wells have been drying up one after another. This is particularly problematic given that groundwater is so-called fossil groundwater in some parts of the world, which cannot be recharged after pumping. Over-pumping groundwater for irrigation can temporarily increase food production, bursting a bubble when groundwater supplies are depleted. With 40% of the world's cereal production coming from irrigated land, the potential reduction in irrigation water could be a major concern. Of the three major cereal-producing countries, roughly one-fifth of US grain production is irrigated, three-fifths in India, and four-fifths in China. The problem of too many wells and their intensive groundwater extraction is not only a problem for agriculture. World cities are at risk. The constant and massive pumping of groundwater from thousands of wells is simply causing the soil to collapse like a dried-up sponge. Beijing is sinking by 11 cm a year, Bangkok by 12 cm, Jakarta and Mexico City by 28-28 cm, and Shanghai is facing a similar problem.

Sustainable crop production means a production technology that does not degrade but preserves (and, if possible, improves) the ecological conditions of production in the long term, maintains and improves the diversity of the natural environment, produces high-quality crop products, and is an economically feasible production method. However, these seemingly clear criteria are challenging to reconcile with large-scale farming practices. Some factors benefit sustainable crop production technology implementation, while others are unfavorable. Therefore, integrating these factors into the production process is very important. We strive to integrate agroecological, bio-genetic, and agrotechnical factors into crop production. To this end, it is essential to have a precise knowledge of the agroecological factors (weather-climate, soil, and topography), not only in terms of average values but also in terms of deviations from them, about weather parameters and the different soil and topography characteristics of the so-called zones within the field. This information and knowledge of the level of agrotechnology used should be used to select the most suitable crop species or hybrid. When designing the agrotechnical elements, it is necessary to know both the site-specific conditions and the specific agrotechnical responses of the variety/hybrid being grown (site- and variety-specific technologies).

4.1 Cereals

Cereal crops are grown almost all over the world, depending on the climate. The diverse climatic requirements and good adaptability of wheat species and varieties have made their spread possible. They are grown in almost all but the most extreme climates (tropics, deserts, and the Arctic). Wheat is the world leader in global area under cultivation and third in global yields, behind maize and rice. In the case of maize, this is due to its high use for animal feed, and in the case of rice, mainly due to rice production in Asian countries for domestic consumption. Wheat and maize account for about 60% of the global cereal crop and are vital for humanity's nutrition. Cereals are our country's most critical and largest arable crop globally. Most cereal crops are primary crops, i.e., among the first to be introduced into cultivation. Cereal crops were first cultivated in different parts of the world around 10-12 thousand years ago, using much more primitive 'technology' than today. Cereal crops' diversity and species richness allow them to be grown effectively under various climatic, soil, and agrotechnical conditions. However, cereals are also widely used in other areas. Cereals are grown in almost every country in the world. The central cereal-producing countries with the most significant area are India, China, the USA, the European Union, Russia, and Australia. In these countries, the proportion of each cereal species varies according to agroecological conditions, market needs, and farming traditions. In the USA, for example, maize has traditionally been the most critical cereal; in China, it is mainly wheat and Rice, while in India, in addition to cereals, sorghum and millet are also very important. The differences between countries are very significant. They are essential for producing animal feed, industrial raw materials, and, more recently, fuels (bioethanol). Cereal crops have a long storage life under the right conditions, and their production is highly mechanizable. These conditions have led to cereals being grown on about 50% of the world's arable land (~700 million ha).

The "green revolution" promoted by the FAO since the 1960s has now achieved significant results. Crop yields have increased, partly through introducing new stress-tolerant varieties/hybrids with higher yield potential and partly through introducing more modern agrotechnical methods. As a result, countries whose populations have grown significantly in recent decades have become self-sufficient in cereals (e.g., India). In recent years, providing sufficient quantity and quality irrigation water for arable crops and improving precipitation utilization through breeding and agrotechnical practices has become increasingly challenging.

The technical conditions for the cultivation of cereal crops have changed considerably over the millennia. Remarkably rapid and significant changes have taken place since the 1950s and 1960s. These changes are continuing today. As a result of technological developments, the role of manual labor in cereal crop production technology has been radically reduced, while the role of skilled labor has increased. Today, cereal cultivation technology is based entirely on modern, high-performance machinery in developed countries. This has led to a significant increase in productivity. Modern machinery can be used for tillage, nutrient supply, sowing, crop protection, irrigation, and harvesting, i.e., the entire cereal production technology is highly mechanized. These observations are valid for cereal production in developed countries (and to a greater or lesser extent for domestic cereal production). Still, in many developing countries, even today, the extensive use of simple machinery, animal traction, and manual labor is very significant. In these countries, however, using the most modern techniques is limited not only by financial constraints but also by production traditions and employment and social considerations.

We strive to harmonize ecological, biological, and agrotechnical factors in cereal crop production technology. The level and quality of this harmonization determine yield, yield security, and yield quality. Knowledge of agroecological conditions is essential in this three-fold framework.

The agroecological conditions for growing cereal crops include climatic, soil, and topography. Weather-climate factors majorly influence the selection of the biological basis and the agro-technology used. The genotype and adaptation to different, usually extreme weather conditions are crucial in choosing a biological background. Weather conditions must be considered when designing agrotechnical operations for crop rotation, tillage, nutrient supply, sowing, crop protection, irrigation, and harvesting. Global climate change phenomena occur globally, with local effects also pronounced in individual regions.

An example is Hungary, where since the beginning of precise meteorological measurements in the 1870s, research has shown (Szász et al.) that the frequency of dry, drought-prone years for crop production has increased from 22.5% to 52%. At the same time, the frequency of wet years hardly changed (22.5% and 21.1% respectively), and the proportion of years considered average weather decreased significantly (from 55.0% to 26.3%). The surveys showed that drought damage was the most frequent (42.4%), while the frequency of damage caused by other meteorological factors (ice, water, and frost) was significantly lower (16-20%).

During their ontogenetic development, cereal plants grow vegetative and generative parts from the sown seed, and after fertilization, they produce a crop. The cereal species show very different developmental stages, with different phenophases and thus different growing seasons (e.g., maize 120-155 days, winter wheat 270-290 days). However, all cereal species have the same growing period, which can be divided into two major developmental stages: vegetative and generative. The length of the vegetative period and the time and duration of each phenophase within a given crop species are influenced and modified to varying degrees by environmental conditions (weather, soil), biological and genetic features (specific characteristics of the variety/hybrid), and various agrotechnical interventions (nutrient supply, sowing, crop protection, irrigation, etc.). Of particular importance are those phenophases that are decisive for yield production. These are called critical phenophases.

Different agrotechnical interventions (agrotechnical optimum) can influence the vegetative and generative performance of the crop species and its productive processes in each phenophase during the growing period of cereal crops. In the case of cereal crops, a distinction is made between biological production (the total aboveground plant mass, i.e., the combined mass of vegetative and generative parts) and proper output (usually the mass of the grain yield). The ratio of the two is called the harvest index (HI).

The following yield components determine the yield of cereal crops:

- the number of ears per unit area (pieces/m2)

- the number of grains per ear of cereals (number of grains per ear)

- the thousand-grain weight (g)

The following yield components determine the grain yield of maize:

- the number of plants per unit area

- the number of corn-cob per plant

- number of grains per corn-cob

- the thousand-grain weight

Cereal crops are utilized in many forms. Their main products and by-products are valuable, and their potential for use is expanding. The main product of cereal crops is - in general - grain, which is mainly used in human nutrition, animal feed, raw material in some industries, fuel (bioethanol), and seeds. Its by-products are also of value, mainly used as litter in livestock production and, to a lesser extent, partly after processing, as animal feed, industrial raw material, and energy. Due to the significant decline in the use of organic fertilizers, particularly manure, cereal crop residues, when worked into the soil, represent an essential source of organic matter for crop production.

The principles of cereal crop production technology have changed fundamentally. Quality is a critical objective in today's and future production technologies. A quantitative approach to cereal crop production characterized the past. The primary aim was to achieve maximum, record yields while ensuring only the minimum quality. In present and future crop technology, the objective is to achieve optimum yields and maximum quality in cereal crops.

4.1.1 Wheat

Wheat (Figure 2) is our largest arable crop in the world. It is grown in nearly 100 countries, mainly in temperate climate zones (oceanic, continental, Mediterranean). Its adaptability is demonstrated by its cultivation in subarctic, subtropical, arid, semi-desert, and desert areas. It cannot be grown only in arctic and tropical regions. Wheat is grown mainly in the temperate zone of the northern hemisphere. Its northern cultivation limits are Lake Winnipeg, the Gulf of Finland, and the middle section of the Lena River, and its southern limits are the Yucatan Peninsula, the Nile Valley, and the Deccan Plateau. In the southern hemisphere, it is grown on the Argentine pampas, as well as in South Africa and Australia. China is the world's largest wheat producer and the only country to produce more than 100 million tonnes annually. Over the last hundred years, the world's wheat production has undergone significant changes. Between 1900 and 1960, the increase in output was driven by extensive sources (increased area sown). In contrast, from 1960 to the present, the increase in volume has been mainly due to an increase in intensity (increase in average yield). Wheat is grown on about 220 million ha, with a total yield of ~680 million tonnes.

The main product of wheat is the grain, which has a particular morphological structure. The endosperm, the carbohydrate-rich component of the grain yield, accounts for the majority (~80%) of the grain yield. Although carbohydrates (~65-70%) comprise the central part of its chemical composition, proteins (~11-18%) are the most important for its utilization. The quantity and quality of their components (albumin, globulin, gliadin, glutelin) determine the quality of wheat products. Its by-product is straw. Both the main and the by-products are used in various ways. In addition to its food, the grain is also extensively used as animal feed. In human diets, cereals (mainly wheat flour) are inversely related to the country's development level. In developed countries, consumption is 60-90 kg per capita per year; in developing countries, 200-350 kg per year.

Figure 2. Wheat (source: internet)

Using site- and variety-specific models is vital in successful wheat production. Under less suitable site conditions, the weather and soil conditions are considered (site-specific model), while under favorable ecological conditions, the specific agrotechnical needs of the wheat variety are incorporated (variety-specific model). In practice, these two models are never used in their pure form, but a particular, complex application can make wheat growing profitable.

Winter wheat is exposed to a wide range of favorable and unfavorable weather events during its 260-290-day growing season. From a crop production point of view, the most essential meteorological factors are temperature, precipitation, and light. These factors have their optimum and minimum-maximum values for the growth and development of wheat, which vary continuously during the growing season, depending on the phenophases. Wheat is a crop with relatively moderate water requirements. This is demonstrated by the fact that it requires 480-550 mm of water over its long growing season, with most of its water uptake occurring from intensive stem growth to fertilization (mid-April to early June). Autumn is a critical period for water supply. Wheat can produce excellent yields in an ideal soil type with high humus content, neutral pH, medium moisture, good water, nutrients, heat, and air management. These are chernozem soils. The main wheat-growing areas in the world are located in landscapes characterized by chernozem soils. However, wheat is also very adaptable to soils.

The broader objective of soil preparation for wheat is to ensure favorable soil conditions for plant development during the growing season and the beneficial effects of other agrotechnical inputs and soil protection. The direct objective of soil cultivation is to provide a suitable seedbed for the wheat when sown.

Nutrient supply has a unique, central place in wheat production, as it directly and indirectly influences the performance of other agrotechnical elements. The wheat nutrient supply system is complex, including fertilization, management of soil nutrient resources, organic fertilization, organic materials recycling, by-product utilization, use of various soil conditioners, and other factors. The essential nutrients for wheat can be divided into macro-, meso- and micro-nutrients. Some macroelements (C, H, O) are taken up by environmental grain (water, air). The soil is the primary source of the other nutrients. Of all the nutrients, the most important is nitrogen, the most reactive nutrient, and its deficiency or excess is immediately visible. Deficient macro-nutrient supply can reduce wheat yields by 10-60% and inadequate meso- and micronutrient supply by 5-15% and 2-10%, respectively, and yield quality.

Integrated pest management is the basis for crop protection in cereals, including wheat. We aim to use various integrated pest management tools to minimize chemical control to increase agronomic and economic efficiency, protect the environment, and meet food safety standards. In the integrated pest management of wheat, chemical treatments should only be considered as a last option if other methods do not or only partially solve pest management problems. Integrated Pest Management for wheat includes the following main areas; the proper selection of the growing area, variety selection, chemical treatments.

Harvesting wheat is a significant part of farming technology. If harvesting is carried out properly, we can harvest the quantity of wheat produced without any loss in quantity or quality. Due to the extremely varied ecological conditions and the contrasting seasons in the northern and southern hemispheres, wheat is harvested every month of the year throughout the world.

No other cereal flour can make bread as digestible and quickly leavened as wheat. The use of grain and its meal also means that it is an excellent quality feed for fodder. The by-products of wheat grinding are also important. The resulting wheat bran is a protein-rich feed material. Its straw is a valuable litter material, which is increasingly used for non-food purposes, e.g., in pulp production or the agro-energy sector.

4.1.2 Maize

Maize (Figure 3) is one of the world's most important commercial crops, and it is used for food, feed, and industrial purposes. It occupies 16% of the world's arable land (161 million ha) and yields 820 million tonnes. Maize is grown in temperate regions with warm summers, mainly in the United States of America, China, the former Soviet Union, Central and Southern Europe, and Northwest Africa. There can be considerable variation in the use of the crop in different regions of the world. More than half of the world's supply of maize is used for animal feed. Maize is now cultivated almost all over the world, thanks to its good adaptive properties and the efforts of a targeted breeding effort. The US is the world's largest maize producer, with an average yield of over 9 t/ha on nearly 35 million ha. It is grown in a large area in China, Brazil, and Mexico, but the average yield is only 3-5 t/ha. France plays the most significant role in Europe, with 1.7-1.8 million ha at 10 t/ha, and Italy, with 1.1 million ha at around 11 t/ha. In both countries, irrigation also plays a significant role in achieving outstanding yields. The uses of maize, both as a primary product and a by-product, vary greatly. In the world, it is mainly used as energy-rich animal feed, but in developing countries and countries with food security problems, around 80-90% of the crop is used for human consumption. Its role in animal feed is mainly energy supply, with a 65-70% starch content. The growing demand for bioethanol production will further increase the area and yield of maize. In the future, the biological potential and the role of maize hybrids will be even more appreciated. In particular, hybrids with good adaptability and agronomic properties can respond to specific ecological and growing conditions.

In maize production, achieving the right quality is essential for the destination of the final product. As a characteristic of the feed value, the protein content varies on average between 7 and 9%. The genotype of the variety mainly determines the protein content but is also influenced by cultivation technology and ecological factors. There is generally a negative correlation between yield and protein content. Maize has the lowest protein content of all cereals and amino acid composition, with an oil content of 3-5 %.

Maize is the most sensitive of all cereals regarding soil quality and growing conditions. Although maize can be grown on different soil types, it can only produce outstandingly good results on deep, easily warmed chernozem, meadow chernozem, and brown forest soils with good water management.

Maize requires most of the mineral elements nitrogen, its potassium requirement is medium, and its phosphorus requirement is moderate. The requirements for Ca and Mg are not negligible. Among microelements, it is sensitive to Zn and Cu deficiencies. In addition to soil, weather, and agronomic factors, maize hybrids can significantly modify fertilizer demand and utilization.

Maize requires deep cultivation, high soil water/air ratio, soil water, and heat turnover. The quality of cultivation affects the yield of maize. Tillage has several objectives, such as the supply of nutrients and fertilizers, the achievement of good soil looseness, weed control, the direct control of pests, and, indirectly, the reduction of their reproduction and living conditions. Tillage allows water permeability and retention, promoting root development and rapid penetration through mechanical improvement. Chemical soil improvers can also be introduced into the soil by tillage.

Plant protection includes prevention, crop rotation, and the use of healthy seeds with good biological value. Soil disinfectants should be used to control soil-dwelling pests and American corn beetle larvae, worked into the soil before sowing. Weed control is the most crucial task in maize plant care and protection. Chemical weed control plays a key role in recognizing and respecting the need for complex work.

Maize can tolerate little water in the first period of its development (40-60 days); the same is true for the post-emergence period. An essential characteristic of maize for cultivation is its outstanding water absorption capacity. Its roots penetrate a deep layer of soil (200-250 cm), absorbing much of the moisture stored in the soil. Except in deplorable conditions, it makes good use of the water it absorbs. Compared to other crops, its peak water demand is long after July and August, so maize is very drought-sensitive.

Figure 3. Corn-maize (source: internet)

Harvesting of grain maize grown for fodder can start after biological maturity. The maize is biologically mature when the incorporation of nutrients into the grain crop is complete and when the maize grain reaches its best nutritional value for the growing conditions. The form of storage influences the time and method of harvesting. Grain maize can be stored at 14-14,5 % moisture content. When harvesting, the grain crop has a higher moisture content than this and requires artificial drying. It is vital to organize the harvesting of maize well to reduce drying costs and preserve the quality of the crop. Efforts should be made to ensure that, as far as possible, maize intended for dry storage in the grain gives off a more significant proportion of its moisture naturally in the field after biological ripening. The stalk, a by-product of maize, can also be used for fodder, heating, or working in the soil to supplement nutrients.

4.1.3 Rye

Rye (Figure 4) is the most important cereal in the cool climate of northern Europe. It is the second bread cereal after wheat in the temperate zone. It is the eighth-largest cereal crop in the world, with an area of around 16-17 million hectares. Most of the area sown to rye is in Europe. It is widely grown in Poland, Italy, Germany, the Czech Republic, Slovakia, and Denmark. Its importance is enhanced because it can be grown on soils with lower fertility where wheat is no longer economically viable. Rye is an undemanding cereal in environmental conditions, and its yield is much lower than wheat. However, rye is the most "universal" cereal. It is used primarily as a bread grain: its flour can be used to make good quality bread, either processed as a pure grain or mixed with wheat. Its by-products are also valuable; its bran is an excellent quality forage for tracts; its straw is suitable for litter and roofing and wrapping saplings. Rye is also a necessary green manure and fodder crop. Pure green rye is the earliest green fodder that can be fed. In sand vineyards and fruit plantations, it is the safest soil protection and green fodder crop to grow.

Figure 4. Rye (source: internet)

4.1.4 Oats

Oats (Figure 5) are a precious feed grain. Oats are the fifth most important cereal in the world in terms of production. In addition, around 15% of the crop is consumed by humans, making it the fourth most crucial food grain after wheat, Rice, and maize. Its starch content is easily digestible and a popular ingredient in infant food, but it is also an essential raw material for the breakfast cereal industry. Oats are well-tolerant of cool climates and are not soil-sensitive. It is mainly grown in the cooler temperate zones, especially in the post-Soviet states, the United States of America, Canada, Australia, and northern Europe. It is primarily grown in cooler, temperate countries. Most oats are grown in North America and Russia. However, oats are not only a valuable feed crop but are also grown for human consumption and food processing because of their specific chemical composition. The feed value of oats is excellent due to their high protein, starch, and fat content. It is rich in calcium and phosphorus, which promotes bone formation. But it also contains vitamin E, which increases the breeding ability of male breeding animals. Thus, oats are mainly used as a feed grain for young and breeding animals. Oats are also suitable for human consumption because of their valuable nutritional value. In addition to oatmeal, oat flour, etc., it is also used to make infant formulae.

Figure 5. Oats (source: internet)

4.1.5 Barley

Barley (Figure 6-7) is a very important and valuable cereal grown in various climates because of its high varieties. The environmental requirements of barley are more modest than those of wheat. It is grown in the temperate zone, its range extending north and south of that of grain. Barley is the fourth most important cereal in the world after wheat, maize, and rice. Most of the crop is used for animal feed; the second largest user is the beer industry, with only a tiny proportion going directly to the kitchen. Most of the beer is brewed from barley, or more specifically from the malt made from it. Barley malt is also the main ingredient in Scotch and Irish whisky/whiskey. In shops, you can buy hulled barley for cooking and, more often, a chopped and polished version of barley, known as barley pearl. In the north, where the summer season is shorter, it is a crucial bread grain; in the south, it is mainly used as a feed grain in warmer countries. It is primarily grown as a brewing barley in temperate regions with warm summers. It is the world's fourth most widely grown cereal, covering 90-95 million hectares. Barley is used for various purposes: as a precious grain feed, especially for pigs; as an essential raw material for brewing beer; and as a barley pearl for human consumption. For animal feed, winter barley is primarily grown. As beer consumption and production grow, so does the demand for malting barley. Spring barley is a valuable feed ingredient; even straw is suitable for animal feed.

Figure 6-7. Barley (source: internet)

4.1.6 Rice

After wheat, rice (Figure 8-9) is the most critical plant food for humanity, as 95% of Rice is used directly for human consumption. The area under cultivation is 130-140 million ha and is native to Southeast Asia. The grain yield of rice is precious, and it is very digestible. Its protein (8-12%) is almost complete compared to animal protein. Its starch is easily digestible and rich in vitamins E and B. As consumer habits evolve, so do food safety requirements. Today's agriculture is undergoing significant diversification in this area. In particular, while bread consumption is slowly declining in developed countries, it remains essential in developing countries. At the same time, wheat flour-based products are replacing rice consumption in some regions of Asia, particularly in urbanizing areas. In some developed countries, bread, especially white bread, has become less popular due to its carbohydrate consumption and low glycemic index. At the same time, fibre-rich flour-based products are gaining popularity.

Figure 8-9. Rice (source: internet)

4.2 Roughages

Economically grown fodder crops, their rational, waste-free use, and responsible management of fodder potential are crucial to the competitiveness of livestock production. Fodder crops are crops grown directly (for green roughage) or indirectly (as fermented and dried fodder) for animal feed. Fodder and roughage grown in arable land and grassland form the forage base for the animals that consume them. Whether green or conserved, quite a wide variety of crops can be used as roughage. All cereals (wheat, rye, barley, oats, triticale) and oilseeds (rape, sunflower) can be used as fodder up to a particular stage of development. The range of leguminous crops (Alfalfa, red clover, lupins, sweet clover, vetches), silage maize, moss, millet, and Sudan sorghum are used explicitly as fodder crops. Farmers choose from a wide range of crops based on the following criteria: considering the species of livestock and their feed requirements, such as soil and climate conditions, technologies to be used, cost price, and land requirements.

The limited land available means that arable fodder production must be subdivided into maincrop and dual cropping. The shorter growing season of fodder crops allows them to be grown as pre- or post-crops and harvest two crops in the same area in the same year. For example, early-harvested purple clover and rape can be used to sow maize for bait-corn, or mohair, millet, and turnip can be used to harvest early-planted linseed in the autumn. The lower yields and, thus, higher costs of forage crops suitable for double cropping can be offset by the land release effect for the production of cash crops.

4.2.1 Alfalfa

Alfalfa (Figure 10-11) is a fibrous fodder crop with a high protein content and yield per unit area. The protein in Alfalfa is of high biological value. In addition to protein, Alfalfa is rich in calcium, other minerals, and vitamins, especially carotene. It is also essential as an energy source; although it has low starch and sugar content, its digestible raw fiber content is critical in ruminant feed. In addition to its direct feeding importance, Alfalfa can be used in biological soil cultivation and also benefits soil fertility. Its well-developed root system enriches the soil with nitrogen and organic matter. Alfalfa is a fully mechanized crop well suited to on-farm cultivation, where its excellent adaptability allows it to be grown safely under both irrigated and dry cultivation. The livestock sector uses Alfalfa in various ways: fresh and silage fodder, dried as hay, green meal, or granules. Alfalfa is grown all over the world because of its importance and adaptability. Most Alfalfa is grown in the USA, but it can be found in every European country where the conditions are appropriate.

Figure 10-11. Alfalfa (source: internet)

4.2.2 Silage maize

Silage maize (Figure 12) is typically used for livestock production, so production costs are of primary importance. It is rarely marketed, and prices vary locally depending on supply and demand. In recent decades, the area sown to silage maize has increased dynamically worldwide as the number of ruminants in all countries where temperatures allow silage maize to be grown. Silage maize is a staple crop for arable mass feed production, where precipitation deficits limit grass production. In the continental regions of Europe, the potential for maize production, and through it silage maize production, and its primarily high quality, represents a significant potential advantage over areas outside the maize belt (e.g., the Netherlands). However, this advantage started to reverse around 2010. One of the biggest challenges facing the production of bulk feed is the contamination of the crop by fungal toxins due to climate change. Fungal strains, previously a problem only in the subtropics, have also appeared in the continental zone. Although they were not yet of significant importance when they appeared, around the turn of the millennium, they caused massive damage to maize crops in 2012, especially in drought-stricken areas.

Figure 12. Silage maize (source: internet)

4.2.3 Grass

The grass (Figure 13) is the cheapest and healthiest source of fiber fodder for ruminants. Grassland management can be either extensive (twice a year), semi-intensive (three times a year), or irrigated-intensive (four times a year), adapted to the needs of livestock production. Dairy production is closely linked to grassland management in Europe's grassland-rich areas. Elsewhere, farmers rely on maize and concentrated feeds. In continental climates, grazing is mainly practiced on smaller farms with lower genetic value.

In the EU Member States, the share of grassland in total agricultural area varies widely, from around 70% in Ireland and 65% in the UK to just over 1% in Finland and none in Malta. In Central and Eastern Europe, the share of grassland is highest in Slovenia, at 57.3%, and lowest in Hungary, at 14.2%. The region's grassland share could be better than that of the leading competitors (Poland, the Czech Republic, and Slovakia). Except in Serbia and Romania, there is a robust correlation between milk production and the size of the grassland area (Figure 6.8). The average number of ruminants per hectare of grassland, expressed in livestock units (LU), is highest in Slovenia, Poland, Croatia, and the Czech Republic. If only dairy cows were considered, the order would be Poland, Croatia, Slovenia, and the Czech Republic. In the case of Hungary, both indicators are lower than the average for the countries in the region, i.e., our grassland has fewer animals. While the length of the grazing period and grassland yield depend on temperature in the northwest, southern and eastern Europe depend mainly on soil moisture, especially in summer. The production potential of grassland in the Carpathian Basin is among the lowest in the European Union. This comparative competitive disadvantage, coupled with the predominance of extensive grassland management and the failure to exploit the potential of ecological conditions, which is modest but not negligible, is reflected in the low milk production per unit of agricultural area.

Figure 13. Grassland (source: internet)

4.3 Oilseeds

In the past, oilseeds were grown mainly for their oil (hence their name); the meal and the husk were merely by-products. Today, however, this has changed. This can be explained by the increased importance of protein in the feed industry, especially after the ban on animal protein in feed in Europe and the significant growth in GDP in emerging countries.

4.3.1 Soybean

Soybean (Figure 14-15) is an annual herbaceous crop of the legume family native to East Asia. The World Food and Agriculture Organization (FAO) classifies it as an oilseed because its oil content can exceed 20%, 95% of which is food oil. 98% of the dry matter after oil production is used in animal feed. Only a tiny proportion (about 2%) of all soybeans produced is consumed directly for human consumption. Soy straw can also be used as feed, bedding, fuel, and fiber.

Of all oilseeds, soybean stands out, accounting for nearly 60% of world vegetable oil production. Soybean supply and demand conditions determine world market trends. However, it is also worth noting that soy is an oilseed and a significant protein crop. Global soybean yields have increased tenfold in fifty years, but production growth has not been uninterrupted. Between 2014 and 2018, global soybean yields increased in three years and decreased once compared to the previous year, rising by around 14% overall, while average yields increased from 2.6 tonnes/hectare to 2.8 tonnes/hectare in five years.

Nearly 86% of production comes from the Americas and 11% from Asia, while Europe's share is only 2%. The largest soy producer is the United States (1/3 of world production), followed slightly behind by Brazil; one in six tonnes of soy is produced in Argentina (16.7%).

The growth in the importance of soybeans has been fuelled mainly by industrial purposes, with the chemical and food industries starting to use them as raw materials due to their high oil content. Nowadays, from a nutritional point of view, soy has become a divisive crop as a food or feed ingredient. Numerous studies have been carried out on its consumption's benefits and potential harms. There currently needs to be a consensus on its assessment. Still, no one side disputes that it has a relatively high crude protein content, consisting of valuable amino acids, and a high oil content, which is not an average fatty acid composition. It contains high levels of omega-6 fatty acid, also known as linoleic acid, which is essential for the human body. Soybean oil is also vital for the feed industry, both as a raw material for producing energy-rich feeds to meet the needs of farm animals and as a feedstuff for improving palatability and granule quality. It is also used extensively in the food industry to produce cooking oil, margarine, and bakery products. In the chemical and pharmaceutical industries, it is used as a starting material for producing surfactants, diluents, emulsifiers, soaps, detergents, plastics, dyes, synthetic oils, and various phosphate concentrates. The soybean oil extracted from the industrial processing of soybeans is widely used, and the extracted soy meal, a by-product, is still one of the most essential feed materials for the animal feed industry, as it is an excellent protein feed. Significant genetic progress in livestock production and the need to meet the nutritional requirements of intensively reared breeds have led to the use of feed components with high biological value, high protein content, and a favorable amino acid composition, increasing the feed's energy content. The use of alternative feed materials is also justified by the recent requirements of the processing industry to feed animal protein-free feed, particularly in the poultry sector. All these criteria are fulfilled by full-fat soybeans processed by heat treatment. The so-called anti-nutritional substances in raw soybeans, which inhibit trypsin function in the digestive tract and thus the utilization of proteins, can be inactivated by heat treatment. It is widely used in feeds for various animal species, primarily poultry, especially turkeys and broilers, but is also excellent as a component of piglet and sow feeds. It is also an effective feed material for other animal species, whether rabbit feed or dairy cattle.

Soybeans are best grown in deep soils with favorable air, heat, and water management and a pH of around neutral. Soils that can successfully grow maize are also suitable for soybeans. Soybean benefits the soil; its roots improve soil structure, and its root and stem residues increase soil organic matter and nitrogen reserves. Soybean is a nutrient-demanding crop. Fertilizers fully meet their nutrient needs. Although soybeans are heat-sensitive, they tolerate early sowing well when the soil temperature at the sowing depth reaches 8 degrees Celsius sustainably. During the flowering period, soybeans require

160-180 mm of water, which can mostly only be met by artificial watering. In irrigated stands, protein content usually decreases, and oil content increases.

When dried and stored as seed, harvesting takes place when the pods are brown, the shape and color of the seeds are typical for the variety, the moisture content of the seeds is 14-17%, and the foliage of the plants has already yellowed and fallen off. Harvesting at 13-14% moisture is ideal. After harvesting, seeds with lower moisture content should be pre-cleaned before drying and then gently dried. Soybeans can be stored at a moisture content below 12%.

Figure 14-15. Soybeans (source: internet)

4.3.2 Sunflower

The most significant tradition of sunflower (Figure 16) cultivation is in Europe, Asia, and the Middle East, where it has spread to the New World. Today, it is also grown in the United States and South America. Since the turn of the millennium, sunflower seeds have accounted for 7-9% of world oilseed production, while its primary product, sunflower oil, has accounted for 8-11% of world vegetable oil production. However, sunflower seeds are not only grown for their oil content. In the United States, for example, a quarter of the total output is sold as birdseed, and 10-20% is consumed directly by the food industry each year, with only the remainder being used for pressing. Sunflower is one of our most critical arable crops because its valuable oil is an essential staple in our diet. Sunflower oil is made up of 85-91% unsaturated fatty acids. The fatty acids are 20-50% oleic acid and 50-70% linoleic acid. Linoleic acid is essential for the human body. Its value is enhanced by vitamins D and E and provitamins dissolved in the fat. Sunflower oil produces margarine, pesticides, detergents, cosmetics, and fine art paints. The so-called oil cake, which remains after the extraction of the sunflower oil content, has a protein content of 48-50% and a high mineral salt content. It plays a vital role in the feeding of our farm animals. Sunflowers are an excellent honey-producing crop, with bees collecting 20-30 kg of nectar per hectare of sunflowers during flowering. The requirements for sunflower varieties are stable yield and yield security; good adaptability to extreme soil and climatic conditions; high oil content of the crop; good shape and position of the plate; uniformity of the sunflower stand; firmness of the stem; high drought tolerance; good nectar and pollen production; high oil quality and protein content; plant disease and insect resistance. Sunflowers have a high water requirement, with most of their water taken from the soil through the roots. After a drier growing season, the soil beneath it dries out. It is a heat-demanding crop. Sunflower crops are artificially dried before harvesting, and harvesting can be done with harvester diggers.

Figure 16. Sunflower (source: internet)

4.3.3 Rapeseed

World rapeseed (Figure 17-18) production has increased twenty-fold in the last five decades. Initially, the increase was slight but steady, and from the 1990s onwards, it was more dynamic but with several declines. Between 2009 and 2013, production fell in three years and rose in only one year compared to the previous year. Over the five years, global production increased by only 0.7%. The continental share is even greater than that of palm oil and soya. Europe accounts for 37% of global production, Asia 34%, and America 25%. The central countries producing rapeseed are Canada (25%), China (20%) and India (10%). In Europe, Germany produces the most (8% of world production), followed by France with 6%. The global average yield mainly was the same between 2014 and 2018: around 2 tonnes per hectare. Rapeseed oil has been used for a long time, all over the world, mainly for lighting. This was the period in its history when its cultivation was most important, as reflected by the size of the area under its production. This role has been lost with the advent of newer and more modern energy sources, but its versatility means it remains an essential arable crop. The production of edible oil is the primary production purpose. The extent to which it is used in food is determined by consumer habits and the quality of rapeseed oil. The development of rapeseed oil refining technology has opened up the possibility of producing margins. It is also used in significant quantities in the production of paint and soap, in the tool and machinery industry, in metallurgy, and in the manufacture of various cosmetics, pharmaceuticals, and explosives. Rapeseed can be used in animal feed for a variety of purposes. The rapeseed and rapeseed meals left over from edible oil production are a high-quality, protein-rich feed. It is also eagerly consumed by animals as green fodder. Its traditional use is as a green manure, where the green parts and the crop's root system are turned into the soil and improve the soil structure. Its cultivation does not require any special machinery, and the agricultural machinery used in cereal production can be used for agronomic tasks.

Figure 17-18. Rape (source: internet)

5. Livestock production

The primary purpose of farm animal production is to provide raw materials for meat, dairy, and egg production, as well as for the textile and leather industries worldwide. For many centuries, livestock farming was an auxiliary sector of crop production - work was done by animal labor. Even today, the two activities are generally linked, with livestock helping crop production, for example, by producing organic manure. However, the link is also reversed, for example, in producing fodder crops. For example, almost half of the agricultural area of the European Union is used directly for livestock production, either as permanent pasture and mowing or as areas sown to fodder crops. In addition, half of the cereal harvest is used for animal feed. In addition, by-products of crops processed by the food industry are often returned to the sector for animal feed. The aim of livestock farming in the narrow sense is, therefore, to produce animal products at a profit. Animals are kept for fundamentally different purposes in different regions of the world. All animal products are used in traditional livestock farming, especially in Africa and Asia. Traditional livestock production is based on locally produced feed. This system had virtually no waste because everything could be used for something. Modern animal farming systems have a primary product, or twin products, to which production is directed. In a dual-purpose system, two products have greater weight: beef and dairy cattle, wool, beef and dairy sheep, and beef and dairy goats. In tropical areas, cattle rearing for meat and pasture is still practiced. In large-scale meat production, the production process includes the operation of slaughterhouses, meat processing, and the rationalization of sales and marketing. The main objective is to maximize profit by reducing costs, often involving minimizing live labor, automating the production process, and standardizing products.

Over the last fifty years, the world's population has grown from three billion to almost seven and a half billion, and global meat consumption has risen from 70 million to over 300 million tonnes. In order to improve the efficiency of animal product production, hybrid animals are typically created through breed improvement to achieve high levels of reproduction and meat production. The production of most staple foods, including meat, has increased faster than population growth. After five decades of steady and uninterrupted growth, world meat production totaled 342 million tonnes in 2018. Of this, pork accounts for 36%, poultry for 35%, and beef for third place at 21%. The dynamics of production were temporarily disrupted by the crisis and animal diseases. The epidemics led to market disruptions and affected international trade and consumption of meat. However, with increased production and globalization of the meat market, access to meat products has become more favorable. With the improvement of the economic situation, there has been a shift in the population's food consumption towards meat products within certain limits. However, there are also significant differences in meat consumption between regions at the global level. People in developed regions consume 80 kg of meat per year, compared with 32 kg in developing countries. World population growth and the rise in per capita GDP, among other factors, are driving an increase in meat production, particularly in developing countries.

In recent years, meat production of all animal species has increased but at a decreasing rate. Pig and poultry meat account for 72% of global meat production, and this share is increasing because pigs, especially poultry and even more so fish, can gain a kg of live weight with less feed (Figure 6.18). World meat production has moved in this direction - and is expected to increase by 60% to 470 million tonnes by 2050. At the same time, per capita meat consumption is projected to rise by 25%, from 42 kg to 52 kg. Poultry and pigs require an average of 2 kg and 3 kg of feed, respectively, while sheep and beef require more than 3 kg and 7 kg of feed per kg of live weight gain, and ruminants must also be fed on roughage. It is no coincidence that the proportion of meat production has changed accordingly over the years, to the detriment of beef.

So the efficiency of producing each type of meat differs significantly: poultry meat requires less energy and water and less grain than beef, but aquaculture has also grown enormously in recent times because fish meat competes not only with beef but also with pork and poultry meat - and better feed sales can save permanent grassland and arable land. In the production of animal products, using the caloric content of energy-bearing feed is not sufficiently efficient. On average, six kg of feed are needed to gain one kg of live weight. A study of the efficiency of food production shows shocking figures: even the most efficiently produced chicken meat can only produce 700 kilocalories of energy from 2 kg of feed, i.e., 8,000 kilocalories.

The growth dynamics in total meat production have been best outperformed by poultry and pork production. This has led to a change in meat production for consumers' benefit.

At the same time, the substantial expansion of poultry promotes cheaper animal protein supplies due to its favorable feed conversion efficiency. Nevertheless, the rise of animal-based diets in the human diet is a cause for concern, as they require more resources (feed, water, land) than the foods they would replace. Livestock farming pollutes the environment through land and water use, nitrogen emissions, and greenhouse gas production. The increase in meat production must, therefore, be designed in such a way as to maximize the resource requirements and environmental impact per unit of product, as this is the only way to sustain the expansion of production in the long term, both economically and environmentally. The place of meat in the diet should also be reconsidered, especially where meat consumption is already high. Ethical considerations and food security challenges, particularly in cereal use, raise the question of whether increasing meat consumption is justified.

As the world's population grows, demand for animal products, and meat in particular, continues to increase, but this is closely linked to the increase in the number of people living at a higher standard of living, particularly in Asia. However, the rapid expansion of the world population and the rapid development of economies in emerging countries in recent decades has significantly increased the production of animal products and made it more efficient.

Almost 10% of world meat production (around 30 million tonnes) is exported to international markets. Poultry meat accounts for 44% of all meat traded worldwide, beef for 27%, pork for 25% and sheep for only 3%. Shortly, the pork share is expected to decrease; sheep meat is expected to stagnate, while beef and poultry meat are expected to increase by almost 4%. China, Japan, Russia, Mexico, and the United States are leading meat importers. Exporters include the United States, Brazil, the European Union, China and Canada.

The world is almost universally turning to meat consumption, which is low in fat and energy but juicy and lean. The distribution of meat types also varies considerably between geographical areas. There are also differences in the technologies, breeds, and intensity of the sector.

Livestock production is much more constrained than crop production, and there is limited scope for changing the pattern of inputs. Livestock yields depend mainly on the breed and species kept and farming techniques, but the influence of weather could be more pronounced. Over the last fifteen years, the production structure of farm crops in most countries with developed agriculture has changed significantly. Large-scale, specialized livestock farms have become the dominant feature and are now geared to mass production and expanding markets. It is well known that the trend in agricultural price relations, over a multiannual period, is one of producer prices generally rising less than inputs. The increasing role of mass production and the extraordinary rate of development in some areas of the world are dictating a significant growth rate in world meat production (despite resource constraints), even today. When analyzing well-functioning economies, the striking differences are not between farming forms or sizes but along different levels of farm management. Success lies in organization, expertise, possible integration, and good management.

Animal product production is an economic activity in which the primary aim is to produce an animal product profitably at the lowest possible cost, considering the behavioral characteristics of the animal species on the farm.

It should be borne in mind that the farming technology of each farm animal species differs in many respects according to their species specificities and their tolerance to technology. A common feature of large-scale farming systems for our farmed animals is that they are under human supervision and within the limits set by the species/breed and the direction of use.

General characteristics of large-scale livestock farming:

- there is one species of animal on one farm
- large number of people
- concentrated, high levels of animal product emissions
- a permanent fence encloses the livestock farm. The so-called black-and-white principle must also be applied to the movement of persons and vehicles. This means that persons arriving on the premises must change their clothes in the white changing room and, after scrubbing, enter the holding area. Vehicular traffic passes through a designated gate where vehicles must pass through a disinfection basin; the vehicles are washed with disinfectant
- the above requirement is for animal health prevention; given the high concentration of animals, introducing any infectious disease into livestock could cause severe economic damage. In order to prevent the emergence of pathogens that pose a risk to the species, preventive medication is administered according to a strict protocol
- a high degree of mechanization, but this varies from one species to another: most poultry are kept in completely enclosed buildings throughout their production period. In swine breeding, some age groups must be housed in enclosed sheds with a run (e.g., pregnant sows). In cattle breeding, meat breeds may be kept outdoors, in pasture-based housing, or semi-enclosed buildings. Wherever possible, farmers shall seek to automate the operation of all elements of automated housing technology (heating, cooling, ventilation, watering, feed distribution)
- run by managers and workers with specialist knowledge: The farm is run by people with a degree in agriculture or animal husbandry. Specialized qualifications are also required to carry out certain specific tasks (e.g., artificial insemination)
- building and herd rotation is a feature of large-scale livestock farms: the all-in-all-out principle is applied: animals are moved from one building to another (to the slaughterhouse for finished products or to another building for breeding animals). They are moved to another building at the same time. There are several reasons for this: firstly,

for reasons of work organization and efficiency; secondly, moving animals together means less breakage for group-housed animals; and thirdly, it is also easier to plan for the highest possible utilization of the buildings on the farm

- It follows from the above that efficient production and a steady income can only be achieved if animals are kept in the buildings all year round. Thus, the animal product per unit area is the highest.

Five Freedoms - Making animal welfare a reality

Large-scale industrial livestock systems have been heavily criticized for poor animal welfare in recent years. It is important to note that food production is about producing animal and plant foods and raw materials. Achieving animal welfare is essential because poor treatment directly affects the animal and indirectly affects the quality of the animal product. Therefore, modern animal farming practices take the needs of animals into account to a large extent and are regulated by law.

To do this, you must comply with the following requirements:

- a life free from thirst and hunger, which is ensured by the constant presence of fresh drinking water and a balanced diet for a healthy life and good physical condition
- a life free from inconvenience, which includes shelter from the elements and other hazards, as well as a comfortable place to rest
- a life free from pain, injury, and disease through preventive measures, rapid diagnosis, and treatment
- natural behavior is achieved by providing sufficient space, appropriate housing conditions, and the presence of conspecifics
- a life free from fear and distress can be achieved by providing housing conditions and treatment that avoid the animal's emotional suffering

Brief description of the farming technology of large-scale livestock production:

Technological characteristics of free-range:

The livestock is kept outdoors all or part of the year, in small to large groups, in an informal way. Open-air, pasture-based, or extensive farming is close to nature. It is particularly beneficial for young, developing animals (heifer rearing or drying out in cattle farming). However, it must also consider the extremes of weather, as the animals are in the open and must be protected from the weather. Since outdoor grazing is mainly used for ruminants (cattle and small ruminants), pasture is primarily a source of fodder for them. The pasture's grass yield and regeneration capacity determine the number of animals that can be kept on it. This also means that pasture maintenance is essential and that a reliable or electric fence must be built. When pasture grass yields do not meet the nutritional needs of the animals, suitable space must be provided for the distribution of supplementary feed. In outdoor grazing, parasites may be present on the pasture vegetation and the soil surface, and care should be taken to prevent parasites. One of the advantages of confined housing is that high-value assets, such as the animals themselves, are kept in a visible location. This asset protection is more challenging to achieve in open-air housing. It is also more difficult to collect data on the microclimate and immediate environment of the animals and the individuals of the animals in an outdoor environment, which is less controllable by humans.

An interesting fact: pigs kept outdoors

Outdoor pig farming is a centuries-old tradition in many European countries. From the Middle Ages onwards, in England, when following was a standard method of resting arable land, pigs were driven onto land not used by crop farmers. This was a mutually beneficial business for both the farmer and the pig farmer: the pigs not only loosened the top 25-30 cm layer of soil (helping to aerate and loosen the soil) but also consumed soil pests and weeds and their seeds - providing what is known as organic crop protection. Thirdly, they replenished the soil's organic matter content with their manure, which also helped to replenish nutrients. This had a measurable positive effect on crop yields the following year. To achieve this positive effect, the density of pigs in the fallow area had to be considered. Nowadays, there are fewer of these pigs in England, but they are kept outdoors for other purposes. This is a more extensive type of farming, with slower growth rates and less sensitivity to environmental influences.

A good example is the Mangalica, one of our most critical Hungarian products, and the Iberian pig in Spain, the raw material for the famous Serrano ham. Customers are willing to pay a higher price for these premium meat products because this farming method focuses on quality rather than quantity. The low level of automation in free-range conditions is also a feature here. The pigs are kept outdoors all year round, and at most (but not permanently), the sows are moved to a sheltered area for farrowing or are housed in a 'farrowing hut' in the pasture. The great advantage of this rearing method is that it allows the pigs to express all their natural behaviors: wallowing (foraging and comfort wallowing) and wallowing. Comfort farrowing is an interesting behavior: it is not foraging but welfare behavior. Pigs have a "sense of smell," which means that when they are wallowing, they not only use their nose to collect odors but also use their mouth slightly open to search for attractive odors that suggest food. Another exciting behavior unique to pigs among our mammalian farm animals is the pre-nesting behavior of sows. A few days before farrowing, the sow finds a secluded place and builds a nest for herself using the materials available to her (dry grass, small twigs, and straw). This activity is carried out in complete concentration; the sow carefully builds the nest, burrows into it, and remains there until the end of the farrowing period. The piglets are thus born in a sheltered, warm place and return with their mother during the first days of lactation. This behavior is so deeply encoded in the sow's genetic makeup that the scratching and curdling movements can be observed even when the sow is kept in a closed enclosure on a grid. In free-range systems, human supervision of pigs is minimal. Therefore, it is not easy to monitor the movements and individual behavior of the herd as it is for ruminants. Care must also be taken to avoid damage to wildlife in pasture pig herds. This includes the entry of boar boars with a high interest in breeding sows into grazing areas and the escape of breeding sows. Wild predators such as foxes can cause economic damage to pig farmers by taking piglets. Diseases spread by feral pigs can also cause problems, such as African swine fever today.

A brief description of the semi-free or semi-enclosed system:

- an open-air runway belonging to a stable building
- the floor is earth or concrete
- a fence encloses the outruns
- animals are free to roam in and out for most of the year
- in cattle and pig farming, mainly in breeding (rearing of breeding pigs and sow breeding)

Characteristics of confinement, intensive systems:

This farming technology is characterized by producing meat products in a "meat factory." The animals are kept in a closed building, which is an entirely human-controlled environment. It follows that farming technology is highly automated, with automatic feeding and feeding systems supplying the animals with nutrients according to parameters defined and expected by man. The automatic watering system ensures that fresh drinking water is always available to the animals, and the design of the waterer and the water flow rate is adapted to the needs, age, and utilization of the livestock species. Automatic ventilation technology ensures that the necessary oxygen is supplied and the air is expelled. The natural-to-artificial light ratio is regulated according to the animal species. These enclosed "artificial" conditions differ significantly from the animal's natural environment. This is where the technological tolerance mentioned earlier is essential. To illustrate this with a simple example: the world's pig farming is dominated by enclosed, intensive farming (95%), in which intensive pig breeds and hybrids (crosses of two breeds) are kept, which can cope well with this farming and intensity and perform well. In the second module, these pig breeds will be presented. However, Mangalica, for example, is an extensive breed. It should not be reared in such conditions because its technological tolerance is very low and becomes aggressive and challenging to handle. The primary justification for intensive, closed production systems is to produce the maximum product per unit area, i.e., the most efficient production. This type of farming achieves the highest yield per unit of input. This is the opposite of free range, where quality is the primary consideration, not quantity. However, other market segments are also targeted by animal products from confined and free-range farming.

Depending on the type of closed production system, the barn floor is covered with litter material (litter and pre-rearing) or lattice and concrete. In laying hen rearing, cage rearing is (still) typical.

The all-in and all-out system technique mentioned in the general characterization of industrial livestock farms can be applied in confined housing. Even in closed housing systems, animals are kept in groups wherever possible to comply with animal welfare rules. Individual housing is possible where justified, for example, in the case of breeding pigs. However, it is not advisable to keep them together, as they may become aggressive to the extent that could cause economic damage (the attacked and weaker animal may die).

Given the high level of automation and the drive for the most efficient production, intensive feeding and lighting programs are typical of confined housing systems. Contentment can minimize animal health risks. This farming method is mainly used for poultry and pigs.

Technology for large-scale cattle farming - dairy cattle farming

Bound hold:

- one of the least natural ways of keeping
- high labour demand
- modern feeding is difficult or not feasible
- given the fact that cattle cannot move freely in this way, the horn does not wear out and must be groomed from time to time

Unbound, low-lying stance:

- more natural than a tied stance
- the need for the animals' natural movement patterns (e.g., standing/lying down, changing position) is usually provided
- well-maintained bedding provides a pleasant lying surface (dry bedding!)
- high litter demand (which increases production costs)
- high manure storage capacity demand (at the same time, soil fertilization, see for the crop sector, the first module, the complexity of agriculture!)
- high labor intensity - although it can be well-mechanized

Casual, relaxed boxing:

- in addition to group paddocks, paddocks with resting stalls are sometimes built
- this reduces the need for bedding straw and is also beneficial for later loose cow rearing
- difficulties with the size (dimensions) of the pens for animals of different weights and ages
- the installation of resting stalls of different sizes can only be justified in large heifer-rearing farms because of their cost
- place a grid or an adjustable neck trellis at the front of the booths to adjust the "length" and size of the booths.

United, notable purpose stables:

- dry cow barn - although the animals are not producing milk, it is not a non-productive period! Regeneration, preparation for calving, and the next lactation
- preparation barn - ensuring the transition between the dry period and the lactation period
- stables - hygiene is the most critical aspect, requires constant supervision - responsible, well-trained staff
- Receiving shed - After calving, maximum comfort and a stress-free environment should be provided, with constant supervision

5.1 Ruminants

5.1.1 Dairy and beef cattle

Most cattle-produced products are essential for human nutrition, which is why they are farmed in every country. Dairy production is the primary form of utilization. Milk production is significant because feed transformation is most beneficial for the well-milking cow. Most (83%) of the world's milk production is from cattle. Buffalo milk production is also significant (13%), while other animal species play a much smaller role in milk production: goats account for 2.3%, sheep for 1.3%, and camels for 0.4% of global production. World production has been increasing slightly for all dairy species in recent years. The cattle population has increased steadily over the past five decades, albeit at a not particularly spectacular rate, and by around 58% globally. The upward trend has been maintained in recent years, although the expansion rate has slowed markedly, reaching only 3% between 2014 and 2018. Today, the world's registered cattle population is close to one and a half billion head.

The proportion of these livestock used for dairy and slaughter production varies between continents and countries. The type of utilization is usually closely linked to the breed. The Holstein-Friesian breed is the most important worldwide in milk production.

However, there are also world breeds of concentrated milk production (the so-called Channel Island breeds) and local landraces of local importance, albeit in smaller quantities. Among these are the Simmental and the Borzderes, essential in Germany, Austria, and Slovenia. While global milk production is determined by the volume of cow's milk, buffalo milk production has doubled in the last two decades, and its share of global milk production has increased and is now stabilizing. Over the past fifty years, total milk production has increased steadily at 2.5 times while cow's milk production has doubled to 2.2 times. This also shows that the most dynamically growing buffalo milk production plays a slightly more significant role in today's global milk production.

Almost two-thirds of world milk production and related milk processing is concentrated in Europe and North America, with higher specific yield levels. Developing countries account for about half of the world's dairy cows but only about 20% of global milk production. The European Union produces a quarter of the world's cow's milk. Milk consumption varies considerably between countries around the world. Per capita milk consumption is exceptionally high in Europe and North America but very low in Asia, although the trend steadily increases here.

Three-quarters of the world's dairy cows are kept in households, typically one or two cows, producing 30-40% of all milk. Larger family farms hold nearly a quarter of the world's cows, averaging 10-30 per farm, producing 40-50% of all milk. Larger farms hold less than 1% of the world's cows but account for nearly a fifth of all milk produced. Between 93-94% of the milk produced is shipped for processing. The remaining 6-7% is used for consumption or sold directly. It can be seen that the trend in farm size growth is also towards commercial farms in the developing world, as they are much more efficient and productive, even than family farms with 10-30 animals. The number of farms is decreasing while the number of cows per farm is increasing.

The importance of the dairy economy, which embodies the triple unit of milk production, processing, and marketing of dairy products, is incomparably more significant than the combined monetary value of the products it markets worldwide. Its outstanding importance is due to its functions in the national economy. First and foremost, it produces the four primary food groups (milk, meat, cereals, fruit, and vegetables) that are the most basic and irreplaceable and are the 'first among equals.' Recent research has also shown that milk and its products are the most decadent foods in so-called health-promoting bioactive ingredients.

Figure 19-20. Dairy cattle (source: internet)

Over the past five decades, world beef and buffalo meat production has more than doubled, with steady, balanced growth. Production has been virtually stagnant in recent years, increasing by only 5% in five years, and in 2018 it reached 67 million tonnes. Beef production is expected to decline slightly in the coming years. America accounts for 46% of world production, a quarter of it comes from Asia, a sixth from Europe, while Africa contributes 9% and Oceania 4%. If we look at the top producing countries, the United States is first with around 18%, Brazil second with 14% and China third with 10%. Argentina also plays a significant role, with an annual production of nearly 3 million tonnes, but is also one of the world's leading producers in consumption. The world's cattle population has grown steadily, at around 1% per year over the past decade. Growth is remarkably rapid in India, Brazil, China, and Ethiopia. India has the largest cattle population. The European Union has 6% of the world's cattle population, and the number of cattle is declining. Cow herds are usually fed on extensive grazing systems, and their offspring are either sold directly for slaughter after weaning or sold to farmers who fatten cattle in a more intensive production system before slaughter. The most recent data shows that the EU suckler cow population has stabilized at around 12.4 million animals over the last two decades. Today, the most common beef breeds are Limousin, Charolais, Aberdeen Angus, Hereford, Blonde d'Aquitaine, and White and Blue Belgian. In the Americas, the most famous of these is the Aberdeen Angus, which has a long tradition of breeding and is usually advertised in restaurants in capital letters when a chef serves it on steak or hamburger.

With an annual production of nearly 12 million tonnes, the United States is the world's largest beef producer. Its economic strength is reflected in the fact that it accounts for only 5% of the world's population but represents 18-20% of global beef production and consumes 17% of world production. On the other hand, it accounts for only around 6% of the world's cattle population. Brazil's cattle population has increased 3.78-fold over the past 50 years, while the last two decades have seen a one-third increase in numbers. Today, it exceeds 213 million head (population 203 million), and annual meat production exceeds 9 million tonnes. By 2018, Brazil's production was close to that of the United States.

In countries outside the United States and the European Union, beef is produced under less intensive conditions, and the animals are not fattened to a high final weight. In the European Union, fattening systems can be divided into two main categories: intensive stall-based and pasture-based fattening systems with winter housing. Within the EU, these systems also differ markedly according to the feedstuffs produced, the climates, and the needs of the different cattle breeds. The main groups of breeds in use are:

- Dairy breeds, where the main product is milk.
- Dual-purpose breeds, where both milk and meat are produced.
- Beef breeds, where the main product is the fattening animal.

Cattle feed utilization for meat production could be more favorable. On average, 11 kg of feed is needed to produce one kg of beef, 7 kg of feed for one kg of pork, and only 4 kg for one kg of poultry. Of course, it must also be considered here whether the animal species is a roughage consumer or a bulk feeder. Beef production is the most energy-intensive of all meat types, generates significant amounts of greenhouse gases, and is the most land-intensive. It is thought-provoking that producing 1 kg of beef requires 16,000 liters of water, several times more than the water required for broiler chicken, pigs, eggs, and dairy production. US researchers have calculated that beef production has almost ten times the environmental impact of other meats.

It should be noted, however, that ruminants convert feed unsuitable for human consumption into human food. Beef production in the European Union is closely linked to the dairy sector, with around three-quarters of beef production coming from dairy herds. In the EU beef production, beef fattening is almost 60% of the time a secondary activity of the dairy sector. In recent years, the increase in specific milk yields has led to a reduction in the number of cows needed, resulting in a reduction in the number of cows and a fall in beef cattle production.

Figure 21. Beef cattle (source: internet)

5.1.2 Sheeps and goats

The global sheep and goat population increased by less than two-thirds between the early 1960s and 2014. Of all farm animal species, their numbers have changed the least. During the five decades under study, fluctuations in the number of animals in the herds were common, while the average annual increase in the number of animals was 3%. There is no evidence of stable numbers, and fluctuations in numbers are still common today. Over the last five years, there has been an overall increase of 6%.

Just over half of the global sheep and goat population (52%) is kept in Asia, 30% in Africa, 7% in Europe, 6% in the Americas and 5% in Oceania. Out of a total of 2.2 billion animals, some 1.2 billion are sheep and 1 billion goats. Nearly a fifth of the sheep flock is also used for dairy production. Again, China is the world leader in this sector, with nearly 400 million heads (18% of the global population), nearly 200 million in India, 110 million in Nigeria, and nearly 100 million in Pakistan. The world sheep population in 2010 was the same as 25 years earlier. Since then, it has remained stable, rising by roughly 1% a year. Meanwhile, the world's goat population has more than doubled in 30 years, and in 2018, it increased (by 2.3% to 23 million head), surpassing 1 billion head. In the two years that followed, the global population increased by 24 million goats.

Small ruminants play an important economic and ecological role in developing countries, especially in the poorest regions. These sectors provide animal protein (meat, milk), textile raw materials (wool, hair), hides, drought resilience, food security, and more stable family livelihoods.

Of the 1 billion sheep in the world, Europe has only 98 million sheep, less than 10% of the total, while the European ewe flock is down to 67 million. The United Kingdom is a significant contributor (35 million head), Spain (15 million head) is also significant, and Ireland is over 300% self-sufficient. However, France's population of 13 million 30 years ago has now declined to 7 million and is stabilizing.

Sheep farming is essential in countries with less favorable geographical conditions and regions where sheep farming and lamb consumption are essential for the population because of historical traditions. Sheep and goat farming is important for people in remote and mountainous regions, where this sector is often the only economic and agricultural prospect. Sheep farming has no negative impact on the environment and protects the countryside. This type of livestock farming plays an important role in preserving the natural environment, including maintaining less fertile land using natural methods and conserving biodiversity. It also helps to protect the landscape. In addition, sheep and goats are usually kept where the land is unsuitable for other agricultural activities.

Besides Greece, Russia is a significant player in European goat farming, with only 10% of the two million animals kept on large farms and 90% grazing in small domestic gardens. The combined production of sheep and goat meat has increased by 2.3 over five decades and has grown by almost 7% in the last five years (Table 6.33). In 2013, the combined global volume of the two types of meat reached nearly 14 million tonnes, of which sheep meat accounted for 8.6 million tonnes and goat meat for 5.4 million tonnes. Production volumes of both types of meat have shown a slightly fluctuating but modestly increasing trend in recent years. China, with a production of more than 2 million tonnes, can essentially claim the lion's share of global sheep meat production, but Australia (500-650 thousand tonnes) and New Zealand (around 450 thousand tonnes) also account for a significant share of the world's sheep meat production. The major sheep meat exporting countries have seen a significant decline in numbers in recent decades, but this has had little if any, impact on the amount of sheep meat they export to the market.

What might be behind this phenomenon?

- improvements in what can be described as conventional livestock farming technology - better grass quality, yields, and plant composition without a significant increase in feed costs;

- taking into account the potential and tools of genetics and research results, the quality of lambs placed on the market has improved, and their carcass weight has increased,

- also, by applying the results of genetic research, the utilizable yield (utilizable yield per ewe) of the commercial flocks has improved,

- in exploiting the potential of breeding and breeding management, also based on the results of research, lamb production has become more continuous, thus reducing the downward pressure on prices due to the seasonality of product weight,

- significant producer organizations have been set up to improve representation, market presence, and product quality,

- marketing work covers both domestic and foreign markets with at least equal importance,

- a permanent presence of representatives in the central export market regions with permanent marketing work (television, radio and newspaper, internet, mobile phone advertising);

- promotions and presence on clothing products,

- excellent packaging, cutting, ready-to-eat products, sizes adapted to the customer (small and large family sizes),

- continuous producer training - through producer organizations and service companies,

- good market and butcher presentation, consistent product presence, and uniform meat quality,

- disease resistance - breeding and prevention, less forced handling,

- individual marking of animals does not increase costs for producers, but farm-level marking (group mark) allows a certain level of traceability,

- origin protection and branding (e.g., "New Zealand lamb"; "Australian lamb" - which also appears in Irish, British, Spanish /Aragon/ and French /Grillon/ lambs; in all cases bringing additional price and revenue to the producer - processor - distributor).

The most essential product of goat farming outside Europe is meat. Goat meat production has also increased very significantly in recent times. Not only has the total amount of goat meat increased, but so has its share in a country's meat consumption.

The second most important product of goat farming is milk. Its production is steadily increasing in both developed and developing countries. Cyprus accounts for 40% of global goat milk production, Greece for 25%, and Turkey and Lebanon for 15-15% each. World goat milk production in 2013 was 17 million tonnes, 2.3% of total world milk production. Asia leads the ranking (10.7 million tonnes), followed by Africa (4.2 million tonnes) and Europe third (2.5 million tonnes). The largest goat milk-producing countries are India (5 million tonnes), Bangladesh (2.6 million tonnes), and Sudan (1.5 million tonnes). The average milk production per animal is 17.8 kilograms.

World wool production is on a declining to stagnating trend.

Since ancient times, goat's milk has been known to be the best substitute for breast milk, as its composition is the most similar to breast milk compared to other kinds of milk. In Western Europe today, goat's milk and its products are so highly valued that they have become a staple food for the well-off. This can be seen from the fact that the market price of goat's milk products is several times higher than that of cow's milk and sheep's milk. Moreover, the products made from it, particularly the various cheeses, can be the most delicious dishes even on festive tables. It is also known that goat's milk can be consumed by people allergic to cow's milk because it lacks the protein component that causes milk allergy in France, where the cult of cheese is the greatest, goat's cheese accounts for more than 40% of the hundreds of cheese varieties. Many of the most expensive French beauty products are also based on goat's milk. Folk medicine worldwide uses goat's milk to cure many ailments.

In these two livestock sectors, it is important to distinguish between two very different sets of conditions: extensive livestock farming, which is more environmentally beneficial but less profitable for those who practice it, and intensive livestock farming systems, which are more competitive but less environmentally sustainable.

Figure 22-23. Sheep and goat (source: internet)

5.2 Monograstrics

5.2.1 Pig

Pig farming is a vital livestock sector worldwide. It has spread mainly to temperate regions of Europe, the United States, and Asia. However, it is also present in areas of high population density. In Africa and the Middle East, it is almost unknown. Even though, in many places, the population does not consume pork because of religious restrictions (Muslims, Israelites), the pig as a meat producer is the most numerous of the significant economic species of animal of global importance. Meat and fat from pig carcasses have long been important in humanity's food supply. The feed requirements for the growth of omnivorous pigs can be met from an extensive range of naturally occurring plant and animal foods. Cereal production's development and current production levels have the most direct impact on pig production. Cereals are the main raw material base for pig production, so a large proportion of the cereals produced and fed to pigs are used for pigmeat production.

Asia increasingly dominates pig farming, now accounting for 60% of the global pig population. Europe accounts for 19%, America 17%, Africa 3%, and Oceania 0.5%. China is the world's largest pig producer, with 48% of the world's pig population, partly because there are around 500 pig breeds in the world, of which China has more than 300. The United States is second in the production ranking with 7%, followed by Brazil with 4%. In the European Union, there are nearly 160 million pigs. Germany is the largest pig producer in the European Union, with only 3% of the world's pig population. At the world level, Vietnam is also worth mentioning, with a pig population approaching that of Germany. The number of pigs in the world's top five pig-keeping countries has remained stable over the last five years.

The way pigs are reared varies worldwide depending on local conditions, mainly due to different feeding options. Pig farming is almost universal in the significant maize-producing regions of the world. In some regions, potatoes are the main feed for pigs. In Western Europe, skimmed milk and grain feed are the most critical feedstuffs. In northern Europe, pigs are raised on barley and potatoes.

Because of their excellent adaptability, pigs are reared under a wide range of conditions, and the efficiency of pigmeat production varies considerably from country to country, from very primitive farming systems to concentrated, industrialized, large-scale pig farms. In addition to East and South Asia's traditional and distinctive pork production, based mainly on domestic and garden waste, the vast swathes of grain-producing states in the United States (Iowa, Illinois, Nebraska) are producing mass quantities of pork based on maize and soya. At the same time, in China, Vietnam, Cambodia, Malaysia, and Indonesia, some traditional pig production is still based mainly on feeding the vegetative parts of plants, i.e., garden waste. In many cases, pig farming in these countries also has a public sanitation function through feeding waste by pigs. In countries with a maritime coastline, so-called 'trash fish,' unfit for human consumption, are also used in pig feed.

Since the turn of the millennium, wholly enclosed, landless housing systems have become widespread in pig production worldwide. These farms are associated with technologies that convert pig manure into bioenergy through fermentation and now produce more than half of the world's pork. Several European countries produce high-quality lean meat for particularly demanding consumers in an organized, mostly intensive production framework - using significant quantities of imported feed. In addition to intensive farming systems, alternative and organic farming systems meet the needs of a specific consumer group.

In Europe, pig production has evolved to adapt to local conditions and has moved from rigid pig farming to intensive farming systems. In the past, pigs in northern European countries, for example, were fed on root crops, barley, and oats in the context of traditional housing and husbandry systems. In these regions, skimmed whey has been fed to pigs as animal protein feed since the last century. The high level of feeding and housing increased yields and improved production efficiency. In continental areas, such as Hungary, grain crops (maize, barley, milling by-products) were the basis of pig production. It is no coincidence that producing edible bacon and lard from fattened fat pigs was the main objective for a long time. Today, meat pigs are playing an increasingly important role in the world's pig population. The number of pigs reared for their fat has declined, and their share of pig production has also fallen in recent years.

Grazing is also a vital feed option for pigs. Pigs can use food on the surface and, by digging, use food in the ground (roots, tubers, worms). The considerable quantities of fodder in permanent and occasional pastures, which would otherwise be wasted, would only have been used with pig production. In most developed countries, however, no significant extensive outdoor pig production exists.

The high genetic variability of pigs is illustrated by the fact that fat-type pigs have been converted into meat-type pigs in less than a century. The proportion of muscle tissue has increased significantly in favor of fat tissue, and the fat content of chops has been reduced by 70%, from 4.6% to 1.2%, in just under a quarter of a century, in line with consumer demand. Despite the hundreds of pig breeds worldwide, the role of local breeds is not dominant, except in China. In Europe and America, five world breeds predominate in intensive slaughter pig production, whether crossbred or hybrid. These are the Great White, the Lapland, the Duroc, the Pietrain, and the Hampshire.

The complex environmental impact of modern pigmeat production (environmental footprint) is the lowest compared to other sectors, alongside poultry production. The by-products that cannot be used for human consumption are also of high value. Pig skins have many uses in the leather industry, and their bones can be used to make glue. The value of pig manure for agriculture is very significant, as well as sustainable agriculture.

Figure 24-25. Intensive pig (hybrid) and Mangalica pig (source: internet)

5.2.2 Poultry

Poultry farming and rearing have been important agricultural sectors for four millennia. The poultry sector has been developing very dynamically for decades. It is a significant contributor to world livestock production and animal food production. In addition to the steadily increasing quantitative indicators, the improvement in genetic values has outperformed all animal species. All poultry species respond well to selection, as the yields of recent decades and the improvement in specific feed use are reasonable indications. Worldwide, the investment in production, new technologies, and products is increasing in most cases. With the generalization of artificial insemination, it is possible to produce large numbers of flocks of the same age, quality, and close relationship simultaneously. This technological element helps poultry farmers to implement uniform breeding and feeding principles. Most poultry species have excellent technological tolerance. By exploiting this, huge stock concentrations could be achieved. It is common practice to place 20,000 broilers in a single airspace, and placing 300,000 laying hens on a single site is not an outlier. The poultry industry's production and consumption of poultry products will continue to grow in the future, but at a differentiated and probably slower rate than has been the case so far. This growth trend has become a worldwide phenomenon since the second half of the 20th century. The main reasons for this phenomenon are the following:

- complex conditions for the safe and continuous production of eggs and poultry meat have been established on all continents - at least to a medium or high standard.

- there are no religious restrictions on the consumption of poultry products, eggs, and meat

- the biological characteristics of poultry breeds, such as high reproductive performance, good adaptability to different environmental conditions, genetic characteristics allowing the production of high-yielding breeds and hybrids, and easy and relatively cheap transport of hatching eggs and day-old animals are all factors that have contributed significantly to the worldwide expansion of poultry farming.

- the production of eggs for human consumption and poultry meat is well served by the feed fed, a sector that is only surpassed in this respect by fish production.

The secure production and supply of poultry feed worldwide have been established, mainly through the surge in cereal and soya yields and the resulting advanced compound feed production.

- increased purchasing power among a broad cross-section of the population and a steady demand for the products produced.

Due to the reproductive biology of poultry species, the poultry sector is the most flexible of the livestock sectors in terms of adapting to market needs. The world's poultry population has increased virtually steadily over the last five decades, with 25.7 billion birds in 2014, more than five times the number in 1961. A slight decline occasionally accompanies this growth, but the trend is unaffected. Over the last five years, there has been an increase of 11.7%.

Around 56% of the world's poultry are kept in Asia, a quarter in the Americas, a tenth in Europe, only 8% in Africa, and less than 1% in Oceania. China, like many other sectors, is the world leader in poultry production, with around 5.5 billion, followed by the United States with 2.2 billion, Indonesia with nearly 2 billion, and Brazil with 1.4 billion.

The United States is the world's largest consumer, although its consumption fell by a few percent at the beginning of the millennium's second decade. This has led to China making headway, making it the world's largest poultry meat consumer globally. Poultry meat consumption is also significant in Brazil and the European Union. In particular, a substantial increase in the consumption of imported poultry meat can be observed in middle and low-income countries. Consumer confidence in the consumption of poultry products, which had been eroded due to previous animal health problems, has been restored by the intense food safety pressure on the poultry sector. Although the dynamic growth of poultry consumption, which had been going on for decades, slowed down in the first decade of the millennium, it is now back on a steady upward trend.

In the coming years, consumer demand for safe, environmentally friendly poultry products will continue to grow, and demand for organic and free-range poultry products is expected to increase. The latter is also because consumers often identify these products with food produced sustainably and environmentally friendly. As this method of meat production could supply at most half of the world's population, global meat demand will continue to be produced on a large scale - but using environmentally friendly and sustainable intensive production methods. Of all animal products produced on an industrial scale, poultry meat is the most environmentally friendly and sustainable intensive production method, which also explains the rise of poultry products. Since domestication, chickens have been reared and kept mainly to produce two products of high biological value: eggs and meat. In addition to the eggs and meat, the main products of chicken farming, the by-products can also be used. The feed industry uses feathers to produce feather meals, and the parts from the slaughterhouse, which are unfit or unfit for human consumption, are further processed to produce, among other things, valuable pet food for hobby and companion animal keepers. The latter activity has now developed into a vital industry. The litter manure from the production of chicken manure and poultry manure contains 13 nutrients that are very important for plant growth. It is no coincidence that poultry manure is widely used to replenish soil energy. Through various physical, chemical, and biological fermentation processes, poultry manure can be processed into valuable animal feed, mainly for ruminants and fish, with appropriate animal protein. Litter manure can also be used effectively for heat production and fermentation to produce biogas and ethanol. It can also be used as a growing medium for intensive mushroom production using appropriate ripening procedures.

Figure 26-27. Broiler chickens (source: internet)

Besides poultry meat production, another critical sector of the poultry industry is table egg production. Hen eggs are one of the most complete and valuable human foods. Unique formulations of feed can be used to modify the composition of eggs to some extent, but these techniques are relatively limited in egg production worldwide. Despite the undisputed nutritional value of the product, egg production is of less international commercial importance than poultry meat production. At the world level, egg production has followed a similar steady growth path over the last five decades as poultry meat production, although the rate of increase has been more modest, less than five times. Egg production and consumption have yet to benefit from the over-emphasis on cholesterol and the eclipse of its ideal biological value. There has been a noticeable positive change in this area in recent years.

Figure 28. Layers for table egg production (source: internet)

Asia (61%) and China (around 40%) are the absolute leaders in global production. The Americas account for less than a fifth of world production, with the United States accounting for 8% and Europe for 15%. India (5%) and Japan (3.4%) are also significant.

The world's chicken population exceeds 21 billion. Asia has the largest chicken population, followed by North and Central America and Europe. The chicken population in Africa and Oceania is relatively tiny. The main growth driver in world egg production is Asia, where production has increased by almost a quarter in a decade (2008-2018). This figure is boosted because egg production in Europe has increased by only 6% over the same period, with minor fluctuations.

Asia accounts for a significant proportion of the world's waterfowl production, but Europe, including France, Poland, and Hungary, has a strong tradition in this poultry sector. Ducks are the largest waterfowl species in the world (1.13 billion), followed by geese (around 300 million), which are much smaller. The three duck species are used for meat, eggs, and liver production worldwide, and their feathers are also valuable. The musk duck is sought after for its red, gamy meat, while the attractively marbled muscles of the mallard duck are a particular delicacy. These two products differ from domestic duck in many respects, particularly in the body's lower fat content and the meat's quality.

Duck livers - fattened muskmelon and mallard duck livers - are a popular ingredient in liver pâtés because they are easier and cheaper to produce than foie gras. Of the three liver-producing birds, musk ox liver is the least valuable. The liver production capacity of the mallard duck is just below that of the goose and is superior to that of the goose in terms of economy. Duck liver production is limited to a few countries, with France leading the way and Hungary being one of the top producers.

Duck farming and duck rearing are practiced worldwide, but their economic importance, husbandry, and primary use can vary from continent to continent and country to country. In Asia, it plays an essential role in producing eggs for human consumption and meat production (the main foodstuff in Southeast Asia). At the same time, it is a diversified product in developed countries. Although the world duck population has grown spectacularly in recent decades, and the genetic ability and production parameters of the hybrids available for the production of ducklings are improving dramatically, the global duck population has declined by about 9% in five years since 2010 before rebounding to increase by 5% in five years. Most ducks (87%) are reared in Asia, with China accounting for nearly 60% of the global population. Some Southeast Asian countries (Vietnam, Indonesia, Malaysia) also have a significant duck population. Their presence totals 7% in Europe, while their presence is negligible in the other continents. World duck meat production was around 4.3 million tonnes in 2013, according to the Food and Agriculture Organization of the United Nations (FAO), rising to 4.5 million tonnes in 2018, of which China produced almost 70% (3 million tonnes). This position has been maintained for five years. France was second (277 thousand tonnes), accounting for only 6% of world duck meat production. This quantity had fallen to 246 thousand tonnes by 2018. EU production was around 480 thousand tonnes in 2013 and 489 thousand in 2018.


Figure 29-30. Duck for fattening and geese for liver production (source: internet)

In most East and Southeast Asian countries, the demand for table eggs is still met mainly by duck eggs; more than half of the duck population is thought to be used for egg production. Eggs are mainly consumed in processed (boiled, salted) form. Duck meat and duck eggs play a much more critical role in the diet of the local population than in Europe or North America. In Southeast Asia, specific, extensive methods of duck farming have developed and are closely integrated with local forms of crop and animal product production.

For several centuries, the predominant husbandry system has been based on transhumance. For generations, the native duck breeds were selected for their ability to gather most of their food in harvested rice paddies, swamps, and canals. A combination of duck and fish production, based on principles similar to the former, is common in many countries. In developed countries, duck, although a popular product, is overshadowed by chicken or turkey. Ducks are mainly used for meat production, less for their animal protein source than their gastronomic specialty. In some European countries, it is used for liver production, but in none, it is used for table eggs. Specific semi-intensive or intensive forms of duck farming, quite different from those in developing countries, are widespread.

Turkey meat is a popular foodstuff worldwide, but it only took off when the conditions for continuous production were created. Turkey meat has become an integral part of a healthy diet. In the Anglo-Saxon countries, there is a centuries-old tradition of seasonal turkey meat consumption, notably at Christmas in Britain and on the fourth Thursday of November, Thanksgiving Day, in the United States. In the USA, nearly one-third of all turkey meat consumed in the country is associated with Thanksgiving, Christmas, and Easter. In terms of per capita consumption of turkey meat, Israel ranks first (13 kg/person/year), the United States of America is second (8 kg/person/year) and Hungary is third (5 kg/person/year).

The world's turkey population has increased by about 2.25 over the last five decades, but it has stabilized over the last 20 years and increased by 4% in the last five years. Of the approximately 467 million birds, two-thirds are kept in the Americas (US, Brazil, and Chile), and nearly a quarter are kept in Europe, while Africa accounts for 5% and Asia for 3%. The annual production of turkey meat is around 5.9 million tonnes and has been on a slight upward trend in recent years. The share of world poultry meat production reaches 7%. The Americas also lead in turkey meat production, with the United States of America accounting for almost 50% of the global share and the most significant consumption. Significantly behind, Brazil is second with a share of around 10% and Germany third with 8%.

Since 2010, the United States has produced the most turkey meat annually. In 2010, 2,560,000 tonnes were produced, and 2018 2,666,000 tonnes left the plants. The EU, the world's second-largest producer of turkey meat, is also showing a steady increase. In 2010, the EU produced 1,852 thousand tonnes of turkey meat, while in 2018, this figure reached 1,945,000 tonnes.

In the same year, Brazil was the third largest producer, with 576 thousand tonnes of turkey meat. Canada followed with 169 thousand tonnes. Russia accounted for 109 thousand tonnes, Mexico for 17 thousand tonnes, South Africa for 5 thousand tonnes, and China for only 3 thousand tonnes.

The share of the United States of America in world turkey meat exports is significant but declining. In 2010, 240 thousand tonnes of turkey meat were exported, which shrank to 219 thousand tonnes in 2018. In 2010, 78 thousand tonnes of turkey meat were exported from Brazil, and in 2018, only 63 thousand tonnes were sold in foreign markets. The largest exporter of turkey meat is the European Union, which exported 513 thousand tons of turkey meat to international markets in 2010, but in 2018 they exported 576 thousand tons.

The amount of turkey produced within the European Union is also highly distributed between Member States. In 2014, 469 thousand tonnes of meat was produced, a slight increase compared to the previous year, and the same trend is still present in 2018. The French were not far behind the Germans, producing 378 thousand tonnes of meat in 2014. They saw a slight increase compared to 2013 but fell back to 367 thousand tonnes in 2018. The Italians contributed 300 thousand tonnes of turkey meat, and the Spanish contributed 210 thousand tonnes to community production in 2018. The former shows a slight downward trend, while the latter shows a marked upward trend. (It should be noted that Spanish production in 2009 was only 28 thousand tonnes.) The notoriously turkey-loving English produced 172 thousand tonnes of turkey meat in 2014, a slight decrease compared to 2013. This decline has continued, with UK production in 2018 at just 157 thousand tonnes. In Eastern Europe, the Poles are the leading players.

Figure 31. Turkey fattening (source: internet)

5.2.3 Aquaculture

Global fish production is around 150 million tonnes per year. This includes fish caught in natural waters, artificially farmed fish, mollusks, and crustaceans for human consumption. Approximately 60% of the total is available through fishing and 40% through farming. In the last two decades, aquaculture fish production has increased significantly, more than threefold.

Due to stagnating marine fish catches, aquaculture production is becoming increasingly important, growing faster than most conventional agricultural sectors over the last decade. Freshwater fish production is significant in Asia but has become a dynamically expanding sector in many world regions. Almost 70% of fish comes from Asian countries and around 80% from developing countries. The primary production objective of freshwater fish farming is the production of fish meat for human consumption. Around three-quarters of world production is used for this purpose. Fish consumption is growing in developed countries as healthy diets become more popular.

At the same time, in many underdeveloped regions of the world, the development of fish farming is the most economical way to produce essential animal protein.

With a large proportion of the catch being taken in Asia, it is unsurprising that most of the top ten fishing nations are from this continent. China is the world's most significant volume fisherman, accounting for one-sixth of global catches in 2011. Three countries from the Americas are in the top ten, while Russia is the only European country in the top ten. However, Russia's catches in the Pacific account for around 70% of its total catches and would not make it into the top ten countries if only its catches in European waters were included. Apart from Russia, Norway and Iceland are the other European countries with annual catches of more than 1 million tonnes. In 2011, the world's largest catch was anchovy from Peru, with 8.3 million tonnes caught.

China, the world's largest producer, accounts for 61% by volume and 50% by value. India is in second place and Vietnam in third, both in terms of volume and value. Norway, the first non-Asian country in terms of volume, is seventh in the ranking. Also in terms of volume, 40% of fish farmed are carp and the most important species farmed is amur (4-4.5 million tonnes). Most fish farming is carried out in freshwater. Around 20 million tonnes of fishery products are produced without feeding. According to FAO data, in 2010, 86% of fisheries production (128.3 million tonnes) was for human consumption, with the largest share of fresh fishery products sold unprocessed. Other uses, which accounted for 14% of production, were mainly for the production of fishmeal, used as feed material, and fish oil. Almost 40% of the global fish production is traded internationally, highlighting the global nature of fish production and trade. As with total trade in goods, China is the world's largest exporter, and the United States is the largest importer.

Figure 32. Aquaculture (source: internet)

5.2.4 Insects

Insects have helped agricultural production for thousands of years. Bees that pollinate crops contribute to proper fertilization and production. Their work collects nectar and pollen, which the beekeeping industry uses. Honey is the only nutrient suitable for human consumption without human processing. The bees themselves do this, and all that is needed to make the honey available to the final consumer is for it to be collected and packaged by humans.

Throughout history, the story of humans and the honeybee have been intertwined. Long sought after for their honey, honeybees have been depicted in ancient cultures and modern religions as a symbol of fertility, industriousness, and cooperation. From prehistoric cave drawings depicting honeygatherers in southern Africa, Asia, Australia, and Europe to modern agriculture heavily relying on the honeybee for their crop pollination services, the honeybee has played a vital role in the development of human traditions, agriculture, and society.

Mesolithic (Stone Age) Honey Gatherers

The earliest recorded record of honey gathering is in the Cave of the Spider (la Cueva de la Araña) near Valencia, Spain. This 15,000-year-old painting shows a woman on a rope ladder gathering honey from a bee nest precariously on the side of a rock cliff. Other drawings show similar depictions of people gathering honey in the wild, where honeybee nests are usually found high up in cavities of trees and the side of rock cliffs.

This method of gathering honey is still practiced in many parts of the world today. Examples include the Bedouin tribes in the Syrian desert, the Veddhas people of Sri Lanka, and the Gurung group of Nepal. For many of these people, gathering wild honey can be a valuable source of income and food and essential to rituals.

Ancient Egyptian Beekeeping

The earliest known form of organized beekeeping occurred in ancient Egypt, where twig and reed-designed hives would allow beekeepers to manage honeybee colonies of a species known today as the Egyptian honeybee. By 1500 BCE, beekeeping was widespread throughout the Nile region, with Egyptian administrators accepting honey presented by farmers as payment. With this, honey was collected in clay vessels and stamped according to their quality and color.

Honeybee products were not only a source of food but also used for medicinal purposes and religious practices. Honey was commonly used as an antiseptic to treat wounds and was a standard offering to the Egyptian Gods. Beeswax was also used for candle making and was a vital substance for mummificacess. According to Egyptian mythology, when Ra, the god of the sun, cried, his tears would turn into bees so that they would bring fertility by pollinating the flowers of the Nile.

Ancient Chinese Beekeeping

While there are no existing records of Mesolithic honeygatherers in China, it is likely that some groups of people did harvest honey from the wild honeybee populations. Three species of honeybee are native to China, including the Dwarf Honeybee (Apis florea), the Giant Honeybee (Apis dorsata), and the eastern honeybee (Apis cerana).

Apiculture has a long history in China, where the native Eastern Honeybee has been used for crop pollination and honey and beeswax production. Although the Eastern Honeybee forages in a smaller area and produces less honey than its European counterpart, beekeeping was well-established and highly profitable by the time of the East Han Dynasty (25-150 CE). With this, early records reference professional beekeepers and an industry forming around honey and wax production for the Emperor, his court, and the educated elites.

Ancient Mayan Beekeeping

Before contact with the Spanish in Mesoamerica during the 16th century, the Mayan culture had existed for over 1,800 years. It is believed the Mayans had an extensively developed knowledge of beekeeping, being able to divide existing hives to increase numbers and taking care not to over-harvest to leave enough honey stores for the honeybees. According to early Spanish accounts, the Mayans had well-established apiaries with hundreds and sometimes thousands of native Stingless Honeybee hives. Each hive was made from hollowed-out logs in the shape of large drums, which were individually carved featuring figures, ornaments, and the sign of the owner of the hives. The bees would enter and leave a hole in the middle of the hive, with round stone discs being used as end-stoppers. These stone discs are the oldest known beekeeping artifacts in the world, dating from 300 BCE to 300 CE.

Figure 33. Madrid Codex shows the juxtaposition of named days and coefficients of the Tzolk'in in the first column of both the upper and lower register. The calendrics introduce two small vignettes showing the good and adverse auguries in the lowest and upper registers. (source: internet)

The honeybee species used by the Mayans is native to Mesoamerica and is called the Stingless Honeybee (Melipona beaches) because it lacks a stinger for defense. The Stingless honeybee typically builds its nests inside cavities of trees, which is observed in tropical rainforests around the Yucatan region of modern-day Mexico and Central America. Sadly, due to severe deforestation, widespread insecticide use, and the arrival of European and African Honeybees, it is estimated that Stingless Honeybee populations have decreased by over 90% in the past two decades.

Beekeeping through the European Middle Ages and Colonial Periods

Honey-gathering practices in medieval Europe were based on early beekeeping practices seeking out large trees with wild resident honeybees. After finding honeybee nests, beekeepers would cut out small sections of the tree to make protective wood panels with flight entrances so that the honeybees could be easily accessible and protected from bad weather and predators. It was also typical for the tops of growing trees to be cut off to reduce their height and thicken their trunks so it was possible to carve out more artificial cavities in the trees. Throughout northern Europe, this model of tree beekeeping was common, leading to the creation of "bee forests," usually owned by the aristocracy or the Church.

Beekeeping in bee forests was essential to local European economies. However, it was a very time-consuming method. Other methods for keeping Honeybees included "log hives" and basket-weave hives known as "skeps." These methods were utilized in many parts of Eastern Europe, especially Germany, Poland, and Lithuania. Log hives were essentially cut rounds from trees containing honeybees' nests. It would be typical for the log hives to be then carved and painted and, in some instances, made into sophisticated human figures.

In contrast, skeps were typically made out of willow wicker or a combination of straw, reeds, and sedges. The benefits of these methods were the ability to have the hives closer to human settlements and transport them to different locations. In Britain, France, and other parts of Western Europe, specially made stone structures called "bee boles" were erected in south or southeast-facing walls near orchards and gardens to protect honeybee hives from wind and rain. During the Age of Discovery and the subsequent colonization of the Americas, ship cargo manifests show that European Honeybees were among the first animals to ship with the early settlers. Honeybees were considered an essential part of colonial life because the bees would produce precious honey and beeswax and pollinate crops.

Moreover, during England's taxation upon American colonists, honey was used instead of highly taxed sugar. With this, honey and beeswax were an essential source of income and for making products like candles, lipstick, shoe polish, and mead. This led to the spread of honeybees across North America, but it was not until 1856 that honeybees arrived on Vancouver Island, among the first places in Western Canada.

Modern Beekeeping

Although honey-hunting from wild bees still exists in some societies, most honey is now harvested from managed and kept bees. Today, the most widely used beehive design is known as the Langstroth hive, which is made up of wooden boxes and individual wooden frames that can easily be removed to harvest honey and wax and inspect the hive's health. The other benefit of the Langstroth hive is that it can be moved easily. It is common for hives to be migratory throughout the growing season to where the best forage and high nectar-yielding flowers are.

In Europe, it is common for beekeepers to move their honeybee hives to the countryside in the latter part of the growing season. In the UK, heather is in bloom, and in France, the lavender fields perfume the air with the sweet smell of lavender flowers. These honey varieties can fetch a high price, making it worthwhile to move the hives. Meanwhile, in Central Asia, Uzbekistan, Kyrgyzstan, Tajikistan, and southern Kazakhstan, migratory beekeepers move their hives on wagon trains, sometimes up to 300 to 400 hives. When nectar-rich flowers bloom in April and May, migratory bee wagons will move to the best areas of wildflower forage.

Figure 34. Modern beekeeping (source: internet)

Threats to Honeybees and Other Pollinating Species

Today, beekeepers have to grapple with significant honeybee diseases, particularly the bee mite known as Varroa destructor, one of the likely contributors to the phenomenon known as Colony Collapse Disorder (CCD). In addition, the widespread use and exposure to pesticides, insecticides, and herbicides not only threatens the life of the individual bee but also the survival of the hive with the contaminated nectar and pollen being shared throughout the hive. The problems of disease and potentially toxic environments are only exacerbated by the lack of nectar-rich forage areas year-round and the movement of bees around for pollination services. Though migratory beekeeping can be successful in terms of greater productivity from the hives, it can be very stressful for the bees, making them more susceptible to disease. Although the migration and spread of the Western honeybee species worldwide has brought about improved commercial honey and crop production, unfortunately, on a commercial scale, it poses a significant threat to the well-being and health of honeybee populations and other pollinating species. In many areas of the world, the arrival of the Western Honeybee displaced and introduced diseases to native bee species. In places like Australia, there is concern that introduced Western Honeybees compete for forage with the native stingless bees and other pollinating species. In Canada and the United States, there is evidence linking the rise of large-scale commercial beekeeping operations and honeybee diseases being transferable to native bumblebee populations. In Central and South America, introducing "Africanized Bees" has decimated native stingless bees that were historically used by the local indigenous people and the Mayan beekeepers of centuries past. As Western honeybees are an introduced species to Vancouver Island and much of the world, it is essential not to rely on them wholly for their pollination services and not forget that other species play an essential role in pollinating our food crops and native plant species. With this, honeybees should be considered a partner with other pollinators in their pollination services. It must support local beekeepers and organic agriculture to overcome these significant threats to honeybees and other pollinating species. By establishing healthy and sustainable areas with a diversity of forage crops and flowers year-round, we can work to overcome the threats of disease and commercial monocultures.

Edible insects are a potential solution to a suite of pressing environmental and human health issues, including climate change, malnutrition, food insecurity, and environmental degradation resulting from agro-industrial production. In 2013, the United Nations Food and Agriculture Organization (FAO) published Edible Insects: Future Prospects for Food and Feed Security, which presented a comprehensive analysis of human consumption of insects globally and advocated for edible insects as a viable future food source. Since this landmark report, media and consumer interest in edible insects has increased. The global edible insect market value is estimated at 406 million USD (2018) and is predicted to increase to over 1.18 billion by 2023 (Statista, 2019). Though the US market is comparatively small (8 million USD), the North American market is predicted to grow more than any other region, with an estimated 28% growth by 2023 (Statista, 2019). In Europe, edible insect production has been hindered by EU legislation requiring 'novel food' authorization since 2018 (2015 for most EU countries). However, approval is expected to be achieved in 2020, notably impacting research and broader consumption.

Past and current research on edible insects recognizes that current food systems, particularly animal agriculture industries, contribute to the climate crisis and that edible insects are a potentially underutilized food source (e.g., Holt, 1885; DeFoliart, 1992; Premalatha et al., 2011; van Huis Arnold et al., 2013). In 2006, the FAO's report, 'Livestock's Long Shadow,' pointed to the global livestock industry as one of the top three contributors to global environmental degradation, citing, for example, that 26% of the earth's land is used for livestock grazing and 70% of agricultural land is used for feed production (Steinfeld, 2006). The livestock industry also accounts for 8% of human freshwater usage and produces 18% of the planet's greenhouse gas emissions, more than the total combined emissions from transportation (Steinfeld, 2006). It is predicted that global meat demand will double by 2050 (Steinfeld, 2006; Davis et al., 2015; Godfray et al., 2018) due to population growth, but mainly due to agro-industrial expansion and high rates of consumption among affluent consumers and new consumers in emerging countries (Myers & Kent, 2013; Hoelle, 2018; Hoelle, 2017). In response to the 'ecological hoofprint' of the livestock industry (Weiss, 2013), researchers and consumers are seeking alternative diets that exert less of an environmental burden, and entomophagy, the consumption of insects, has become an essential option among educators, governments, individuals, and NGOs (Prather and Laws 2018, NACIA 2019, Mason et al. 2018). Compared to livestock production, many insects have a high feed-conversion efficiency, can be reared on organic waste, emit few greenhouse gases, and take a small fraction of the feed, water, and space required by traditional livestock (van Huis et al., 2013; van Huis & Oonincx, 2017; Govorushko, 2019). With the shift to industrialization, it is difficult to know how sustainable edible insects will be, but some studies have given us an idea (Halloran et al., 2017; Wegier et al., 2018). In a life cycle assessment conducted by Halloran et al. (2017), a mass-reared cricket production site in Thailand was found to have less environmental impact than broiler chicken farms, and they predicted that impact would continue to lessen as cricket farms are further industrialized.

Moreover, many edible insects present comparable nutritional benefits to animal-based foods, with a high potential for addressing global food and nutrition insecurity (Murefu et al., 2019; van Huis, 2020; Imathiu, 2020). Though macro and micronutrients vary greatly by insect, some species contain as much protein as beef, more iron than spinach, as much vitamin B12 as salmon, all nine amino acids, as well as high calcium, Omega 3, and fiber contents (van Huiset al, 2013; Govorushko, 2019; Imathiu, 2020). Insects have been eaten extensively throughout human existence and are commonly consumed by over 2 billion people today, despite the near absence in the global North (van Huiset al, 2013; Lesnik, 2019). Since the FAO report, and especially in the past ten years leading up to 2018, edible insect research has increased substantially, coinciding with a growing acceptance of insect consumption and the appearance of insect food products in the market. Edible insect research before 2018 may be best characterized by disproving the assumptions associated with insect food aversion. These articles worked to establish credibility for edible insect acceptance in the West and urged for further research into the feasibility of insects as food. In the past two years (2017–2018) of edible insect scholarship, they have presented responses to these research calls and increased studies on the impact of insect consumption on the human body (e.g., digestibility, toxicity, allergenicity, biome impacts), on the environment (e.g., livestock comparison, rearing potential, carbon footprint, policy) and, most notably, on edible insect industrialization practices.

In recent years, there has been a dramatic increase in edible insect publications. As this review showed, most articles published in 2018 focused on one of five scales of the industrialization process. We are generally concerned with increasing food safety and productivity in industrial edible insect operations, guided by an overarching interest in increasing consumer acceptance. While research before 2018 often presented edible insects as a somewhat homogenous food source, current research shows the extensive variability of rearing requirements, nutrition profiles, and tastes of different edible insect species. Likewise, present and future insect consumers were also described in greater detail in this research sample, with sociocultural factors (e.g., gender, age, socioeconomic status) found to be influential in determining consumers' willingness to eat insects in both areas familiar and unfamiliar with insect consumption. Despite the increase of edible insect industrialization publications, we found that specified information on rearing practices was more widely published when framed as economic or food security 'development' but was more closely guarded by Western industry due to the increasing competitiveness of the market.

A critical finding in this literature review was related to sustainability, specifically the relationship between a) the industry's intent and b) the interests of consumers, which is predicted to impact c) industrialization practices. As evidenced by the articles on product development and consumer acceptance, consumers were motivated more by short-term factors like price, taste, and availability than environmental sustainability (Berger et al., 2018; House, 2018). As the edible insect industry continues to find that sustainability is not the main factor influencing insect consumption or acceptance, continued industrialized insect rearing may forgo some sustainability measures, and the efficacy of insects as a sustainable food alternative could be diminished. Furthermore, the future sustainability impact of insect rearing has yet to be discovered. There still needs to be more knowledge concerning almost every aspect of production: from suitable species, their housing and requirements, and potential for accidental release' (Berggren et al., 2019). Research on this topic has increased substantially, yet a focused, transparent, and standardized approach to data gathering is necessary to substantiate sustainability claims. The applied research dedicated to improving industrial production could benefit from social science and humanities research on the social, moral, and ethical aspects of an industry that has the opportunity to produce on a large scale without the associated environmental impacts of other livestock industries.

Despite challenges to the sustainability of edible insect industrialization, the research and overall global increase in the mass rearing and consumption of insects still indicate the potential for a radical change in eating practices. Changes are driven by the social and environmental values of producers, governments, and consumers and are aimed at dismantling deeply rooted cultural food aversions in favor of more sustainable and ethical consumption practices. Despite sustainability being widely professed as the rationale for increasing edible insect industrialization, the relationship between consumer choice, industrialization practice, and environmental impact will continue to be an area for research, as evidenced by Halloran et al. (2018) and Berggren et al. (2019). These topics intersect at the level of legislation set to change in the EU in 2020, which could facilitate broader global acceptance of insect production and consumption. Suppose edible insects are to be more available, affordable, and, therefore, more widely consumed and accepted as a food source. In that case, the scaling-up of industrialization is necessary, but to what extent will these increased industrialization efforts be at the cost of sustainability goals? The fact that increased production and increased acceptance are inextricable raises more questions: Can edible insects remain a 'sustainable' alternative if they are industrially reared and inserted into the global food system? What role may localized or home-based edible insect-rearing initiatives play in broader industrialization and food security efforts (Dickie et al., 2019)? These and other vital questions, unique to edible insects as a food source, were beyond the scope of most of the articles in this review.

The promise of edible insects as a global food source is coming to fruition through research to increase efficiency, safety, and acceptance. However, the fact that very few articles addressed the social and environmental impacts and that many were oriented to industrialization by private companies is also cause for reflection. As evidenced by this review, notable future research is necessary to discern the verifiable environmental impact of insect rearing (e.g., energy usage and waste production) alongside measurable evidence for the social impact of edible insect production and consumption. It would be an overgeneralization to position edible insect industrial projects as part and parcel of the same environmentally destructive agro-industrial complex to which many aim to provide an alternative, such as edible insect advocacy groups oriented to food security and many for-profit edible insect companies operating in a much 'greener' mode of production. However, the benefits of edible insect industrialization, with the ability to produce mass quantities of food under strict safety guidelines and with limited environmental impact compared to other agro-industrial regimes, must remain in focus as the industry expands and insects gain consumer acceptance.

Figure 35. Insects for food and feed (source: internet)

6. Artificial intelligence and machine learning

In this chapter, I will briefly describe some machine learning methods, algorithms, considered a subfield of artificial intelligence, that play an essential role in applying precision agriculture technologies.

6.1 The relation between artificial intelligence and machine learning

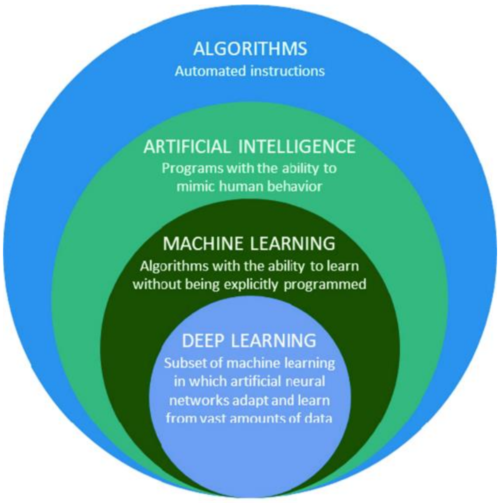


Figure 36. Relation between algorithms, AI, ML, and DL (source: internet)

Artificial intelligence and machine learning are very much related. According to McCarthy (2007), one of the field's founders, AI is "the science and engineering of making intelligent machines, brilliant computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to biologically observable methods." This is pretty generic and includes multiple tasks such as abstractly reasoning and generalizing about the world, solving puzzles, planning how to achieve goals, moving around in the world, recognizing objects and sounds, speaking, translating, performing social or business transactions, creative work (e.g., creating art or poetry), and controlling robots. Moreover, the behavior of a machine is not just the program's outcome; it is also affected by its "body" and the environment it is physically embedded in. To keep it simple, however, if somebody can write a very clever program with human-like behavior, it can be AI. However, only if it automatically learns from data is it ML. ML is the science "concerned with constructing computer programs that automatically improve with experience" (Mitchell, 1997). So, AI and ML are both about constructing intelligent computer programs, and DL, being an instance of ML, is no exception. Deep learning (LeCun et al., 2015; Goodfellow et al., 2016), which has achieved remarkable gains in many domains spanning from object recognition, speech recognition, and control, can be viewed as constructing computer programs. Namely, it involves programming layers of abstraction differently using reusable structures such as convolution, pooling, autoencoders, and variational inference networks. In other words, we replace the complexity of writing algorithms that cover every eventuality with the complexity of finding the proper general outline of the algorithms—in the form of, for example, a deep neural network—and processing data. By the generality of neural networks—they are general function approximators—training them is data-hungry and typically requires large labeled training sets. While benchmark training sets for object recognition store hundreds or thousands of examples per class label, creating labeled training data is the most time-consuming and expensive part of deep learning (DL) for many AI applications. Playing video games may require hundreds of hours of training experience and expensive computing power. In contrast, writing an AI algorithm that covers every eventuality of a task to solve, say, reasoning about data and knowledge to label data automatically (Ratner et al., 2016; Roth, 2017) and, in turn, make, for example, DL less data-hungry–is much manual work, but we know what the algorithm does by design and that it can study and that it can more easily understand the complexity of the problem it solves. When a machine has to interact with a human, this seems to be especially valuable.

This illustrates that ML and AI are similar but different. Artificial intelligence is about problem-solving, reasoning, and learning in general. Machine learning is specifically about learning—learning from examples, from definitions, from being told, and from behavior. The easiest way to think of their relationship is to visualize them as concentric circles, with AI first and ML sitting inside (with DL fitting inside both) since ML also requires writing algorithms covering every learning process eventuality. The crucial point is that they share the idea of using computation as the language for intelligent behavior. What kind of computation is used, and how should it be programmed? This is not the right question. Computation neither rules out search, logical, probabilistic, and constraint programming techniques nor (deep) (un)supervised and reinforcement learning methods, among others, but does, as a computational model, contain all of these techniques.

Using computation as the common language, we have come a long way, but the journey ahead is still long. Today's intelligent machines are far from the breadth and depth of human intelligence. In many real-world applications, as illustrated by AlphaGo and the Allen AI Science Challenge, whether problem formulation falls neatly into complete learning is still being determined. The problem may have a significant component, which can be best modeled using an AI algorithm without the learning component. However, there may be additional constraints or missing knowledge that takes the problem outside its regime, and learning may help to fill the gap. Similarly, programmed knowledge and reasoning help learners fill their gaps.

In 1959, IBM published a paper in the IBM Journal of Research and Development with an, at the time, obscure and curious title. Authored by IBM's Arthur Samuel, the paper invested the use of machine learning in the game of checkers "to verify the fact that a computer can be programmed so that it will learn to play a better game of checkers than can be played by the person who wrote the program." Although it was not the first publication to use the term "machine learning" per se, Arthur Samuel is widely considered the first person to coin and define machine learning in the form we know today. Samuel's landmark journal submission, Some Studies in Machine Learning Using the Game of Checkers, is also an early indication of homo sapiens' determination to impart our system of learning to artificial machines. In his paper, Arthur Samuel introduces machine learning as a subfield of computer science that allows computers to learn without being explicitly programmed. Almost six decades later, this definition remains widely accepted. Although not directly mentioned in Arthur Samuel's definition, a key feature of machine learning is the concept of self-learning. This refers to applying statistical modeling to detect patterns and improve performance based on data and empirical information without direct programming commands. Arthur Samuel described this as the ability to learn without being explicitly programmed. However, he does not infer that machines formulate decisions without upfront programming. On the contrary, machine learning is heavily dependent on computer programming. Instead, Samuel observed that machines do not require a direct input command to perform a set task but input data.

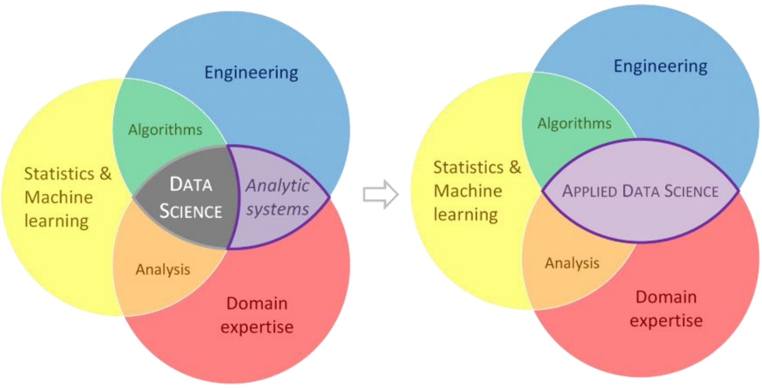


Figure 37. The role of data science and its relation to other disciplines (source: internet)

As can be seen from the figure above, applied data science includes statistical and machine learning methods, algorithms applied jointly with engineering, and domain-specific knowledge and experience. In this book, the latter implies the active involvement of professionals with expertise in a particular aspect of the complex food production system in developing and applying precision agriculture solutions.

Machine learning can be divided into three areas, depending on the nature of the digital data prepared in advance: supervised, unsupervised, and reinforcement learning methods. In many cases, the data analysis generated in a precision agriculture project is based on one algorithm and a combination of several algorithms (assemble learning). These learning paradigms form the backbone of many AI applications, from simple recommendation systems to complex autonomous vehicles. Understanding the intricacies of these methods is essential for anyone delving into artificial intelligence.

6.2 Supervised learning

Supervised learning is akin to a scenario where a student learns under the guidance of a teacher. The 'teacher' here is the labeled dataset, which provides the machine with both the input and the correct output. Supervised learning algorithms aim to create a mathematical model that can make predictions or decisions by learning the relationships between the given input and output.

How Supervised Learning Works

a.) Data Acquisition and Labeling: The first step involves collecting and labeling data as the training set.

b.) Model Training: The algorithm 'learns' from this training set. It attempts to understand the patterns or relationships between the input and output.

c.) Algorithm Tuning: Various parameters of the model are adjusted to improve its accuracy and reduce errors.

d.) Evaluation: After training, the model is tested using a separate dataset, the testing set, to evaluate its predictive power and accuracy.

Applications of Supervised Learning

- Image Recognition: Used in facial recognition systems and image classification tasks.

- Speech Recognition: Powers voice-controlled devices and applications.

- Fraud Detection: Employed in financial services to detect unusual patterns signaling fraudulent activities.

- Predictive Analytics: Used to forecast sales, weather conditions, and stock market trends.

- Email Filtering: Classifies emails into spam and non-spam categories.

- Medical Diagnosis: Assists in diagnosing diseases based on patient data.

- Language Translation: Translates text or speech from one language to another.

Challenges in Supervised Learning

a.) Data Labeling: Extensive and accurate data labeling is required, which can be time-consuming and expensive.

b.) Generalization: The model may not perform well on new, unseen data if overfitting occurs.

c.) Scalability: Managing large volumes of data and the computational resources required for training can be challenging.

Algorithms:

- regression (linear, logistic)
- support vector machine (SVM)
- Bayes(ian) model
- classification (KNN, decision tree, random forest...)
- artificial neural networks (ANN)

6.2.1 Neural networks:

Behind the scenes – human neural networks (HNN)

Neurons are nerve cells that send messages all over the human body, allowing it to do everything from breathing to talking, eating, walking, and thinking. Until recently, most neuroscientists (scientists who study the brain) thought we were born with all the neurons we would ever have. As children, we might grow some new neurons to help build the pathways—called neural circuits—that act as information highways between different areas of the brain. However, scientists believed that once a neural circuit was in place, adding new neurons would change the flow of information and break the brain's communication system (source: https://www.ninds.nih.gov). The central nervous system (which includes the brain and spinal cord) is made up of two basic types of cells:

- Neurons, the nerve cells that send and receive signals
- Glia, cells that provide structure in the brain

In some parts of the brain, there are many more glia than neurons, but neurons are the key players in the brain. Neurons are information messengers. They use electrical and chemical signals to send information between different areas of the brain and between the brain, the spinal cord, and the entire body. Everything we think, feel, and do would be impossible without the work of neurons and their support cells, the glial cells called astrocytes and oligodendrocytes. A neuron has three essential parts: a cell body and two branches called an axon and a dendrite. Within the cell body is a nucleus, which controls the cell's activities and contains the cell's genetic material. The axon looks like a long tail and sends messages from the cell. A dendrite looks like the branch of a tree and receives messages from the cell. Neurons communicate by sending chemicals, called neurotransmitters, across a tiny space called a synapse between the axons and dendrites of nearby neurons.

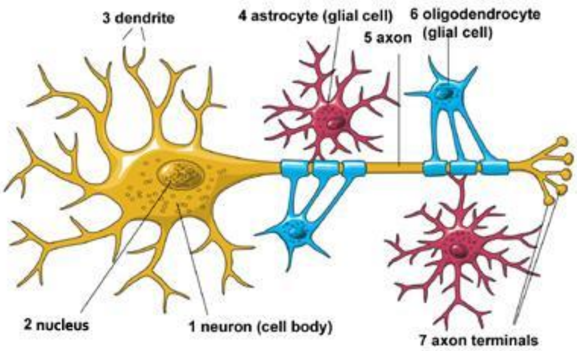


Figure 38. Structure of a human brain cell (source: internet)

There are three kinds of neurons:

- *Sensory neurons* carry information from the sense organs (such as the eyes and ears) to the brain.

- *Motor neurons* control voluntary muscle activity, such as walking and talking, and carry messages from nerve cells in the brain to the muscles.
- *Other neurons*, all of which are called interneurons.

Scientists think neurons are the most diverse kinds of cells in the body. Within these three kinds of neurons are hundreds of different types, each able to send and receive messages differently. How these neurons communicate by making connections makes us unique in thinking, feeling, and acting.

Artificial neural networks (ANN)

Similar to neurons in the human brain, artificial neural networks are formed by interconnected neurons, also called nodes, which interact with each other through axons, called edges. In a neural network, the nodes are stacked in layers and generally start with a broad base. The first layer consists of raw data, such as numeric values, text, images, or sound, divided into nodes. Each node sends information to the next layer of nodes through the network's edges. Each edge has a numeric weight (algorithm) that can be altered and formulated based on experience. If the sum of the connected edges satisfies a set threshold, known as the activation function, it will activate a neuron at the next layer. However, the activation will not be triggered if the sum of the connected edges does not meet the set threshold. This results in an all-or-nothing arrangement.

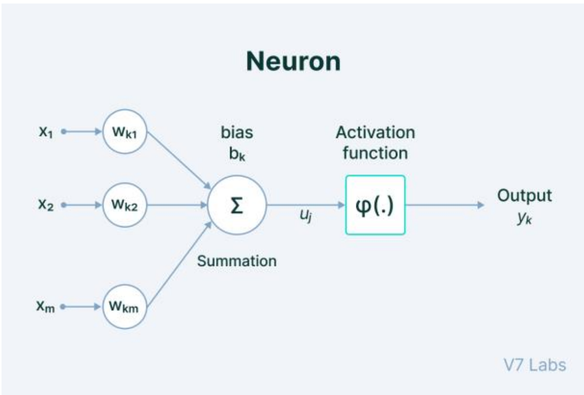


Figure 39. Structure of artificial neural network (source: internet)

Input (x): the set of features fed into the model for the learning process. For example, the input in object detection can be an array of pixel values about an image.

Weight (w): Its primary function is to assign importance to features that contribute more to learning. It introduces scalar multiplication between the input value and the weight matrix. For example, a negative word would impact the decision of the sentiment analysis model more than a pair of neutral words.

Transfer function: The transfer function combines multiple inputs into one output value so that the activation function can be applied. It does this by simply summating all the inputs to the transfer function.

Activation Function: It introduces non-linearity in the working of perceptrons to consider varying linearity with the inputs. Without this, the output would be a linear combination of input values and would not be able to introduce non-linearity in the network.

Bias: The role of bias is to shift the value produced by the activation function. Its role is similar to that of a constant in a linear function. When multiple neurons are stacked together in a row, they constitute a layer, and multiple layers piled next to each other are called a multi-layer neural network. We have described the main components of this type of structure below.

To train the network through supervised learning, the model's predicted output is compared to the actual output (that is known to be correct), and the difference between these two results is measured and is known as the cost or cost value. The purpose of training is to reduce the cost value until the model's prediction closely matches the correct output. This is achieved by incrementally tweaking the network's weights until the lowest possible cost value is obtained. This process of training the neural network is called back-propagation. Rather than navigate left to right like how data is fed into a neural network, back-propagation is done in reverse and runs from the output layer on the right towards the input layer on the left. One of the downsides of neural networks is that they operate as a black box. While the network can approximate accurate outcomes, tracing its structure reveals limited or no insight into the variables that impact the outcome. For example, when using a neural network to predict the probable outcome of a Kickstarter (the world's largest funding platform for creative projects) campaign, the network will analyze variables such as campaign category, currency, deadline, and minimum pledge amount. However, it cannot specify their relationships to the final outcome. Moreover, two neural networks with different topologies and weights can produce the same output, making tracing variable relationships to the output even more challenging. Examples of non-black-box models are regression techniques and decision trees (source: Data Science for absolute beginners)

A typical neural network can be divided into input, hidden, and output layers. Data is first received by the input layer, where broad features are detected. The hidden layer(s) then analyze and process the data. Based on previous computations, the data becomes streamlined through the passing of each hidden layer. The final result is shown as the output layer. The middle layers are considered hidden because, like human vision, they covertly break down objects between the input and output layers. For example, when humans see four lines connected in the shape of a square, we instantly recognize those four lines as a square. We do not notice the lines as four independent lines with no relationship to each other. Our brain is conscious only of the output layer.

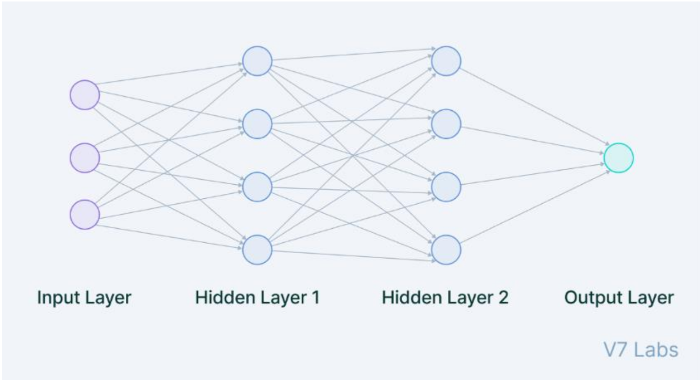


Figure 41. Layers with nodes in artificial neural network (source: internet)

Neural networks work much the same way: They break down data into layers and examine the hidden layers to produce a final output. While there are many techniques for assembling the nodes of a neural network, the simplest method is the feed-forward network. In a feed-forward network, signals flow only in one direction, and there is no loop in the network. The most basic form of a feed-forward neural network is the perceptron.

6.3 Unsupervised learning

This ML method is like a student learning without direct supervision. The algorithm is given data without explicit instructions on what to do with it. It must find structure and patterns within the dataset on its own. This form of learning is ideal for exploratory data analysis, pattern discovery, and finding hidden structures in data.

How Unsupervised Learning Works

a.) Data Collection: The process starts with collecting data that is not labeled or classified.

b.) Pattern Recognition: The algorithm tries to find patterns, groupings, or correlations within the dataset.

c.) Model Adjustment: The model is tweaked to ensure it accurately finds these patterns or groupings.

d.) Interpretation of Results: The outcomes are interpreted to provide insights or to make decisions based on the discovered patterns.

Applications of Unsupervised Learning

- Market Basket Analysis: Understanding customer buying habits by finding associations between purchased items.
- Clustering: Used in customer segmentation to target marketing more effectively.
- Anomaly Detection: Identifying fraudulent transactions in banking or defects in manufacturing
- Dimensionality Reduction: Reducing the number of variables in high-dimensional data, often used in genomics.

- Association Mining: Finding rules that capture relationships between variables in large databases, such as retail.
- Natural Language Processing: Used for topic modeling and understanding document similarities.
- Social Network Analysis: Identifying communities and influencers within networks.

Challenges in Unsupervised Learning

a.) Data Interpretation: The absence of labeled data interprets the results as more subjective and challenging.

b.) Algorithm Selection: Choosing the correct algorithm and parameters is often more complex than supervised learning.

c.) Evaluating Performance: Traditional performance metrics like accuracy are not applicable without labeled data.

Algorithms

- clustering (k-means)
- association algorithms, social network analysis
- reinforcement learning is the third and most advanced algorithm category in machine learning. Unlike supervised and unsupervised learning, reinforcement learning continuously improves its model by leveraging feedback from previous iterations. This differs from supervised and unsupervised learning, which reach an indefinite endpoint after a model is formulated from the training and test data segments. Reinforcement learning can be complicated and is best explained through an analogy to a video game. As players progress through the virtual space of a game, they learn the value of various actions under different conditions and become more familiar with the field of play. Those learned values then inform and influence a player's subsequent behavior, and their performance immediately improves based on their learning and experience. Reinforcement learning is very similar, where algorithms are set to train the model through continuous learning. A standard reinforcement learning model has measurable performance criteria where outputs are not tagged—instead, they are graded. A concrete algorithmic example is Q-learning: the machine learns to associate the action that generates or maintains the highest Q-level with a given state. It learns through a sequence of random actions under different conditions (states). The machine records its outcomes (rewards and punishments) and how these affect its Q level and stores these values to inform and optimize its future actions.

Hyperparameters are parameters whose values control the learning process and determine the values of model parameters that a learning algorithm learns. Hyperparameters are used by the learning algorithm when it is learning, but they are not part of the resulting model. At the end of the learning process, we have the trained model parameters, which we refer to as the model. The hyperparameters that were used during training are not part of this model. We cannot, for instance, know what hyperparameter values were used to train a model from the model itself; we only know the model parameters that were learned. Algorithm selection is essential in forming an accurate prediction model but deploying an algorithm with a high accuracy rate can be difficult. The fact that each algorithm can produce vastly different models based on the hyperparameters provided can lead to dramatically different results. They are the algorithm's settings, similar to the controls on the dashboard of an airplane or the knobs used to tune radio frequency—except hyperparameters are lines of code. A constant challenge in machine learning is navigating underfitting and overfitting, which describe how closely the model follows the actual patterns of the dataset. To understand underfitting and overfitting, we must first understand bias and variance. Bias refers to the gap between the predicted value and the actual value. With high bias, our predictions will likely be skewed in a specific direction away from the actual values. Variance describes how scattered the predicted values are. Bias and variance can be best understood by analyzing the following visual representation.

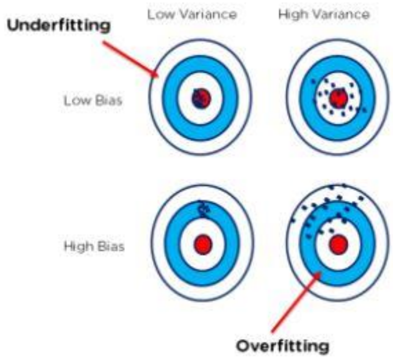


Figure 42. Overfitting and underfitting (source: internet)

Imagine that the center of the target (Figure 42), or the bull's eye, perfectly predicts the correct value of the model. The dots marked on the target represent an individual realization of the model based on the training data. In some instances, the dots will be densely positioned close to the bull's eye, ensuring that predictions made by the model are close to the actual data. In other cases, the training data will be scattered across the target. The more the dots deviate from the bull's eye, the higher the bias and the less accurate the model will be in its overall predictive ability. In the first target, we can see an example of low bias and low variance. Bias is low because the hits are closely aligned to the center, and there is low variance because the hits are densely positioned in one location. The second target (located on the right of the first row) shows a case of low bias and high variance. Although the hits are not as close to the bulls-eye as in the previous example, they are still near the center, and the bias is relatively low. However, there is high variance this time because the hits are spread out from each other. The third target (located on the left of the second row) represents high bias and low variance, and the fourth target (located on the right) shows high bias and high variance. Ideally, we want a situation with low variance and bias. However, a trade-off between optimal bias and variance is more common. Bias and variance both contribute to error, but it is the prediction error that we want to minimize, not bias or variance specifically.

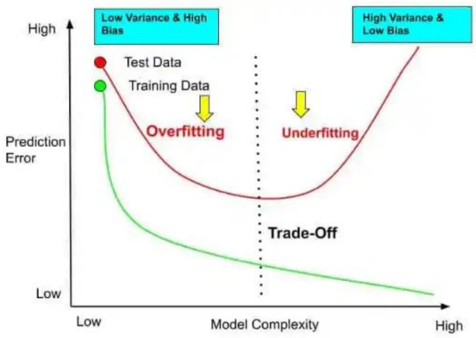


Figure 43. The trade-off between over- and underfitting (source: internet)

In Figure 43, we can see two lines moving from left to right. The line above represents the test data, and the line below represents the training data. From the left, both lines begin at a point of high prediction error due to low variance and high bias. As they move from left to right, they change to the opposite: high variance and low bias. This leads to low prediction error in the case of the training data and high prediction error for the test data. An optimal balance of prediction error between the training and test data is in the middle of the chart. This is a typical case of a bias-variance trade-off.

6.4 Confusion matrix

A confusion matrix is a performance evaluation tool in machine learning. It represents the accuracy of a classification model by displaying the number of true positives, true negatives, false positives, and false negatives. This matrix aids in analyzing model performance, identifying misclassifications, and improving predictive accuracy.

A confusion matrix is an N x N matrix used for evaluating the performance of a classification model, where N is the total number of target classes. The matrix compares the actual target values with those the machine learning model predicted. This gives us a holistic view of how well our classification model performs and what kinds of errors it makes.

For a binary classification problem, we would have a 2 x 2 matrix, as shown below, with four values:

Predicted Values		Actual Values: Positive (1)	Actual Values: Negative (0)
	Positive (1)	TP	FP
	Negative (0)	FN	TN

Figure 44. Confusion matrix (source: internet)

The target variable has two values: Positive or Negative

The columns represent the actual values of the target variable

The rows represent the predicted values of the target variable

True Positive (TP): The predicted value matches the actual value, or the predicted class matches the actual class. The actual value was positive, and the model predicted a positive value. True Negative (TN): The predicted value matches the actual value, or the predicted class matches the actual class. The actual value was negative, and the model predicted a negative value.

False Positive (FP) – Type I Error: The predicted value was falsely predicted. The actual value was negative, but the model predicted a positive value. It is also known as the type I error.

False Negative (FN) – Type II Error: The predicted value was falsely predicted. The actual value was positive, but the model predicted a negative value. Also known as the type II error.

Let me explain these four categories to make it clear with the following figure:

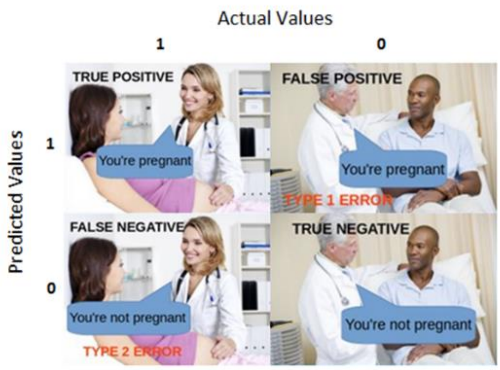


Figure 45. Easy understanding of confusion matrix (source: internet)

Important equations:

- Recall = TP/TP+FN; how many we predicted correctly
- Precision = TP/TP+FP; we have predicted as positive; how many are positive
- Accuracy: from all the classes (positive and negative), how many of them we have predicted correctly
- F-measure = 2 x recall x precision/recall + precision; F-score helps to measure Recall and Precision at the same time

6.5 Brief description of the most commonly used algorithms

Linear regression

One of the simplest algorithms in machine learning is regression analysis, which determines the strength of a relationship between variables. Linear regression comprises a straight line that splits the data points on a scatterplot. The goal of linear regression is to split the data to minimize the distance between the regression line and all data points on the scatterplot. This means that if we draw a vertical line from the regression line to each data point on the graph, the aggregate distance of each point would equate to the smallest possible distance to the regression line. The technical term for the regression line is the hyperplane. Another important feature of regression is the slope, which can be conveniently calculated by referencing the hyperplane. As one variable increases, the other variable will increase at the average value denoted by the hyperplane. The slope is, therefore, very useful in formulating predictions. The closer the data points are to the regression line, the more accurate the final prediction. If there is a high degree of deviation between the data points and the regression line, the slope will provide less accurate predictions.

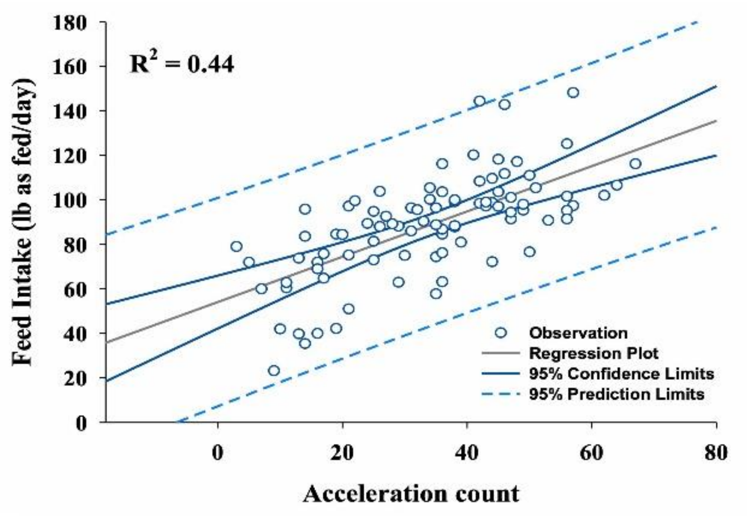


Figure 46. Linear regression (source: internet)

Figure 46 shows the application of linear regression in a research study on a dairy cattle farm. An accelerometer (independent variable) was attached to the cows' heads, and its value was used to determine feed intake (dependent variable).

Logistic regression (Figure 47)

The sigmoid function produces an S-shaped curve that can convert any number and map it into a numerical value between 0 and 1, but it does so without ever reaching those exact limits. A typical application of the sigmoid function is found in logistic regression. Logistic regression adopts the sigmoid function to analyze data and predict discrete classes in a dataset. Although logistic regression shares a visual resemblance to linear regression, it is technically a classification technique. Whereas linear regression addresses numerical equations and forms numerical predictions to discern relationships between variables, logistic regression predicts discrete classes. Logistic regression is typically used for binary classification to predict two discrete classes. To do this, the sigmoid function is added to compute the result and convert numerical results into an expression of probability between 0 and 1. In a binary case, 0 represents no chance of occurring, and 1 represents a certain chance. The degree of probability for values between 0 and 1 can be calculated according to how close they rest to 0 (impossible) or 1 (particular possibility) on the scatterplot. Given its strength in binary classification, logistic regression is used in many fields, including fraud detection, disease diagnosis, emergency detection, loan default detection, and spam emails, which are used to identify specific classes. Logistic regression with more than two outcome values is multinomial logistic regression.

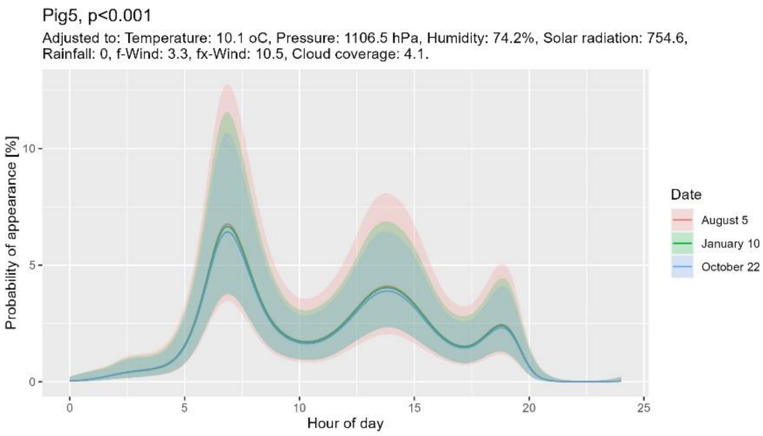


Figure 47. Logistic regression in pig monitoring in outdoor farming

In one of our precision livestock research studies, as shown in the figure above, the data were analyzed using logistic regression. We detected the presence of pigs on pasture in a given area of the field. Data were collected using RFID technology. The binary dependent variable was the appearance of the sow (0 and 1 value), and the independent variables were the effect of weather parameters and time of day. Three days were selected to plot the probability of sow appearance at the experimental site.

Support vector machine (SVM)

As an advanced category of regression, support vector machine (SVM, Figure 48.) resembles logistic regression but with stricter conditions. To that end, SVM is superior at drawing classification boundary lines. A support vector machine (SVM) is a supervised learning algorithm used in machine learning to solve classification and regression tasks; SVMs are particularly good at solving binary classification problems, which require classifying the elements of a data set into two groups.

A support vector machine algorithm aims to find the best possible line, or decision boundary, that separates the data points of different data classes. This boundary is called a hyperplane when working in high-dimensional feature spaces. The idea is to maximize the margin, which is the distance between the hyperplane and the closest data points of each category, thus making it easy to distinguish data classes. SVMs help analyze complex data that a simple straight line cannot separate. Nonlinear SMVs do this by using a mathematical trick that transforms data into higher-dimensional space, where it is easier to find a boundary.

To do this, SVMs use a kernel function. Instead of explicitly calculating the coordinates of the transformed space, the kernel function enables the SVM to implicitly compute the dot products between the transformed feature vectors and avoid handling expensive, unnecessary computations for extreme cases. SVMs can handle both linearly separable and non-linearly separable data. They use different kernel functions, such as the linear kernel, polynomial kernel, or radial basis function (RBF) kernel. These kernels enable SVMs to capture complex relationships and patterns in the data effectively. During the training phase, SVMs use a mathematical formulation to find the optimal hyperplane in a higher-dimensional space, often called the kernel space. This hyperplane is crucial because it maximizes the margin between data points of different classes while minimizing classification errors. The kernel function plays a critical role in SVMs, making it possible to map the data from the original feature space to the kernel space. The choice of kernel function can significantly impact the performance of the SVM algorithm; choosing the best kernel function for a particular problem depends on the characteristics of the data.

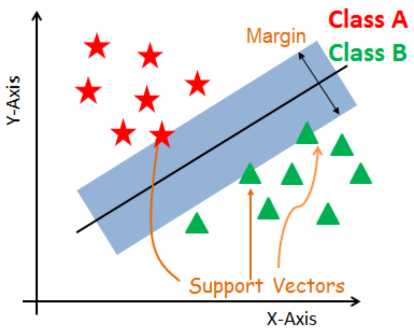


Figure 48. Sample of support vector machine (source: internet)

Support vector machines have different types and variants that provide specific functionalities and address specific problem scenarios. Here are two types of SVMs and their significance (Figure 49):

- Linear SVM: Linear SVMs use a linear kernel to create a straight-line decision boundary that separates different classes. They are effective when the data is linearly separable or when a linear approximation is sufficient. Linear SVMs are computationally efficient and have good interpretability, as the decision boundary is a hyperplane in the input feature space.
- Nonlinear SVM: Nonlinear SVMs address scenarios where the data cannot be separated by a straight line in the input feature space. They achieve this by using kernel functions that implicitly map the data into a higher-dimensional feature space, where a linear decision boundary can be found. Popular kernel functions used in this type of SVM include the polynomial kernel, Gaussian (RBF) kernel, and sigmoid kernel. Nonlinear SVMs can capture complex patterns and achieve higher classification accuracy than linear SVMs.

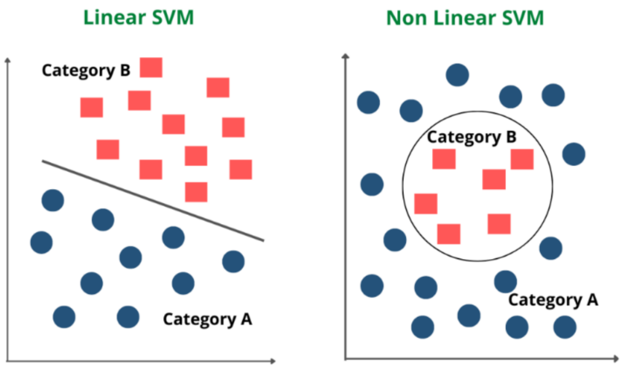


Figure 49. The difference between linear and nonlinear SVM (source: internet)

Advantages of SVM:

Effective in high-dimensional spaces: High-dimensional data refers to data in which the number of features is larger than the number of observations, i.e., data points. SVMs perform well even when the number of features is larger than the number of samples. They can handle high-dimensional data efficiently, making them suitable for applications with many features.

Resistant to overfitting: SVMs are less prone to overfitting than other algorithms, like decision trees -- overfitting is where a model performs exceptionally well on the training data but becomes too specific to that data and cannot generalize to new data. SVMs' use of the margin maximization principle helps generalize unseen data well.

Versatile: SVMs can be applied to both classification and regression problems. They support different kernel functions, enabling flexibility in capturing complex relationships in the data. This versatility makes SVMs applicable to a wide range of tasks.

Effective in cases of limited data: SVMs can work well even when the training data set is small. Support vectors ensure that only a subset of data points influences the decision boundary, which can be beneficial when data is limited.

Ability to handle nonlinear data: SVMs can implicitly handle non-linearly separable data using kernel functions. The kernel trick enables SVMs to transform the input space into a higher-dimensional feature space, making it possible to find linear decision boundaries.

Disadvantages of SVM:

Computationally intensive: SVMs can be computationally expensive, especially when dealing with large data sets. The training time and memory requirements increase significantly with the number of training samples.

Sensitive to parameter tuning: SVMs have parameters such as the regularization parameter and the choice of kernel function. The performance of SVMs can be sensitive to these parameter settings. Improper tuning can lead to suboptimal results or longer training times.

Lack of probabilistic outputs: SVMs provide binary classification outputs and do not directly estimate class probabilities. Additional techniques, such as Platt scaling or cross-validation, are needed to obtain probability estimates.

Difficulty interpreting complex models: SVMs can create complex decision boundaries, especially when using nonlinear kernels. This complexity may make it challenging to interpret the model and understand the underlying patterns in the data.

Scalability Issues: SVMs may face scalability issues when applied to large data sets. Due to memory and computational constraints, training an SVM on millions of samples can become impractical.

k-Nearest Neighbors and K-means algorithms

Clustering analysis is a supervised and unsupervised learning technique. As a supervised learning technique, it classifies new data points into existing clusters through k-nearest neighbors (k-NN). As an unsupervised learning technique, it is applied to identify discrete groups of data points through k-means clustering.

Using the k-Nearest Neighbours algorithm, the data points on the scatter plot are grouped into clusters, and then a new data point whose class is unknown is added to the graph. We can predict the class of the new data point based on its relationship to the existing data points. We need to set the value "k" to determine how many data points we want to assign to the new data point classification. If k is set to 3, k-NN will only analyze the relationship of the new data point to the three nearest data points (neighbors).

Although generally a highly accurate and simple technique to learn, storing an entire dataset and calculating the distance between each new data point and all existing data points does place a heavy burden on computing resources. Thus, k-NN is generally not recommended for use with large datasets. Another potential downside is that applying k-NN to high-dimensional data (3-D and 4-D) with multiple features can be challenging. Measuring multiple distances between data points in a three- or four-dimensional space is taxing computing resources and complicating accurate classification. Reducing the total number of dimensions through a descending dimension algorithm such as Principle Component Analysis (PCA) or merging variables is a common strategy to simplify and prepare a dataset for k-NN analysis.

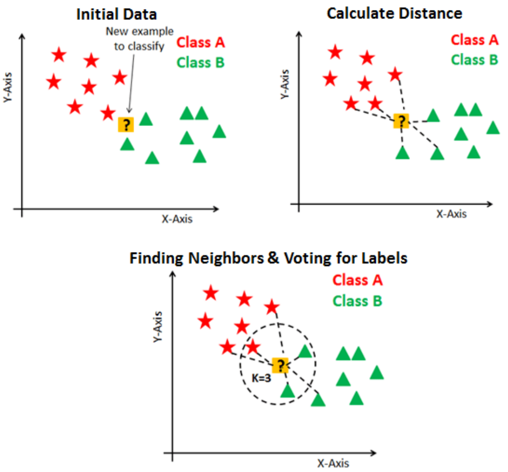


Figure 50. KNN-algorithm (source: internet)

With the K-means algorithm, we can divide the procedure into three steps:

- step 1: Divide the data into k clusters and select the centroids of the k clusters. Centroids can be chosen randomly, meaning we can nominate any data point on the scatterplot to act as a centroid. Each data point can only be assigned to one cluster without overlap. Each cluster is discrete;
- step 2: The average of all data points in each cluster is summed;
- step 3: These x and y values are inserted to update the centroid coordinates. In setting k, striking the correct number of clusters is essential. As "k" increases, clusters become smaller, and variance falls. However, the downside is that neighboring clusters become less distinct as "k" increases. A more simple and non-mathematical approach to setting k is applying domain knowledge.

Once we reach the stage where the data points no longer have extended change clusters, we have the final set of clusters after the centroid coordinates are updated. The k-means algorithm helps cluster animal species or segment the housing market.

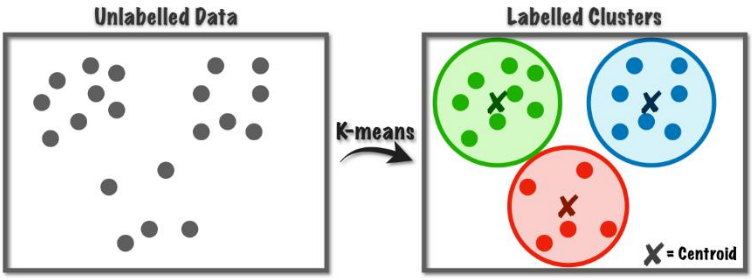


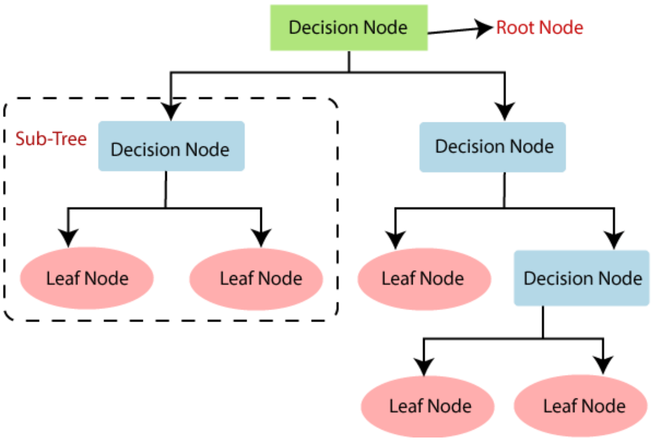
Figure 51. K-means algorithm (source: internet)

Decision trees and random forests

As a supervised learning technique, decision trees are used primarily for solving classification problems, but they can also be applied to solve regression problems; classification trees can use quantitative and categorical data to model categorical outcomes. Regression trees also use quantitative and categorical data but instead model quantitative outcomes. Decision trees start with a root node, which acts as a starting point (at the top) and is followed by splits that produce branches. The statistical/mathematical term for these branches is edges. The branches then link to leaves, also known as nodes, which form decision points. A final categorization is produced when a leaf does not generate any new branches, which results in what is known as a terminal node. They also produce a neat visual flowchart for the presentation of the results. Decision trees are built by first splitting data into two groups. This binary splitting process is then repeated at each branch (layer). The aim is to select a binary question that best splits the data into two homogenous groups at each branch of the tree, such that it minimizes the level of data entropy at the next. Entropy is a mathematical term that explains the measure of variance in the data among different classes. We want the data at each layer to be more homogenous than the last.

A caveat to remember when using decision trees is their susceptibility to overfitting. The cause of overfitting, in this case, is the training data. Considering the patterns in the training data, a decision tree is precise when training the first round of data. However, the same decision tree may then fail to predict the test data, as there could be rules that it is yet to encounter or because the training or test data were not representative of the entire dataset. Moreover, because decision trees are formed from repeatedly splitting data points into two partitions, a slight change in how the data is split at the top or middle of the tree can dramatically alter the final prediction. This can produce a different tree altogether! The offender, in this case, is our greedy algorithm. From the first split of the data, the greedy algorithm focuses on picking a binary question that best partitions data into two homogenous groups. It is oblivious to the future repercussions of its short-term actions.

Random Forests: Rather than striving for the most efficient split at each round of recursive partitioning, an alternative technique is to construct multiple trees and combine their predictions to select an optimal path of classification or prediction. This involves a randomized selection of binary questions to grow multiple decision trees.



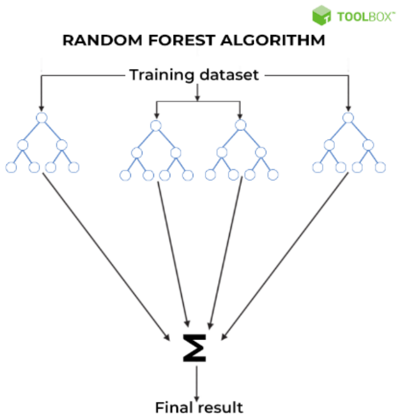


Figure 52-53. Decision tree and random forest (source: internet)

The key to understanding random forests is first to understand bootstrap sampling. There is little use in compiling five or ten identical models—some variation must be added. This is why bootstrap sampling draws on the same dataset but extracts a different variation at each turn. Hence, in growing random forests, multiple varying copies of the training data are first run through each tree. For classification problems, bagging undergoes a voting process to generate the final class. The results from each tree are compared and voted on to create an optimal tree to produce the final model, known as the final class. For regression problems, value averaging is used to generate a final prediction.

Ensemble modeling

One of the most effective machine learning methodologies is ensemble modeling, also known as ensembles. Ensemble modeling combines statistical techniques to create a unified prediction model. Combining estimates and following the crowd's wisdom, ensemble modeling performs a final classification or outcome with better predictive performance. Ensemble models can be classified into various categories, including sequential, parallel, homogenous, and heterogeneous. For sequential ensemble models, the prediction error is reduced by adding weights to classifiers that previously misclassified data. Gradient boosting and AdaBoost are two examples of sequential models. Conversely, parallel ensemble models work concurrently and reduce error by averaging. Decision trees are an example of this technique.

Ensemble models can also be generated using a single technique with numerous variations (known as a homogeneous ensemble) or through different techniques (known as a heterogeneous ensemble). An example of a homogeneous ensemble model would be numerous decision trees working together to form a single prediction (bagging).

Meanwhile, an example of a heterogeneous ensemble would be the usage of k-means clustering or a neural network in collaboration with a decision tree model.

Naturally, it is essential to select techniques that complement each other. Neural networks, for instance, require complete data for analysis, whereas decision trees can effectively handle missing values. Together, these two techniques provide added value over a homogeneous model. The neural network accurately predicts the majority of instances that provide a value, and the decision tree ensures that there are no "null" results that would otherwise be incurred from missing values in a neural network. The other advantage of ensemble modeling is that aggregated estimates are generally more accurate than any single estimate.

Bagging, as we know, is short for "boosted aggregating" and is an example of a homogenous ensemble. This method draws upon randomly drawn datasets and combines predictions to design a unified model based on a voting process among the training data. Expressed in another way, bagging is a unique process of model averaging. Random forest, as we know, is a famous example of bagging.

Boosting is a popular alternative technique for forming a final model by addressing errors and data misclassified by the previous iteration. Gradient boosting and AdaBoost are both famous examples of boosting.

A bucket of models trains numerous different algorithmic models using the same training data and then picks the one that performed most accurately on the test data.

6.6 Computer vision

Computer vision is a field of study within artificial intelligence (AI) that focuses on enabling computers to Intercept and extract information from images and videos like human vision. It involves developing algorithms and techniques to extract meaningful information from visual inputs and make sense of the visual world.

Given that precision agriculture solutions are increasingly based on a database of digital images and video from cameras, the importance of computer vision technologies is growing. In this section, I will present the general characteristics of computer vision, which I will refer to in the following chapter on precision agriculture technologies.

Thanks to advances in artificial intelligence and innovations in deep learning and neural networks, the field has taken great leaps in recent years and has been able to surpass humans in some tasks related to detecting and labelling objects. One of the driving factors behind the growth of computer vision is the amount of data we generate today, which is then used to train and improve computer vision.

Along with a tremendous amount of visual data (more than 3 billion images are shared online daily), the computing power required to analyze the data is now accessible. As the field of computer vision has grown with new hardware and algorithms, so has the accuracy rates for object identification. In less than a decade, today's systems have reached 99 percent accuracy from 50 percent, making them more accurate than humans at quickly reacting to visual inputs.

Early experiments in computer vision started in the 1950s, and by the 1970s, it was first used commercially to distinguish between typed and handwritten text. Today, the applications of computer vision have grown exponentially.

How Does Computer Vision Work?

One of the major open questions in Neuroscience and Machine Learning is: How exactly do our brains work, and how can we approximate that with our own algorithms? There are very few working and comprehensive theories of brain computation, so despite the fact that Neural Nets are supposed to "mimic the way the brain works," nobody is quite sure if that's actually true.

The same paradox holds true for computer vision—since we don't know how the brain and eyes process images, it's difficult to say how well the algorithms used in production approximate our own internal mental processes.

On a certain level, Computer vision is all about pattern recognition. So, one way to train a computer to understand visual data is to feed it images—lots of images, thousands, millions if possible—that have been labeled and then subject those to various software techniques or algorithms that allow the computer to hunt down patterns in all the elements that relate to those labels.

So, for example, if we feed a computer a million images of a cat, it will subject them all to algorithms that let them analyze the colors in the photo, the shapes, the distances between the shapes, where objects border each other, and so on so that it identifies a profile of what "cat" means. When finished, the computer will (in theory) be able to use its experience if fed other unlabelled images to find the cat's ones. Each pixel's brightness is represented by a single 8-bit number ranging from 0 (black) to 255 (white). Computers usually read color as a series of 3 values — red, green, and blue (RGB) — on that same 0–255 scale. Now, each pixel has three values for the computer to store in addition to its position. Much memory is required for one image, and many pixels are required for an algorithm to iterate over. However, to train a model with meaningful accuracy, especially when talking about Deep Learning, we usually need tens of thousands of images, and the more, the merrier.

The Evolution Of Computer Vision

Before the advent of deep learning, the tasks that computer vision could perform were minimal and required a lot of manual coding and effort by developers and human operators. For instance, if we wanted to perform facial recognition, we would have to perform the following steps:

Create a database: We had to capture individual images of all the subjects we wanted to track in a specific format.

Annotate images: Then, for every individual image, we would have to enter several key data points, such as the distance between the eyes, the width of the nose bridge, the distance between the upper lip and the nose, and dozens of other measurements that define each person's unique characteristics.

Capture new images: We would have to capture new images, whether from photographs or video content. And then, we had to go through the measurement process again, marking the critical points on the image. We also had to factor in the angle the image was taken.

After all this manual work, the application could finally compare the measurements in the new image with the ones stored in its database and tell us whether it corresponded with any of the profiles it was tracking. Very little automation was involved, and most of the work was done manually. Moreover, the error margin was still significant. Machine learning provided a different approach to solving computer vision problems. With machine learning, developers no longer need to code every rule into their vision applications manually. Instead, they programmed "features," smaller applications that could detect specific image patterns. They then used statistical learning algorithms such as linear regression, logistic regression, decision trees, or support vector machines (SVM) to detect patterns, classify images, and detect objects. Machine learning helped solve many historically challenging problems for classical software development tools and approaches. For instance, years ago, machine learning engineers could create software that could predict breast cancer survival windows better than human experts. However, building the software's features required the efforts of dozens of engineers and breast cancer experts and took much time to develop.

Deep learning is a fundamentally different approach to machine learning. It relies on neural networks, a general-purpose function that can solve any problem represented by examples. When we provide a neural network with many labeled examples of a specific kind of data, it can extract common patterns between those examples and transform them into a mathematical equation that will help classify future pieces of information.

For instance, creating a facial recognition application with deep learning only requires us to develop or choose a preconstructed algorithm and train it with examples of the faces of the people it must detect. Given enough examples (lots of examples), the neural network can detect faces without further instructions on features or measurements. Deep learning is a very effective method for doing computer vision. In most cases, creating an excellent deep learning algorithm comes down to gathering a large amount of labeled training data and tuning the parameters, such as the type and number of layers of neural networks and training epochs. Compared to previous types of machine learning, deep learning is easier and faster to develop and deploy.

Most current computer vision applications, such as cancer detection, self-driving cars, and facial recognition, use deep learning. Thanks to the availability and advances in hardware and cloud computing resources, deep learning and deep neural networks have moved from the conceptual realm into practical applications.

Applications Of Computer Vision

- Computer vision is one area in Machine Learning where core concepts are already being integrated into major products that we use every day.
- CV In Self-Driving Cars: computer vision enables self-driving cars to make sense of their surroundings. Cameras capture video from different angles around the car and feed it to computer vision software, which then processes the images in real-time to find the extremities of roads, read traffic signs, and detect other cars, objects, and pedestrians. The self-driving car can then steer its way on streets and highways, avoid hitting obstacles, and (hopefully) safely drive its passengers to their destination.
- CV In Facial Recognition: computer vision also plays a vital role in facial recognition applications, the technology that enables computers to match images of people's faces to their identities. Computer vision algorithms detect facial features in images and compare them with databases of face profiles. Consumer devices use facial recognition to authenticate the identities of their owners. Social media apps use facial recognition to detect and tag users. Law enforcement agencies also use facial recognition technology to identify criminals in video feeds.
- CV In Augmented Reality & Mixed Reality: computer vision also plays a vital role in augmented and mixed reality, the technology that enables computing devices such as smartphones, tablets, and smart glasses to overlay and embed virtual objects on real-world imagery. Using computer vision, AR gear detects objects in the real world to determine the locations on a device's display to place a virtual object. For instance, computer vision algorithms can help AR applications detect planes such as tabletops, walls, and floors, an essential part of establishing depth and dimensions and placing virtual objects in the physical world.
- CV In Healthcare: computer vision has also been an essential part of advances in health tech. Computer vision algorithms can help automate tasks such as detecting cancerous moles in skin images or finding symptoms in X-ray and MRI scans.

Challenges of Computer Vision

Inventing a machine that sees like we do is a deceptively tricky task, not just because it is hard to make computers do it but because we are still determining how human vision works in the first place. Studying biological vision requires understanding the perception organs, like the eyes, and the interpretation of the perception within the brain. Much progress has been made, both in charting the process and in discovering the tricks and shortcuts used by the system, although, like any study involving the brain, there is a long way to go.

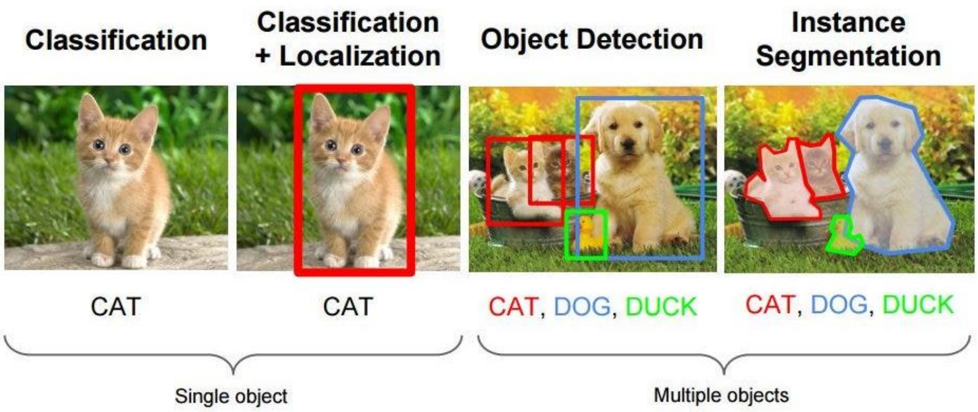


Figure 54. Computer vision tasks (source: internet)

Many popular computer vision applications involve trying to recognize things in photographs; for example:

Object Classification: What broad category of object is in this photograph?

Object Identification: Which type of a given object is in this photograph?

Object Verification: Is the object in the photograph?

Object Detection: Where are the objects in the photograph?

Object Landmark Detection: What are the critical points for the object in the photograph?

Object Segmentation: What pixels belong to the object in the image?

Object Recognition: What objects are in this photograph, and where are they?

Outside of just recognition, other methods of analysis include:

Video motion analysis uses computer vision to estimate the velocity of objects in a video or the camera itself. In image segmentation, algorithms partition images into multiple sets of views.

Scene reconstruction creates a 3D model of a scene inputted through images or video.

In image restoration, noise such as blurring is removed from photos using Machine Learning based filters.

Any other application that involves understanding pixels through software can safely be labeled as computer vision.

Figure 55. The market size of computer vision techniques (source: internet)

Challenges in computer vision:

- Computer vision is changing how machines understand images. However, it faces several challenges, including ensuring data quality, processing data quickly, the effort needed for labeling data, scaling, and addressing privacy and ethical issues. Effectively addressing these challenges will ensure computer vision's advancement aligns with both tech progress and human values.
- Quality of Raw Material

This addresses the clarity and condition of input images or videos, which is crucial for system accuracy. Specific challenges include poor lighting, obscured details, object variations, and cluttered backgrounds. Enhancing input quality is vital for the accuracy and reliability of computer vision systems:

- Enhanced Image Capture: Use high-quality cameras and adjust settings to optimize lighting, focus, and resolution.
- Preprocessing: Apply image preprocessing methods like normalization, denoising, and contrast adjustment to improve visual clarity.
- Data Augmentation: Increase dataset diversity through techniques like rotation, scaling, and flipping to make models more flexible
- Advanced Filtering: Use filters to remove background noise and isolate essential features within the images.
- Manual Inspection: Continuously review and clean the dataset to remove irrelevant or low-quality images.

Real-Time Processing

Real-time processing in computer vision requires powerful computing to analyze videos or large image sets for immediate-action applications quickly. This includes interpreting data instantly for tasks like autonomous driving, surveillance, and augmented reality, where delays can be critical. Minimizing latency and maximizing accuracy is critical for the need for a fast, accurate algorithm in live scenarios:

- Optimized Algorithms: Develop and use algorithms specifically designed for speed and efficiency in real-time analysis.
- Hardware Acceleration: Use GPUs and specialized processors to speed up data processing and analysis.
- Edge Computing: Process data on or near the device collecting it, reducing latency by minimizing data transmission distances.
- Parallel Processing: Implement simultaneous data processing to improve throughput and reduce response times.
- Model Simplification: Streamline models to lower computational demands while maintaining accuracy.

Data Labelling

Labeling images manually for computer vision demands significant time and labor, with the accuracy of these labels being critical for model reliability. The extensive volume creates a significant bottleneck in advancing computer vision applications. Embracing automation and advanced methodologies in data labeling is critical to creating adequate datasets:

- Automated Labelling Tools: Use AI to auto-label images, reducing manual effort and increasing efficiency.
- Crowdsourcing: Use crowdsourced platforms to distribute labeling tasks among a large pool of workers.
- Semi-Supervised Learning: Minimize labeling by combining a few labeled examples with many unlabelled ones.
- Active Learning: Prioritize labeling of the most informative data that benefits model training, optimizing resource use.
- Quality Control Mechanisms: Establish robust quality control checks for accurate label verification, mixing automation with expert human review.

Scalability

Scalability in computer vision faces challenges like adapting technologies to new areas, needing large amounts of data for model retraining, and customizing models for specific tasks. To advance scalability across diverse industries, we need to focus on efficiency at each stage:

- Adaptable Models: Create models that can easily adjust to different tasks with minimal retraining.
- Transfer Learning: Use pre-trained models on new tasks to reduce the need for extensive data collection.
- Modular Systems: Design systems with interchangeable parts to easily customize for various applications.
- Data Collection: Focus on efficiently gathering and labeling data needed for retraining models.
- Model Generalization: Improve models' ability to perform well across diverse data sets and environments.

6.7 Ethical and Privacy Concerns

These issues highlight the need for careful surveillance and facial recognition handling to safeguard privacy. Solving these challenges requires clear rules for data use, openness about technology applications, and legal support.

- Data Protection Policies: Establish strict guidelines for collecting, storing, and using visual data to ensure privacy.
- Transparency: Communicate to users how their data is being used and for what purpose, fostering trust.
- Consent Mechanisms: Ensure individuals provide informed consent before their data is captured or analyzed.
- Legal Frameworks: Create robust legal protections that define and enforce the ethical use of computer vision technologies.
- Public Dialogue: Involve the community in discussions about the deployment and implications of computer vision to address societal concerns and expectations.

7. Digitalization in crop and livestock farming

7.1 Principles of Precision Agriculture

The critical principle of precision agriculture is to ensure plant and animal well-being under large-scale conditions by minimizing damage to the natural environment and maximizing economic efficiency and social utility by applying digital technologies throughout the entire food supply chain.

Information technology provides the opportunity to understand complex agricultural processes better and discover and analyze internal, often hidden, interrelationships and interactions. This chapter establishes the link between agricultural production, briefly described in the first chapter, and the data science methods presented in the second chapter. I aim to explain only some of the precision agriculture technologies developed and applied in the world so far, as new innovative solutions are being created daily in crop and livestock production. Instead, I will focus on the conditions for applying digital methods and the standard features of the production process of crop and livestock products from an IT point of view. For this purpose, I refer to the figure below (Figure 56).

Figure 56. Data process procedure (source: internet)

This figure illustrates the steps that can be taken to explain and sketch a solution to a practical agricultural problem. These will be discussed in detail below, linking food production and IT.

7.1.1 Business understanding

It all starts with business understanding: the first step is understanding the problem's importance. Why is this difficult for the farmer? This is far from trivial because even when reading the first chapter, the reader may have felt that there are too many and too complex interrelated processes, whether we are talking about crop production or livestock farming. There are factors that the two sectors have in common, such as the fact that they work with living organisms that vary in time, space, and organizational structure, that they have different utilization patterns, that they are automated and mechanized, and, last but not least, that their production is geared towards generating income. At the same time, the production of plant and animal products differs in many areas, e.g., farm animals can change position and location, plants can only change position; animals make sounds, plants do not; animal production requires year-round continuous work, and thus resource use, crop production is seasonal, with peaks in work. In arable crops, the largest share of five economic crops is grown in Hungary, in addition to rice at the global level. Except for the latter, the land suitable for large-scale arable crops is rotated, i.e., the crops are grown in a predetermined order. The main reason is that each crop's nutrient requirements (macro, meso, and micro-nutrients) are different and must be considered to maintain soil fertility. However, this does not represent a significant heterogeneity for precision arable farming solutions. For yield mapping, cameras mounted on drones can be used, irrespective of the crop species, sensors placed on the ground among the crop or in the soil, or meteorological stations near the field to collect parameters on the crop and the environment.

In contrast, in large-scale livestock farming, there are considerable variations, which significantly impact the problems that can be solved with the support of digitalization. Cattle, pigs, poultry, sheep, and goats are the most common among large-scale livestock. They represent a very different set of problems due to their farming practices and behavioral patterns, specific to each species, breed, or hybrid created by crossing breeds (the latter for pigs and poultry, which are multiparas). This is discussed in more detail in the sub-section on precision livestock farming.

7.1.2 Data understanding

In providing IT solutions to practical problems in crop and livestock production with the expected added value, the scope of the digital data to be collected needs to be well defined. Practical and theoretical knowledge of agriculture is essential, as is knowledge of IoT tools that can be applied in a resource-efficient way.

Digital data collection can be divided into two broad categories: environmental data collection and animal or crop data collection. Environmental factors significantly impact a living organism's performance due to the interaction between the environment and the animal or plant. Just as a plant or animal affects its environment, so does a plant or animal affect its environment. Examples of environmentally focused data include temperature, humidity, air movement rate, lighting (natural or artificial), quantity and quality of nutrients, and water required to develop the living organism. These data sources can be further specified according to whether we discuss crop or livestock production. Plant- and animal-centered data collections are already very different. Still, there is one thing they have in common for precision technologies: the smaller the unit of data we can analyze, the better we can determine the performance of the population and make predictions for the future.

Data collection tools can be divided into three groups, regardless of whether we are talking about animals or plants, and independently of the production technology. These are sensors, microphones, and cameras. It is clear that microphones have no role in crop production, as plants are incapable of producing sound, but they can be of great importance in animal farming.

Sensors can help collect data about the environment and the living organism in crop production by placing sensors in the crop (e.g., soil sensors) and animal farming by attaching sensors to the animal's body (e.g., a bolus in the rumen of ruminants). These transmit numerical data continuously throughout the production period, and the technology's success depends on the receiving unit's effectiveness and durability (e.g., the reader in the case of RFID).

Microphones can be used in livestock farming in closed housing systems, such as pig or poultry farming. Their primary role is to detect abnormal sounds emitted by livestock. However, their use is made difficult by the high noise level caused by the housing equipment in the building, which must be isolated and separated from the sounds made by the animals.

Cameras are playing an increasing role in precision crop and livestock production systems. They can detect the development of living organisms and any variations (in color and shape) in their development. In large-scale animal farming, in addition to collecting images, video analysis from the cameras allows the identification of dynamic animal behaviors and the detection of abnormalities.

The smallest data collection unit in crop production is the land plot unit subdivided into fields producing the same crop. Since soil can be heterogeneous within a single hectare of land, it is necessary to break it down into smaller sub-units that can still be treated as a single unit.

In livestock farming, the smallest unit is the individual animal for which specific data can be collected. The smaller the body size of a farm animal, the more difficult it is to monitor them individually and mark them with individual identity tags. In addition, the different body structure is also a determining factor: all our farm animals except poultry and insects can have an IoT device with a digital signal attached to their ears, while for poultry, we can only monitor them individually using cameras. In order to create a credible, accurate database that provides valuable information for farmers, a thorough knowledge of the behavior of the farm animal species is required. With this, it is possible to establish the range of data required effectively and false correlations and findings will be made in the analyses.

The figure 57. below shows the features that digital data must have before it can be analyzed.

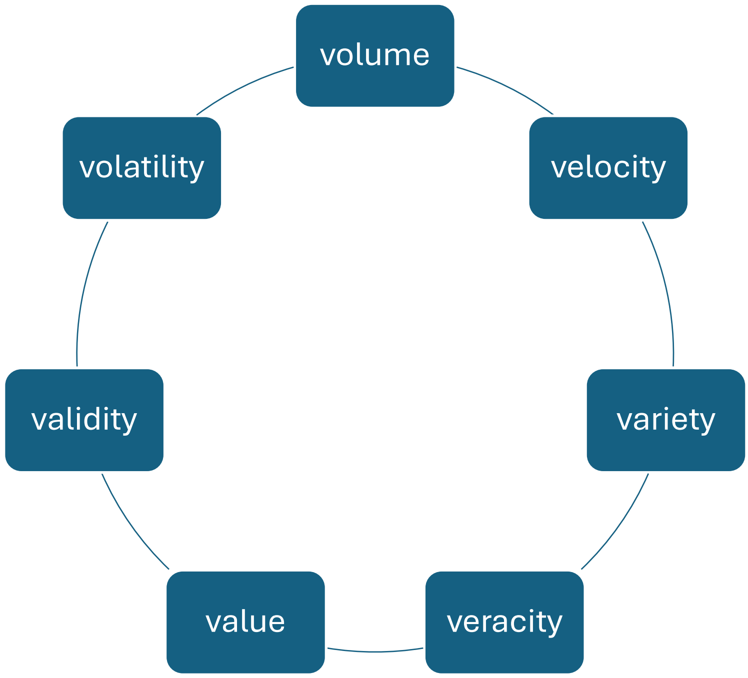


Figure 57. 7Vs in data characteristics (source: internet)

Volume is the total number of data sets that should be analyzed and processed. As data sets grow in size and complexity, they regularly exceed terabytes and petabytes. Because of the sheer amount of data, new processing methods that are different from those used for regular data storage and analysis must be created.

Velocity: how quickly data can be processed and made available.

Variety refers to the range of data sources and types. The form, origin, and format of data can differ according to the source, and the data may be structured or unstructured. In order to extract beneficial insights from big data, it is essential to comprehend the many types of data, where they came from, and how they relate across datasets.

Veracity refers to its accuracy and credibility. Data must be validated to verify that it genuinely reflects critical business operations and that any data processing, modeling, or evaluation has no impact on the data's accuracy. This is also the reliability or integrity of the data a business receives and processes to draw relevant insights.

Value is the end game. After addressing volume, velocity, variety, variability, veracity, and visualization—which takes a lot of time, effort, and resources— we want to be sure our organization is getting value from the data.

Validity refers to the erroneous information we must remove or fix during the data transformation. Data accuracy is directly connected to how much time we will spend cleaning our data.

Volatility: One has to prepare for data volatility, especially in production systems. Data that should "never" be missing disappears, and numbers suddenly contain characters.

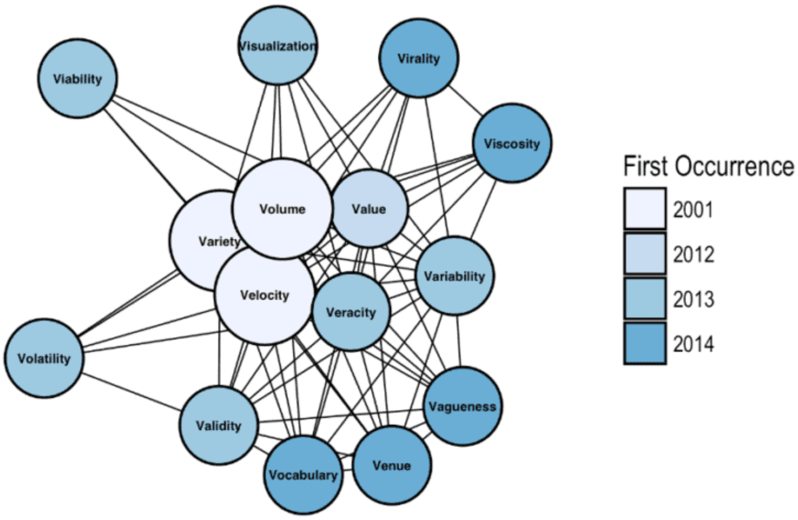


Figure 58. Importance of occurrence of characteristics of data (source: internet)

7.1.3 Data preparation

Data preparation is the process of preparing raw data so that it is suitable for further processing and analysis. Key steps include collecting, cleaning, and labeling raw data into a form suitable for machine learning (ML) algorithms and then exploring and visualizing the data. Data preparation can take up to 80% of the time on an ML project. Using specialized data preparation tools is essential to optimize this process.

Clean data

Cleaning data corrects errors and fills in missing data to ensure data quality. After clean data, we must transform it into a consistent, readable format. This process can include changing field formats like dates and currency, modifying naming conventions, and correcting values and units of measure so they are consistent.

Label data

Data labeling is the process of identifying raw data (images, text files, videos, etc.) and adding one or more meaningful and informative labels to provide context so an ML model can learn from it. For example, labels might indicate if a photo contains a bird or car, which words were mentioned in an audio recording, or if an X-ray discovered an irregularity. Data labeling is required for various use cases, including computer vision, natural language processing, and speech recognition.

Validate and visualize

After data is cleaned and labeled, ML teams often explore it to ensure it is correct and ready for ML. Visualizations like histograms, scatter plots, box and whisker plots, line plots, and bar charts are all valuable tools to confirm that data is correct. Additionally, visualizations also help data science teams complete exploratory data analysis. This process uses visualizations to discover patterns, spot anomalies, test a hypothesis, or check assumptions. Exploratory data analysis does not require formal modeling; data science teams can use visualizations to decipher the data.

The preparation of the database, based on digital data collected in crop and livestock production, includes, among other tasks, fractionating data series resulting from sensor calibration, analyzing outliers, and filling in missing data, for example, by interpolation.

7.1.4 Data modeling

When examining the cleaned and prepared-for-analysis database, the data science model best suited to answer the original question should be selected. Depending on what data the original database contains will largely determine the method of analysis.

Crop simulation models or crop growth models are formal representations of mathematical algorithms that describe the interaction of crops with the environment. Crop models are dynamic models (time is a factor) and usually simulate daily interactions between the soil, plants, weather, and crop management.

7.1.5 Evaluation

After running the data science models, the results should be evaluated. It is at this stage that the interdisciplinary nature of precision agriculture becomes clear once again: agricultural specialists, computer scientists, electrical engineers, veterinarians in the case of animal technology, and, where possible, farmers jointly evaluate the results obtained, the internal patterns and relationships found by the model. This is essential because no single discipline alone can determine a model's outcome. The accuracy of the estimation, the database size, the precision, and the interpretation of outliers are all essential criteria. For example, only a specialist in a particular field can make a professional judgment on the result of a computer vision estimation of the weight of pigs in a herd or the significance of spots on the leaves of a plant population. Given that many problems can be solved using IT, it is inappropriate to list them here. However, they all share that only solutions with a high added value for agriculture can be taken to the next, final phase, i.e., the practical application in farming.

Most precision agriculture technologies aim to integrate their results into agricultural production and support economic and production decisions. They can be used to create recommender, decision-making, and support systems.

7.1.6 Deployment

The development of precision farming technologies will only be successful if farmers apply them in their daily farming practices. There needs to be more than a suitable IT solution; at this stage, it is essential to know the cost of implementation and maintenance and the potential for increased income. Simplicity and understanding of how to use it are also essential. It should be easy to integrate into day-to-day farming practices; for example, cameras for monitoring bird behavior in poultry farming should be fixed to the supporting structure of the shed so that they do not interfere with cleaning tasks. At the same time, they should be able to withstand the harsh conditions of the housing environment.

7.2 Precision crop farming

The practical technological unit of crop production is the field. Nowadays, each technological operation (e.g., tillage, fertilization, sowing, crop protection, irrigation, harvesting) is carried out schematically on a given field, even though there are subfields with significantly different characteristics within the field, especially when the field is more significant. Implementing different agrotechnical interventions within the field that are adapted to the site and local conditions is advisable. Precision farming is an excellent solution to this problem. Precision crop production is a process in which agrotechnical operations are varied and adapted to the agroecological conditions within the field. This requires on-farm positioning (GPS systems), on-farm soil and yield maps, specific agrotechnology developed based on these maps, and the operational software that allows this. At present, the technical conditions for precision farming are mainly in place. However, developing operating software that allows for technically correct implementation adapted to the individual subfields within the field is much more complex. Precision farming can be used effectively in soil cultivation, nutrient management, sowing technology, crop protection, irrigation, and harvesting. It increases agronomic and economic efficiency, reduces input use, improves crop quality, and reduces negative environmental impacts.

7.3 Precision livestock farming

The primary objective of precision livestock technologies is to use machine learning methods to analyze a database of digital data to identify intrinsic patterns and relationships in animal-environment interactions, thereby reliably supporting livestock farmers to make more effective economic decisions and optimize their production resources.

phenotype = genotype x environment

First, let me explain what the elements of the above formula mean:

- *phenotype*: the set of observable characteristics of an individual that results from the interaction of genotype and environment. In simple terms, it is the appearance of the animal that we see and the production results that we expect from it. Its body weight, body shape, leg structure, temperament, egg production, quality and quantity of wool, and so on. It is important that the keeper sells the phenotype, and the farmer is paid according to the yield. This is what we want to maximize.

- *genotype*: the genetic structure of a particular (animal) organism, inherited from its parents and consciously modified by man through artificial selection and high breeding to achieve the expected production level. In large-scale animal production, this is taken to be 100%, which can be surpassed by biotechnological processes (where the genetic structure of an individual is artificially "touched" and altered at specific points to achieve higher performance potentially. This includes GMO organisms; GMO: genetically modified organisms). Practical large-scale livestock production systems aim to achieve this 100% to maximize the revenue from producing a given livestock product.

- *environment*: the environment or circumstances in which a human, animal, or plant lives or functions. Our farm animals are kept in basically two types of housing: free range or confined. We aim to create the optimal housing environment according to the animal's genetics (i.e., genotype), age, utilization direction, and sex. A perfect housing environment to achieve 100%, i.e., maximum phenotypic performance, is impossible in practice. However, it must be pursued, and information technology and precision farming technologies can significantly help achieve this. The document accompanying the next module gives a detailed description of the housing environment. The most important environmental parameters are briefly described here:

- *housing systems*: For outdoor animals, this is the pasture area, a source of feed and living space for ruminants and, for pigs, as a single cavity stomach and litter-consuming animal, mainly living space. In the case of confined housing, the enclosure's flooring, the building's insulation, the condition and integrity of the internal housing technology, and its design, according to the species' needs, are decisive.

- *microclimate*: grazing livestock are in the open air and are therefore more exposed to weather conditions. Temperature, humidity, air pressure, cloud cover, solar radiation, wind speed, precipitation, and the microclimate they create influence the well-being of the animals. Humans do not influence these, but can indirectly help to cope with their extremes. On the other hand, in the case of livestock kept in closed housing conditions, mainly pigs and poultry, humans can have complete control over the microclimate. This can be achieved by using automatic ventilation technology, which uses sensors to monitor the enclosure (temperature, humidity, ammonia levels, air flow) and uses the data collected to adjust the microclimate of the barn to the values previously set by man.

- *forage*: In a pasture-based system typical of ruminant farming, the grasses of the pasture are the primary source of forage. The animals can freely move around the pasture areas and choose the plants they eat. In the case of poorer grass yields, it may be necessary to supplement with preserved fiber fodder (hay) and forage for abstracts. In confined housing, as with microclimate control, this is under human control: the animals can eat what and as much fodder as the human feeds them through an automatic or semi-automatic feeding system

- *drinking water*: Fresh grass consumed on the pasture can satisfy a large part of ruminants' water needs, and humans can help provide more by placing watering troughs. In confined housing, drinking water, like feed, is delivered to the animals through a closed system in the correct quantity and quality.

- *caretakers, workers*: In both types of farming, although to a different extent, there is a human presence around the animals. Animals in confined housing interact less with caretakers, whereas in confined housing, the frequency of contact may be daily. The personality of the caretaker and the way he or she treats the animals are essential factors in successful and profitable pet food production. Today, the problem of securing the correct quantity and quality of human resources in large-scale livestock farming is becoming increasingly important, and IT can help in this respect. In particular, cameras can be placed in the enclosure of a closed barn to replace (but not replace!) the farmer's garbage.

- *the presence of mates*: As living organisms, animals interact not only with their environment but also with their fellow animals. The species-specific and individual characteristics of the farmed animal species determine whether they can be kept and, if so, in what group. Group housing is essential for pigs because they are social animals and seek each other's company. In large-scale poultry farming, humans take advantage of the imprinting phenomenon characteristic of birds. This means that in a flock of poultry hatched from eggs and settled at day old, the birds look to their peers as role models and learn basic behavioral patterns from each other (as they do from their mothers in nature). Without this, they will die. However, group housing can also be dangerous: an aggressive individual in a group of pigs may injure their mates, who cannot escape because of the limited space available. A definite ranking is established between pigs, which remains unchanged (between the same individuals) until the end of the animals' cohabitation.

- *sudden noises and lights*: These are particularly common in confined housing. This is compounded by the technological tolerance mentioned above: the operation of automatic housing technology involves noise (fans, heat blowers, noise from the feed intake system, etc.). A regular and constant noise level does not disturb the animals, and the noise of the feeder signals to them that feeding time is approaching. However, sudden, unexpected, and loud noises can startle them. The same applies to sudden light intensity.

7.4 Precision Livestock Farming research in practice

7.4.1 Practical aspects of image processing-based bird weight measurement (https://doi.org/10.3390/agriculture12111869)

Introduction

Main Characteristics of Precision Livestock Farming

Precision Livestock Farming (PLF) tools can help to provide evidence-based strategies to improve facility design and farm management [1]. Several scientific articles have recently been published on precision farming methods in large-scale livestock production. However, these have yet to be disseminated in a natural farming environment. However, these technologies could contribute to achieving several Sustainable Development Goals (SDGs) [2]. PLF technologies focus on individual animals with data collection and analysis. It gives added value from the evaluation of the results, which helps farmers increase their livestock income and reduce the negative environmental impact of their farming [3]. The cost of inputs (e.g., feed, drinking water, energy, drugs, and human labor) used in the production of animal products can be optimized, and the production conditions can be monitored continuously. A transparent animal product value chain can be achieved from the very first step - from the farms.

The efficient and effective use of PLF depends on the quality of the data collected by digital devices (Internet of Things, IoT). To solve the practical problem defined previously, the most appropriate data collection tool must be found for the specificities of the farm animal species and the farming method [4]. In most cases, cameras can be used in commercial-intensive poultry housing systems to collect the right amount and quality of individual data. One of the main challenges in the poultry sector is that a poultry house can contain tens of thousands of birds simultaneously, making it challenging to distinguish animals from each other and their environment. During rearing, the birds also change in body size, feather cover, and color. The processing of images and videos of birds is part of machine learning within data science. Machine learning also involves several processes and methods, including neural networks. Neural networks are a subset of computer vision systems specifically designed to analyze visual data. Image recognition is one of the tasks in which deep neural networks (DNNs) excel.

The main components of a computer vision system include cameras, recording units, processing units, and models. In the application of a computer vision system, animals (e.g., cattle, sheep, pig, poultry, etc.) are monitored by cameras installed at fixed locations, such as ceilings and passageways, or onto mobile devices like rail systems, ground robots, and drones. Recording units (e.g., network video recorders or digital video recorders) collect images or videos at different views (top, side, or front view) and various types (e.g., RGB, depth, thermal, etc.). Recordings are saved and transferred to processing units for further analysis. Processing units are computers or cloud computing servers [5]. Challenges in processing images and videos collected from animal environments are associated with inconsistent illuminations and backgrounds, variable animal shapes and sizes, similar appearances among animals, animal occlusion and overlapping, and low resolutions [6].

Main characteristics of the poultry and foie grass sector

The EU produces approximately 90% of the world's foie grass. China, the United States, and Canada are the other principal producing countries. Approximately 117 979 tons of foie grass were produced in the European Union in 2021. The largest waterfowl populations are in France and Hungary (68.6 and 20% respectively) [7]. The bird flu outbreaks in recent years have caused severe economic damage to waterfowl farmers. This is because, traditionally, waterfowl are reared in semi-free or free-range systems. Poultry reared in this way are at increased risk of infection with avian influenza. The loss of production caused by such an epidemic is a severe economic loss for farmers; therefore, confinement techniques are increasingly used. Large-scale, commercial waterfowl farming technology is mainly like that used for broiler chicken houses. Thousands of birds are kept in closed buildings under automated and controlled housing conditions.

The Hungarian poultry sector is characterized by integrated production. The integrator company provides the farmer with the day-old ducks, the feed, and the advisory service. The company delivers the animals for slaughter according to an integrator contract with the producer at the contractually agreed price. The maximum mortality rate and the expected average weight of the slaughtered animals are fixed in the contract [8]. In practice, the integrator's consultants check the weight of the birds once during the production period by representative manual weighing (10-15 ducks per weighing). The cooperation with the integrator creates a predictable framework for duck farmers, and the right carcass quality is in the interest of both the integrator company and the duck farmer. Optimizing the amount of inputs used (feed, drinking water, energy, litter, medicine, human resources) and minimizing mortality and animal health risks are in the farmer's best interest. This includes automating keeping technology, making high use of existing equipment, and increasing the speed of production rotation (shortening the length of service periods). The more efficiently the farmer produces ducks, the more profitable his business is. Poultry weight provides information about growth, and the feed conversion efficiency of the flock can be calculated. Whether the aim is fattening or liver production, the weight and health of the ducks and geese are equally essential indicators for farmers. In large-scale poultry production, weight monitoring of the birds is carried out either manually at some points a couple of times during the rearing period or by using digital scales placed among the birds in the building. In both cases, the average of the individual weight measurements of a small proportion of the flock is used to determine the average weight of the whole flock. When large numbers of birds' weights have to be obtained, the conventional method is labor-intensive and time-consuming [9], stressful to birds [10], [11], subject to transcription errors, and prone to human errors [12]. It might be helpful to automatically collect simultaneous information about the growth trend of all the birds to identify deviations from the expected growth trend [13], [9] and also to have details about the health and welfare status of the animals [14].

Our ongoing research is being conducted on a large-scale duck fattening farm in Hungary. Our research objective is to estimate the individual weights of the birds in their live state by machine learning analysis of camera images. This paper presents our experience using the IT method in a natural farming environment. Our research aimed to find the best method to determine the individual weight of ducks, a waterfowl species, using machine learning methods already known and applied in other research fields.

Materials and Methods

The experiment was conducted on fattening ducks housed in a commercial, intensive, indoor system between September 2020 and September 2021 on a private farm in the south-eastern part of Hungary (Kiskundorozsma, GPS coordinates: 46.2667,19.9167). Our study collected data on duck individuals during the whole fattening period. One period takes seven weeks, followed by a two-week service period before a second fattening period. Our experiment included three periods between September 2020 and September 2021. The farmer fattened ducks all year round except during the winter months. At the end of each fattening period, the slaughtered animals are transported to a slaughterhouse under contract with the farm. The poultry farm complies with the current legislation on keeping farm animals, animal welfare, and environmental protection, which is continuously monitored by the authorities (Hungarian State Treasury Agriculture and Rural Development Agency, National Food Chain Safety Office, and Hungarian Association of Poultry Farmers). The poultry farm is run as a family farm; there are no permanent employees; the farmer and his family live 5 km away from the farm and have been fattening ducks for ten years.

Animals

Cherry Valley ducks are fattened on the poultry farm where our research occurred. The Cherry Valley duck is a commercial cross of Pekin ducks, one of the significant duck crosses used for commercial duck meat production in Hungary. It has a high growth rate and reaches a market live weight of 3.45 kg at 42 days of age with an FCR of 1.92 in the case of a medium-sized commercial duck (Figure 59). In one period, 8,000 ducks are fattened and housed in one building from one day to two weeks, and then the ducks are divided into four buildings. Thus, 2000 ducks are fattened in one building until the end of the seven-week fattening period.

Figure 59. Cherry Valley duck

The stock density was 23 ducks/m2 during the two weeks of pre-rearing and 5.7 ducks/m2 post-rearing.

Housing and Management Conditions

The duck fattening farm keeps the ducks in four foil tents, each 7 meters wide, 50 meters long, and 2.5 meters high. Hungary has a long tradition of fattening ducks, and the so-called foil tents are popular among duck farmers. The farmers attach the multi-layered, insulated foil to the metal frame structure (Figure 2). The two shorter sides of the foil tents can be partially folded down. This simple building is perfect for the needs of ducks. In the foil tent where our experiment was carried out, the ventilation technology consisted of a single-phase exhaust fan with a 42,000 m3/h capacity installed at one of the end walls. The air entered the barn naturally, without automatic control.

A simple controller using a temperature sensor switched the fan on and off. In winter, a natural gas space heater ensured the right temperature, while in summer, a low-pressure cooling system was used. The latter consisted of a plastic pipe system attached to the longitudinal side wall of the barn and water spray valves. The ducks were supplied with water by two watering lines with valves placed along the length of the shed. A water pressure regulator at the end of the watering lines regulated the water pressure. The design of the watering valves was adapted to the water requirements of the ducks. The feeding technology consisted of a feed bin, a spiral dry feed filling tube with an overhead track located at the shorter end of the building, close to the outer wall of the barn, from which the feed was delivered through plastic pipes connected to a plastic barrel in the foil tent. A feed level sensor in the last drop tube started and stopped the feed from the silo bin. The plastic barrels fed the feed into the rubber trays below (Figure 61). As the integrator company recommended, the commercial feed was fed to the ducks. Two-phase feeding was used on the farm. The ducks were fed starter feed for the first two weeks and finishing feed for the last five weeks.

Lighting was natural for most of the fattening period, with supplementary artificial lighting consisting of fluorescent tubes attached to the frame system of the foil house used during the shorter days in autumn and spring. The foil tents had no concrete floor, and litter straw was placed on the ground, adding a new layer of straw every three or four days. After the birds had been transported to the slaughterhouse, a two-week break was observed when the foil tents were cleaned, disinfected, rested, and prepared for the next flock of ducks by veterinary regulations. On average, they complete 5-6 fattening periods in a year. In the on-farm circumstances, the birds' weights were only representative, and no bird weighing was possible in the foil tents. Data on bird weights were obtained manually by rounding up and weighing 10-15 ducks per fattening period. The foil tents did not have separate service facilities; these were located approximately 10 m away from the tents in a small building.

Figure 60. The commercial Hungarian duck fattening farm with foil tents (experimental site)

Figure 61. A look inside the barn (experimental location)

Data collection

The main parts of the system are:

- The data collection system (weight sensors and cameras).
- The module for manual and automatic processing of the incoming data.

- The database of the processed data.

Using the database created by combining the images and the measured weight data, a machine learning-based weight estimation system was developed, which can replace the hanging bird scales. The development of filtering components that run on the data is also presented. The aim was to be easily integrated into the pipeline and usable in other machine vision and learning-based projects, customizable depending on the task.

For image data collection, we used outdoor IP66 security IP cameras in the stables, equipped with built-in infrared illumination, fixed dome (BOSCH Felxidome IP 5000i IR NDI-5503-AL 5Megapixel resolution, BOSCH Felxidome IP 3000i IR NDE-3502-AL 2Megapixel resolution) or Bullet camera (BOSCH DINION IP 3000i IR NBE-3502-AL 2Megapixel resolution).

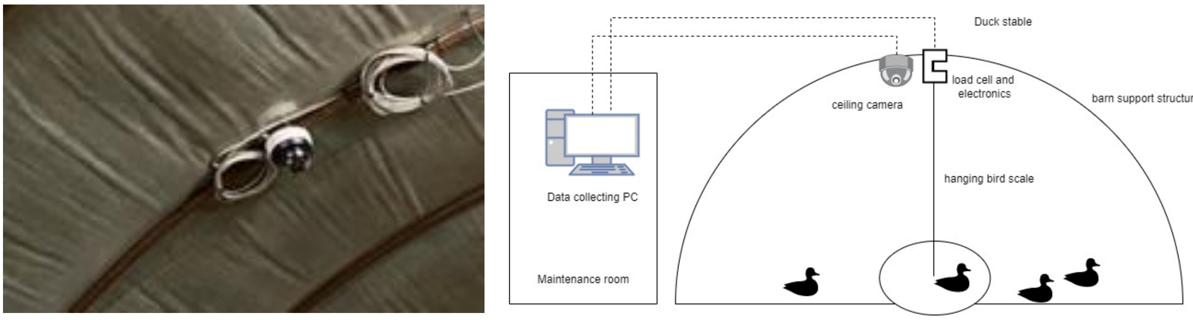


Figure 62. Attaching the camera above the bird scale to the frame of the foil tent

Time stamps were assigned to the collected data so that we could later match the bird weight data - camera image pairs.

The weight data were collected using a bird scale consisting of a suspended weighing plate and a digital data collection unit. The cameras were connected to a Gigabit Ethernet switch in the barn using Cat5e UTP wiring, and power was also supplied over this wire using PoE (Power Over Ethernet) technology. The cameras and the scale were fixed to the barn's metal roof structure using a loose bolt connection and HILTI tape (Figure 4).

The measured weight data were transmitted to the data collection computer via the RS485 interface. The camera images were processed, and a PC with an Intel Core i3 processor recorded the measured weight data. The data were recorded in rotation so that new data overwrote the oldest one, thus providing the capacity needed to store data and images continuously. Data processing was performed on an Asus ROG Zephyrus M15 GU502LW Notebook (Intel(R) Core (TM) i7-10875H CPU @ 2.30GHz, 32 GB RAM, NVIDIA GeForce RTX 2070 with Max-Q Design).

Several considerations had to be analyzed in selecting the development environment. Python was selected because it has several libraries that make development much more accessible in deep learning, big data, and image processing. One of our primary input data was images of the OpenCV library, Tensorflow, and Keras, which contain various essential predefined machine learning models. In addition, the Pandas module helped us in data processing, and for annotation, we used the publicly available LabelMe software.

Processing and verifying data

The first task was to create a database for training and validating the weight estimation system by assembling and saving the bird images and the weight data associated with them.

The measured weight data was labeled with a timestamp and then uploaded to the database. In the next step, we filtered those time stamps for which the system recorded no weights. The measurements of the same timestamp were averaged, and the relevant data were saved in a separate file.

Since the weighing pan's size, material, and shape and the birds' willingness to step on the scales greatly influence the accuracy of the measurement, different materials and shapes of weighing pan were tested. They were contaminated at different rates, and their material and color made them quickly or more challenging to detect and accurately segment birds on the weighing pan. The height of the scale was also significant, as the high hanging point caused the adult, heavier birds to "swing" into the scale when it was used, causing the birds - especially the larger ones - to avoid the scale so that very few measurements were taken and only with smaller weight's ducks. We, therefore, moved the weighing platform closer to the barn floor. The optimal height had to be found because if a pan were too low, the birds would carry littering straws to the weighing pan, causing inaccurate measurements.

Figure 63. The weighing plates used in the experiment

Initially, we used the round weighing plate to detect the weight of the ducks. This scale is widely used in broiler fattening, but our practical experience has shown that the square scale is a better solution for recording individual weight data for waterfowl. Another problem was the removal of feces (guano) and other dirt accumulating on the weighing platform because the weight of these increased the measured weight data, generating a continuous error in the measurements (Figure 5). An automatic zeroing solution was developed, similar to the zeroing of conventional hanging bird scales. The applied system detects when no birds are on the scale based on the camera image and performs the zeroing of the scale at these moments by setting an offset zeroing value corresponding to the idle weight currently measured. The validation of the camera images to be included in the neural network training dataset was performed in several steps. First, we selected images from the collected ones where exactly one bird is visible on the scale, the scale appears to be at rest, and the bird is on the scale with its entire body. As a result, only those images were left in the image database that could serve as input to the feature extractor component. The other images of birds not on the measuring plate with their whole body were removed from the database.

Initially, we started with 45000 color images, of which 2500 color images were extracted into the teaching set. This was further reduced to 800 due to faulty masks. For example, if there were many ducks in one place, the edges of the scale were not visible because of the birds, the reference point was wrong, or there were no ducks on the scale in the evaluated picture. We also removed images from the database where only the head or the breast was placed on the scale. These cases were filtered during the comparison with the weighted data, and the measured values that did not match the development curve of the ducks and the corresponding images were also removed from the cleaned database. Consecutive, identical images were averaged. We also discarded images that did not carry additional information (e.g., ducks sleeping on the scales, multiple images of the same duck in the same pose, etc.)

Finally, we selected 449 images where a duck was visible on the scale. From these 225 pictures, the scale data was also correct, and the area had been filtered. The original 3072x1728 resolution camera images were downscaled to 40%, resulting in a resolution of 1228x691.

We also used a segmentation component to process the images as annotated data were needed to train the neural network. To reduce the workload of the learning process, we added instance segmentation to the image sorting software, which selected the birds on the scale.

This segmentation represented the polygons containing the scale and the birds in our experiment. To generate this, we used the publicly available LabelMe software [15], which offered several inclusion shapes, including the polygon with the unique shape we needed. LabelMe software has the additional advantage of supporting the concatenation and conversion of the files generated during annotation into more popular formats.

The multi-stage processing pipeline was built in the following way: The first step was to segment the camera images using an instance segmentation neural network, which located the bird in the image and created a mask. If the network recognized a duck on the scale with at least 80% surety, the next step was the pose estimation and computation of the features used in the weight estimation: the mask area and the ellipse area fitted to the mask. These two main features were passed to the weight estimator, and two further features were calculated: wingtip length and spine length. As a last step, the weight of the duck the camera saw was estimated by the used model (Figure 64).

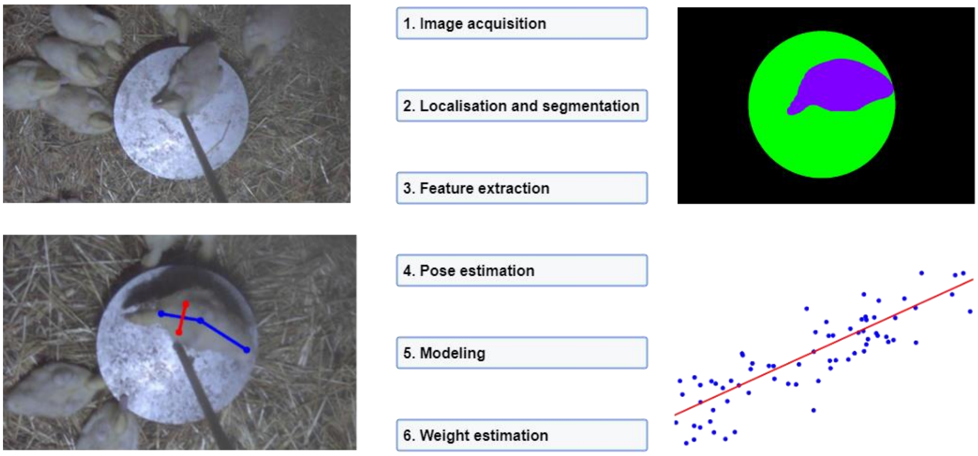


Figure 64. Image processing flowchart

The first step of image processing was instance segmentation, localization, and pixel-level classification of objects in the image. The first component of localization was responsible for object detection. Its accuracy was crucial, as it also played an essential role in data processing as a filter to remove objects with no interest to us. In addition, derived properties were also computed from it (Figure 64).

To accomplish this task, semantic segmentation was not suitable. It could not ensure that animals were visible one by one in isolation, so mask overlapping could easily occur, leading to incorrect measurement. In the training, only single ducks on the scale were used, so no instance segmentation was needed.

Instance segmentation could also be used as object detection since, for the resulting mask, one can easily find a bounding box, and some neural networks require a bounding box to be found before the mask is computed. Of course, the animal mask, which contained additional information, could also be helpful during the weight estimation. We tested yolact [16] and Mask R-CNN [17], but due to learning difficulties, we decided to use the latter, whose implementation is available in the detectron2 library.

In the next step of image processing, the feature set we were looking for was extracted from the images. In our first approach, this was the area occupied by the animal on the scale, performed by an instance segmentation component responsible for detecting the scale and birds.

As a starting point, we investigated the accuracy of inferring the animal's weight based on the area occupied by the bird alone, which was tested by linear regression.

However, more than this solution alone needed to provide sufficient accuracy; further steps had to be taken to increase the weight estimation's reliability by determining the animals' position. So, a pose estimation was performed, and a bounding box was derived from the results. This was a more inaccurate solution in terms of localization. However, the information from the position of the body parts was needed anyway, so it was appropriate to try it as an object detection component. The DeepLabCut [18] (DLC) model was investigated for pose detection. DeepLabCut is an open-source software that helps explicitly estimate animal poses. It even has a separate library (DLC live) that allows the processing of live recordings. DLC was used to determine the neck point, dorsal midpoint, and tail point, and from these, we calculated the bird's backbone. The straight line connecting the wing points must intersect the bird's backbone, and the backbone must be within the previously defined contour. The mask was inaccurate if the bird's backbone fell outside this. Processing the image with DLC in this way increased the accuracy. The occasional wing flapping in waterfowls was a problem because it significantly increased the area occupied by the bird in the camera image used for weight estimation. Therefore, we also calculated a wing length from the left- and right-wing flap points to correct the models. Ducks with outstretched necks covered more significant areas, reducing accuracy, so we calculated bird areas without heads and necks. The mask R-CNN [17] is a convolutional neural network (CNN), a state-of-the-art solution for image segmentation. Using the MASK-RCNN, instance segmentation could be performed, which returned a bounding box and the mask. (In principle, a bounding box can be derived from the DLC feature, but the MRCNN-derived bounding box was more reliable in our experience; the DLC was helpful in cases where no object detection was available in the system.)

MASK-RCNN detects objects in an image and generates a high-quality segmentation mask for each instance. As a starting point, we used MaskRCNN to obtain the duck mask and calculated the headless bird area from it and the area of the ellipse fitted to the mask as a feature for the weight estimator module. The two files (images and weight data) were stitched together along the timestamps as a final step in the learning process. We calculated average weights and areas as a function of the animals in the image. We performed an outlier filtering on them, as all the training samples that were very different from the others could be considered noise (e.g., a corrupted measurement) and deleted.

We did this by expecting an ellipse with a larger area for the larger mask, so we sorted the feature pairs in ascending order by area. The expected result was that the ellipse area was also monotonically increasing. In cases where this was not true, it was worth examining the mask output because a measurement error was suspected.

As a result, four [i3] features could be used in the weight estimation: the area of the bird mask, the fitted ellipse, and the distance between the ridge and wingtip points provided by the DLC.

The above features were given as input parameters for several models. Since the body size and weight of the birds under study increased monotonically during rearing, if the feature - the area of the bird - was more prominent, the weight was also more significant. Therefore, the linear regression model was well suited. In addition, the Random Forest and Multilayer Perceptron (MLP) models were also tested, and the accuracies of the weight estimates obtained with these models were compared. The Multilayer perceptron is an artificial neural network with three or more perceptron layers. These layers are- a single input layer, one or more hidden layers, and a single output layer of perceptrons. Random forest is a Supervised Machine Learning Algorithm used widely in Classification and Regression problems - it builds decision trees on different samples and takes their majority vote for classification and average in case of regression; this reduces the risk of over-learning.

Development of filter modules and detection of repetitive data

The functionality of the automated image processing pipeline has been improved by adding filter modules that could be loaded separately. Several filters were created as stand-alone modules that could be inserted into the processing pipeline, improving accuracy.

The detection of repetitive images was necessary for three reasons:

(i) Bias is introduced into the data set, which gives more chance to the neural network to learn specific patterns;

(ii) impair the generalization ability of the network, i.e., it would be less accurate on data that the algorithm has not yet seen;

(iii) and the required storage capacity could be minimized.

Developing a solution that scaled well on large data sets was important. One possible solution was to look at different hash values, which had several advantages:

- it was not sensitive to different resolutions and aspect ratios;

- changing the contrast did not change the hash value or changed it only slightly, so the hash values of very similar images would be closer to each other;

- the method worked quickly.

To make a difference, we counted more images as repeats and made two changes: We used grayscale images and did not look for a complete match between hash values, allowing for slight variations.

Another approach was to pre-filter the images, i.e., to decide before writing them to disk whether a given frame has had new information. For this, we used a motion detector.

Blurred image detection

During the measurements, we noticed that either motion or dirt (e.g., spider eggs, cobwebs, insects) on the camera or in front of the camera caused some blurred images, which were unsuitable for analysis because they greatly impaired the automatic segmentation.

Therefore, we implemented edge detection on the input images because the fewer edges found, the more likely they are to become blurred. We used the Laplace operator and ran it on three test images overlaid with an artificial blur. The first image remained sharp everywhere, the second image blurred the scale area, and the last image blurred everywhere. The result of the run is shown in Figure 65.

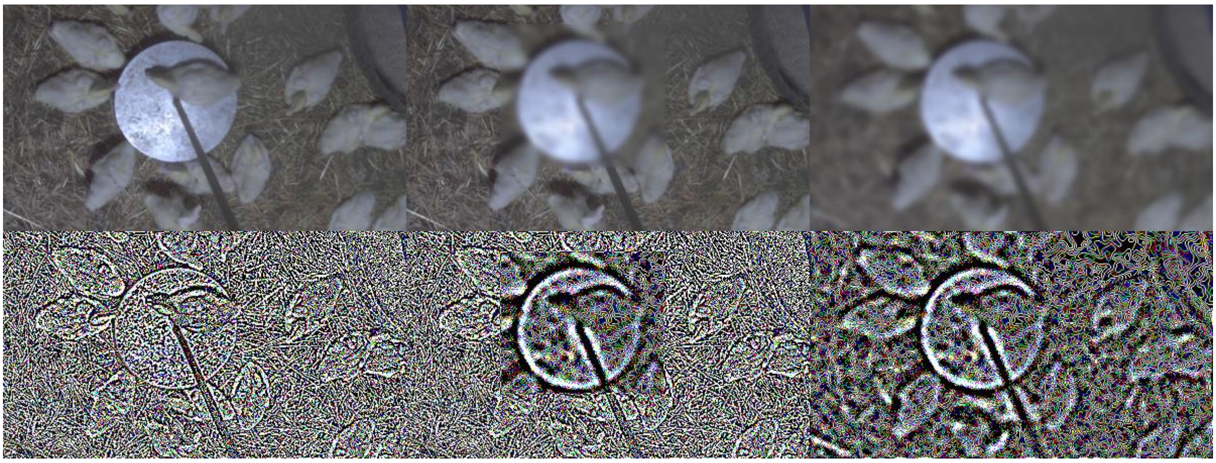


Figure 65. Artificial blur

To obtain a metric for assessing blurriness, we looked at variance. A more considerable variance value was related to the number of detected edges, so the smaller the variance, the more blurred the image was. These variances were as follows for the above pictures:

- sharp image 1141;
- blurred scale 1063;
- blurred image 26.

The results showed that this method found highly blurred images, but the difference between the first two images was not significant. Important information was that the area of the balance was blurred (Figure 65).

Another approach to detecting blur was the use of a Fourier transform. This algorithm detected only high blur, so it did not solve the problem. In practice, this was not a big problem since it could be placed in the pipeline after instance segmentation, which explicitly gave the area of the scale. This way, we could detect blur in the area with a lot of information and extract these images before feature computation.

However, this method can perform segmentation on the right images. It was important that the filtering could dash. Overcoming this disadvantage, the Fourier transform solution worked almost as well. Speed was the deciding factor between the two solutions. The measurements showed that the previously presented variance analysis-based solution evaluated an image for half the time. The advantage of the FFT was that finding a threshold was less complicated; it could even work in other environments without change, but due to the near-fixed location, it did not currently balance half the speed. The noise/error of the balance mask was filtered by calculating the center and area of the polygon and replacing it with a circle of equal area and center. This worked quite well, but unfortunately, it did not apply to non-circular scales.

Results

Table 1 gives the weight estimation error for the models tested. The accuracy of the models was tested using Mean Absolute Error (MAE), which gives the average magnitude of the difference between the fact and the estimated values. The MAE values were calculated for each individual, and all errors calculated for each bird were summarized.

Table 1. Results on weight estimation

	IQR 0 filter					
	Linear Regression		Random Forest		MLP	
	population	individual	population	individual	population	individual
Ellipse	-681	79,1	-1500	76,4		
Area	-575,00	77,60	-900,00	64,00		
Area & Ellipse	-570,00	77,60	-950,00	86,76	153,00	143,60

	IQR 0.2 filter					
	Linear Regression		Random Forest		MLP	
	population	individual	population	individual	population	individual
Ellipse	-1059,88	110,97	-1700	97,77		
Area	-1068,00	110,60	-3350,00	105,00		
Area & Ellipse	-1058,00	110,47	-2150,00	78,88	-2216,00	114,00

The Interquartile range (IQR) is an example of a trimmed estimator, defined as the 50% trimmed central range (the difference between the first and third quartile), which enhances the accuracy of dataset statistics by dropping lower contribution, outlying points (denoted by "IQR 0 filter" In table 1). The data in the "IQR 0.2 filter" columns are shifted by 0.2 times IQR (the lower bound is Q1-IQR*0.2, and the upper bound is Q3+IQR*0.2) to investigate the impact on the models of the lower filtering efficiency expected in a real-life application environment. In this case, the best values were obtained by combining ellipse and occupied area, as expected, i.e., ellipse fitting is an additional technique that increased the accuracy of the estimation in practice.

The estimation of bird weights could be traced back to a classification problem if we do not aim to determine the exact weight of birds but group them into groups of 50 grams. In this case, there was no regression problem, but a classification problem evolved, which was easier to solve. This approach improved the accuracy of the estimation by 10-15 grams on average by removing erroneous values because the 50-gram class breakdown allowed us to perform meaningful filtering.

Figure 66 shows the 50-gram class data (the y-axis shows the area counted from the mask, and the x-axis shows the weight data sent by the scale) and the distribution of the area occupied by the birds on the image.

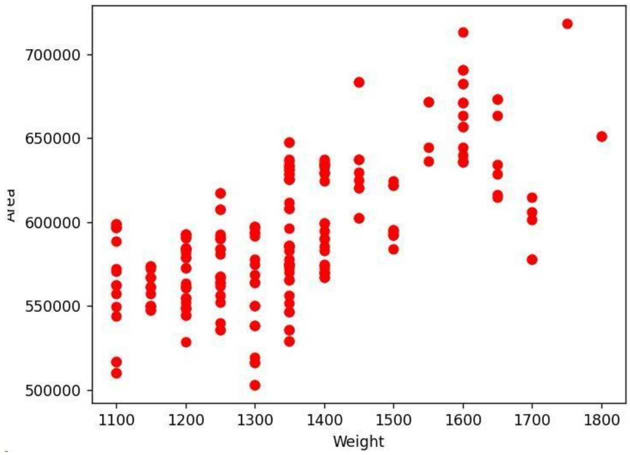


Figure 66. Distribution of bird weights and area occupied

When evaluating the accuracy of the weight estimation, it should be considered that the weight of the birds examined in compiling Table 1 ranged from 1100 to 1800 grams, so the MAE was converted to a percentage value of 3.55-9.54%.

Discussion

PLF techniques in large-scale confined poultry flocks focus on the position of birds within the house and the relationships that can be identified in their behavioral patterns. By analyzing the typical behavioral patterns of birds, the aim is to identify differences that are primarily useful for the early detection of diseases. The estimation of individual weights of birds is addressed in a limited literature. There is yet to be a widespread method in practice to estimate the individual weights of birds by analyzing camera images. In our research under actual farming conditions, we encountered several practical problems in data collection and processing. Data collection devices (the cameras) that can be used in practice should be able to withstand contamination and should not interfere with the work processes involved in poultry housing systems. At the same time, they must provide accurate and reliable results.

In addition to the accuracy of the weight estimation, the complexity and resource requirements of the used model are also necessary. The applied bird weight estimation system should use the lowest possible power hardware components due to cost and the feasibility problems in field (barn) conditions of the intensive air cooling required for the powerful GPU. The regression and random forest models work well on the CPU and do not require a costly and cooling-intensive GPU. The advantage of the regression model was that it did not require training, which saved significant cost and time in case of possible future additions of new species to the weight estimation system. The MLP did not perform well. However, it had the highest resource requirements among all tested models; we did not even perform a single feature-based visualization; we only performed the weight estimation from the Area & Ellipse ensemble, but it still performed worse than the other two. As a further simplification, we used a classification method for the weight estimation (classification problem) by grouping the birds into groups of 50 grams. A significant value was that it happened in a natural farming environment, so we had to adapt it to the daily farming practices. Several articles study single animals or use animals observed in "boxes" under highly controlled conditions [19], [20], [21]. In comparison, the natural farming environment poses several practical problems that significantly affect the system's accuracy.

Using cameras in commercial poultry meat production, we found several interesting and valuable research results, which focus on welfare status [22], [23], identifying biomechanical variables of broiler chickens during feeding [24], lameness of broilers [25], and counting the broiler flock [26]. However, there is scant literature on the study of waterfowl.

Research using 3D sensors, mainly in cattle farming [27], [28], can be found in the literature, but this practical application is not a cost-effective solution in poultry farming. A 3D camera-based system that could determine the weight of several broilers at once or predict the weight of an individual broiler was developed [29]. They stated that machine vision and SVR could promisingly estimate the weight of life broiler chickens. In [30], they used available Kinect cameras, but the method's applicability has yet to be tested with the newer versions of Kinect. Our solution did not use 3D sensors and relied on cheap off the self-components.

Conclusions

Precision farming technologies are also becoming increasingly important in the poultry sector. Their primary aim is to maximize the profitability of livestock production. This requires digital data on individual animals and the production data collected. Whatever technology is used in large-scale livestock production, we must consider the conditions under which animals are kept. These will significantly impact the range of tools that can be used, how data is collected and processed, and how the results are fed back to the farmer. In this paper, we present a methodology to determine the individual weight of ducks by camera imaging. The proposed solution is suitable for replacing conventional poultry weighing scales, providing similar accuracy at a lower cost. This article summarizes the practical experience of our research in a natural farming environment.

We were looking for the simplest practical solution regarding required IT devices and data analysis methods/computing requirements. Regular cleaning of the camera lenses is essential for the solution's applicability. Advantageously, the artificial illumination is roughly constant for waterfowl, predominantly kept in closed conditions. The effectiveness of the weight estimation algorithm is highly dependent on the accuracy of the data used for training, and it is. Therefore, the image and weight data used in constructing the training data set should be appropriately filtered and validated.

7.4.2 The potential of RFID technology for tracking Mangalica pigs in the extensive farming system – research from Hungary

Introduction

In the context of livestock farming, sustainability implies models that provide the farmer with a reasonable and stable income without significant negative side effects on the environment and that is acceptable to society by ensuring both animal and farmer welfare [1]. The pig industry is characterized overwhelmingly by closed, intensive, and large-scale housing systems worldwide that can supply large meat processing facilities [2], [3]. This highly human-controlled farming system allows the collection of a large amount of data describing the farming environment and monitoring the pigs individually [4] by sensor, camera, and microphone systems that belong to IoT (Internet of Things) devices placed in the housing area and on the animals [5]. The IoT is a network of interconnected devices that communicate, sense, and interact with internal and external environments via embedded technology [6]. They are completed by the data analysis using machine learning and traditional statistical methods, which aim to achieve environmental and economic sustainability. Bercksman [7] summarised the critical features of Precision Livestock Farming (PLF) technologies, which focus on animal health and welfare. One of the less studied areas of PLF applications in pig farming is pasture-based pig management systems, where pigs are on pasture all year round or at specific periods of the year, both during breeding and fattening periods [8]. In this housing system, pigs have more opportunities to express their species-specific behaviors, improving animal welfare and allowing for sustainable pig production. This ecological housing fits into the five freedoms framework defined by the Farm Animal Welfare Council in 1979 [9] and also falls in line with the U.N. Sustainability Development Goals [10]. In order to achieve sustainable management in these extensive pig production systems, the housing environment and continuous monitoring of the animals are important factors. However, heat stress can also be a problem in free-range conditions, especially in the summer.

In contrast to confined housing, outdoor pigs are kept in larger territory, close to nature, where they have more opportunities to protect themselves from weather extremes. One behavioral solution is wallowing, when pigs cover their bodies with mud or lie in the water to reduce their exposure to excessive heat and sunlight [11]. Under all conditions, the pig farmer has an outstanding need to get up-to-date information on the health of the pigs and the risk of heat stress. Pigs, unlike ruminants, use pasture primarily as a living space rather than a feed source. Their species-specific behaviors (rooting and wallowing) mean they heavily use the grazing land. At the same time, the premium meat quality of the free-range pig breeds (Mangalica, Iberian pig) is of gastronomic value [12]. However, due to the species-specificity of pigs and the extensive keeping system, digital and IoT data collection devices are limited. Most of the digital tools provided for ruminants on pasture are not available for the pigs, although the demand for this technology from organic pig farmers is growing. One possible solution is Radio Frequency Identification (RFID) technology, which has been used for many years in confined pig farming (mainly to monitor eating and drinking habits and to sort pigs according to specific criteria [13], [14] RFID technology is also used to monitor ruminants on pasture [15], and can therefore be adapted to some extent to outdoor pig farming.

The RFID technology is a powerful tool for tracking the location of objects in real-time. Identification is achieved using radio frequency waves, and the information is stored on a data carrier (electronic memory chip). RFID exists in a full range of systems and components which operate on the same technological principle [16]. The reader transmits information in two directions between the medium and the receiver [17]. The tag comprises a small chip, an antenna, and a memory. Some types of RFID tags allow repeated recording of data. Tags are attached to observed objects, either on the object's surface or outside the object, for identification purposes [18]. This technology belongs to the group of automatic identification technologies. RFID technology works by contactless transmission of information from tags. The advantages of RFID technology include providing information over long distances (several meters) and ease and affordability of the way to identify, track, and monitor livestock, thus improving the traceability of animals along the supply chain and reading rates of more than 100 tags per second. The adoption of RFID technology in practical farm management has allowed the development of asset-tracking software, where daily records on individuals (e.g., medical treatments, growth values, pedigree, reproductive performance) are automatically collected and stored [19]. RFID readers are also capable of collecting climatic data. In large-scale pig farming, passive RFID tags are mainly placed in ear tags [20]. The signals transmitted by the antenna are collected by the readers and transmitted using a 2G/3G/4G/5G network to a data storage server, where the data is analyzed. Information of interest to pig farmers can be displayed graphically [21]. RFID technology has several limitations in outdoor circumstances. One of the disadvantages is that the presence of animals outside the operational range of the readers cannot be recorded, extreme weather conditions affect the continuous transmission of data, tags placed in pigs' ears can be lost or torn out of each other's ears during social interactions (playing or fighting). These factors and events can make the output data of an RFID system unreliable.

In our seven-month-long study, we tested the applicability of passive RFID technology in free-range pig housing conditions, collecting data on sows' appearance and weather parameters at the wallow.

Material and methods

The investigation was conducted on sows housed in an outdoor, pasture-based system on a private farm in Hungary between June 2020 and January 2021. Data collection devices were installed at the pilot site in June 2020, and data collection started in July 2020 [22]. The free-range pig farm is primarily used for the reproduction of gilts and fattening pigs. It complies with the current E.U. [98/58/E.C.] legislation on the keeping of farm animals, animal welfare, and environmental protection, which is continuously monitored by the authorities (Hungarian State Treasury Agriculture and Rural Development Agency, National Food Chain Safety Office and Hungarian National Association of Mangalica Breeders).

Animals

The Mangalica is a native Hungarian pig race that was the most typical breed locally until the middle of the last century. It is a fat-type, curly-haired swine with relatively low reproductive capacity but strong motherliness and very good adaptability to extensive housing conditions [23]. During the research, RFID data was collected on twenty sows with an average age of 5 years (Figure 67).

Figure 67 Experimental Mangalica sows' group

The sows farrow once a year, and the boars are with the sows all year round, except in September and October (to avoid winter farrowing). Therefore, the timing of first and subsequent insemination is challenging to determine. When the expected day of farrowing approaches, the farmer shepherds the sows to the enclosed farrowing building next to the pasture, where they farrow in individual boxes with concrete floors covered with straw and return to the pasture with their piglets two to three weeks after farrowing. The piglets are separated from the sows when they are 8-10 weeks old. Then, the sows are replaced in the paddock for regular keeping; the piglets go to the pasture section for fattening. At the beginning of our study, the sows were just after farrowing and were kept on the pasture for the entire experiment period. Data acquisition occurred from July 2020 to January 2021, so data on weather and sows' presence on the pasture were collected for three seasons (summer, autumn, and winter).

Housing

In one group, the sows were kept on approximately four hectares of fenced land, about 270 x 150 meters, all year round. The layout of the land is shown in Figure 68.

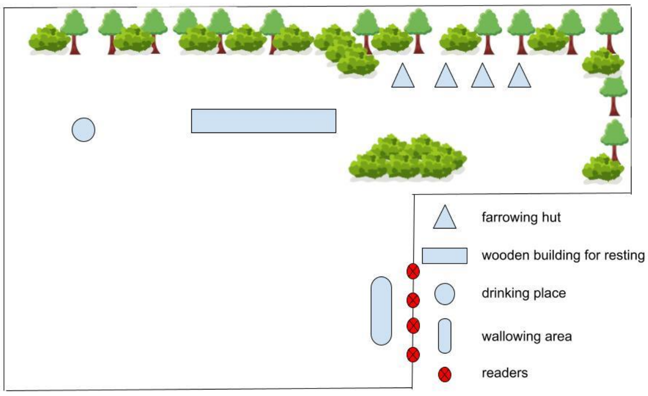


Figure 68 The layout of the pasture

The site had a 2.5 x 10 m wooden resting building and a drinking trough filled with a water tank. The four farrowing huts' sizes were 2.5 x 2.5 m each and were located in the upper part of the pasture, close to the tree line. About 20% of the area was covered with shrubs and lower trees. In the central part of the area, there was a Jimsonweed (Datura stramonium L.) plant, which pigs used as a resting and walking place. The rest of the area was not covered with vegetation, leaving only the greenery that was not eaten or destroyed by the sows. Near the fence, from inside the area, a wallow was created by the pigs themselves, whose shape and depth varied according to their activity. Its largest size was in the autumn, with a 5 m length, 2.5 m width, and 45 cm depth. The farmer occasionally filled water into this pit, and rainwater flowed there too. When the wallow was full of water, the sows could sink entirely to the muddy water if they lay laterally, and a maximum of five sows could lay into the water at once. The wallow was far from the watering trough, the feeding area, and the resting place. The pasture had a slope of 5%. The fenced pasture had a free place for pigs to use; the sows could use the wallowing area at any time. In the early morning, animals were fed once a day with no commercial feed ingredients (fruits, vegetables, corn), and they had ad libitum access to drinking water. Farm workers entered the pasture only when it was necessary, so there was no human presence most of the time.

Data acquisition about the presence of sows and weather parameters

Environmental recording

The Hungarian Meteorological Service (OMSZ) provided relevant data from the nearest meteorological stations: solar radiation, precipitation, and wind data daily and cloud coverage data monthly. The precipitation data were obtained from Nagymaros (47°46'18''N, 18°56'44''E, altitude above mean sea level: 106.0 m; distance from the experimental site: 3.7 km) the solar radiation data were collected in Püspökszilágy (47°29'52.8 "N, 19°19'36” E, altitude above mean sea level: 239.0 m; distance from the experimental site: 31 km). Wind speed values were obtained from Nagy-Hideg Hill (47°56'11''N, 18°55'18"E, altitude above mean sea level: 854.8 m; distance from the experimental site: 22 km). Cloud cover data were obtained from the Pestszentlőrinc meteorological station (47°29'52.8 "N 19°02'23.7 "E, altitude above mean sea level: 125 m; distance from the experimental site: 35 km). Four RFID readers fixed to the fences near the wallow recorded hourly temperature, humidity, and air pressure values, resulting in 5040 values for each of the three parameters (15120 records in total).

In order to describe environmental effects on animals, several indices, including the Temperature-Humidity Index (THI), have been developed and are widely applied [24]. We calculated the THI for each day of the study to estimate the level of potential heat stress. The aggregation of ambient temperature and relative humidity causes it:

$$\text{THI} = [0.8 \times \text{ambient temperature}] + [(\%\ \text{relative humidity} \div 100) \times (\text{ambient temperature} - 14.4)] + 46.4,$$

where ambient temperature is in degrees Celsius and relative humidity is in % [25].

For descriptive purposes, the mean, maximum, minimum, and standard deviation of temperature, relative humidity (R.H.), and THI were calculated for all study days for all three seasons of the behavioral observation period (Table 2).

		summer (from 1st of July to 31th of August)	**autumn** (from 1st of September to 30th of November)	**winter** (from 1st of December to 31th of January)	**whole experimental period**
experimental days (n)		58	90	60	
temperature	mean ambient ± SD	22.39 ± 7.26	10.74 ± 7.96	1.55 ± 3.6	11.58 ± 10.37
(C°)	maximum ambient	40.04	35.44	15.07	40.04
	minimum ambient	5.88	-8.73	-15.09	-15.09
humidity	mean ambient ± SD	45.81 ± 24.29	71.38 ± 23.66	81.38 ± 13.99	66.78 ± 25.74
(% RH)	maximum ambient	99.71	99.99	99.98	99.99
	minimum ambient	0.51	1.25	30.92	0.51
THI	mean ambient ± SD	71.7 ± 8.61	53.78 ± 12.75	38.24 ± 6.66	54.69 ± 16.17
	maximum ambient	8.4	84.9	62.6	88.37
	minimum ambient	46.4	19.9	10.5	10.54

Table 2 THI with calculation based on ambient temperature and relative humidity at the wallow

Sows' presence in the wallowing area

Their passive RFID tags detected sows' presence at the wallowing site at the pasture fence (Figure 2) 24 hours/day. They recorded the arrival and leave of 20 sows, each with one RFID tag (diameter 30 mm, circle shape, frequency 850 MHz UHF). The detection distance of the RFID readers with antennas (type: SR450, square-shaped, 45 cm x 45 cm, Inpinj R2000 chip, model: CF-MU804, HDMI, USB 2.0, Ethernet, optional WLAN or 3G, Ethernet network connection, 10Mbit, Four 50 Ω/RPTNC connectors supporting four monostatic antennas, Read Rate: up to 750 tags/second using high-performance settings) was 4.5 meters. The detection distance was tested in the field before the study's launch. The readers were set up so the entire wallowing area could be observed. The size of the wallow varied during the experimental period, depending on the weather conditions. In total, the RFID readers recorded the presence of sows 9070 times during the study.

At the start of the study, after testing the RFID devices, plastic ear tags containing the passive RFID tags were attached to both ears of each 20 Mangalica sows. Data on sows' presence at the wallow were collected for the entire study period, from 1 July 2020 to 26 January 2021. We added the weather parameters collected directly by the RFID readers to the data series on sows' presence,

The data collection period was divided into 10-minute time intervals. A sow's appearance was considered for the analysis if it spent at least this interval at the wallowing site. In addition, we cleaned the dataset of sows that appeared less than 200 times at the experimental site. Half the initial number of sows, ten sows, participated during the entire study. Thus, our dataset prepared for analysis recorded ten sows' appearances representing at least 200 occasions over the entire study period.

Statistical analysis

As the presence of a pig at a certain time was a binary variable (yes/no indicator), logistic regression was used for its multivariable modeling, with temperature, air pressure, humidity, and hour-of-the-day as the covariates. All covariates were continuous variables, and they were expanded with splines to allow for a potential non-linear effect (hour-of-the-day was expanded with cyclic spline) [26]. Note that this model is static and is not corrected for autocorrelation.

To present the regression results, we selected one day from each of the three seasons (5 August, 22 October, and 10 January) when the seasonal weather could be considered typical. We fixed the values of independent variables except the hour of the day. All independent variables were fixed at the average values for the study period on the selected three days. We also presented the likewise partial effect of a few critical meteorological variables.

Calculations were carried out under the R statistical environment version 4.2.0 [27], using the mgcv package version 1.8-40 [28] and gratia version 0.7.3 [29]. The full analysis script is available at https://github.com/tamas-ferenci/MangalicaStatistics.

Results

Environmental conditions

The environmental conditions during the study are summarized in Table 1. The highest THI value (71.7 ± 8.6) was measured in the summer period, while in the autumn, the THI's average value was 53.8 ± 12.7. However, the most considerable variation in THI values was observed in this season because the weather in early autumn was similar to the conditions in the summer season. At the same time, the temperature in November was much lower. During the winter period, we had the lowest THI values (38.2 ± 6.66) due to the low temperatures. The THI for the whole seven-month study was 54.7 ± 16.2.

Looking at the temperature data, the average was 22.4 ± 7.3°C during the summer, but there were sweltering days with temperatures reaching 40°C. However, on cooler summer days, the temperature was 6 °C in the morning hours. Four hot days were recorded during this season, with maximum temperatures between 37 and 40 °C (in mid-July and August). These days, the humidity was extremely low (between 0.56 and 12.6%). The lowest temperatures were recorded at night and dawn (between 3 and 6 am). The average temperature during the autumn was 10.7 ± 7.9 °C. The measured daytime temperatures in early autumn were 34-35 °C on five days. This was the so-called "Indian summer," which is typical of the Hungarian climate in early/mid-September. These values were coupled with higher humidity values (between 8.5 and 11.8%) compared to the relative humidity of hot summer days. However, temperatures can drop drastically in a few days during this season. From mid-October to the end of November, there were sixteen days when the daily average temperature fell below 0 °C. In October, this was observed during the night, at dawn, and in November, even in the early evening. The average temperature during winter was 1.55 ± 3.6°C, which was in line with the season. The daytime maximum and minimum temperatures were the lowest during this season (15.1 and -15.1 °C, respectively). The lowest temperatures were recorded in January (between -15 and -12 °C) at night and dawn.

The average humidity over the study period was 66.8 ± 25.7%. During the summer, the average humidity was 45.8 ± 24.3%, with the lowest recorded value of 0.5%. Humidity values between 0.5 and 5% belonged to temperatures above 35°C, with little cloud cover on these days (5, 10, 15, and 30 July and 1, 12, and 13 August, the values were 0.0, 0.5, 2.2, 2.2, 0.0, 1.4, 1.8%, respectively). However, on three days (11, 15, and 17 August), the air humidity reached 95-99.7%. The cloud cover value on these days was significant (4.8, 4.5, 4.5), and the amount of precipitation was also considerable (2.4, 19.2, 16.9 mm). A summer thunderstorm caused the higher air humidity. During the autumn, the decreasing daily mean temperature was accompanied by increasing humidity values (71.4% ± 23.66), with a maximum value of 99.9% and a minimum value of 1.25%, similar to the summer period. On thirty-four days, the humidity values ranged between 95 and 99.9%, and the temperature was between -1.7°C and 12.1°C. There was one day with higher precipitation (13 October, 47.6 mm) when the daily temperature dropped from 6.1°C in the morning to 5.1°C at the end of the day. The warm weather of the autumn period gradually changed to early winter conditions, with humidity levels ranging from 95 to 99.9% on one-third of the autumn period (thirty-four days out of ninety). The highest average humidity values (81.4 ± 13.9%) were recorded in winter, with a minimum value not falling below 30.9%. On twenty-five of the sixty days studied, the measured humidity was between 95 and 99.9%, with a gradual temperature decrease from 7.2°C to -4.4°C. On the coldest days, the temperature at night did not rise above -10°C (on two days, 17 and 18 January), with humidity ranging between 68.6% and 76.6%.

Patterns of the probability of sows' appearance by the time of day and weather parameters

During the study period, we covered three seasons. Given that each season brought different weather conditions and day lengths, we examined the probability of sows' appearance at wallowing sites on a selected day in each season, controlling for meteorological parameters (Figure 69). The graphs for each sow in Figure 69 show the effect of a given weather parameter on the appearance of the sows. The grey bar above and below the black line shows the confidence interval of the weather parameter's effect on the sow's appearance, within which the independent variable had an effect. We also examined the significance of the covariates in the logistic regression; Table 3 shows these for a few selected explanatory variables. The results for each independent variable were as follows.

Analyzing the impact of hour-of-the-day, it can be concluded that the wallowing activity of each sow differed and that two or three daily activity peaks were observed. Two activity peaks were seen for sows 7, 11, 14, 15, 17, 18, 19, and 20; three were characteristic for sows 5 and 12. The first activity peak was observed in the early morning for all sows. It was related to the feed distribution time, around 6:30 am daily. Afterward, the sows went to the wallow, where they spent approximately one hour. The second peak of their activity was observed in the afternoon (between 2 pm and 5 pm), except for sow no. 20, which returned to the observation area later than the others, rather in the evening. The effect of hour-of-day on the likelihood of appearing at the wallow showed a significant correlation for all sows, with no statistically significant values found for all except no. 20 ($p=0.115$).

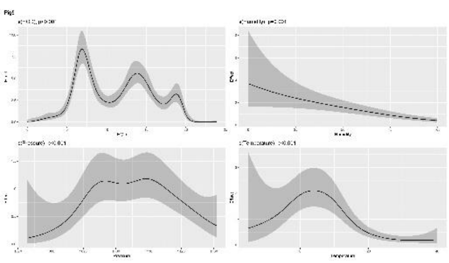
The first of the weather parameters to examine was the effect of air humidity on the probability of the appearance of individual sows at the wallowing site. The data clearly show that the probability of sows' appearance at the study site decreases as air humidity increases. Analyzing the behavior individually, it can be concluded that there was a significant correlation between the probability of sows' presence and the air humidity. A vital significance (p-value between 0.001 and 0.005) was observed for the probability of appearance of sows 5, 7, 15, and 18, with 5% significance values for the probability of appearance of sows 11, 12, 17, and 19. No significant effect was found between sows 14 and 20, nor was the air humidity value.

The following weather parameter was the air pressure, the effect of which was analyzed on the probability of sows' appearance at the wallowing site. No significant effect was found between the probability of emergence of 12 sows and the air pressure values (p=0.96). The effect for 7 and 20 sows and the air pressure was significant at a 5% significance level, while a p-value of 1% or below was found for seven sows. Referring to Figure 3, we found that air pressure values had less effect on the probability of sows appearing at the wallow.

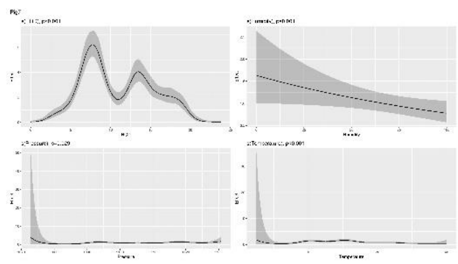
The third weather parameter studied was temperature. We found a significant correlation between the effect of the temperature and the probability of sows being at the wallow for all sows ($p<0.001$). Analyzing the graphs in Figure 3, we concluded that the probability of sows' appearance at the wallowing site is reduced with increasing temperature. Although the relationship between sows' appearance and temperature values was significant, the temperature had only minimal effect (for sows 7, 11, 12, and 20) on the probability of sows' appearance at the wallowing site. Uniquely, the probability of the appearance of sow 14 increased with rising temperature. The temperature influenced the occurrence of five sows; they (sows 5, 15, 17, 18, and 19) visited the wallowing site most frequently between 0 and 20 °C.

Examining the effects of the four independent variables, we found that the effect of hour-of-day was observed for all sows studied. A trend in the effect of humidity on sow emergence can be observed for most sows: as humidity increases, sows' appearance at the wallowing site decreases. No such trend in sow emergence is observed for air pressure or temperature.

Sow5

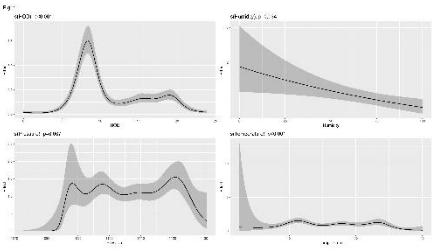


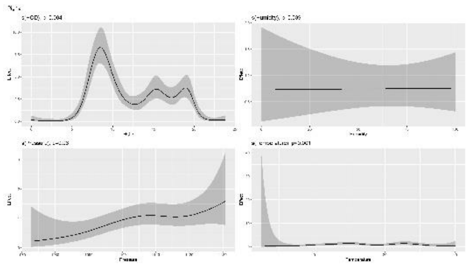
Sow7



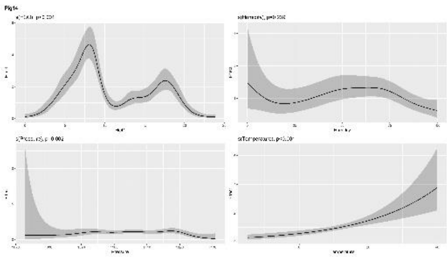
Sow11

Sow12

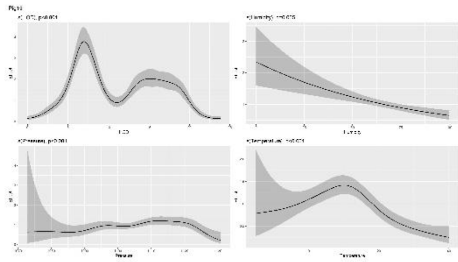




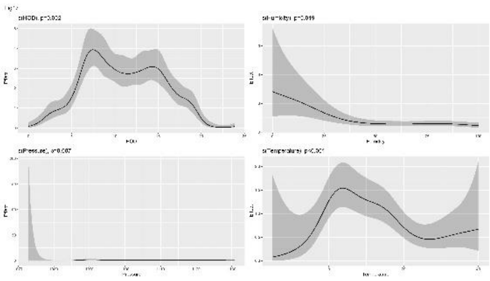
Sow14



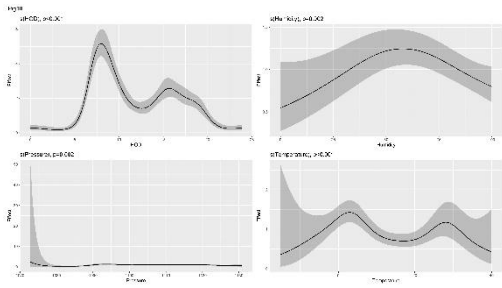
Sow15



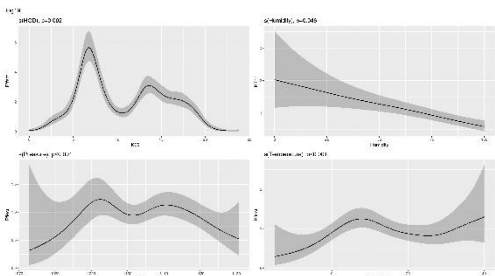
Sow17



Sow18



Sow19



Sow20

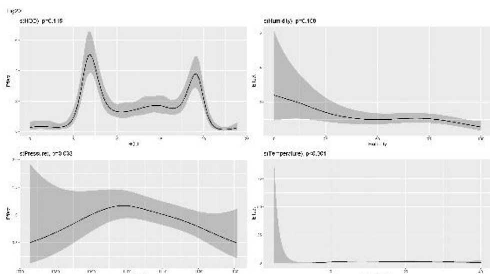


Figure 69 Evolution of the daily activity curves for each sow at the wallow, effect of time-of-the-day, humidity, air pressure and temperature

Sow nr.	HOD	Humidity	Air pressure	Temperature
Pig5	p<0.001	p=0.001	p<0.001	p<0.001
Pig7	p<0.001	p=0.001	p=0.029	p<0.001
Pig11	p<0.001	p=0.034	p=0.007	p<0.001
Pig12	p=0.004	p=0.039	p=0.96	p<0.001
Pig 14	p<0.001	p=0.239	p=0.002	p<0.001
Pig15	p<0.001	p=0.005	p<0.001	p<0.001
Pig17	p=0.002	p=0.019	p=0.007	p<0.001
Pig18	p<0.001	p=0.002	p=0.002	p<0.001
Pig19	p=0.002	p=0.046	p<0.001	p<0.001
Pig20	p=0.115	p=0.108	p=0.033	p<0.001

Table 3 The strength of correlation between sow appearance and each of the independent variables (hour-of-day, humidity, air pressure and temperature)

Discussion

By analyzing the data, we determined whether the correlations and findings obtained would provide helpful information to the farmer regarding the group of free-range breeding sows. RFID technology allows the recording of animal appearance in the area monitored by the RFID reader; in this study, we did not assign any other data related to breeding and housing to the concerned sow. One of the objectives of this study was to confirm or refute the applicability of RFID technology in outdoor pig housing conditions and, by assigning weather parameters to the sows' appearance dataset, to obtain information on whether a correlation can be established between the appearance of breeding sows at the wallowing site and the three studied weather parameters.

Our research confirmed that the risk of heat stress increases on hot days in summer and early autumn. It can be assessed by calculating the Temperature-Humidity Index, for which the required temperature and humidity values were recorded directly at the experimental site. We did not include the THI as an independent variable in our regression model since this indicator is calculated from the temperature and humidity values, and these two weather parameters were already included in the model as independent variables. A similar study was conducted by [30], who investigated the behavior of pregnant and lactating sows kept outdoors during the summer. Their experiment demonstrated that sows preferred to use the wallow in hot summer weather conditions (especially at high THI values). In our research for several days during the summer months, THI values ranged from 83.2 to 88.4, indicating a risk of heat stress. It could be hazardous for breeding animals because prolonged heat stress negatively affects their reproductive performance, which reduces the profitability of the pig farms [31]. Under free-range conditions, pigs can reduce their body temperature by seeking shade or wallowing. Wallowing is a species-specific behavior of pigs and plays a vital role in thermoregulation and defense against parasites and their social relationships. However, this behavior is not essential for subsistence but has a positive impact on the welfare of the pigs. Considering this aspect, we were curious to know at which time of day the sows visit this place. It was one of the reasons for placing the RFID readers around the wallowing area. The other parameters characterizing the current weather (solar radiation, cloud cover, precipitation, wind speed) were obtained from the meteorological stations nearest the research site. Given that the pig farm is located in the mountains and the meteorological stations are several kilometers away, these values could only be considered in our conclusions with some bias. Although they were included in the dataset model, this study did not investigate their impact.

In the data set analysis, we found that the most robust statistical relationship was between the presence of sows and the temperature ($p<0.001$). However, the temperature did not affect the probability of being at the wallowing site for a subgroup of the sows. However, when examining the effect of hour-of-day, we found that the two peaks of activities reported in the literature were observed with these animals [32], [33]. The first activity peak was likely due to the morning feed distribution, after which most sows visited the wallowing area. By analyzing the relationship of humidity values with sows' attendance, it can be concluded that as humidity values increase, the probability of individual sows' attendance at the wallow decreases. Examining the air pressure values, we found that although there was a significant relationship between the sow emergence and the air pressure values, this weather parameter was less influential.

Our research confirmed the literature sources on the limitations of RFID technology [19]. These include the limitation of the reading distance, which meant that we could only monitor sows for as long as the reader could detect the tag in their ears. Second, attention must be paid to replacing ear tags lost by animals. In our research, we placed ear tags in both ears of the sows to ensure that the animal could be identified when one was lost. In addition, stormy weather caused problems with data transmission twice, but the reader stored the data for three days, so there was no data loss.

Moreover, the RFID readers require electricity, which is not evident under free-range conditions. In our field research, this location was the wallowing area, where we supplied the readers with electricity most cost-effectively. One of the advantages of RFID technology is that it has worked reliably, with the readers withstanding the varying weather conditions for months [34]. Neither the ear tags nor the other elements of the RFID technology devices disturbed the sows in any noticeable way. Inserting the ear tags was a relatively straightforward process for the farmer, as it is required to identify breeding pigs uniquely. The RFID tags were the same size and material as conventional ear tags. Another possibility to be highlighted is to add to the RFID tag's identification number additional husbandry and breeding data that the farmer considers essential, thus creating an even more complex dataset based on the unique identification of the animals and offering a helpful precision livestock technology.

Unlike cattle, the RFID tag used in pigs is primarily passive [35]. The reason for this is that the size and weight of the ear tag are determining factors in pigs. The pigs' behavior is characterized by their daily social interactions, which can be either fighting or playing, where pigs can bite into each other's ears with their teeth and tear out the ear tag. In free-range conditions, pigs are less frequently checked by their farmers, so individual identification, completed with digital data collection, is essential for monitoring the health of the pig herd.

Our study has several limitations that should be considered when conducting another similar study in the future. First, installing RFID technology in other parts of the pasture area used by pigs (resting area, watering area, feeding area) is worthwhile. On the other hand, installing cameras can provide more accurate information on the monitored sites. In this case, the cost of installing cameras and the infrastructure needed should be considered. Third, it is also worthwhile to associate breeding and husbandry data with the RFID identification of the sows.

In our future work, we aim to extend the study to monitor more animals in various settings. To this end, a combination of RFID and CCD-based monitoring technologies can be developed [36].

Conclusion

Providing animal welfare and increasing consumer expectations are critical elements of sustainable pig farming. Despite its limitations, RFID technology represents a step forward in using precision methods compared to current practices under sustainable outdoor pig farming circumstances.

With this research, we have achieved our goal of demonstrating that RFID technology can reliably collect individual data on free-range pigs, given its limitations, and that the data can be transmitted to the data analysis server via a 4G network without any problems. RFID technology should be considered one of the digital data collection methods for sustainable pig farming.

7. References

Chapter 4 and 5:

https://mek.oszk.hu/01200/01216/01216.htm#9; https://dtk.tankonyvtar.hu/bitstream/handle/123456789/8751/0010_1A_Book_09_Gabonanovenyek_termesztese.pdf?sequence=2&isAllowed=y

https://mersz.hu/horvath-komarek-a-vilag-mezogazdasaganak-fejlodese/

https://magyarszoja.hu/tudastar/szoja/; https://szojaextruder.hu/a-szojatermesztesben-hato-agrotechnikai-elemek/; https://gyorgymag.hu/csillagfurt; http://www.pkkft.axelero.net/nvt/cikkek/bojsza_lucerna.htm

https://queenbeefarms.ca/pages/the-history-of-beekeeping

M. Wade and J. Hoelle. A review of edible insect industrialization: scales of production and implications for sustainability Environ. Res. Lett. 15 (2020) 123013, DOI 10.1088/1748-9326/aba1c1

Chapter 6:

Kersting, K. (2021) Machine Learning and Artificial Intelligence: Two Fellow Travelers on the Quest for Intelligent Behavior in Machines.
https://doi.org/10.3389/fdata.2018.00006

https://www.v7labs.com/blog/neural-network-architectures-guide

https://towardsdatascience.com

simplilearn.com

mltut.com

techtarget.com

https://www.elderresearch.com/blog/the-42-vs-of-big-data-and-data-science

https://aws.amazon.com/what-is/data-preparation/

Chapter 7:

precision poultry research:

[1] Li, G.; Ji, B.; Li, B.; Shi, Z.; Zhao, Y.; Dou, Y.; Brocato, J. Assessment of layer pullet drinking behaviors under selectable light colors using convolutional neural network. Comput. Electron. Agric. 2020, 172, 105333.

[2] Wolfert, S. and Isakhanyan, G. Sustainable agriculture by the Internet of Things – A practitioner's approach to monitor sustainability progress. Computers and Electronics in Agriculture. 2022. doi: 10.1016/j.compag.2022.107226

[3] Kashiha, M., Pluk, A.; Bahr, C.; Vranken, E.; Berckmans, D. Development of an early warning system for a broiler house using computer vision, Biosystems Engineering, Volume 116, Issue 1, 2013, Pages 36-45, ISSN 1537-5110, https://doi.org/10.1016/j.biosystemseng.2013.06.004

[4] Alexy, M. and Haidegger, T. Precision Solutions in Livestock Farming –feasibility and applicability of digital data collec-tion. In: Szakál, Anikó (editor) IEEE 10th Jubilee International Conference on Computational Cybernetics and Cyber-Medical Systems ICCC 2022 Budapest, Magyarország : IEEE Hungary Section. 401 p. pp. 233-238. , 6 p.

[5] Li, G.; Huang, Y.; Chen, Z.; Chesser, G.D., Jr.; Purswell, J.L.; Linhoss, J.; Zhao, Y. Practices and Applications of Convo-lutional Neural Network-Based Computer Vision Systems in Animal Farming: A Review. Sensors 2021. 21,1492. https://doi.org/10.3390/s21041492

[6] Okinda, C.; Nyalala, I.; Korohou, T.; Okinda, C.; Wang, J.; Achieng, T.; Wamalwa, P.; Mang, T.; Shen, M. A review on computer vision systems in monitoring of poultry: A welfare perspective. Artif. Intell. Agric. 2020, 4, 184–208.

[7] https://www.eurofoiegras.com/, 2022

[8] Lourens, A.; Heerkens, J.L.T.; Star, L. Sensors and techniques to monitor and improve welfare and performance in poultry chains. In: Precision technology and sensor applications for livestock farming and companion animals. Edited by van Erp-van der Kooij, E. Wageningen Academic Publishers. 2021. ISBN 978-90-8686-364-8 pp. 131-165

[9] Mollah, Md.B.R.; Hasan, Md.A.; Salam, Md. A.; Ali, Md. A. Digital image analysis to estimate the live weight of broiler, Computers and Electronics in Agriculture. 2010. Volume 72/1, pp. 48-52, https://doi.org/10.1016/j.compag.2010.02.002

[10] Newberry, R.C., Hunt, J.R., Gardiner, E.E. Behaviour of roaster chickens towards an automatic weighing perch. British Poultry Science 1985. 26, 229–237.

[11] Doyle, I., Leeson, S. Automatic weighing of poultry reared on a litter floor. Canadian Journal of Animal Science 1989. 69, 1075–1081.

[12] Feighner, S.D., Godowsky, E.F., Miller, B.M. Portable microcomputer based weighing systems: applications in poultry science. Poultry Science 1986. 65, 868–873

[13] Fontana, I., Tullo, E., Butterworth, A., Guarino, M. Broiler vocalisation analysis used to predict growth. In: Spink, A.J., Loijens, L.W.S., Woloszynowska-Fraser, M., Noldus, L.P.J.J. (Eds.), Proceedings of Measuring Behavior 2014. Wageningen, The Netherlands

[14] Fontana, I.; Tullo E.; Butterworth, A.; Guarino, M. An innovative approach to predict the growth in intensive poultry

farming. Computers and Electronics in Agriculture. 2015. 119. pp. 178-183. doi: 10.1016/j.compag.2015.10.001

[15] Russell, B.C., Torralba, A., Murphy, K.P. et al. LabelMe: A Database and Web-Based Tool for Image Annotation. Int J Comput Vis 77, 157–173. 2008. https://doi.org/10.1007/s11263-007-0090-8

[16] D. Bolya, C. Zhou, F. Xiao, és Y. J. Lee, „YOLACT: Real-time Instance Segmentation”, arXiv, arXiv:1904.02689. 2019. doi: 10.48550/arXiv.1904.02689.

[17] K. He, G. Gkioxari, P. Dollár and R. Girshick, "Mask R-CNN," 2017 IEEE International Conference on Computer Vision (ICCV). 2017. pp. 2980-2988, doi: 10.1109/ICCV.2017.322.

[18] A. Mathis et al., „Markerless tracking of user-defined features with deep learning", arXiv, arXiv:1804.03142. 2018. doi: 10.48550/arXiv.1804.03142.

[19] Guo, Y.; Chai, L.; Aggrey, S.E.; Oladeinde, A.; Johnson, J.; Zock, G. A Machine Vision-Based Method for Monitoring Broiler Chicken Floor Distribution. Sensors. 2020. 20, 3179. https://doi.org/10.3390/s20113179, Fang, C.; Zheng, H.; Yang, J.; Deng, H.;

[20] Zhang, T. Study on Poultry Pose Estimation Based on Multi-Parts Detection. Animals. 2022. 12/1322. https://doi.org/10.3390/ani12101322)

[21] Cheng Fang, Tiemin Zhang, Haikun Zheng, Junduan Huang, Kaixuan Cuan, „Pose estimation and behavior classifica-tion of broiler chickens based on deep neural networks", Computers and Electronics in Agriculture, Vol. 180. 2021. 105863, ISSN 0168-1699, https://doi.org/10.1016/j.compag.2020.105863.

[22] Dawkins, M.S.; Sain, R.; Merelie, K.; Roberts, S.J. In search of the behavioural correlates of optical flow patterns in the automated assessment of broiler chicken welfare. Applied Animal Behaviour Science. 2013. 145:44-50.

[23] Pereira, D.F.; Miyamoto, B.C.B.; Maia, D.D.N.; Tatiana Sales, G.; Magalhaes, M.M.; Gates, R.S. Machine vision to identify broiler breeder behaviour. Computers and Electronics in Agriculture. 2013. 99:194-199.

[24] Mehdizadeh, S.A.; Neves, D.P.; Tscharke, M.; Naas, I.A.; Banhazi, T.M. Image analysis method to evaluate beak and head motion of broiler chickens during feeding. Computers and Electronics in Agriculture. 2015. 114:88-95.

[25] Aydin, A. Development of an early detection system for lameness of broilers using computer vision. Computers and Electronics in Agriculture. 2017. 136:140-146.

[26] Cao, L.; & Xiao, Z.; Liao, X.; Yao, Y.; Wu, K.; Mu, J.; Li, J.; Pu, H. Automated Chicken Counting in Surveillance Camera Environments Based on the Point Supervision Algorithm: LC-DenseFCN. Agriculture. 2021. vol. 11(6), pages 1-15.

[27] Schlageter-Tello A, Bokkers EAM, Groot Koerkamp PWG, Van Hertem T, Viazzi S. Comparison of locomotion scoring for dairy cows by experienced and inexperienced raters using live or video observation methods. Animal Welfare. 2015. 24:69–79

[28] Van Hertem, T.; Viazzi, S.; Steensels, M.; Maltz, E.; Antler, A. Automatic lameness detection based on consecutive 3D-video recordings. Biosyst. Eng. 2014. 119:108–16

[29] Amraei, S.; Mehdizadeh, S.A.; Sallary, S. Application of computer vision and support vector regression for weight pre-diction of live broiler chicken. Engineering in Agriculture Environment and Food. 2017. 10:266-271.

[30] Mortensen, A.K.; Lisouski, P.; Ahrendt, P. Weight prediction of broiler chickens using 3D computer vision. Computers and Electronics in Agriculture. 2016. 123:319-326.

outdoor precision pig research:

[1] ten Napel J, van der Veen AA, Oosting SJ, Koerkamp PWGG. A conceptual approach to design livestock production n systems for robustness to enhance sustainability. Livest. Sci. 2011; 139, 150–160. https://doi.org/10.1016/j.livsci.2011.03.007

[2] Mason A, de Medeiros Esper I, Korostynska O, Cordova-Lopez LE, Romanov D, Pinceková M, Bjørnstad PH, Alvseike O, Popov A, Smolkin O, Manko M. RoBUTCHER: A novel robotic meat factory cell platform. The International Journal of Robotics Research. 2024:02783649241234035.

[3] Takacs K, Mason A, Cordova-Lopez LE, Alexy M, Galambos P, Haidegger T. Current Safety Legislation of Food Processing Smart Robot Systems–The Red Meat Sector. Acta Polytechnica Hungarica 2022; 19, 249-267.

[4] Banhazi TM, Black JL. Precision Livestock Farming: A Suite of Electronic Systems to Ensure the Application of Best Practice Management on Livestock Farms, Australian Journal of Multi-Disciplinary Engineering, 2009; 7:1, 1-14, DOI: 10.1080/14488388.2009.11464794

[5] Tzanidakis C, Simitzis P, Arvanitis K, Panagakis P. An overview of the current trends in precision pig farming technologies, Livest. Sci. 2021; 249, 104530, ISSN 1871-1413, https://doi.org/10.1016/j.livsci.2021.104530.

[6] Lee I, Lee K. The Internet of Things (IoT): Applications, investments, and challenges for enterprises. Bus. Horiz. 2015; 58, 431–440. DOI: 10.1016/j.bushor.2015.03.008

[7] Berckmans D. 2014. Precision livestock farming technologies for welfare management in intensive livestock systems. Rev Sci Tech. 2014; 33(1):189-96. doi: 10.20506/rst.33.1.2273. PMID: 25000791.

[8] Edwards, S.A. 2005. Product quality attributes associated with outdoor pig production. Livestock Prod. Sci. 94. 5-14. doi:10.1016/j.livprodsci.2004.11.028

[9] Anonymous. Second Report on Priorities for Research and Development in Farm Animal Welfare; Department of Environment, Food and Rural Affairs, Farm Animal Welfare Council (FAWC) 1993; London, UK

[10] Haidegger T, Mai V, Mörch CM, Boesl DO, Jacobs A, Khamis A, Lach L, Vanderborght B. Robotics: enabler and inhibitor of the sustainable development goals. Sustainable Production and Consumption. 2023 Dec 1;43:422-34.

[11] Bracke MBM. Review of wallowing in pigs: Description of the behaviour and its motivational basis, Appl. Anim. Behav. Sci. 2011. https://doi.org/10.1016/j.applanim.2011.01.002.

[12] Horrillo A, Gaspar P, Muñoz Á, Escribano M, González E. Fattening Iberian Pigs Indoors vs. Outdoors: Production Performance and Market Value. Animals, 2023; 13, 506. https://doi.org/10.3390/ani13030506

[13] Reiners K, Hegger A, Hessel EF, Böck S, Wendl G, Van den Weghe HFA. Application of RFID technology using passive H.F. transponders for the individual identification of weaned piglets at the feed trough. Computers and Electronics in Agriculture 2009; 68 (2), 178–184.

[14] Voulodimos AS, Patrikakis CZ, Sideridis AB, Ntafis VA, Xylouri EM. A complete farm management system based on animal identification using RFID technology. Computers and Electronics in Agriculture 2010; 70 (2), 380–388

[15] Munak A. Preface. In: CIGR Handbook of Agricultural Engineering Volume VI. Information Technology. 2006; CIGR

[16] Umble EJ, Haft RR, Umble MM. Enterprise resource planning: "Implementation procedures and critical success factors". European Journal of Operational Research. 2003; 146(2), 241–257

[17] Hou H, Ma L. The relationship management of information system outsourcing provider perspective. Industrial Engineering and Engineering Management (IE&EM). 2010; 1760-1763.

[18] Drazic I, Barisic B, Mujakovic N, Čep R. Implementation of Shishkin Mesh in the Modelling of Spring-Mass System. Sborník vědeckých prací Vysoké školy báňské - Technické univerzity Ostrava. Řada strojní, Vysoká škola báňská – Technická univerzita Ostrava. 2010; 49-52, ISSN 1210-0471

[19] Ruiz-Garcia L, Lunadei L. The role of RFID in agriculture: Applications, limitations and challenges, Computers and Electronics in Agriculture. 2011; 79/1, 42-50, ISSN 0168-1699, https://doi.org/10.1016/j.compag.2011.08.010.

[20] Maselyne J, Saeys W, Briene P, Mertens K, Vangeyte J, De Ketelaere B, Hessel EF, Sonck B, Van Nuffel A. Methods to construct feeding visits from RFID registrations of growing-finishing pigs at the feed trough, Computers and Electronics in Agriculture. 2016; 128, 9-19, ISSN 0168-1699, https://doi.org/10.1016/j.compag.2016.08.010.

[21] Alexy M, Horvath T, Reich C, Felfoldi J, Tarcsi A. Adaption of data-intensive monitoring and tracking systems in outdoor pig production for better decision-making, literature review, and project idea. In Precision Livestock Farming 2019-Papers Presented at the 9th European Conference on Precision Livestock Farming, ECPLF, 684-691. 2019.

[22] Alexy M, Horvath T. Tracing the Local Breeds in an Outdoor System – A Hungarian Example with Mangalica Pig Breed. Tracing the Domestic Pig [Working Title] [Internet]. 2022 Jan 6; Available from: http://dx.doi.org/10.5772/intechopen.101615

[23] Egerszegi I, Ratky J, Solti L, Brussow KP Mangalica—An Indigenous Swine Breed from Hungary (Review). Archives Animal Breeding, 2003; 46, 245-256. https://doi.org/10.5194/aab-46-245-2003

[24] Wegner K, Lambertz C, Das G, Reiner G, Gauly M. Effects of temperature and temperature-humidity index on the reproductive performance of sows during summer months under a temperate climate. J. Anim Sci, 2016; 87: 1334– 1339. doi: 10.1111/asj.12569.

[25] Mader TL, Davis MS, Brown-Brandl T. Environmental factors influencing heat stress in feedlot cattle. J. Anim. Sci. 2006; 84, 712-719. DOI: 10.2527/2006.843712x

[26] Harrell J, Frank E. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. Springer, 2015

[27] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, 2022; Vienna, Austria. URL https://www.R-project.org/

[28] Wood S.N. Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC. 2017

[29] Simpson G. _gratia: Graceful ggplot-Based Graphics and Other Functions for GAMs Fitted using mgcv_. R package version 0.7.3, 2022; https://gavinsimpson.github.io/gratia/

[30] Baert S, Aubé L, Haley DB, Bergeron R, Devillers N. To wallow or nurse: Sows housed outdoors have distinctive approaches to thermoregulation in gestation and lactation. Appl. Anim. Behav. Sci., 2022; https://doi.org/10.1016/j.applanim.2022.105575

[31] St.Pierre NR, Cobanov B, Schnitkey G. Economic Losses from Heat Stress by U.S. Livestock Industries1, J. of Dairy Sci., 2003; 86, E52-E77, ISSN 0022-0302, https://doi.org/10.3168/jds.S0022-0302(03)74040-5

[32] Dailey JW, McGlone JJ. Oral/nasal/facial and other behaviors of sows kept individually outdoors on pasture, soil, or indoors in gestation crates. Appl. Anim. Behav. Sci. 1997; 52, 25-43. https://doi.org.10.1016/S0168-1591(96)01099-4.

[33] Alexy, M. Results of pasture-based gilts and breeding sow housing system. PhD-thesis. University of Debrecen, Centre of Agricultural Sciences, 2004

[34] Morrison MJ, Curkendall LD. Apparatus and Method for Reading Radio Frequency Identification Transponders used for Livestock Identification and Data Collection, edited by U.S. Patent. United States, Aginfolink Holdings, 2001

[35] Alexy M, Haidegger T. Precision Solutions in Livestock Farming–feasibility and applicability of digital data collection. In 2022 IEEE 10th Jubilee International Conference on Computational Cybernetics and Cyber-Medical Systems (ICCC), pp. 000233-000238. IEEE, 2022.

[36] Mason A, Haidegger T, Alvseike O. Time for change: the case of robotic food processing [industry activities]. IEEE Robotics & Automation Magazine. 2023 Jun 14;30(2):116-22.

[37] Szabo S, Alexy M. Practical Aspects of Weight Measurement Using Image Processing Methods in Waterfowl Production. Agriculture, 2022; 12, 11:1869.