



TOWARDS IMPROVEMENT OF **RUMINANT** BREEDING
THROUGH **GENOMIC** AND EPIGENOMIC APPROACHES

Grant agreement number: 101000226

H2020 – Research and Innovation Action

Deliverable 3.1

Report on the effect of heat stress on production and fertility traits in dairy cattle

Due date of deliverable: M24 – May 2023

Actual submission date: M28 – September 2023

WP concerned	WP3	
Lead partner	Jeremie VANDENPLAS	WR
Author(s)	Sophie AGUERRE Sophie MATTALIA Julie PROMP Roxane VALLEE Didier BOICHARD Beatriz CASTRO-DIAS-CUYABANO Aurélie VINET María Jesús CARABAÑO Clara DIAZ Manuel RAMON Mario CALUS	IDELE IDELE IDELE IDELE INRAE INRAE INRAE INIA-CSIC INIA-CSIC IRIAF WU



	Han MULDER	WU
Version	V1	
Dissemination level	PU: Public	

Call: H2020-SFS-2020-2

Topic: SFS-13-2020 | Genome and epigenome enabled breeding in terrestrial livestock

Start date of the project: June 1st, 2021

Duration: 60 months

End date of the project: May 31st, 2026

Project ID: 101000226

Table of Contents

Summary	5
1 Introduction	7
2 Description of dairy cattle and meteorological data.....	7
2.1 Breeds, national phenotypic data, genotypes and pedigree	7
2.1.1 Breeds.....	7
2.1.2 Milk production and udder health traits.....	7
2.1.3 Fertility trait	7
2.2 Meteorological information.....	8
3 Choice of heat stress indicators.....	8
4 Effect of THI at a population level.....	10
4.1 Methods.....	10
4.1.1 Production traits and SCS	10
4.1.2 Conception rate.....	11
4.2 Results and discussion	11
4.2.1 Milk production and SCS.....	11
4.2.2 Conception rate.....	13
5 Effect of heat stress at an individual level.....	17
5.1 Estimation of genetic parameters at the country level	17
5.1.1 Methods	17
5.1.2 Results and discussion.....	20
5.2 Meta-analyses: genetic correlations among countries.....	27
5.2.1 Methods	27
5.2.2 Results and discussion.....	28
6 Heat tolerance at a genome level	34
6.1 Material and methods	34
6.2 Results and discussion	35
7 General discussion and conclusions.....	40
8 References.....	42
Appendices.....	44
Appendix 1: List of abbreviations	45
Appendix 2: General description of the phenotypes used to define heat stress indicators.....	46
Appendix 3: Statistical analyses of weather data.....	51
Appendix 4: Software used in the different studies.....	56

Appendix 5: Genetic parameters in second lactation.....	57
Appendix 6: Comparison of estimated genetic correlations between countries (RUMIGEN vs routine Interbull international evaluations)	58
Appendix 7: Manhattan plots.....	59
Appendix 8: Study of effect of THI at the population level in France: comparison of THI effects at the regional level.....	62

Summary

One of the main objectives of RUMIGEN is to provide breeding tools for dairy cattle adapted to future environmental conditions, and thus to adapt the selection criteria for taking climate change into account. For this purpose, Institut de l'Elevage (Idele, France), Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE, France), Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA-CSIC, Spain), Instituto Regional de Investigación y Desarrollo Agrolimentario y Forestal de Castilla-La Mancha (IRIAF, Spain), Wageningen University (WU, the Netherlands) and Wageningen Research (WR, the Netherlands) collaborate in WP3 entitled "Impact of trade-offs on productive lifespan" to define several traits related to heat tolerance. More specifically, the aims of this deliverable D3.1 are 1) to define heat stress indicators, 2) to investigate the phenotypic impact of heat stress on the Dutch, French and Spanish Holstein cows, as well as on the Dutch Meuse-Rhine-Yssel (MRY) cows and the French Montbéliarde cows, for milk production, udder health, and fertility traits, 3) to estimate variance components and breeding values for heat tolerance traits, 4) to estimate genetic correlations among countries for some heat tolerance traits using a meta-analysis, and finally 5) to identify possible genome regions impacting heat stress.

To achieve these aims, large-scale performance, pedigree and genotype data from commercial herds provided by Base de Données Zootechniques Nationale (BDZN, France), CRV (The Netherlands), and Confederación de Asociaciones de Frisona Española (CONAFE, Spain), were first combined with meteorological data provided by Météo France (Safran Data base; France), Koninklijk Nederlands Meteorologisch Instituut (KNMI, The Netherlands), and National Meteorological Agency (AEMET, Spain), respectively. After investigating multiple heat stress indicators (e.g., temperature, relative humidity) in each country, a temperature-humidity index (THI) averaged over three days on and previous test dates for milk production and udder health traits, and averaged over eight days on and after the insemination date for fertility traits were chosen for investigating the impact of heat stress in dairy cattle at the population and individual levels. The THI summarizes temperature and relative humidity in a single index. Considering the relative humidity in a heat stress indicator is important because high humidity worsens the effect of high/low temperatures.

Analyses at a population level showed that increasing THI had a negative impact on all studied traits, especially for milk production and fertility traits. However, differences of magnitude of the estimated effects and of THI thresholds varied across traits and countries, more than across breeds. Within a country, similar effects and THI thresholds were observed between breeds. For example, moderate THI, with a THI threshold as low as 50, had a stronger impact on the milk production traits and somatic cell score (SCS) of French Holstein and Montbéliarde cows than for the other populations. It was also observed that loss in conception rate (CR) associated with heat stress was substantially larger for the French and Dutch Holstein cows than for the Spanish Holstein and French Montbéliarde cows. These differences across countries could be explained by several factors, such as acclimation of cows to chronic heat stress during summer and heat abatement in the barns, and would require future investigations relying on currently unavailable farming system information.

In a third step, genetic variance components and breeding values in the five dairy cattle populations were estimated using reaction norm models. Such models allow the prediction of breeding values for the production level and the slope of decay of the different traits. The trends for heritabilities and genetic correlations were similar across countries and breeds for the different milk production traits, SCS and CR, even if some differences could be observed. Overall, heritabilities were quite stable along the THI gradient for milk production traits at 150 days-in-milk, and increase slightly for SCS and markedly for CR with increasing THI. Estimated genetic correlations between different THI values were almost all higher than 0.9 for milk production traits and SCS, for the different breeds

and countries, suggesting no, or at most weak, genetic-by-THI interactions. In the other hand, estimated genetic correlations for conception rate between neutral and heat stress conditions decreased with increasing THI, especially for France and Spain for which genetic correlations lower than 0.50 were estimated.

Meta-analyses of results obtained in the previous step for the first-parity Holstein populations in each country showed a very good consistency of the traits based on levels under thermo-neutral and heat stress conditions. These results are encouraging for future international genetic evaluations of the Holstein populations, because regions that currently do not experience heat stress conditions will benefit from information already collected in other countries that often suffer from heat stress. However, these meta-analyses also showed the limits of their applications on low heritable traits such as the slopes of decay. Future studies should investigate other models and approaches to better measure the decline in performances due to heat stress and to take advantage of the information available in multiple countries.

GWAS results were based on de-regressed proofs of bulls and integrated information from the whole population in each country. In contrast to females, bulls have daughters with performances in variable conditions and specially in variable THI, allowing to define different traits under neutral conditions and heat stress and estimate breeding values for these different traits. Results highlighted the existence of specific regions for production levels as well as genomic regions shared with slopes of decay, our indicator traits of heat stress. For udder health all analyses failed to detect QTL and for fertility traits no putative genes were identified, probably because of a too limited power.

1 Introduction

One of the main objectives of RUMIGEN is to provide breeding tools adapted to future environmental conditions, and thus to adapt the selection criteria for taking climate change into account. For this purpose, France (Idele, INRAE), Spain (INIA, IRIAF) and the Netherlands (WU and WR) collaborate in WP3 (“Impact of trade-offs on productive lifespan”) to define several traits related to heat tolerance. To achieve this purpose, large-scale data from commercial farms were combined with meteorological data. The aims of this deliverable D3.1. are 1) to define heat tolerance indicators, 2) to investigate the phenotypic impact of heat stress on the Dutch, French and Spanish Holstein cows, as well as on the Dutch Meuse-Rhine-Yssel (MRY) cows and the French Montbéliarde cows, for milk production, udder health, and fertility traits, 3) to estimate variance components and breeding values for heat tolerance traits, 4) to estimate genetic correlations among countries for some heat tolerance traits using a meta-analysis, and finally 5) to identify possible genome regions impacting heat tolerance.

2 Description of dairy cattle and meteorological data

2.1 Breeds, national phenotypic data, genotypes and pedigree

2.1.1 Breeds

The breeds investigated in this study are a cosmopolitan breed, Holstein, and two local breeds, Montbéliarde and MRY. Data for Holstein are available in France, Spain, and the Netherlands. Data for Montbéliarde are only available in France, and for MRY only in the Netherlands.

2.1.2 Milk production and udder health traits

These studies focused on the performances in first and second lactations on the following traits: milk yield (MY, in kg/d), fat yield (FY, in g/d or kg/d), protein yield (g/d or kg/d), fat content (FC, in %), and protein content (PC, in %) and somatic cell score (SCS; defined as $SCS = 3 + \log_2(SCC/100,000)$, with SCC being somatic cell counts in cells/ml). Test-day records of French, Spanish and Dutch Holstein cows, of French Montbéliarde cows and of Dutch MRY cows, were extracted from the respective genetic evaluation data bases (BDNZ, France; CONAFE, Spain; and CRV, the Netherlands). It is worth noting that the records for the Netherlands were extracted for Dutch herds having a majority of cows genotyped.

For all countries, the extracted datasets used for this deliverable covered approximately a period of 10 years, starting from the 2010s. Detailed description of the data and edits in each country is presented in Appendix 2. Briefly, for the period starting around 2010 until around 2020, the datasets for the milk production traits (MY, PY, FY, PC, FC, and SCS) included first and second-parity cows in France and in the Netherlands, and first to third parity cows in Spain. Around 49, 19.8, and 7.5 million records from 7, 1, and 0.5 million Holstein cows from France, Spain, and the Netherlands, respectively. A total of 37 million test-day records from 1.6 million French Montbéliarde cows were available, as well as 165 thousand test-day records from 14 thousand Dutch MRY cows. More details on the data sets can be found in Appendix 2.

2.1.3 Fertility trait

The fertility trait chosen to assess the impact of heat stress on the fertility of dairy cattle is the conception rate (CR) defined as the success (0/1) of the first artificial insemination (AI). Phenotypes were extracted from the French, Spanish, and Dutch national data bases, similarly to the phenotypes for milk production traits.

Briefly, all datasets covered approximately a period of about 10 years, starting from the 2010's. For the Holstein populations in France, Spain, and the Netherlands, the datasets included around 3.4, 1.4, and 0.8 millions of Holstein cows, respectively. Around 650 thousand CR status of French Montbéliarde cows were available, and 17 thousand CR status were available for the Dutch MRY cows. More details on the data sets can be found in Appendix 2.

2.2 Meteorological information

Meteorological data were provided by Météo-France (Safran database) for France, by the National Meteorological Agency (AEMET) for Spain, and by the Koninklijk Nederlands Meteorologisch Instituut (KNMI) website for the Netherlands. Meteorological records were available for 1,993 Spanish and 34 Dutch weather stations distributed throughout each national territory. In France, the information resulted from a meteorological model characterizing each 8 x 8 km square (9,892 squares) covering the whole country. Each herd was associated to the closest weather station or square, according to its (partial) ZIP code. The average distance between a farm and a weather station was equal to 8.4 km in Spain, to 14.6 km in the Netherlands, and 5.7 km in France.

Average, minimum and maximum daily temperatures (in degrees Celsius) and daily mean relative atmospheric humidity (in percent) were available, or computed from available records, in each country. All countries computed daily temperature-humidity indices using the formula proposed by (National Research Council, 1971):

$$THI = (1.8 \cdot T + 32) - (0.55 - 0.0055 \cdot RH) \cdot (1.8 \cdot T - 26),$$

with T being the average daily temperature (degrees Celsius) and RH being the average daily relative humidity (percent).

Statistics on meteorological information are provided in Appendix 3 for each country, with details on several regions for France and Spain. They show that the panel of climates covered by the study was broad, and that enough phenotypic records corresponding to high temperatures or THI ($THI \geq 60$ or even ≥ 72) could be found in each country.

3 Choice of heat stress indicators

In order to select a heat stress indicator common to the three countries, several parameters were investigated, such as the daily mean, minimum and maximum temperatures and THI averaged over a period of up to 7 days preceding a given test-date for milk production traits and SCS, and before or after the first AI for CR.

First, it was observed that the different parameters summarizing daily mean, minimum and maximum temperatures and THI averaged over a period of up to 7 days for a given day were all highly correlated (see Appendix 3 for more details). For example, daily THI and average daily temperatures would provide the same amount of information, as the Pearson correlation between both parameters was equal to 1.00. The Pearson correlations between daily mean and minimum or maximum temperatures were also high in France and in the Netherlands (e.g., 0.94 and 0.97, respectively, between minimum and average, and between maximum and average temperatures in the Netherlands), while lower correlations were observed in Spain (i.e. in the ranges [0.79 - 0.90] and [0.89-0.93]). Furthermore, Pearson correlations between 3-day averaged THI and other THI averaged on different periods, were on average all larger or equal to 0.97 for the Netherlands, to 0.91 for France, and to 0.85 for Spain (across all regions).

Second, the effect of these indicators was estimated on several traits and breeds at the population level using models presented in Section 2.4. Figure 1 shows an example of results obtained with the French Montbéliarde population on MY in first and second lactations and Figure 2 shows results obtained with Spanish data on CR of Spanish first-parity Holstein cows. Both figures illustrate that the trends were the same, whatever the heat stress indicator, which is in agreement with the high Pearson correlations observed among all heat stress indicators.

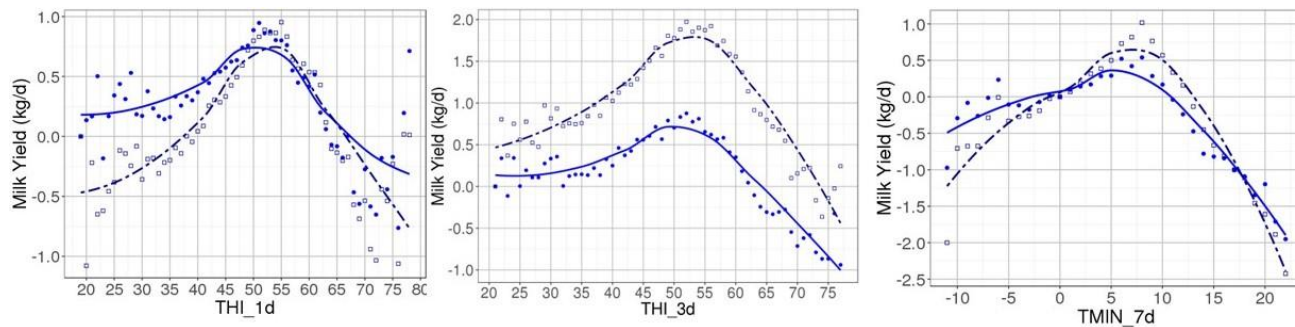


Figure 1. Heat load phenotypic response for different heat stress indicators on daily milk yield of the French Montbéliarde breed (solid lines: first parity cows; dotted lines: second parity cows; THI_1d: daily temperature-humidity index (THI) calculated at the day of record; THI_3d: average THI calculated during the period [Day-2; Day of record]; TMIN_7d: averaged minimum temperature recorded during the period [Day-6; Day of record]).

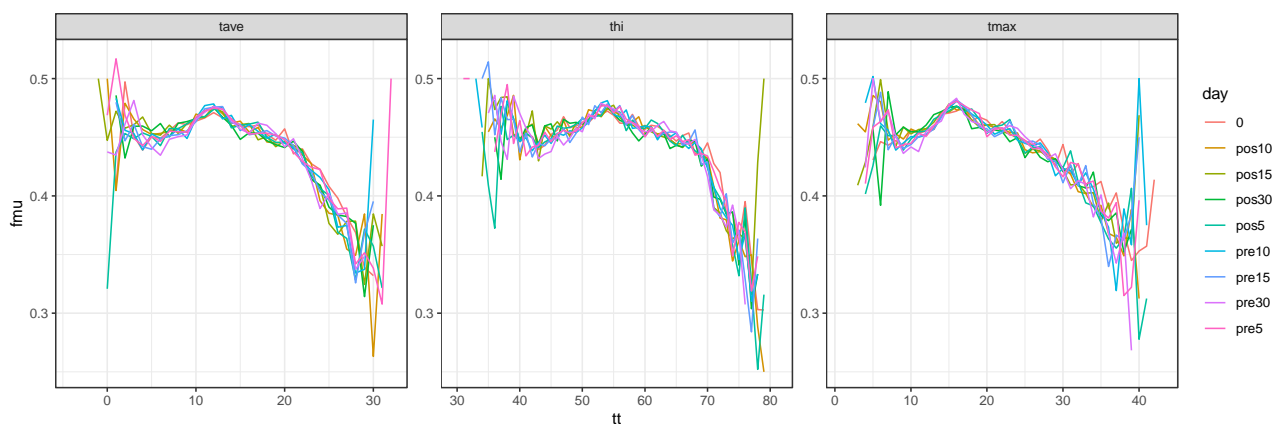


Figure 2. Heat load phenotypic response curves for different heat stress indicators on conception rate of the first artificial insemination of Spanish first-parity Holstein cows between 2010 and 2021. Heat stress indicators were average (tave) and maximum (tmax) temperatures and temperature-humidity index (THI) averaged for different periods (0,5 10, 15 and 30 day before and after day of artificial insemination).

Finally, based on the above results, on the meteorological information available in each country, and considering that relative humidity is important because high humidity worsens the effect of high/low temperatures, daily mean THI was preferred over temperature because it summarizes both temperature and relative humidity in an index. Therefore, THI was used as heat stress indicator in the next studies.

For production traits and SCS, the results showed that the different values of THI averaged over up to 7 days were highly correlated to a 3-day averaged THI. As a long period could include more fluctuations between weather situations and within period, the 3-day averaged THI was finally chosen as a compromise regarding the metabolism of dairy production. This indicator has been used in the next studies presented in this report.

For CR, two periods (before and after AI) might have an impact on the result, and the results were very consistent across all tested periods. Furthermore, a Spanish study showed that the effect of a THI value measured before an AI was not significant anymore when THI values for both periods before and after AI were simultaneously fitted in the model (Figure 3). These results were also confirmed in France. Due to all these reasons, the retained indicator for further analyses on fertility traits was the THI averaged during a period including the day of AI and the following week (hereafter referred to as 8-day averaged THI; THI_8day).

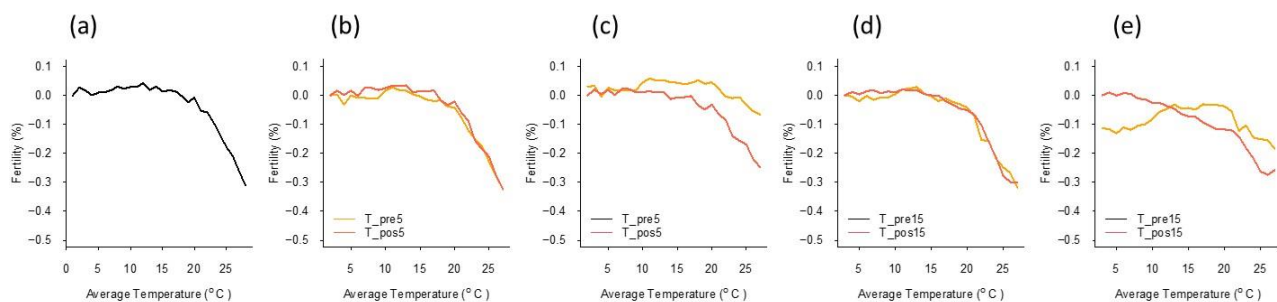


Figure 3. Heat load phenotypic response curves on conception rate of the first artificial insemination of Spanish first-parity Holstein cows using (a) average temperature of the day of insemination; (b) average temperature for a period of ± 5 days of insemination modeled separately; (c) as b, but modeled jointly; (d) average temperature for a period of ± 15 days of insemination modeled separately; (e) as d, but modeled jointly.

4 Effect of THI at a population level

4.1 Methods

This section presents the models used to estimate the effect of THI at the population level. Software used to estimate variance components and solutions are described in Appendix 4. For all the analyses, pedigree information of phenotyped cows was traced back for 3 generations.

4.1.1 Production traits and SCS

The phenotypes (MY, FY, PY, FC, PC and SCS test-day records) of the first and second lactations were analysed in separate studies for each breed and population. As mentioned above, the heat stress indicator was the 3-day averaged THI for production traits and SCS. It is worth noting that the Netherlands used an averaged THI using the 3 days before the record, while France and Spain used an average THI using the day of record and 2 days before. For each combination of trait, breed and parity, the effect of the 3-day averaged THI on test-day performances at the population level was estimated using the following model:

$$y = Xb + Z_a a + Z_p p + e$$

where \mathbf{y} , \mathbf{b} , \mathbf{a} , \mathbf{p} and \mathbf{e} are the vectors of phenotypes, fixed effects, random additive genetic effects, random permanent environment effects and the random residuals, respectively, and \mathbf{X} , \mathbf{Z}_a , and \mathbf{Z}_p are the incidence matrices for the listed effects.

The fixed effects were almost the same in all countries, except for the gestation stage that was included in the French and in the Dutch model, but not in the Spanish one. However, their combination differ from one country to the other as follows:

- France: herd-year, 3-day averaged THI, days-in-milk (DIM), gestation stage, month of calving and age at calving;
- The Netherlands: herd-year, 3-day averaged THI, DIM, gestation stage, age at calving–year of calving– season of calving;
- Spain: herd-year-season of calving, 3-day averaged THI, DIM and age at calving.

As in the previous models, the fixed effect for 3-day averaged THI was fitted as classes corresponding to one degree of THI.

4.1.2 Conception rate

Conception rate of the first-parity cows were analysed in separate studies. As mentioned earlier, the heat stress indicator was the 8-day averaged THI (THI_8day), with the first day being the day of the artificial insemination. For each population, the effect of THI on phenotypic performances at the population level was estimated using the following model:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_a\mathbf{a} + \mathbf{e}$$

where \mathbf{y} , \mathbf{b} , \mathbf{a} and \mathbf{e} are the vectors of phenotypes, fixed effects, random additive genetic effects of the sire (in the French and Dutch models), or of the cow (in the Spanish model) and the random residuals, respectively, and \mathbf{X} and \mathbf{Z}_a are the incidence matrices for the listed effects. In the Dutch model, a random effect for service sire was also fitted.

The fixed effects were almost the same in all countries, except the effect of the day of week and of sexed semen, that are only included in the French and Dutch models. However, their combination differ from one country to the other as follows:

- France: herd-year, THI_8day, season, DIM, week day, and age at calving;
- The Netherlands: general mean, herd-year, THI_8day, year-month, week day, age at calving, and sexed semen (Y/N);
- Spain: herd-year, THI_8day, DIM, and age at calving.

The fixed effect for THI_8day was fitted as classes corresponding to one degree of THI_8day as well.

4.2 Results and discussion

4.2.1 Milk production and SCS

The estimated effects of the 3-day averaged THI on MY, FY, PY, FC, PC, and SCS for first-parity cows of the three breeds and of the three countries are presented in Figure 4. The estimated 3-day averaged THI effects on MY, FY and FC, and SCS in first and second lactations are compared in Figure 5 for French Montbéliarde and Dutch Holstein cows. Estimated THI effects presented on both figures were averaged using a rolling basis of 5 THI units to smooth the fluctuations.

It is worth noting that Dutch results of estimated effects of a high THI must be interpreted carefully since the Netherlands are exposed to oceanic climatic conditions, with moderate daily variations.

Only 66 days with a daily THI over 72 were observed between 2010 and 2019. However, THI values over 60 were observed on average 120 days each year during the same decade, which means that the effect of moderate heat stress can be assessed using Dutch data. The frequency of days with high THI was much higher in France and in Spain. For example, between 88 and 345 days with a THI over 72 were observed during within 5 years in France, depending on the region.

Table 1 shows the estimated heat stress thresholds (change points) and slopes of decay obtained by the R package “Segmented” (Muggeo, 2008). Three change points were searched by the algorithm and those, followed by decay in production, or increase in SCS, were considered as heat stress thresholds. Up to two heat stress thresholds varied across combinations of traits, breeds and countries. All these results showed that increasing THI had a negative impact on all studied traits, but more prominent for milk production than for SCS. More specifically, the estimated decrease in FY between 50 and 70 units of THI (i.e., 9 and 22.5 °C of daily mean temperature for a relative humidity of 70%) corresponds to 11% of the daily production of first parity French Holstein cows. For SCS, thresholds were lower (between 41 and 58), indicating that udder health might be affected by high heat loads at lower THI values than milk production. The pattern of the curves of THI effects, thresholds and decays, depended on the trait: for MY, SCS and PY, the THI effects were almost stable below 50 (with some exceptions such as the French Holstein cows for PY), while the decline in FY associated with increasing THI became much steeper for THI higher than 50. The changes in estimated THI effects on milk and solid yields resulted in a negative effect of increasing THI all along the THI scale for the ratios (FC and PC).

Some differences of magnitude of the estimated effects (represented by slopes of decay; Tables 1-2) and of THI thresholds were observed between countries (Table 1). Within a country, similar effects and THI thresholds were observed between breeds, particularly for the yields and for SCS. For MY, FY and PY, moderate THI, with a THI threshold as low as 50, had a stronger impact on the performances of French cows than for the other populations. For SCS, the effect of increasing THI over a value of 50 was stronger for the first parity French cows. Overall, the Spanish population of Holstein cows showed the highest threshold for milk production traits, with 3-day averaged THI values ranging from 52 for FC to 72 for MY. Moreover, under heat stress, it can be observed that slopes of decay under moderate (THI=70) and more severe (THI=75) heat stress tended to be smallest for the Spanish Holstein population (Table 2). This is an indication of a better adaptation to high heat loads for the Spanish population. The fact that most dairy farms adopt heat abatement strategies (mainly force ventilation and sprinklers) during summer, and that heat stress is chronically happening all along the summer in Spain, with a likely acclimation of animals, might explain the observed differences.

The thermo-neutral conditions observed in France were relatively low when compared to the literature (Carabaño et al., 2017) and to the results obtained in this study for Spain and the Netherlands (Figures 4-5). These differences should be interpreted with caution, as they may be explained by many different reasons. First, different farming systems in the three countries may explain partly these lower thresholds for France. Indeed, the French farming systems are often based on pasture, resulting in French cows often exposed to outside temperatures (even if they are kept inside during very hot periods). On the other hand, the Spanish farming systems mostly keep animals in the barns and potentially use heat abatement devices and mitigation practices. Furthermore, most of the Dutch herds investigated in this study are equipped with automatic milking systems, with less grazing and potentially equipped with barns more adapted to heat than the average Dutch farms. Thus, French observations may better reflect the actual effect of outside THI, as a larger proportion of cows are actually more exposed to outdoor meteorological conditions than in the two other countries. More information on the farming system and on the equipment of each herd would be required to disentangle their impact from the effect of THI on milk production. Second, changes in

feeding ration may also explain partly lower thresholds for France, particularly during the spring. This season is associated with moderate THI (that is around 50-55) as well as with changes of feeding systems. Therefore, the effect of THI on milk production might be confounded with the effect of the transition period from indoor feeding rations to pasture, that is known to have a negative impact on the lactation curve (Hartwiger et al., 2018). A study conducted on the performances of the French Holstein cows showed that the estimated effects of moderate THI values (between 50 and 60) were more negative in pasture regions (Normandy) than in regions with a majority of indoor-feeding systems (Pays de la Loire, mainly based on silage and concentrates), which would support this explanation (see Appendix 8 for more details). Based on these reasons, and without any other information on housing or feeding systems, we should carefully interpret the decrease in milk production observed at moderate THI as a direct welfare indicator.

4.2.2 Conception rate

The estimated effects of THI_8day on CR for first-parity cows of Holstein and Montbéliarde in the three countries are presented in Figure 6. Results for MRY were inconclusive and sensitive to the other fixed effects in the model, most likely due to the relatively low number of records, and the very low number of records subject to heat stress. Analyses for MRY were therefore not further pursued, and also not presented here. Estimated THI effects presented in Figure 6 were averaged using a rolling basis of five THI_8day units to smooth the fluctuations. The results showed for all the populations a thermo-neutral region where no response to increases in heat load is observed followed by a heat stress region where the negative impact of THI_8day can be observed. However, for the different populations, different heat stress thresholds for THI and slopes of decay were observed, although to a smaller extent to differences in the pattern of heat stress response in milk production traits observed for the same populations.

Table 3 presents the estimated heat stress thresholds (change points) and subsequent slopes of decay under heat stress for each population, estimated with the R package “Segmented” (Muggeo, 2008). Estimated change points varied from 61 to 64 for heat stress threshold across populations. Our estimates of heat stress thresholds are similar to that reported by (Gernand et al., 2019) for pregnancies per AI and (Biffani et al., 2016) for non-return rate, and smaller to the thresholds estimated by (Ravagnolo and Misztal, 2002) and (Santana et al., 2017) also for non-return rate. Compared to the heat stress thresholds obtained for milk production traits, the estimated heat stress thresholds for CR were higher for the French Holstein and Montbéliarde populations, similar for the Dutch Holstein population, and lower for the Spanish Holstein population.

Loss in CR associated with heat stress, depicted by the slopes of decay, was substantially larger for French and Dutch Holstein cows (around 1% decrease per THI unit above the threshold) than in Spanish Holstein and French Montbéliarde populations (around 0.5%) (Figure 6; Table 3). Acclimation of cows to chronic heat stress during summer and heat abatement in the barns in the Spanish population might explain these results. In the case of the French Montbéliarde, lower productive levels than Holstein cows and a long background of selection on fertility in a context of a strongly seasoned reproduction period might result in less compromised energy balance and better fertility under heat stress.

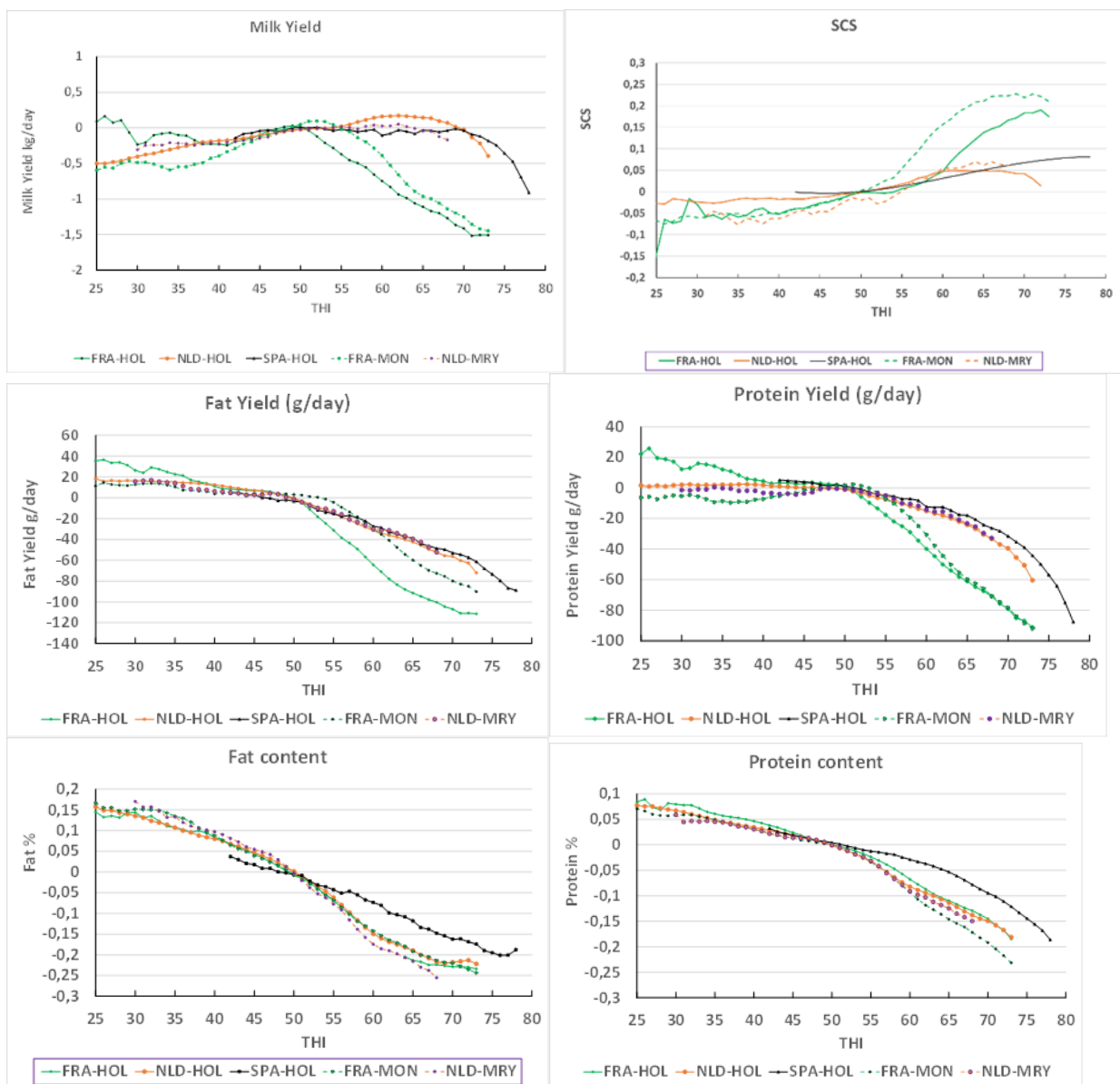


Figure 4. Effect of 3-day averaged temperature-humidity index (THI) on milk yield, fat yield, protein yield, fat content, protein content and somatic cell score (SCS) on first-parity cows, according to the country (FRA=France; NLD=the Netherlands; SPA=Spain) and to the breed (HOL=Holstein; MON=Montbéliarde, MRJ=Meuse-Rhine-Yssel).

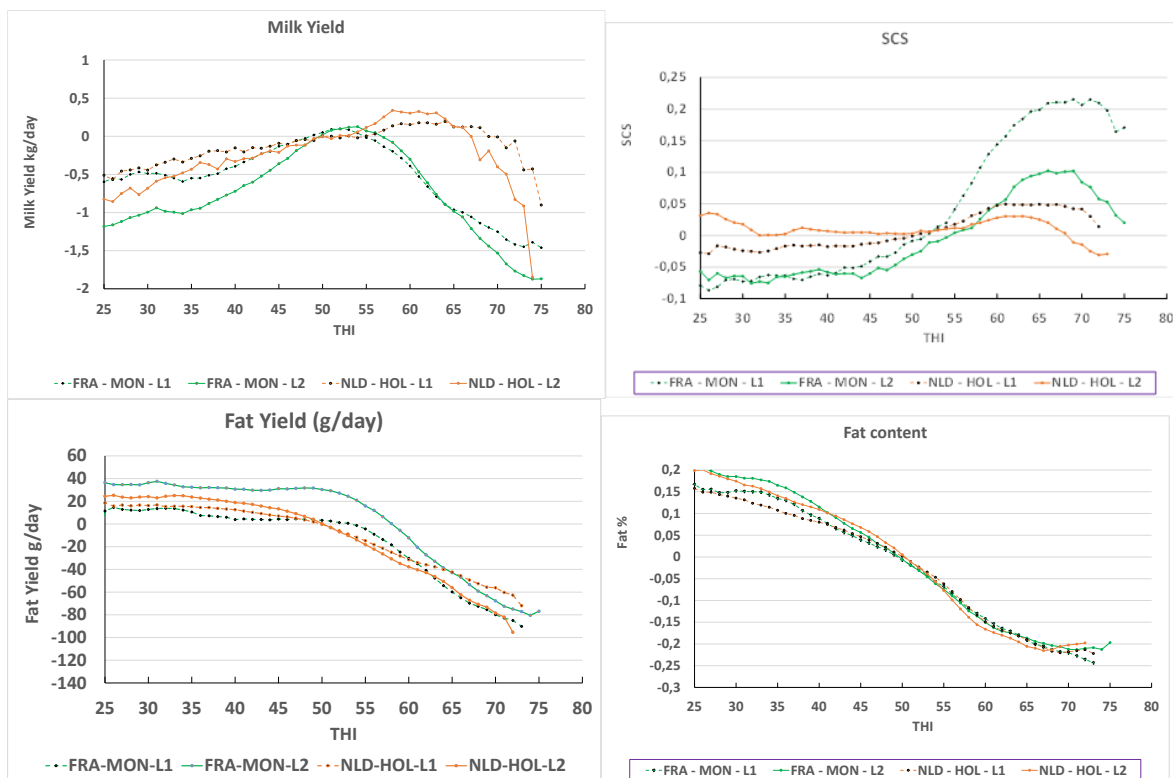


Figure 5. Effect of 3-day averaged temperature-humidity index (THI) on milk yield, fat yield, fat content, and somatic cell score (SCS) on French Montbéliarde and Dutch Holstein cows in first (L1) and in second (L2) lactation (FRA=France; NLD=the Netherlands; HOL=Holstein; MON=Montbéliarde).

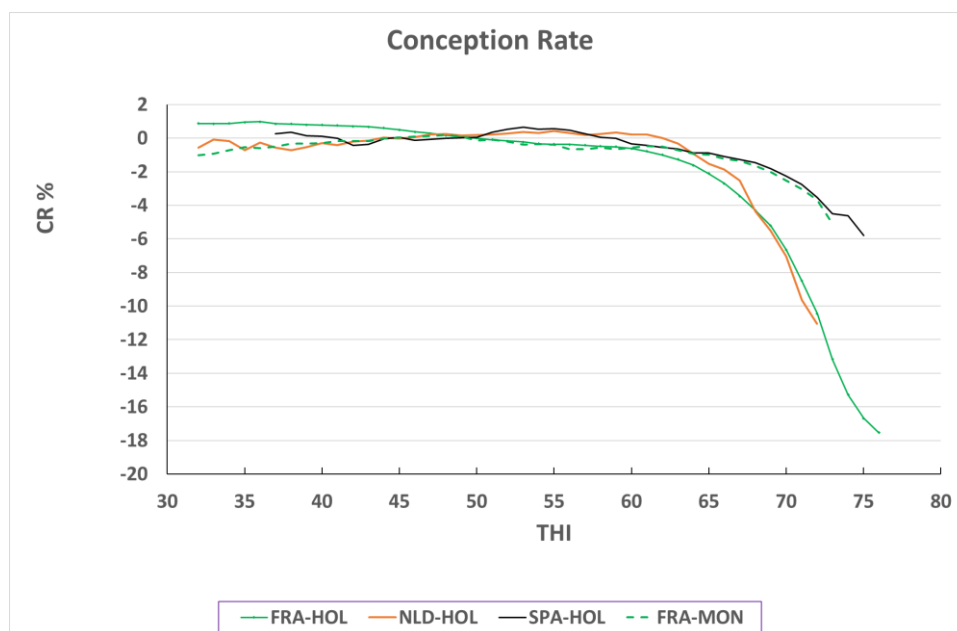


Figure 6. Effect of 8-day averaged temperature-humidity index (THI) on conception rate (CR) on first-parity cows, according to the country (FRA=France; NLD=the Netherlands; SPA=Spain) and to the breed (HOL=Holstein; MON=Montbéliarde).

Table 1. Estimated heat stress thresholds (THI₀) and subsequent slopes of decay (Slp) obtained from the response curve of milk production traits and somatic cell score (SCS) to the temperature-humidity index (THI) for the Holstein (HOL) breed in France (FRA), The Netherlands (NLD) and Spain (SPA) and for the Montbéliarde (MON) and Meuse-Rhine-Yssel (MRY) breeds.

Trait		HOL-FRA	HOL-NLD	HOL-SPA	MON-FRA	MRY-NLD
Milk yield (g/d)	THI ₀₁	52	66	72	49	59
	Slp ₁	-79.61	-60.55	-88.35	-13.01	-25.14
	THI ₀₂	-	72	77	55	-
	Slp ₂	-	-191.09	-287.30	-82.72	-
Fat yield (g/d)	THI ₀₁	49	45	70	51	50
	Slp ₁	-5.74	-2.67	-4.73	-4.69	-2.59
	THI ₀₂	-	72	78	-	68
	Slp ₂	-	-8.82	-1.40	-	-7.32
Protein yield (g/d)	THI ₀₁	35	66	70	48	58
	Slp ₁	-0.55	-3.81	-5.05	-1.56	-2.03
	THI ₀₂	50	71	76	55	65
	Slp ₂	-4.05	-10.06	-14.06	-4.80	-3.30
Fat content (%)	THI ₀₁	42	46	52	43	47
	Slp ₁	-0.013	-0.014	-0.008	-0.012	-0.016
	THI ₀₂	-	64	-	63	59
	Slp ₂	-	-0.004	-	-0.007	-0.012
Protein content (%)	THI ₀₁	50	48	61	49	45
	Slp ₁	-0.008	-0.007	-0.007	-0.008	-0.005
	THI ₀₂	72	72	75	57	52
	Slp ₂	-0.018	-0.014	-0.017	-0.011	-0.009
SCS	THI ₀₁	58	44	53	48	41
	Slp ₁	0.016	0.003	0.004	0.014	0.006
	THI ₀₂	-	-	72	-	-
	Slp ₂	-	-	0.001	-	-

Table 2. Estimated slopes of decay in milk production traits and somatic cell score (SCS) under moderate (THI=70) and more severe (THI=75) heat stress for the Holstein (HOL) breed in France (FRA), The Netherlands (NLD) and Spain (SPA) and for the Montbéliarde (MON) and Meuse-Rhine-Yssel (MRY) breeds.

Trait	THI	HOL-FRA	HOL-NLD	HOL-SPA	MON-FRA	MRY-NLD
Milk yield (g/d)	70	-36.38	-68.81	-10.42	-72.94	-85.65
	75	42.32	-232.82	-99.66	-15.98	-
Fat yield (g/d)	70	-2.18	-3.49	-3.42	-2.55	-8.43
	75	3.21	-10.95	-5.12	-8.71	-
Protein yield (g/d)	70	-3.24	-4.47	-2.85	-4.44	-3.88
	75	-3.57	-12.09	-6.61	-3.02	-
Fat content (%)	70	-0.004	-0.003	-0.009	-0.006	-0.024
	75	0.009	-0.008	-0.006	-0.010	-
Protein content (%)	70	-0.009	-0.008	-0.008	-0.011	-0.004
	75	-0.022	-0.017	-0.011	-0.014	-
SCS	70	0.014	-0.007	0.003	0.002	0.028
	75	0.057	0.000	0.002	0.017	-

Table 3. Estimated heat stress thresholds and subsequent slopes of decay obtained from the response curve of conception rate (CR) to the temperature-humidity index (THI) for the Holstein breed in France, The Netherlands and Spain, and for the Montbéliarde breed.

Breed	Country	Heat stress threshold (THI)	Slope of decay (%CR/unit THI)
Holstein	France	63	-1.29
	The Netherlands	62	-1.04
	Spain	64	-0.51
Montbéliarde	France	61	-0.48

5 Effect of heat stress at an individual level

5.1 Estimation of genetic parameters at the country level

5.1.1 Methods

5.1.1.1 Milk production traits and SCS

To study the genetic-by-THI (GxTHI) interactions, all countries used similar sets of data and the same statistical methods as in the analyses at the population level (see Appendix 2 for more details about data sets and Appendix 4 for used software). For each population, a test-day random regression model was used to estimate GxTHI interactions and has the following form:

$$y = Xb + Z_a Qa + Z_p Qp + e$$

where \mathbf{y} is the vector of test-day records, \mathbf{b} is the vector of fixed effects, \mathbf{a} is the vector of random additive genetic regression coefficients, \mathbf{p} is the vector of random permanent environmental regression coefficients, and \mathbf{e} is the vector of random residual effects. The matrices \mathbf{X} , \mathbf{Z}_a and \mathbf{Z}_p are incidence matrices relating the test-day records to the corresponding effects. Matrix \mathbf{Q} is a matrix containing the covariates of normalized Legendre polynomials for both THI_3day and DIM. The random effects \mathbf{a} and \mathbf{p} are assumed to follow a multivariate normal distribution with a zero mean and a symmetric covariance matrix. Random regression coefficients for THI_3day and DIM are assumed to be correlated for both additive genetic and permanent environmental effects. The residual variances are assumed to be heterogeneous and defined for different combinations of classes of DIM and/or of THI_3day. The normalized DIM Legendre polynomials were standardized between 7 and 305 DIM for France, 20 and 305 DIM for Spain, and 5 and 305 DIM for the Netherlands. The normalized THI_3day Legendre polynomials were standardized between THI_3day values of 20 and 70 for France, 40 and 80 for the Spain, and 10 and 83 for the Netherlands.

The fixed effects were almost the same in all countries, and were defined as follows:

- France: herd-test-day of record, DIM, gestation stage, month of calving and age at calving;
- The Netherlands: herd-test-date, DIM, gestation stage, age at calving-year of calving-season of calving, age at calving-year of calving-season of calving-lactation stage;
- Spain: herd-year-season of record, age at calving, and DIM.

5.1.1.2 Conception rate

To assess GxTHI interactions on CR, the general model used by all countries was as follows:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_a\mathbf{Q}\mathbf{a} + \mathbf{e}$$

where \mathbf{y} is the vector of records for success or failure of the first AI in the first lactation, \mathbf{b} is the vector of fixed effects, \mathbf{a} is the vector of random additive genetic regression coefficients, and \mathbf{e} is the vector of random residual effects. The incidence matrices \mathbf{X} and \mathbf{Z}_a relate the records to the corresponding effects. The matrix \mathbf{Q} (defined below) is the matrix containing covariates for THI_8day. The random effects \mathbf{a} and \mathbf{e} are assumed to follow a multivariate normal distribution with a zero mean and a symmetric covariance matrix. Random regression coefficients for THI_8day are assumed to be correlated. The residual variances are assumed to be heterogeneous and defined for different combinations of classes of THI.

Due to the low proportion of performances at high THI in the Netherlands and in France and to the low heritability of CR, both countries used a sire model, while Spain used an animal model. Moreover, the three countries used a broken line model for estimating GxTHI interactions. France also used a random regression model applied on the French Holstein population. It is worth noting that the regression coefficients over the threshold were associated to the THI_8day values in France, while they were associated to $(\text{THI}_{8\text{day}} - 60)$ in the Netherlands (with 60 being the threshold defined in the Dutch study) and to $(\text{THI}_{8\text{day}} - 66)$ in Spain.

The environmental effects included in the models are almost the same in the three countries. The main difference is on herd-year, that is a fixed effect in the French and Dutch models and a random effect in the Spanish one. The included fixed effects are:

- France: herd-year of AI, month-year of AI, week day, age at AI, interval between calving and AI;
- The Netherlands: regression on $(\text{THI}_{8\text{day}} - 60)^2$, herd-year of AI, month-year of AI, week day, age at calving, sexed semen (Y/N);
- Spain: age at calving, days from calving to AI and THI (class).

5.1.1.3 Analyses

For milk production traits and SCS in all countries, the genetic variances and covariances at each DIM and at each THI value were estimated by pre- and post-multiplying the (co)variance matrix \mathbf{G} by the corresponding DIM- and THI_3day-coefficients of normalized Legendre polynomials as follows:

$$\sigma_{g,ij}^2 = \mathbf{q}'_{ij} \mathbf{G} \mathbf{q}_{ij} \text{ and } \sigma_{ij,kl} = \mathbf{q}'_{ij} \mathbf{G} \mathbf{q}_{kl}$$

where $\sigma_{g,ij}^2$ is the genetic variance evaluated at the i -th DIM and at a value of THI_3day equal to j , $\sigma_{g,ij,kl}$ is the genetic covariance between the combination of the i -th DIM and a value of THI equal to j and the combination of the k -th DIM and a value of THI_3day equal to l , \mathbf{q}_{ij} (\mathbf{q}_{kl}) is the vector of covariates of normalized Legendre polynomials computed for the $i(k)$ -th DIM and a value of THI_3day equal to $j(l)$.

Genetic correlations across DIM and THI_3day were computed as:

$$r_{ij,kl} = \frac{\sigma_{g,ij,kl}}{\sqrt{\sigma_{g,ij}^2 \sigma_{g,kl}^2}}$$

where $r_{ij,kl}$ is the genetic correlation between the combination of the i -th DIM and a value of THI_3day equal to j and the combination of the k -th DIM and a value of THI_3day equal to l .

Heritabilities for all combinations of DIM and THI were computed as:

$$h_{ij}^2 = \frac{\sigma_{g,ij}^2}{\sigma_{g,ij}^2 + \sigma_{p,ij}^2 + \sigma_{e,ij}^2}$$

where $\sigma_{p,ij}^2$ and $\sigma_{e,ij}^2$ are the variances for the permanent environment and residual effects, respectively, for at the i -th DIM and at a value of THI_3day equal to j . For CR, genetic correlations between different THI levels, and heritabilities across THI levels were computed in a similar way, albeit that there was no permanent environment effect for CR.

Finally, estimated breeding values (EBVs) for the animal m for different values of DIM and THI_3day were computed as:

$$\hat{\mathbf{u}}_{m,ij} = \mathbf{q}'_{ij} \hat{\mathbf{a}}_m$$

where $\hat{\mathbf{u}}_{m,ij}$ is the estimated breeding value for the animal m at the i -th DIM and a value of THI_3day equal to j , and $\hat{\mathbf{a}}_m$ is the vector of the random regression coefficients estimated for the animal m .

For example, EBVs at 150 DIM (hereafter called as levels) under thermo-neutral conditions (i.e. at THI_3day = 50) and under heat-stress conditions (i.e. at a THI_3day value higher than the heat stress threshold) were computed for the estimation of genetic correlations across countries in Section 6. These THI_3day values were defined based on the results obtained from the analyses at the population level (see Section 4 for more details).

Similarly, the slope of the regression on (standardized) THI_3day for the animal m was computed using the first derivative as:

$$\frac{\partial \hat{\mathbf{u}}_{m,ij}}{\partial \text{THI}} = \delta \mathbf{q}'_{\text{THI},ij} \hat{\mathbf{a}}_m$$

where the vector $\delta \mathbf{q}'_{\text{THI},ij}$ contains the first derivative of covariates of the standardized Legendre polynomials.

For example, EBVs for slopes under heat-stress conditions (i.e. at a THI_3day value higher than the heat stress threshold) were computed for the estimation of genetic correlations across countries in Section 6. These THI_3day values were defined based on the results obtained from the analyses at the population level (see Section 4 for more details).

The genetic variance of the slope needed to compute reliabilities of the EBVs in Section 6 was computed as $\sigma_{s,ij}^2 = \delta \mathbf{q}'_{THI,ij} \mathbf{G} \delta \mathbf{q}_{THI,ij}$.

5.1.2 Results and discussion

5.1.2.1 Milk production and SCS

Random regression models used to investigate genotype-by-THI interactions allowed us to estimate additive genetic variances, heritabilities and genetic correlations across all combinations of DIM and THI_3day for each trait within each country. Figure 7 shows the estimated genetic standard deviations across THI_3day at 150 DIM (that is, in the middle of the lactation curve) for MY, FY, PY and SCS for first-parity cows in France, Spain and the Netherlands. The trends were similar across countries and breeds for the different production traits and SCS, even if some differences in terms of levels and shapes of genetic standard deviation could be observed between countries and breeds. Briefly, additive genetic standard deviations were the largest for the production traits for values of THI_3day lower than about 50, and decreased with increasing values of THI_3day. For SCS, they had a different trend, because lowest additive genetic variances are observed at the lowest THI_3day values (i.e., THI_3day<50), and increased with increasing values of THI_3day. Similar trends could be observed for the second-parity cows (see Appendix 5 Figure 5.1 for more details).

The evolution of heritability estimates with THI_3day are shown in Figure 8 for milk production traits and SCS in first-parity cows for all countries and breeds. The heritabilities for production traits were quite stable along the THI_3day gradient, with a slight tendency of decreasing with higher values of THI_3day. Only first-parity MRY cows showed an increasing trend for higher values of THI_3day for FY and PY. For SCS, the trend in heritability along the THI_3day gradient depicted an increased pattern with increasing THI_3day values.

The evolution of estimated genetic correlations between different THI_3day values are shown in Figures 9 and 10 for production traits and SCS in first-parity cows for all countries and breeds. Briefly, most of them between different THI_3day values, within trait, parity and DIM, were higher than 0.90, and none of them were lower than 0.75. Estimated genetic correlations lower than 0.90 were observed for THI_3day values higher than 70 with a distant THI_3day value, e.g., a THI_3day of 50 as shown for PY in Figure 9. The lowest genetic correlations, around 0.80, were obtained for PY in Spain, in both L1 and L2. For all traits, the estimated genetic correlations between the thermo-neutral THI of 50 and the heat stress THI of 70 were lower in L2 than in L1, and in most cases, reached the lowest values just after the lactation peak (i.e., after 60 DIM) and tended to increase with DIM (results not shown).

Table 4 shows the rank correlations of the top 100 Holstein bulls on EBVs estimated for levels on MY, FY, PY and SCC under thermo-neutral conditions (THI=50) and under heat stress conditions (that is, THI=65 in France, THI=68 in the Netherlands and THI=75 in Spain). The rank correlations for the top 100 bulls for levels of each trait are high, but different than the unity, which means that some re-ranking are expected between the best AI sires.

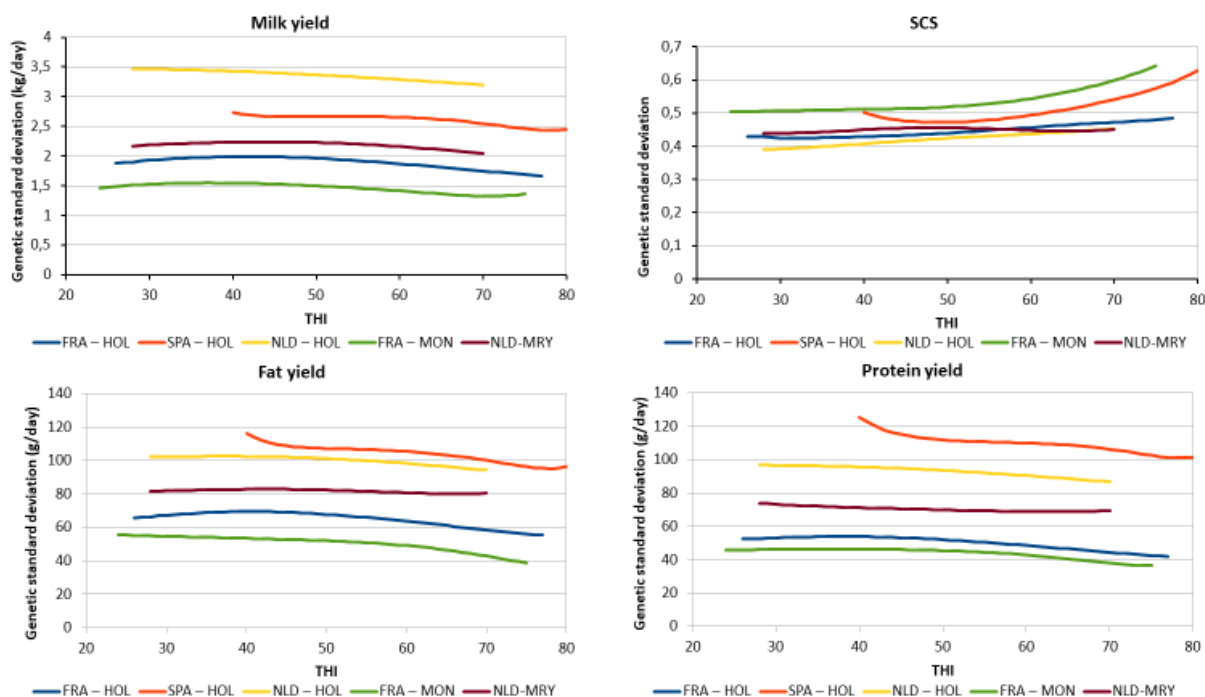


Figure 7. Evolution of additive genetic standard deviations for milk yield, fat yield, protein yield, and somatic cell score (SCS) with 3-day average temperature-humidity index (THI) in first-parity cows, according to the country (FRA=France; NLD = the Netherlands; SPA = Spain) and to the breed (HOL=Holstein; MON=Montbéliarde, MRY=Meuse-Rhine-Yssel). Results are given at 150 days-in-milk.

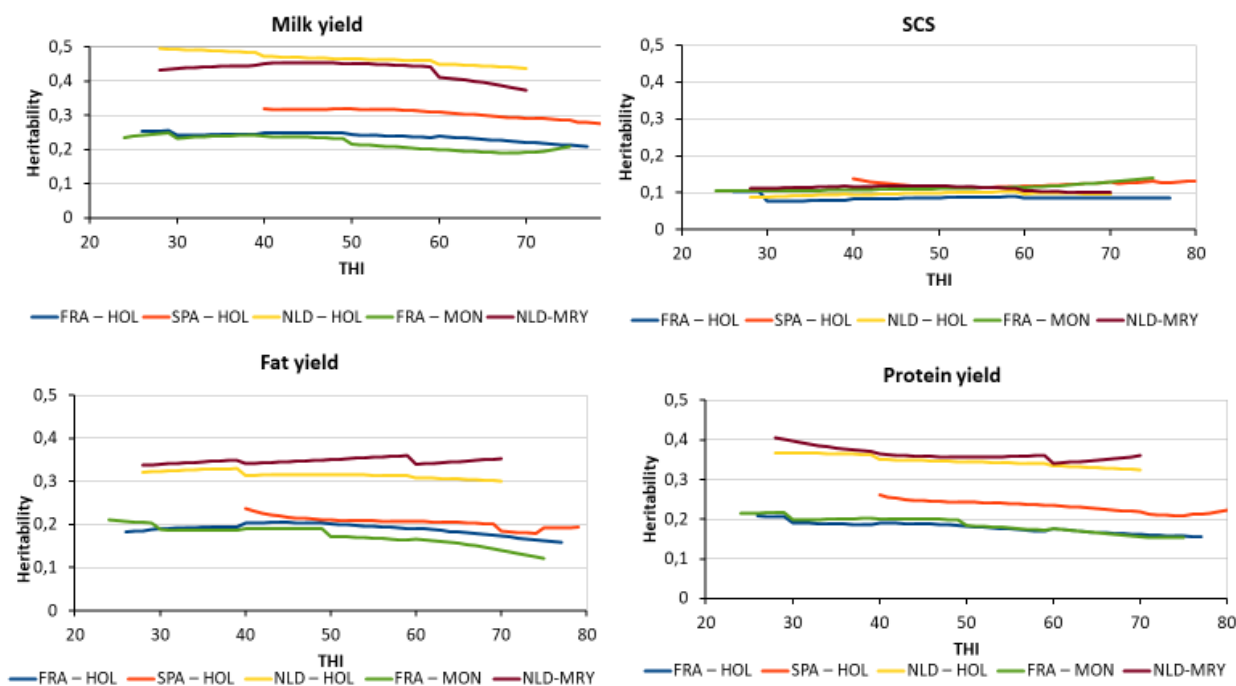


Figure 8. Evolution of heritability estimates for milk yield, fat yield, protein yield, and somatic cell score (SCS) with 3-day average temperature-humidity index (THI) in first-parity cows, according to the country (FRA=France; NLD=the Netherlands; SPA=Spain) and to the breed (HOL=Holstein; MON=Montbéliarde, MRY=Meuse-Rhine-Yssel). Results are given at 150 days-in-milk.

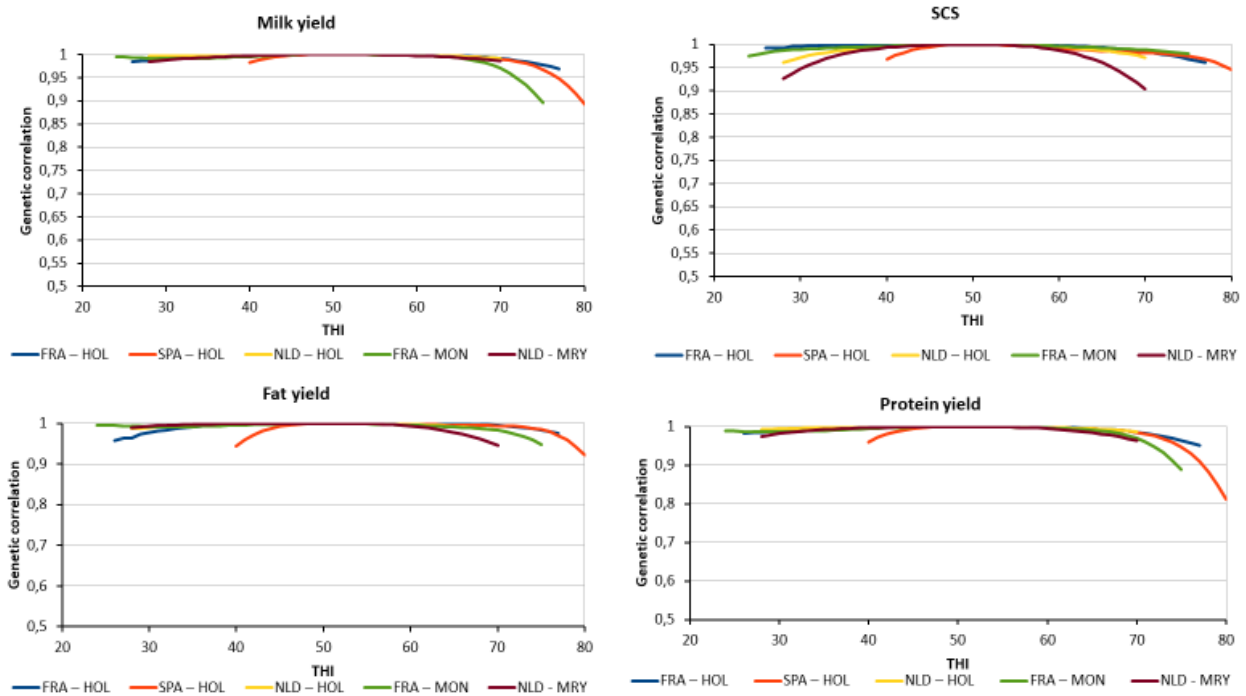


Figure 9. Evolution of genetic correlation estimates for milk yield, fat yield, protein yield, and somatic cell score (SCS) between a 3-day average temperature-humidity index (THI) of 50 and other THI values in first-parity cows, according to the country (FRA=France; NLD=the Netherlands; SPA=Spain) and to the breed (HOL=Holstein; MON=Montbéliarde, MRY=Meuse-Rhine-Yssel). Results are given at 150 days-in-milk.

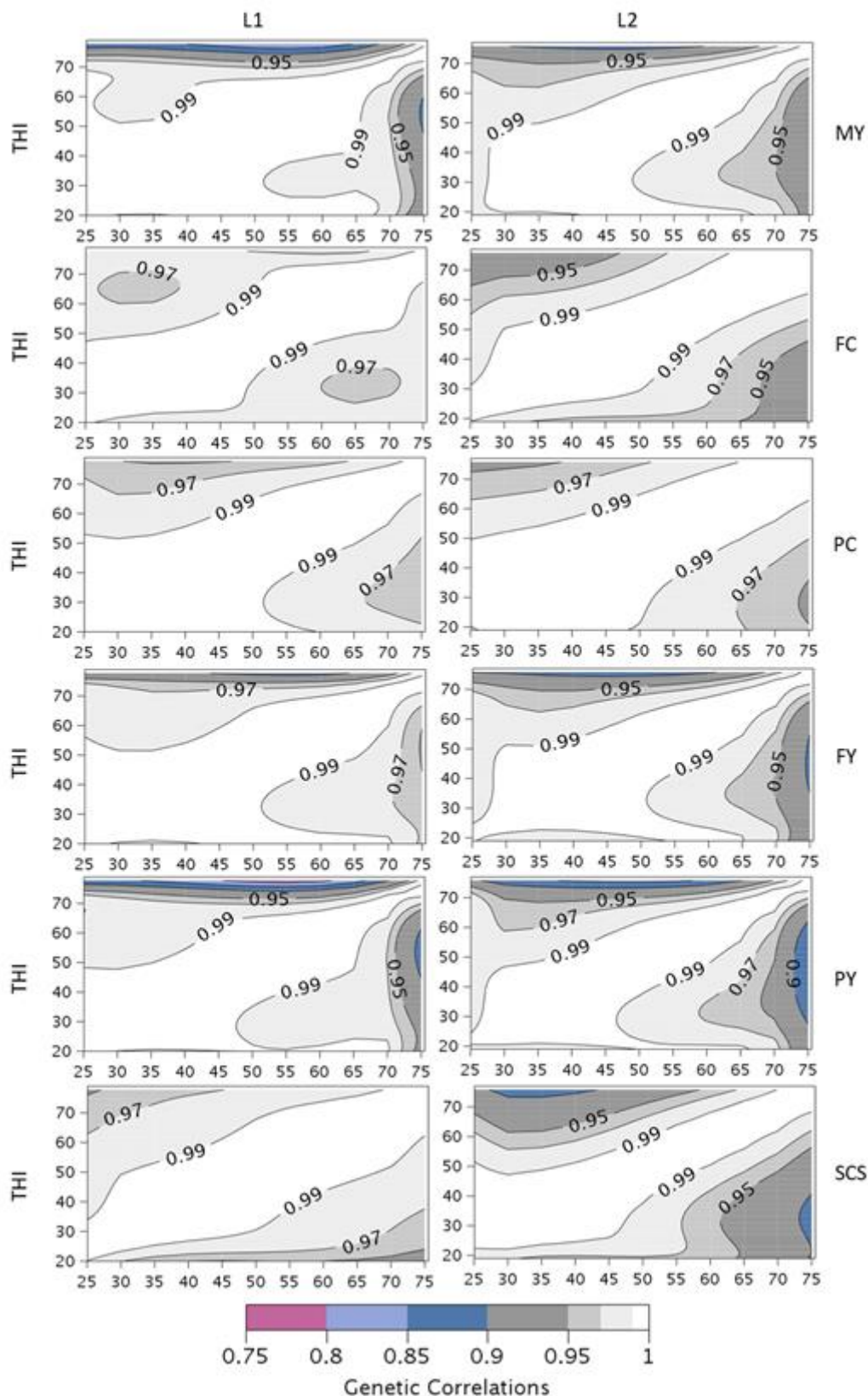


Figure 10. Estimated genetic correlations within trait at different temperature-humidity index (THI) for milk yield (MY), somatic cell score (SCS), fat and protein contents (FC and PC) and fat and protein yields (FY and PY) for first (L1) and second lactation (L2) French Montbéliarde cows. Results are given at 150 days-in-milk.

Table 4. Rank correlations of top 100 Holstein sires with at least 20 daughters with records during the first parity based on estimated breeding values at 150 days-in-milk and a temperature-humidity index (THI) equal to 50 and at a heat stress THI.

Country	Heat stress THI	Milk yield	Fat yield	Protein yield	SCS
France	65	0.89	0.87	0.90	0.92
The Netherlands	68	0.86	0.86	0.81	0.97
Spain	75	0.88	0.98	0.92	0.97

5.1.2.2 Conception rate

Figure 11 shows the estimated genetic standard deviations along the gradient of THI_8day for CR for first-parity cows in France, Spain and the Netherlands. Even if some differences in terms of levels of genetic standard deviation can be observed between countries and breeds, the trends were similar across countries for CR. Briefly, additive genetic standard deviations increased with increasing values of THI_8day from the value of THI_8day defined as threshold (i.e., 62 for France, 66 for Spain, and 60 for the Netherlands) and the heritability estimates followed the same shape as the genetic variances (Figure 12). However, in the case of France, the reaction norm model with Legendre polynomials was preferred to the broken line model, that inflated too much the genetic variances at high THI. Therefore, the French EBVs used in next studies were calculated using a reaction norm model.

The evolution of estimated genetic correlations at different THI_8day values are shown in Figure 13 for CR in first-parity cows for all countries. Briefly, for all countries, genetic correlations between thermo-neutral conditions (e.g., at a THI_8day of 50) and THI_8day higher than the change point decreased with increasing values of THI_8day. However, differences in decays could be observed across countries with the steepest decay for first-parity Holstein cows in Spain and the lowest one for the Netherlands. It can be also observed in France and Spain that genetic correlations were lower than 0.50 between the thermo-neutral conditions and a THI_8day higher than 75. Even at lower THI_8day, moderate genetic correlations with thermo-neutral conditions could be observed, especially for the Holstein populations. For example, genetic correlations between a THI_8day of 50 and of 70 varied between 0.72 for France and 0.83 for Spain and the Netherlands.

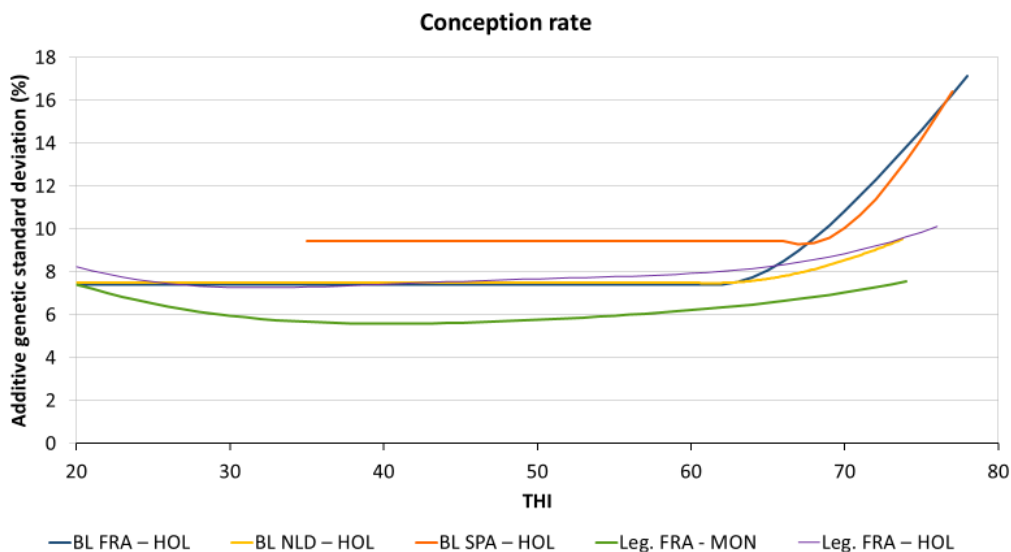


Figure 11. Evolution of additive genetic standard deviations for conception rate (CR, in %) with 8-day averaged temperature-humidity index (THI) in first-parity cows, according to the country (FRA=France; NLD=the Netherlands; SPA=Spain), to the breed (HOL=Holstein; MON=Montbéliarde), and to the model (BL=Broken line model; Leg. = reaction norm model with Legendre polynomials).

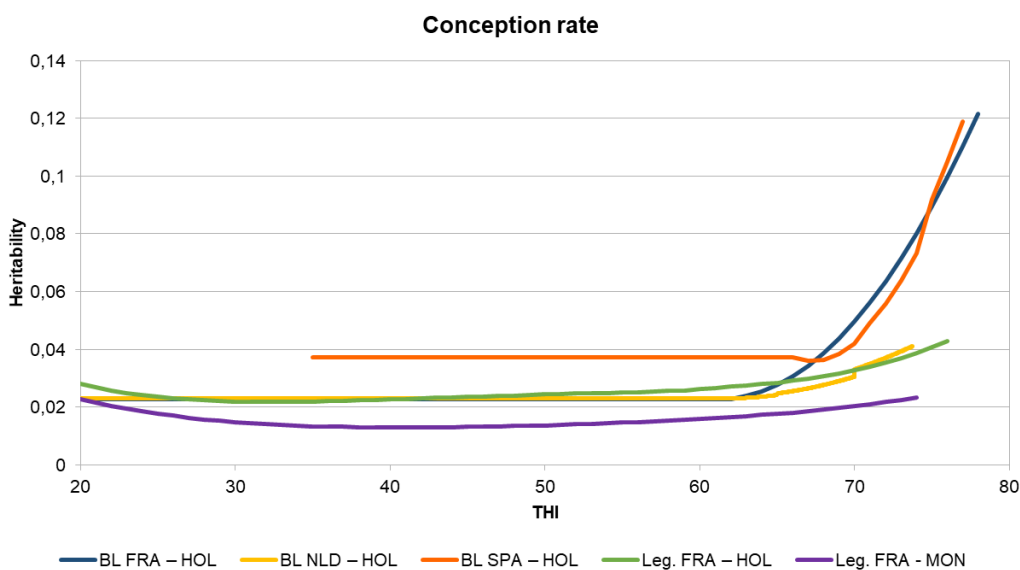


Figure 12. Evolution of heritability for conception rate (CR, in %) with 8-day averaged temperature-humidity index (THI) in first-parity cows, according to the country (FRA=France; NLD=the Netherlands; SPA=Spain), to the breed (HOL=Holstein; MON=Montbéliarde), and to the model (BL=Broken line model; Leg.=reaction norm model with Legendre polynomials).

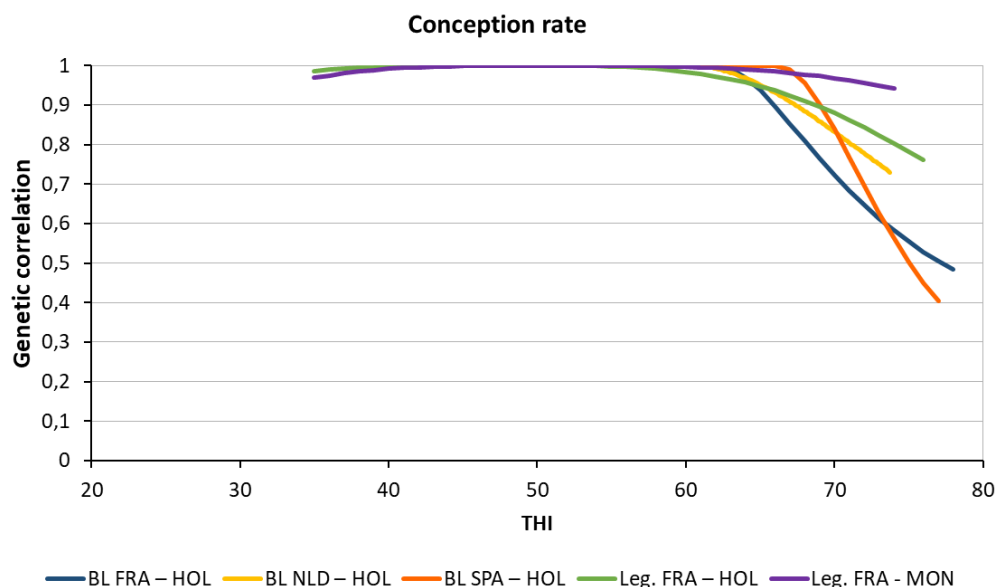


Figure 13. Evolution of genetic correlations for conception rate (CR) between a 8-day averaged temperature-humidity index (THI) of 50 and other THI values in first-parity cows, according to the country (FRA=France; NLD=the Netherlands; SPA=Spain) and to the breed (HOL=Holstein; MON=Montbéliarde).

5.2 Meta-analyses: genetic correlations among countries

5.2.1 Methods

The aim of these meta-analyses was to estimate genetic correlations among countries for the Holstein breed for different traits related to heat tolerance. The phenotypes analysed were deregressed proofs (DRPs) of sires. The DRPs are an estimation of mean performances of daughters corrected for all important environmental effects in each country and for half of the genetic value of their dam. Each DRP was associated with equivalent record contributions (ERCs) as an indicator of its accuracy.

First, EBVs and associated reliabilities were calculated at the national level, using the national data described in Appendix 2 and the models described in Section 5.1. Then the DRPs and ERCs were derived from the national EBVs and reliabilities, using a one-bull-at-a-time deregression approach (VanRaden et al., 2014). This approach assumes that EBVs were obtained with a single-trait model. The heritabilities used for the deregression were based on results of Section 5.1 for the DRPs on levels. For the slopes, heritability is not available because the residual variance cannot be estimated. Therefore, heritabilities of slopes were assumed to be equal to 0.10 for MY, FY and PY, 0.03 for SCS and 0.01 for CR.

Considered traits were levels under thermo-neutral and heat stress conditions, and slope traits under heat stress conditions. As above, MY, FY, PY, SCS and CR were analyzed. The study of levels at heat stress conditions was limited to milk production and SCS. Thermo-neutral conditions were defined as being equal to a THI of 50 for all countries, while heat stress conditions were defined specifically for each country (THI equal to 65, 68, and 77, respectively for France, the Netherlands, and Spain). These values were defined based on the results obtained from the analyses at the population level (see Section 4 for more details). Preliminary studies were performed using higher reference THI levels for France and the Netherlands; however as the results were almost the same,

these values of THI were kept as reference, because less accurate pseudo-phenotypes could be obtained at higher THI due to a smaller amount of performances. The level traits were computed at 150 DIM, while the slope traits were defined as the first derivative of the level at heat stress conditions.

For each country and for each trait, the analyses included DRPs and ERCs of all Holstein bulls with at least 20 daughters with performances and a reliability of at least 0.25. It is worth noting that no selection was performed regarding the number of daughters with performances in heat stress conditions. The pedigree of each bull was traced back on 3 generations.

Based on national DRPs and ERCs, genetic correlations between countries were estimated using the Multiple Across Country Evaluation (MACE) approach (Schaeffer, 1994). In this study, it consisted in a pedigree-based animal model with each country considered as a separate trait. The model used was a multiple trait model, as follows:

$$\mathbf{y} = \mathbf{Xc} + \mathbf{Za} + \mathbf{e}$$

where \mathbf{y} is the vector of DRPs from each country, \mathbf{c} is the vector of country of evaluation fixed effects, \mathbf{a} is the vector of random additive genetic effects in all participating countries, and \mathbf{e} is the vector of residual effects. It is assumed that $\text{Var}(\mathbf{a}) = \mathbf{G}_0 \otimes \mathbf{A}$ and $\text{Var}(\mathbf{e}) = \mathbf{R}_0 \otimes \mathbf{D}$, where \mathbf{G}_0 and \mathbf{R}_0 are the genetic and residual matrices of (co)variances between countries, \mathbf{A} is the pedigree-based relationship matrix and \mathbf{D} is a diagonal matrix with diagonal elements corresponding to the inverse of ERCs. The matrices \mathbf{X} and \mathbf{Z} are incidence matrices that relate phenotypes to the corresponding effects.

The software used to estimate variance components and to compute DRPs and ERCs are described in Appendix 3. It is worth noting that no genetic group was included for unknown parents, neither for the deregression, nor for the estimation of genetic parameters.

5.2.2 Results and discussion

5.2.2.1 Genetic correlations among countries for level traits

Figure 14 shows the number of sires associated with pseudo-information (DRPs and ERCs) in each country for milk yield (levels) and SCS traits (slopes) in first-parity Holstein populations. Briefly, for levels on milk yield (and similarly for other milk production traits; results not shown), 7932 French sires, 3624 Spanish sires and 2281 Dutch sires had an EBV with a reliability of at least 0.25 and associated to at least 20 daughters with recorded milk performances. Among them, the French population shared 1022 and 478 common sires with the Spanish and Dutch populations, respectively, for the milk production traits and SCS. The Spanish and Dutch populations shared 463 sires. In total, 328 sires were common to all three countries. Only 75 Spanish bulls did not reach a reliability of 0.25 for milk production traits and were removed from the sample for this reason. None of these 75 bulls were common with the other countries. The same situation was observed on the levels on SCS and on the slopes of milk production traits. For the slopes on SCS, respectively 1763, 605 and 1735 of French, Dutch and Spanish bulls were removed because of a too low reliability (Figure 14). However, the number of common bulls remained high, with 222 common bulls between the three countries and 813 bulls having more than 20 daughters in at least two countries.

The elimination of bulls due to too low reliability had much more impact on CR, as shown in Figure 15. For the levels on CR, links between countries remained strong enough to estimate genetic parameters, with 275 common bulls between the three countries and 1037 having pseudo-information in at least two countries. For the slopes on CR, the situation was different, as the heritability of these traits was very low (about 0.01) and as the number of cows with performances under heat stress was limited, resulting in EBVs with very low reliabilities. Therefore, after edits, only

29 Dutch bulls were associated with an EBV with a reliability of at least 0.25, and were therefore not considered for further meta-analyses on CR slopes. However, after edits, 485 French bulls, 354 Spanish bulls, and 102 common bulls between France and Spain satisfied all criteria, and these numbers were high enough to estimate the genetic correlations between France and Spain (Figure 15).

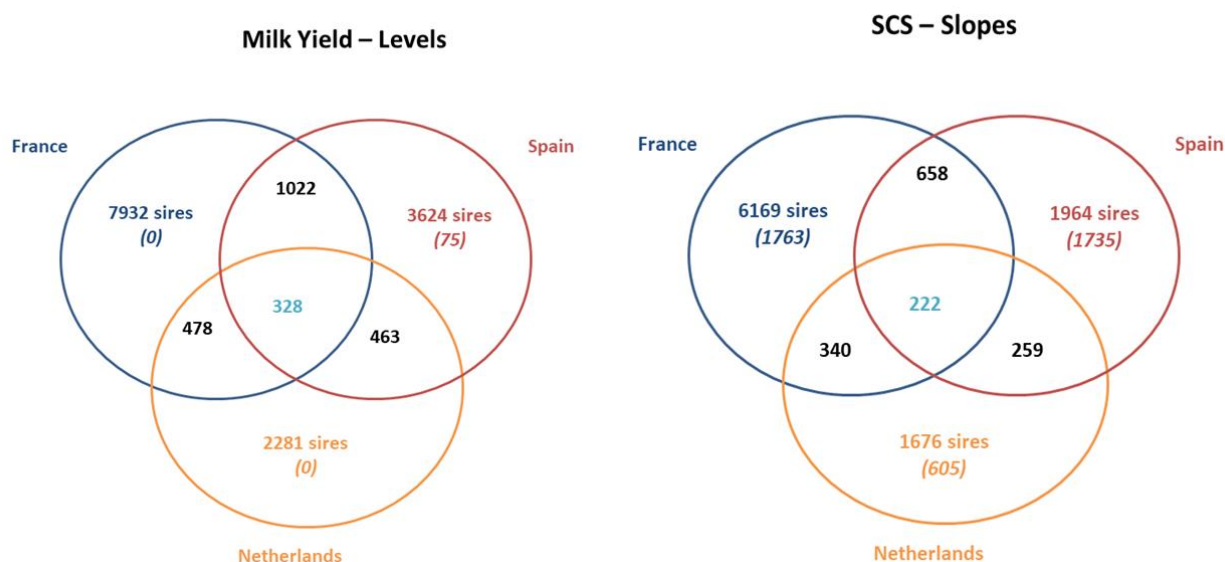


Figure 14. Numbers of common sires among France, Spain and the Netherlands for levels on milk yield and for slopes on somatic cell score (SCS) and used to estimate the genetic correlations across countries. Numbers of bulls removed due to too low reliability are within brackets.

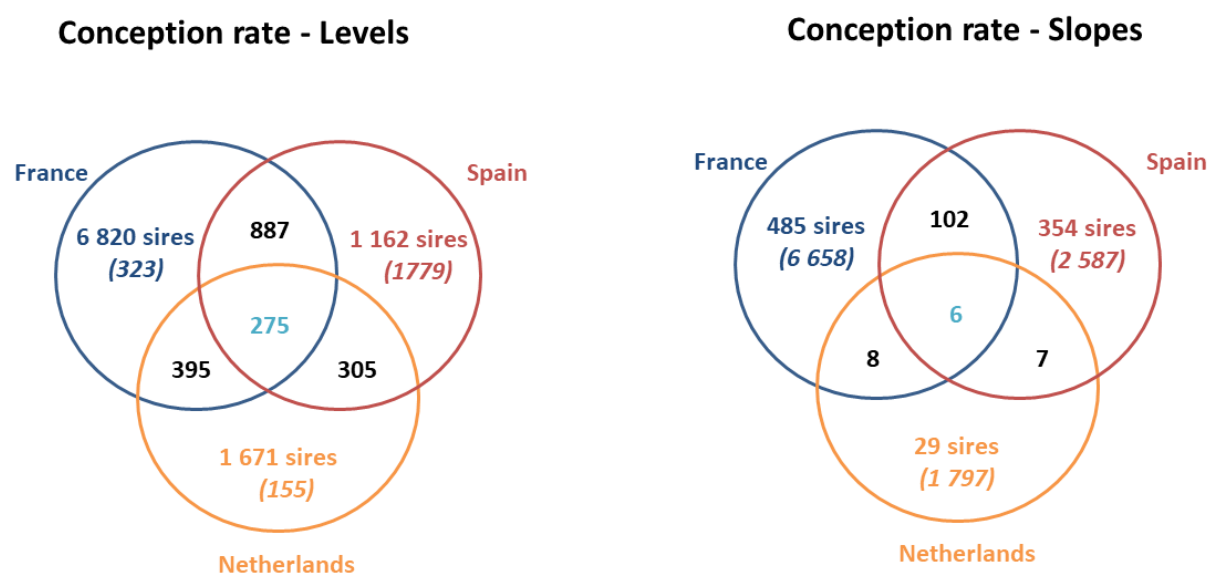


Figure 15. Numbers of common sires among France, Spain and the Netherlands for levels and for slopes on conception rate and used to estimate the genetic correlations across countries. Numbers of bulls removed due to too low reliability are within brackets.

Tables 5 and 6 present estimated heritabilities and genetic correlations among countries for the different level traits under thermo-neutral and heat stress conditions. Estimated genetic correlations among countries for all level traits were high for production traits and for SCS, and for both thermo-neutral and for heat stress conditions. In thermo-neutral conditions, estimated genetic correlations were between 0.89 and 0.97. They were close to those estimated by Interbull in the current MACE evaluations for Holstein bulls and even slightly higher for production traits (see Appendix 6 for more details). Estimated genetic correlations among the three countries were moderate to high for CR (i.e., between 0.63 and 0.87; Table 6). The estimated genetic correlation between France and the Netherlands (0.77) was lower than the corresponding genetic correlation provided by Interbull (0.92). No genetic correlation estimated by Interbull was available between Spain and the two other countries because Spain does not participate to the current Interbull evaluations within the same group of fertility traits as France and the Netherlands. Moreover, the estimated genetic correlations between countries under heat stress conditions were slightly lower than under thermo-neutral conditions, but still high for milk production and SCS traits (between 0.83 and 0.97). This shows a very good consistency of results of genetic evaluations on levels of milk production traits and of SCS under heat stress conditions. Lastly, the estimated heritabilities were lower than those estimated at the national level, particularly for Spain and the Netherlands, except in the case of SCS for France, where the heritability estimated with MACE was surprisingly high (0.34; Table 5).

Table 5. Estimated heritabilities (diagonal) and genetic correlations (off-diagonal) among countries for the level traits under thermo-neutral and heat stress conditions for milk production and udder traits. Standard deviations are within brackets.

Trait ¹	Country ²	Thermo-neutral conditions			Heat stress conditions		
		FRA	SPA	NLD	FRA	SPA	NLD
MY	FRA	0.18 (<0.01)	0.96 (0.01)	0.94 (0.02)	0.18 (<0.01)	0.92 (0.01)	0.89 (0.02)
	SPA		0.19 (<0.01)	0.92 (0.02)		0.25 (0.01)	0.86 (0.02)
	NLD			0.25 (0.02)			0.25 (0.02)
FY	FRA	0.25 (<0.01)	0.97 (0.01)	0.90 (0.02)	0.25 (<0.01)	0.97 (0.02)	0.87 (0.03)
	SPA		0.13 (<0.01)	0.89 (0.02)		0.13 (<0.01)	0.85 (0.04)
	NLD			0.21 (0.01)			0.20 (0.02)
PY	FRA	0.15 (<0.01)	0.96 (0.01)	0.89 (0.02)	0.17 (<0.01)	0.89 (0.01)	0.81 (0.02)
	SPA		0.12 (<0.01)	0.90 (0.02)		0.17 (<0.01)	0.83 (0.02)
	NLD			0.20 (0.01)			0.21 (0.01)
SCS	FRA	0.34 (0.01)	0.95 (0.02)	0.88 (0.02)	0.30 (<0.01)	0.96 (0.02)	0.88 (0.02)
	SPA		0.08 (<0.01)	0.89 (0.03)		0.09 (<0.01)	0.88 (0.03)
	NLD			0.15 (<0.01)			0.16 (<0.01)

¹MY = Milk yield; FY = Fat yield; PY = Protein yield; SCS = Somatic cell score.

²FRA = France; SPA = Spain; NLD = the Netherlands.

Table 6. Estimated heritabilities (diagonal) and genetic correlations (off-diagonal) among countries for the level traits under thermo-neutral conditions for conception rate. Standard deviations are within brackets.

Country ¹	Thermo-neutral conditions		
	FRA	SPA	NLD
FRA	0.02 (<0.01)	0.87 (0.03)	0.77 (0.05)
SPA		0.02 (<0.01)	0.63 (0.07)
NLD			0.04 (<0.01)

¹FRA = France; SPA = Spain; NLD = the Netherlands.

5.2.2.2 Genetic correlations among countries for slope traits

Genetic parameters and genetic correlations among countries for slopes on milk production traits and SCS could not be estimated accurately and were therefore not presented in this report. First, heritabilities for slope traits on milk production estimated with the MACE approach ranged from 0.04 (for PY in the Netherlands) to 0.93 (for FY in France) across countries and traits, resulting in unreliable results. A similar range was observed on SCS, with heritabilities varying from 0.003 (for the Netherlands) to 0.90 (for Spain). Further investigations showed that in many cases, the residual variances seemed to be strongly under-estimated, leading to an under-estimation of the phenotypic variance and thus to an over-estimation of the heritability. Second, estimated genetic correlations for milk production traits and SCS were low and unexpected, with estimates varying between 0.04 and 0.53 for MY, -0.17 and 0.09 for FY, 0.03 and 0.30 for PY, and between -0.87 and 0.24 for SCS. These estimates would suggest that the traits are completely different from one country to another, and even in strong opposition in the case of SCS between Spain and the Netherlands. All these results are not credible, as the traits were defined using the same approach in the 3 countries. Surprisingly, the result obtained on the slope for CR between France and Spain seemed to be much more realistic, with heritabilities close to their expectations (0.02 for France and 0.05 for Spain), and with an estimated genetic correlation between the two countries of 0.84.

5.2.2.3 Discussion

For level traits, the estimated genetic correlations among countries for milk production traits and SCS under thermo-neutral and heat stress conditions were higher than 0.85 and consistent with those obtained by Interbull with the same approach (Appendix 6). These results validated the approach used to estimate the correlations, even if some variations could be observed.

A slight decrease in estimated genetic correlations was observed for level traits under heat stress for milk production traits and SCS. Several non-exclusive hypotheses could explain these results. First, as mentioned in Section 4, the main farming conditions of each country are not the same. If the environment is defined by these farming conditions, the slight drop in genetic correlations between thermo-neutral and heat stress conditions could be interpreted from a physiological point of view as a different impact of heat stress on gene expression according to the environment. The

second hypothesis is connected to the fact that a large part of the performances used to estimate variance components and breeding values were recorded under thermo-neutral conditions (i.e. at moderate THI). Therefore, the slight drops in estimated genetic correlations could be due to less accurate national EBVs used in the MACE analysis. At this stage, the latter seems the most likely hypothesis, as the genetic correlations for levels within country were all very close to 1 (see Section 5.1).

The estimated genetic correlations for slopes for production traits and SCS were disappointing. Most of them were very low and some of them were even strongly negative, which is not credible from a biological point of view. Some additional analyses were carried out at different THI in order to look at the impact of the THI chosen for the slope on the genetic parameters. The variances varied a lot at different THI within countries; in particular the residual variances tended to increase a lot when THI increased. For example, the residual variance for MY increased from 0.9 to 6.9 (kg/THI)² with French data when the chosen THI changed from 60 to 70, and from 4.9 to 8.6 (kg/THI)² with Dutch data when the chosen THI changed from 68 to 70. These results indicate that our approach (MACE correlations based on pseudo-information) was not able to capture the genetic (co)variability of slopes across countries.

These issues for estimating genetic correlations among countries for slope traits may be explained by several reasons related to the use of DRPs as dependant variable for the MACE procedure. First, experiences in deregression showed that very low accuracies adversely affect the quality of DRPs, with very extreme values of some DRPs. In practice, Interbull Centre solves this problem by including only bulls with a higher minimum number of daughters with performances in the Interbull evaluations for low heritability traits than for the others (e.g., 50 daughters for mastitis traits and for maternal calving traits, instead of 10 daughters for the other traits; www.interbull.org). Similarly in our analyses, we required that all bulls included in the meta-analyses should have at least 20 daughters with performances and a reliability of at least 25%. However, these requirements were probably not strong enough for the deregression of EBVs on slopes. Second, the heritability of the slope was not estimated, and guessed and reasonable values were considered. However, these values can be different (probably overestimated) from their true values. Heritability plays an important role when the reliability of the proofs is low and our strategy may have affected the results. Third, the deregression process based on a single-trait approach might not be appropriate for slope traits. Indeed, when EBVs are calculated using multi-trait models, the consequences are limited until the amount of information brought by predictors (i.e. correlated traits) is much less important than the amount of information brought by the studied trait itself. In our case, the heritability of the slopes is low and only very few sires had many daughters with at least one record measured in heat stress conditions because of the low frequency of days above heat stress and also because of the infrequent (monthly) milk recording. Therefore, most of the EBVs on slopes were estimated indirectly from the levels at lower THI, making the deregression and the estimation of standard errors very difficult. (Bohmanova et al., 2008) were confronted to the same issue when they compared heat stress EBVs of sires evaluated in North-eastern and South-eastern regions of the USA. They indicated that EBVs estimated in both regions would be consistent, but only for sires having a very large number of daughters (about 1000 in each region). Finally, an additional potential issue for estimating genetic correlations among countries for slope traits is that these traits were defined as the first derivative of the standardized Legendre polynomials at a given THI. Such values might be too sensitive to small errors in the estimation of the rate of decline of the level under heat-stress conditions.

In conclusion, the meta-analyses showed a very good consistency of the traits based on levels and thus even at high THI. This is very encouraging and re-enforces all the other results summarized in this report regarding these traits. Today, there are still few performances in heat stress conditions in

the Northern part of Europe. Moreover, the results in Section 5.1 showed that even if very few genetic-by-THI interactions exist on levels on milk production traits and SCS, reranking of the bulls with the highest EBVs can be observed between thermo-neutral and heat stress conditions. These re-rankings highlight the importance to prepare new breeding tools adapted to global warming, even in the countries currently less impacted by heat stress. Our results show also that valuable predictions under heat stress conditions could be obtained through international evaluations, that would result in Northern countries benefiting from the information that already exists in the Southern countries. However, these meta-analyses showed also the limits of the studies on low heritable traits such as the slopes, if there are not enough data at high THI. In our study the DRPs on slopes, as they were calculated, were not able to capture the genetic (co)variability associated to these traits. Therefore, other models and approaches should be investigated to better measure the decline in performances due to heat stress and to take advantage of the information available in multiple countries.

6 Heat tolerance at a genome level

6.1 Material and methods

The aim of these studies was to identify genomic regions underlining and explaining the genetic variation of heat tolerance through genome-wide analysis studies (GWAS). The GWAS analyses were performed for milk production, udder health and fertility traits for the Holstein breed in France, Spain and the Netherlands. Within each country, these analyses used the same phenotypes as for the MACE studies at the individual level (see Section 5 for more details), that is DRPs and ERCs for the level and slope traits at high THI. The THI level was defined in each country as follow: levels at THI 50 and slopes at THI 65 for France, THI 68 for the Netherlands and THI 75 for Spain. Only DRPs with a reliability higher or equal than 0.25 were used for GWAS. The number of genotyped sires with pseudo-phenotypes available varied across countries and traits. Table 7 shows the number of genotyped sires kept for GWAS in each country.

France has more than three times data than Spain and The Netherland. The number of French sires involved in the analyses varied between 5435 for levels and production traits and 1692 for the slope of SCS. In the Spanish case, figures varied between 1667 for levels of production traits and 188 for the slope of conception rate. Finally number of sires for the Netherlands varied between 1332 for level and slope of production traits and 1004 for level of conception rate. The slope for CR was not considered for GWAS in the Netherlands because only 22 genotyped sires with pseudo-phenotypes could be identified for the CR slope trait. Genotypes in France, Spain and the Netherlands included 53469 SNPs, 46277 SNPS, and 76313 SNPs, respectively. The cow (*Bos taurus*) genome assembly ARS-UCD1.2 (Medrano, 2017; Rosen et al., 2018) was used for the different analyses in the three countries.

The model used for GWAS was the same in all countries and consisted of a mixed linear model that included an overall mean, the effect of the tested SNP and a polygenic effect to account for population stratification and familial relationships. The polygenic effects were estimated based on the genomic relationship matrix computed out of all SNPs for each country. The analyses were performed with the software GCTA (Yang et al., 2011) modified for supporting pseudo-phenotypes in form of DRPs and ERCs. In the first place, nominal p -values were adjusted by multiple testing using Benjamini and Hochsberg (1995). Positional candidate genes were those located in a window of 500×10^3 base pairs around each significant SNP. They were found using GALLO R procedure (Fonseca et al., 2020). We also found some putative genes near significant SNPs according to a Bonferroni adjustment (α/N with N being the number of SNPs available in each country after quality

control and $\alpha=0.05$ the type I error). Thus, the thresholds for each country were $9.35 * 10^{-7}$ for France, $10.8 * 10^{-6}$ for Spain, and $6.55 * 10^{-7}$ for the Netherlands. Thresholds differ for the same α , because the number of tests or SNP differ from country to country.

Table 7. Number of sires with genotype and pseudo-phenotype available for genome-wide association analysis per first-parity trait and country.

Country	Traits ^a	MY	FY	PY	CR	SCS
FRA	L50	5435	5435	5435	4803	5435
	S65	5293	5293	5293	-	1692
SPA	L50	1667	1543	1667	294	1608
	S75	1543	1490	1523	188	942
NLD	L50	1332	1330	1330	1004	1330
	S68	1332	1330	1330	-	1015

^aL:level; S=slope; MY:milk yield; FY=fat yield; PY:protein yield; CR:conception rate; SCS: somatic cell score.

6.2 Results and discussion

Manhattan plots for the results of the GWAS analyses on the levels and slopes for production, udder health and fertility traits in each country are shown in Appendix 7. The number of significant SNPs at a false discovery rate (FDR) of 0.05 (Benjamini and Hosberg, 1995), of positional candidate genes as well as the putative genes considering a Bonferroni approach for fat yield, milk yield, protein yield are shown in Tables 8, 9 and 10, respectively. For CR, only five significant SNPs were identified in the French population (results not shown here). The windows around those SNPs contained 41 positional candidate genes. Provided the much larger amount of French data (Table 7), it is not that surprising that results are more numerous in terms of number of candidate regions, number of SNPs, number of positional and putative number of genes for the French population, because it is expected that the power of detection in the French data set is much higher than for the other two data sets. On the other hand, since the analysis has been run independently and since the Spanish population has been exposed to more heat stress, common results to more than one population provide more evidence to the obtained results in the populations involved.

Table 8. Number of significant SNPs for level (L) and slopes (S) associated to fat yield (FY), number of positional candidate genes as well as putative candidate genes based of a Bonferroni adjustment approach.

Country	Pseudo-phenotype ^a	Number of SNP ^b	Number of positional candidate genes ^c	Putative candidates genes ^d
FRA	L50	75	171	CCNT1, ABCC9*, DGAT1, CSN1S2*, VPS13B*, GHR*
	S65	54	191	ABCC9*, CSN1S2*, DGAT1*, HSF1*, VPS13B*, GHR*
SPA	L50	69	150	DGAT1*
	S75	33	70	HSF1*
NLD	L50	15	92	DGAT1*
	S68	-	-	-

^aL:level and S=slope at 65, 75 and 68 for France, Spain and the Netherlands, respectively;

^b padj <0.05 p-values adjusted by Benjamini and Hochberg procedure;

^c Number of candidates genes in a 0.5 Mb window around each significant SNP (padj <0.05);

^d Putative genes under Bonferroni significant level (0.05/N with N being the number of SNPs).

*Already found in the literature.

The number of significant SNPs as well as of positional candidate genes varied according to the country and trait. In general, we have found the least number of SNPs associated to slopes, as we expected. Indeed, the power of detection for slopes is smaller than for levels given that variability associated to slopes is smaller than that associated to levels. In general, the most interesting results are related to the identification of potential QTLs of pleiotropy. In our study, to be considered as indicating a region of QTLs of pleiotropy, a SNP should have a significant estimated effect both on levels of production as well as slopes of production and this in at least two of the three populations. The most significant region appears in BTA14. It is a QTL common among LFY, LMY and LPY in the three populations in a window that goes from 483971bp to 700772bp. This region provides a signal in the slopes for fat and milk yields in the French population, for fat yield in the Spanish population and for milk and protein yield in the Dutch population. Nevertheless, no SNP in BTA14 were found associated with the slope for protein yield (S65PY) in the French population. In that specific window, we can find DGAT1 as well as HSF1 as putative genes for levels and slopes, respectively. These genes have been already identified in a pleiotropic QTL involved on the level of milk production traits (Tribout et al., 2020; Cheruiyot et al., 2020; Teng et al., 2023) as well as on slopes of decays of milk production traits (Cheruiyot et al., 2020). In the panel of SNPs used to perform these GWAS analysis, there are several SNPs located within each gene (2 in DGAT1 and 4 in HSF1). Those SNPs are significant for all traits under the Bonferroni adjustment and in very high LD. Thus, assessing true pleiotropy or separate causality between DGAT1 and HSF1 is really challenging. Interestingly, DGAT1 and HSF1 are always found together in mammals what might indicate a selective advantage. On the other hand, it is well known that HSF1 is a transcription factor

that gets activated under stress inducing the transcriptions of HSP to protect cells from endogenous and exogenous stress. This is the reason why it seems a clear candidate gene for slopes provided that they are indicators of heat stress. However, because HSF1 also plays a relevant role in many physiological pathways under distinct physiological conditions (Vihervaara and Sistonen, 2014) we could hypothesize that its role on the slope of decay is more related through the impact of heat stress on the metabolism of the animal rather than the stress in itself. We cannot forget that our heat stress indicators measured the impact of heat stress on production, udder health and conception rate rather than measuring the physiological stress in itself. Among the pathways where HSF1 has a role, we can find the “Insuline signalling path” which is very affected by heat stress in dairy cows (Baumgard and Rhoads, 2013).

Tables 8, 9 and 10 also show other putative candidate genes associated with production levels and slopes of milk production traits, such as MGST1, ABCC9, CSN1S2, or GHR. All of them have been associated to level of production traits in the literature (e.g., Tribout et al, 2020; Teng et al., 2023) as well as slopes of decay related to milk production traits (Cheruiyot et al., 2020; 2022). Even though seven SNPs were significantly associated to CR, we could not identify any clear putative candidate genes so far.

Table 9 Number of significant SNPs for level (L) and slopes (S) associated to milk yield (MY number of positional candidate genes as well as putative candidate genes based of a Bonferroni adjustment approach.

Country	Pseudo-phenotype ^a	Number of SNP ^b	Number of positional candidate genes ^c	Putative candidates genes ^d
FRA	L50	71	201	CCNT1, ABCC9*, DGAT1*, CSN1S2*, VPS13B*, GHR*,HSF1
	S65	37	103	ABCC9*, CSN1S2*, DGAT1*, VPS13B*, GHR, HSF1*
SPA	L50	77	96	DGAT1*
	S75	-	-	-
NLD	L50	57	131	DGAT1*
	S68	-	-	-

^aL:level and S=slope at 65, 75 and 68 for France, Spain and the Netherlands, respectively.;

^b padj <0.05 p-values adjusted by Benjamini and Hoschberg procedure;

^c Number of candidates genes in a 0.5 Mb window around each significant SNP (padj <0.05);

^d Putative genes under Bonferroni significant level (0.05/N with N being the number of SNPs).

*Already found in the literature.

Table 10 Number of significant SNPs for level (L) and slopes (S) associated to protein yield, number of positional candidate genes as well as putative candidate genes based of a Bonferroni adjustment approach.

Country	Pseudo-phenotype ^a	Number of SNP ^b	Number of positional candidate genes ^c	Putative candidates genes ^d
FRA	L50	28	85	CCNT1, ABCC9*, DGAT1*, CSN1S2*, VPS13B*, GHR*
	S65	11	49	ABCC9*, CSN1S2*, VPS13B*, GHR*
SPA	L50	31	73	DGAT1*
	S75	-	-	-
NLD	L50	12	77	DGAT1*
	S68	4	67	HSF1*

^aL:level and S=slope at 65, 75 and 68 for France, Spain and the Netherlands, respectively;

^b padj <0.05 p-values adjusted by Benjamini and Hochberg procedure;

^c Number of candidates genes in a 0.5 Mb window around each significant SNP (padj <0.05);

^d Putative genes under Bonferroni significant level (0.05/N with N being the number of SNPs).

*Already found in the literature.

Table 11 summarizes results from the Manhattan plots across countries. For each chromosome and trait, we have included countries where genomic regions have been identified. Somehow, Table 11 supports the idea of having the least number of false positive region results for those that appear in more than one population, because they have been found in independent analyses. Thus, results in one data set validate results in the other ones. In addition to the QTL in BTA14, there are other regions of potential pleiotropy affecting levels for milk production traits and/or slopes and are common to at least two countries. Those regions are located in BTA6, BTA18, BTA23. In the French data, we have identified other regions of potential pleiotropy affecting levels as well as slopes in BTA2, BTA3, BTA4, BTA5, BTA11, BTA15, BTA19, BTA20. These results provide genomic evidences of the magnitude of estimated genetic correlations between level and slopes associated to traits subject to our study. The estimates of genetic correlation between the level at 150 days of lactations under neutral conditions and the slopes at THI of 65 in the French case varies between -0.37 and -0.03, for PY and CR, respectively; at THI of 75 in the Spanish population ranged between -0.33 for MY and -0.27 for SCS; and finally at THI of 68 of The Netherland varied between -0.37 and -0.01 for MY and SCS as well. One interesting aspect to highlight is that when we compared the common QTL regions between level and slopes, those QTLs always included the total set of markers associated to slopes but, the opposite was never observed. These findings suggest that there are more genes and more gene actions associated to QTLs of levels of production than those associated to slopes and therefore, those mechanisms that get affected by heat stress. Whether or not this is true and it is not an artefact provoked by the differential power of detection needs to be elucidated.

In addition to candidate regions of slopes associated to levels as well, we also found QTLs in BTA13, BTA27, BTA28 and BTA29 only linked to slopes. Finally, some genomic regions coding for miRNA have been identified in all populations for most traits (results not shown in tables). This is an interesting finding that needs to be followed up, because miRNAs are regulating gene expression and are being identified as molecular biomarkers of heat stress process (Zhao et al., 2022).

Table 11. Countries where genomic regions by chromosome and trait (level and slope of decay for *fa* (FY)t, milk (MY) and protein yield (PY), conception rate (CR) and somatic cell score (SCS) have been found statistically significant.

	Traits	1	2	3	5	6	7	11	13	14	15	16	18	19	20	23	25	27	28	29
L E V E L	FY		SP	ND	FR, SP	FR	FR			FR, SP, ND	FR			FR			SP			
	MY		FR	SP, ND	FR	FR, ND				FR, SP, ND	FR, SP			SP	FR, ND	SP	FR			
	PY	SP		FR	FR	FR				FR, SP, ND			FR, SP			FR, SP				
	CR							FR				FR	FR			FR				
	SCS																			
S L O P E	FY	FR	FR	FR	FR			FR	SP	FR, SP		FR	FR, SP	FR	FR	FR		FR	FR	
	MY				FR	FR				FR, ND					FR	ND			ND	FR
	PY		FR		FR			FR		ND	FR	FR			FR	FR			FR	
	CR																			
	SCS																			

The conclusion of this genomic analysis generates a need for a more deeply investigation of our results. Some open questions need to be elucidated. In this first attempt, we did not identify any clear new putative genes associated to levels and slopes besides those contained in Table 8 to Table 10 that have been already described in the literature. However, we are using slopes of decay of production and it is reasonable to identify common regions between level and slopes because our pseudo-phenotypes measure the impact of heat stress on production traits as opposed of having a direct physiological measurement of stress. Nevertheless, we were expecting to find genes with gene actions involved on mechanism of heat dissipation, because somehow, the cows that do not get very affected by heat stress may also be those cows that have a better ability to dissipate heat. Unfortunately the results of our study were limited by the small amount of data at high THI, leading to low accuracy of the slopes. This reduced the GWAS detection power for such traits. The very low heritability of CR has limited our power of detection as well, although this trait is probably the most interesting in terms of heat stress because of its increase in heritability and of its relatively low genetic correlation between neutral conditions and heat stress show its sensitivity to heat stress. Finally, several improvements can be suggested to deep into our findings in the future. The first one would be to perform an enrichment analyses to see what genes and/or functions, to gain a deep insight of our results. We could also explore other approaches such using cow data (instead of DRPs of sires), other pseudo-phenotypes, multi-traits models versus single traits to try to separate pleiotropy from LD, longitudinal analysis to evaluate if candidate regions for level and slopes change depending upon the stage of lactation.

7 General discussion and conclusions

In this study, we defined heat stress indicators for milk production, udder health and fertility traits. Based on our results, on the meteorological information available in each country, and considering that relative humidity is important because high humidity worsens the effect of high/low temperatures, the daily mean temperature-humidity index was chosen as heat stress indicator, as it summarizes both temperature and relative humidity in a single index. Various period lengths summarizing the variation of THI can be found in the literature for milk production traits, such as daily mean THI for the fourth day before a test day (Bernabucci et al., 2014), the daily THI of the test day (Bohmanova et al., 2008), or even the average daily mean THI of the four days before the test-day and of test-day (Nguyen et al., 2016). However, based on our analyses using dairy cattle data of three breeds and in three countries, we observed that different values of THI averaged over up to 7 days were highly correlated to a 3-day averaged THI, which was finally chosen as heat stress indicator for milk production and udder health traits. For CR, heat in periods before and after AI may impact the conception. However, similarly to milk production traits, our results were very consistent across all tested periods, and the THI averaged during a period including the day of AI and the following week was chosen as heat stress indicator for fertility traits. These two heat stress indicators have been used in subsequent analyses of milk production, udder health and fertility traits in France, Spain and the Netherlands.

At a population level, differences in thermo-neutral regions and in impact of heat stress were larger between countries than between breeds within a country, especially for the milk production and udder health traits. However, these observations must be interpreted with care. Indeed, heat stress thresholds ranged from THI values of 50 in France to THI values of 70 in Spain for milk production and udder health traits. While THI values of 70 agree with the literature (Carabaño et al., 2017), the thermo-neutral regions observed in France were relatively low and could be explained by multiple factors, such as differences in farming systems. Such differences were smaller with CR for which estimated heat stress thresholds varied between 61 and 64 across populations and agreed with values reported in the literature (Biffani et al., 2016; Gernand et al., 2019). Future studies are required to investigate the impact of farming systems on heat stress in dairy cattle.

At an individual level, reaction norm models using Legendre polynomials were used to estimate the trajectory of genetic parameters of milk production and udder health traits along the THI gradient in three breeds and countries. Broken-line models were used for fertility traits. Additionally, to genetic parameters, such models allow us to estimate breeding values for a general level of production under different conditions, and for heat stress sensitivity. From these models, no, or at most weak, genetic-by-THI interactions were observed for milk production and udder health traits in all countries and for all breeds, which is in agreement with results in the literature (Brügemann et al., 2012; Carabaño et al., 2017; Cheruiyot et al., 2021). On the other hand, some genetic-by-THI interactions were observed for CR, especially in France and Spain, as genetic correlations between thermo-neutral and (severe) heat stress conditions (i.e. with a THI higher than 75) were lower than 0.50. These results suggest that fertility traits may be more suitable for identifying variability in response to heat stress than production traits. Future research will evaluate if breeding programs should consider indexes for a general level of production under different conditions, and/or for heat stress sensitivity (related to slopes) to increase both production and resilience of dairy cattle in the context of climate change.

Meta-analyses of results obtained for the first-parity Holstein populations in each country showed a very good consistency of the traits based on levels for both thermo-neutral and heat stress conditions. These results are encouraging for future international genetic evaluations of the Holstein

populations, because regions that currently experience heat stress conditions will benefit from information already collected in other countries that often suffer from heat stress. However, these meta-analyses also showed the limits of their applications on low heritable traits such as the slopes. Future studies should investigate other models and approaches to better measure the decline in performances due to heat stress and to take advantage of the information available in multiple countries.

GWAS results were based on de-regressed proofs of bulls and integrated information from the whole population in each country. In contrast to females, bulls have daughters with performances in variable conditions and specially in variable THI, allowing to define different traits under neutral conditions and heat stress and estimate breeding values for these different traits. Results highlighted the existence of specific regions for production levels as well as genomic regions shared with slopes of decay, our indicator traits of heat stress. For udder health all analyses failed to detect QTL and for fertility traits no putative genes were identified, probably because of a too limited power. Future research should focus on methods to extract more information from current data as well as to understand the molecular mechanisms underlined gene actions of common genes associated to levels and slope of decays.

8 References

- Baumgard, L. H., and R. P. Rhoads, 2013. Effect of Heat stress of Postabsorptive metabolism and energetics. *Annual Review of Animal Bioscience*. 1: 311-337. doi.org/10.1146/annurev-animal-031412-103644.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B* 57: 289-300
- Bernabucci, U., S. Biffani, L. Buggiotti, A. Vitali, N. Lacetera, and A. Nardone. 2014. The effects of heat stress in Italian Holstein dairy cattle. *Journal of Dairy Science* 97:471–486. doi:10.3168/jds.2013-6611.
- Biffani, S., U. Bernabucci, A. Vitali, N. Lacetera, and A. Nardone. 2016. Short communication: Effect of heat stress on nonreturn rate of Italian Holstein cows. *Journal of Dairy Science* 99:5837–5843. doi:10.3168/jds.2015-10491.
- Bohmanova, J., I. Misztal, S. Tsuruta, H.D. Norman, and T.J. Lawlor. 2008. Short Communication: Genotype by Environment Interaction Due to Heat Stress. *Journal of Dairy Science* 91:840–846. doi:10.3168/jds.2006-142.
- Brügemann, K., E. Gernand, U. König von Borstel, and S. König. 2012. Defining and evaluating heat stress thresholds in different dairy cow production systems. *Archives Animal Breeding* 55:13–24. doi:10.5194/aab-55-13-2012.
- Carabaño, M.J., M. Ramón, C. Díaz, A. Molina, M.D. Pérez-Guzmán, and J.M. Serradilla. 2017. BREEDING AND GENETICS SYMPOSIUM: Breeding for resilience to heat stress effects in dairy ruminants. A comprehensive review1. *Journal of Animal Science* 95:1813–1826. doi:10.2527/jas.2016.1114.
- Cheruiyot, E.K., M. Haile-Mariam, B.G. Cocks, I.M. MacLeod, R. Xiang, and J.E. Pryce. 2021. New loci and neuronal pathways for resilience to heat stress in cattle. *Sci Rep* 11:16619. doi:10.1038/s41598-021-95816-8.
- Cheruiyot E.K., Nguyen T.T.T, Haile-Mariam M, Cocks B.G., Abdelsayed M, and J. E. Pryce 2020. Genotype-by-environment (temperature-humidity) interaction of milk production traits in Australian Holstein cattle. *J Dairy Sci*. 2020;103(3):2460-76. doi: 10.3168/jds.2019-17609
- Gernand, E., S. König, and C. Kipp. 2019. Influence of on-farm measurements for heat stress indicators on dairy cow productivity, female fertility, and health. *Journal of Dairy Science* 102:6660–6671. doi:10.3168/jds.2018-16011.
- Fonseca, P.A.R., Suarez-Vega, A., Marras, G. and Angela Cánovas. 2020. GALLO: An R package for genomic annotation and integration of multiple data sources in livestock for positional candidate loci. *GigaScience*, 9, 2020, 1–9. doi: 10.1093/gigascience/giaa149
- Hartwiger, J., M. Schären, S. Potthoff, L. Hüther, S. Kersten, D. Von Soosten, A. Beineke, U. Meyer, G. Breves, and S. Dänicke. 2018. Effects of a Change from an Indoor-Based Total Mixed Ration to a Rotational Pasture System Combined With a Moderate Concentrate Feed Supply on Rumen Fermentation of Dairy Cows. *Animals* 8:205. doi:10.3390/ani8110205.
- Medrano, J.F. 2017. The new bovine reference assembly and its value for genomic research. Pages 161–166 *Proc. Assoc. Advmt. Anim. Breed. Genet.*

- Meyer, K. 2007. WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *J. Zhejiang Univ. - Sci. B* 8:815–821. doi:10.1631/jzus.2007.B0815.
- Misztal, I., S. Tsuruta, D.A.L. Lourenco, I. Aguilar, A. Legarra, and Z. Vitezica. 2014. Manual for BLUPF90 family of programs.
- Muggeo, V.M.R. 2008. segmented: An R Package to Fit Regression Models with Broken-Line Relationships. *R News* 8:20–25.
- ten Napel, J., J. Vandenplas, M. Lidauer, I. Strandén, M. Taskinen, E. Mantysaari, M.P.L. Calus, and R.F. Veerkamp. 2021. MiXBLUP 3.0.1 manual - V3.0.
- National Research Council. 1971. *A Guide to Environmental Research on Animals*. National Academy of Sciences, Washington, DC.
- Nguyen, T.T.T., P.J. Bowman, M. Haile-Mariam, J.E. Pryce, and B.J. Hayes. 2016. Genomic selection for tolerance to heat stress in Australian dairy cattle. *Journal of Dairy Science* 99:2849–2862. doi:10.3168/jds.2015-9685.
- Ravagnolo, O., and I. Misztal. 2002. Effect of Heat Stress on Nonreturn Rate in Holsteins: Fixed-Model Analyses. *Journal of Dairy Science* 85:3101–3106. doi:10.3168/jds.S0022-0302(02)74397-X.
- Rosen, B., D. Bickhart, R. Schnabel, S. Koren, C. Elsik, A. Zimin, C. Dreischer, S. Schultheiss, R. Hall, and S. Schroeder. 2018. Modernizing the bovine reference genome assembly. *Page Proc World Congr Genet Appl Livest Prod*.
- Santana, M.L., A.B. Bignardi, G. Stefani, and L. El Faro. 2017. Genetic component of sensitivity to heat stress for nonreturn rate of Brazilian Holstein cattle. *Theriogenology* 98:101–107. doi:10.1016/j.theriogenology.2017.04.052.
- Schaeffer, L.R. 1994. Multiple-Country Comparison of Dairy Sires. *Journal of Dairy Science* 77:2671–2678. doi:10.3168/jds.S0022-0302(94)77209-X.
- Teng J., Wang, D., Zhao, Ch., Zhang, X., Chen Z., Liu, J., Sun, D., Tang, H., Wang, W., Li, J., Mei, Ch., Yang, Z., Ning, Ch, and Qin Zhang. 2023. Longitudinal genome-wide association studies of milk production traits in Holstein cattle using whole-genome. *J. Dairy Sci.* 106:2535–2550. doi.org/10.3168/jds.2022-2227
- Tribout T., Croiseau P., Lefebvre R., Barbat A., Boussaha M., Fritz S., Boichard D., Hoze C., and Sanchez M.P. 2020. Confirmed candidate causative variants for milk production, udder health, and udder morphology in dairy cattle. *Genet Sel Evol*, 52, 55. doi: 10.1186/s12711-020-00575-1.
- VanRaden, P.M., M.E. Tooker, J.R. Wright, C. Sun, and J.L. Hutchison. 2014. Comparison of single-trait to multi-trait national evaluations for yield, health, and fertility¹. *Journal of Dairy Science* 97:7952–7962. doi:10.3168/jds.2014-8489.
- Vihervaara, A., and L. Sistonen. 2014. HSF1 at a glance. *Journal of Cell Science*, 127(2), 261–266. <https://doi.org/10.1242/jcs.132605>
- Yang, J., S.H. Lee, M.E. Goddard, and P.M. Visscher. 2011. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 88:76–82. doi:10.1016/j.ajhg.2010.11.011.
- Zhao, Y., Wei, Q., Chen, T., Xu, L., Liu, J., Zhang X., Han, G. and Qing Ma. 2022. Identification and characterization of heat-responsive miRNAs and their regulatory network in maize. *Plant Growth Regul* 96, 195–208 (2022). <https://doi.org/10.1007/s10725-021-00769-7>

Appendices

Appendix 1: List of abbreviations

AI: Artificial Insemination

BL: Broken Line model

CR: Conception Rate (here: Conception rate at the first insemination)

DIM: Days-In-Milk

EBV: Estimated Breeding Value

FRA: France

FY: Fat yield

FC: Fat Content

GWAS: Genome Wide Association Study

HOL: Holstein

LMY, LFY, LPY, LSCS, LCR: Levels of performances (Milk, Fat, Protein, Somatic Cell Scores and Conception rate, respectively). A number following the abbreviation corresponds to the THI level, eg. L65MY = Level of Milk Yield at THI 75

MACE: Multiple Across Country Evaluation

MON: Montbéliarde

MRY: Meuse-Rhine-Yssel

MY: Milk Yield

NLD: The Netherlands

PC: Protein Content

PY: Protein Yield

QTL: Quantitative Trait Locus

SCC: Somatic Cell Counts (in cells/ml)

SCS: Somatic Cell Score

SMY, SFY, SPY, SSCS, SCR: slope of decay of performances (Milk, Fat, Protein, SCS and Conception Rate at 1st AI, respectively). A number following the first S corresponds to the THI level, eg. S65MY = slope for Milk Yield at THI 75

SNP: Single Nucleotide Polymorphism

SPA: Spain

THI: Temperature-Humidity Index

Appendix 2: General description of the phenotypes used to define heat stress indicators

France

Records from cows with unknown parents or with missing data for any of the traits were excluded. Test-day records with extreme values (<15 or >90) for FC and PC were deleted.

Performances and pedigree information of Holstein and Montbéliarde cows were extracted from the French national database. Data for the years 2010 to 2020 were available. This complete period was used to estimate breeding values necessary for the meta-analyses and GWAS, while the period was restricted to the last 5 years to estimate the genetic parameters.

Regarding production traits and SCS, only cows with age at calving between 23 and 42 months and 35-60 months for first and second parity, respectively, were included. Short lactations (<160 days) were deleted and lactation length was limited to 305 days.

Records from cows with unknown parents or with missing data for any of the traits were excluded. Test-day records with extreme values (<15 or >90) for FC and PC were deleted.

The genetic parameters and effects at the population level were estimated from a sample of herds randomly selected in the whole population with performances between 2016 and 2020. These samples corresponded to 301,078 test-day performances from 39,469 Montbéliarde cows in 1st lactation and 292,648 performances from 37,593 Montbéliarde cows in 2nd lactation, 489,814 performances of 71,958 Holstein cows in 1st lactation.

Tables 2.1 and 2.2 present the size of the datasets available corresponding to the periods 2010-2020 and used for the estimation of breeding values.

Table 2.1. Description of the datasets in the French Holstein breed.

Trait ¹	Period	Number of herds	Number of (test-day) records	Number of animals with records
MY, FY, PY, FC, PC, SCS (parity 1)	2010-2020	44,784	Total (L1+L2)= 145,139,089	5,753,268
MY, FY, PY, FC, PC, SCS (parity 2)	2010-2020	44,784		4,442,242
Conception rate (Parity 1)	2010-2020	35,296	3,351,068	3,351,068

¹MY = milk yield; FY = fat yield; PY = protein yield; FC = fat content; PC = protein content; SCS = somatic cell score.

Table 2.2. Description of the datasets in the French Montbéliarde breed.

Trait ¹	Period	Number of herds	Number of (test-day) records	Number of animals with records
MY, FY, PY, FC, PC, SCS (parity 1)	2010-2020	20,306	Total (L1+L2) = 36,965,004	1,297,979
MY, FY, PY, FC, PC, SCS (parity 2)	2010-2020 2016-2020	20,306	L2: 457,516	1,008,834 58,229
Conception rate (Parity 1)	2010-2020	9,683	649,814	649,814

¹MY = milk yield; FY = fat yield; PY = protein yield; FC = fat content; PC = protein content; SCS = somatic cell score.

Spain

Spanish data set consisted of Holstein dairy cattle milk production data and meteorological data for the 2011 to 2020 period. Production data was provided by CONAFE breeders association. Data used was daily milk, fat and protein yields (Kg/d) and fat and protein contents (%) gathered from 4608 herds. Table 2.3 shows a description of the data used. Only data from cows with sire and dam known, obtained within 20 and 305 days in milk, with associated meteorological information from weather stations at a distance inferior to 20 km and within the levels established by the breed association for fat content [1.5, 7.0] and protein content [1.0, 5.0] were considered for further analyses. In order to avoid estimation for extreme meteorological data for which a small amount of information was available, data associated with THI values outside the interval [40,80] were discarded. In addition, extreme data for yields (outside the range between percentiles 1 and 99) or data from a combination of herd-year and season of test-day with less than 6 observations were removed. For the variance components estimation, further edits were applied. Cows were required to have at least three records per lactation with at least one of them being associated with an average daily temperature over 22°C, to ensure that all animals had experience at least a moderate level of heat stress.

For the analysis of conception rate, data from artificial inseminations in the same period as for production traits were also provided by CONAFE. Edits included also the requirement of known sire and dam and maximum distance to weather stations of 20 km, apart from checks on plausible intervals between successive calving dates, and between calving and insemination dates.

Table 2.3. Description of the datasets in the Spanish Holstein breed used for the estimation of breeding values.

Trait ¹	Period	Number of herds	Number of test-day records	Number of animals with records
MY, FY, PY, FC, PC, SCS (parity 1, 2, 3)	2011-2020	4604	19,833,118	1,016,403
Conception rate (parity 1, 2, 3)	2010-2020	8159	7,001,096	1,390,502

¹MY = milk yield; FY = fat yield; PY = protein yield; FC = fat content; PC = protein content; SCS = somatic cell score.

The Netherlands

Only records from 100% Holstein cows and at least 87.5% Meuse-Rhine-Yssel (MRY) cows with both parents known were kept. Only Dutch herds that could be associated with a weather station were selected. Records outside the range of 3-day averaged THI equal to 10 and 3-day averaged THI equal to 83 were deleted. Each unit of the 3-day averaged THI was required to be associated with at least 400 test-day records.

For milk production and udder health traits, test-day records for milk yield (kg), fat yield (kg), protein yield (kg) and somatic cell score (SCS; defined as $SCS = 3 + \log_2(SCC/100,000)$, with SCC being somatic cell counts in cells/ml) of Holstein and (at least 7/8) MRV cattle after (and including) 2010 were extracted from the database used for the Dutch-Flemish genetic evaluations of production traits. Also, only records between 5 and 305 DIM from first- and second-parity Holstein and MRV cows, that calved between 23 and 42 months for first-parity cows and between 35 and 60 months for second-parity cows were selected. In addition, only records with observations that were less than 100 kg for milk yield, less than 10% fat and protein, and less than 10 million SCC were used.

Finally for production traits and SCS, each contemporary group (i.e., the effect “herd x year of test” for the impact of THI on production traits and the effect “herd x test-date” for genetic by THI analyses) and each class “parity x age at calving (in months) x year x season of calving x lactation stage” were required to have at least five test-day records. Similar data edits were used for conception rate (at least 50 observations per THI level, five cows per herd-year, and five daughters per sire).

The pedigrees were traced back three generations from Holstein and MRV cows with records. The pedigrees included about 830 thousand animals for first-parity Holstein datasets, about 730 thousand animals for second-parity Holstein datasets, about 21 thousand animals for first-parity MRV datasets, and about 18 thousand animals for second-parity MRV datasets.

Tables 2.4 and 2.5 present the size of the datasets available for the estimation of genetic parameters and breeding values for each trait and each breed.

Table 2.4. Description of the datasets for the Dutch Holstein breed available for the analyses at the population and individual levels.

Trait	Period	Number of herds	Number of (test-day) records	Number of animals with records
Fat content (parity 1)	2010-2022	1581	3,876,021	492,666
Milk yield (parity 1)	2010-2022	1581	3,910,411	493,262
Protein content (parity 1)	2010-2022	1581	3,876,333	492,671
Somatic cell score (parity 1)	2010-2022	1580	3,447,210	474,273
Conception rate (parity 1)	2010-2022	1565	417,548	417,548
Fat content (parity 2)	2010-2022	1580	3,218,762	413,683
Milk yield (parity 2)	2010-2022	1580	3,247,362	414,259
Protein content (parity 2)	2010-2022	1580	3,219,081	413,687
Somatic cell score (parity 2)	2010-2022	1580	2,832,756	397,126
Conception rate (parity 2)	2010-2022	1564	342,783	342,783

Table 2.5. Description of the datasets for the Dutch Meuse-Rhine-Yssel breed available for the analyses at the population and individual levels.

Trait	Period	Number of herds	Number of (test-day) records	Number of animals with records
Fat content (parity 1)	2010-2022	87	88,392	11,969
Milk yield (parity 1)	2010-2022	87	89,049	12,005
Protein content (parity 1)	2010-2022	87	88,401	11,970
Somatic cell score (parity 1)	2010-2022	87	81,599	11,554
Conception rate (parity 1)	2010-2022	80	9289	9289
Fat content (parity 2)	2010-2022	86	66,777	9704
Milk yield (parity 2)	2010-2022	86	67,334	9727
Protein content (parity 2)	2010-2022	86	66,790	9704
Somatic cell score (parity 2)	2010-2022	85	58,760	9085
Conception rate (parity 2)	2010-2022	83	7664	7664

Appendix 3: Statistical analyses of weather data

Data

Meteorological data were provided by Météo-France (Safran database) for France, by the National Meteorological Agency (AEMET) for the period 2000 to 2020 for Spain, and were extracted from the Koninklijk Nederlands Meteorologisch Instituut (KNMI) website from 1990 for the Netherlands. Weather records were available for 9,892 French, 1,993 Spanish and 34 Dutch weather stations distributed throughout each national territory. Each herd was associated to the closest weather station, according to the ZIP code of the farm. The average distance to the associated weather station was of 8.4 km in Spain, 14.6 km in the Netherlands, and less than 5.7 km in France (grid of 8x8 km).

Average, minimum and maximum daily temperatures (in degrees Celsius) and daily mean relative atmospheric humidity (in percent) were available, or computed from available records, in each country. Additional information, such as daily measures of precipitation (rain/snow in mm), wind (m/s), radiation (J/cm²), and evapotranspiration (mm) for France, and minimum temperature at 10 cm above surface (in 0.1 degrees Celsius), minimum and maximum relative atmospheric humidity (in percent) for the Netherlands, are also available.

All countries compute daily THI values using the formula proposed by (National Research Council, 1971), as:

$$THI = (1.8 \cdot T + 32) - (0.55 - 0.0055 \cdot RH) \cdot (1.8 \cdot T - 26),$$

with T being the average temperature (degrees Celsius) and RH being the average daily relative humidity (percent). As only average daily relative humidity was available in France, only daily average THI will be reported.

Weather data come from substantially different French and Spanish climatic regions. For this reason, descriptive statistics on daily mean, minimum and maximum temperatures (in degrees Celsius), daily mean relative atmospheric humidity (RH; in percent) and daily mean THI, for 5 different French regions (North: Ile de France, Nord Pas de Calais, Picardie; North East: Champagne Ardennes, Bourgogne, Franche-Comté, Alsace, Lorraine; North West : Bretagne, Basse Normandie, Haute Normandie, Pays de la Loire, Centre; South East : Auvergne, Languedoc Roussillon, Provence-Alpes-Côte d'Azur, Rhône-Alpes; South West : Aquitaine, Poitou Charente, Limousin, Midi-Pyrénées) and for four Spanish regions (northern coast (Cordillera Cantabrica), Continental (north-center of Spain), Mediterranean (Mediterranean coast) and South), and for the Netherlands are shown in Table 3.1 for different periods. Table 3.1 also reports the number of days with a daily mean THI > 60 or a daily mean THI > 72 measured in at least one of the weather stations for each region or country.

Table 3.1. Descriptive statistics for temperature and relative humidity in France, Spain and the Netherlands for different periods.

Country	Area	Period ¹	Temperature			RH			Number of measures	
			Min.	Average	Max.	Min.	Average	Max.	THI ≥ 60	THI ≥ 72
France	North	Winter 2016-2020	-11.5	4.9	19.1	40.4	87.4	99.9	0	0
	North East		-19.2	3.6	20.0	18.1	86.6	100	0	0
	North West		-12.1	6.1	22.3	42.8	86.9	100	1	0
	South East		-25.4	4.1	25.3	12.4	78.5	100	20	0
	South West		-16.6	6.3	25.7	11.7	84.3	99.9	21	0
	North	Summer 2016-2020	2.8	18.7	41.4	34.3	73.8	99.7	446	88
	North East		2.0	19.6	41.2	28.7	70.5	99.6	451	129
	North West		4.6	18.6	41.2	25.6	75.1	99.4	452	112
	South East		-3.1	19.9	43.0	23.8	65.9	99.7	460	345
	South West		-2.5	20.3	41.8	21.4	72.4	99.5	456	197
Spain ²	Northern coast	May-Sept 2010-2020	12.6	17.6	22.6	57.7	76.2	100	154	8
	Continental	May-Sept 2010-2020	10.8	18.0	25.2	34.9	60.7	100	153	36
	Mediterranean	May-Sept 2010-2020	14.3	20.5	26.8	42.6	65.8	100	154	89
	South	May-Sept 2010-2020	14.7	21.5	28.8	35.1	57.5	100	154	104
The Netherlands		Spring 2010-2019	-6.20	9.62	25.20	30.00	76.60	100.00	156	1
		Summer 2010-2019	8.00	17.46	31.20	30.00	76.85	98.00	804	63
		Autumn 2010-2019	-4.20	11.18	25.50	41.00	85.05	100.00	241	2
		Winter 2010-2019	-16.60	3.94	14.30	45.00	86.87	100.00	0	0
		Spring 2020-2022	-0.50	8.99	21.00	34.00	72.87	99.00	20	0
		Summer 2020-2022	9.00	17.83	28.70	38.00	77.40	99.00	174	16
		Autumn 2020-2022	-0.20	11.84	25.30	48.00	84.80	100.00	57	0
		Winter 2020-2022	-8.30	5.34	15.40	40.00	86.57	100.00	0	0

¹Spring = March, April, May; Summer = June, July, August; Autumn = September, October, November; Winter = December, January, February.

²Average values of daily maximum, average, and minimum temperatures and relative humidity for the months of May through September and for the period 2010-2020. For the number of days with THI above 60/72, average across years of number of days in the period of May-Sept. 2010-2020.

Consistency between different indicators

For the period 2010-2022 and across all the 34 selected Dutch weather stations, the average Pearson correlations was equal to 0.94 between the daily mean temperature and the minimum temperature, and 0.98 between the daily mean temperature and the maximum temperature. Although France and the Netherlands reported similar correlations between different measures of temperatures, Spain reported quite different values (Tables 3.2 to 3.5). Indeed, for the period 2010-2020, the Pearson correlations were in the range [0.79 - 0.90] between the daily mean temperature and the minimum temperature, and in the range [0.89 - 0.93] between the daily mean temperature and the maximum temperature.

Furthermore, the daily mean THI were averaged over a period of up to 7 days preceding a specific date. Pearson correlations between 3-day averaged THI and other averaged THI were on average all larger or equal to 0.97 for the Netherlands, to 0.91 for France, and to 0.85 for Spain (across all regions). Finally, for both France and Spain, the correlations between daily mean temperature and THI were equal to 1.00.

Table 3.2. Pearson correlations among minimum, average, maximum temperature and temperature-humidity index (THI)¹ in France.

	Tave_1D	Tmin_1D	Tmax_1D	THlave_1D
Tave_1D	1.00	0.93	0.97	>0.99
Tmin_1D		1.00	0.84	0.93
Tmax_1D			1.00	0.97
THlave_1D				1.00

¹Tave_1D= Average temperature ; Tmin_1D= Minimum temperature ; Tmax_1D : maximum temperature; THlave_1D= average THI (1 day).

Table 3.3. Pearson correlations among temperature-humidity index (THI) computed on the test-day, 3 and 7 days before test-days¹ in France.

	THlave_1D	THlave_3D	THlave_7D
THI1_J0p	1.00	0.96	0.91
THI1_3Jp		1.00	0.96
THI1_7Jp			1.00

¹THI1ave_1D = average THI (on the test-day); THlave_3D = average THI during 3 days before the test-day; THlave_7D = average THI during 3 days before the test-day.

Table 3.4. Pearson correlations among temperature (minimum/average/maximum) and temperature-humidity index (THI) for the May-September period for each climatic region in Spain.

	Northern Coast				Continental			
	Tmax	Tmin	Tave	THlave	Tmax	Tmin	Tave	THlave
Tmax	1.00	0.43	0.89	0.87	1.00	0.56	0.92	0.93
Tmin		1.00	0.79	0.81		1.00	0.84	0.86
Tave			1.00	1.00			1.00	0.99
THlave				1.00				1.00
	Mediterranean				South			
	Tmax	Tmin	Tave	THlave	Tmax	Tmin	Tave	THlave
Tmax	1.00	0.67	0.92	0.90	1.00	0.66	0.93	0.86
Tmin		1.00	0.90	0.92		1.00	0.89	0.90
Tave			1.00	0.99			1.00	0.98
THlave				1.00				1.00

Table 3.5. Pearson correlations among temperature-humidity index (THI) values for the day of recording to the 7 days before for May-September period for each climatic region in Spain.

	Northern Coast						Continental					
THI ₀	THI ₀₋₁	THI ₀₋₂	THI ₀₋₃	THI ₀₋₅	THI ₀₋₇	THI ₀	THI ₀₋₁	THI ₀₋₂	THI ₀₋₃	THI ₀₋₅	THI ₀₋₇	
THI ₀	1.00	0.95	0.90	0.85	0.79	0.75	1.00	0.97	0.92	0.88	0.83	0.79
THI ₀₋₁		1.00	0.97	0.93	0.86	0.82		1.00	0.98	0.95	0.89	0.84
THI ₀₋₂			1.00	0.98	0.92	0.87			1.00	0.99	0.94	0.89
THI ₀₋₃				1.00	0.97	0.92				1.00	0.97	0.93
THI ₀₋₅					1.00	0.98					1.00	0.98
THI ₀₋₇						1.00						1.00

Conclusion

In conclusion, based on the different measurements available in each country, the daily mean THI, that is based on daily mean temperature and relative humidity, can be computed in each country and used in different analyses. Other measurements are only available in some countries, or cannot be used due to lack of information at the herd level. Furthermore, based on the above results, it can be concluded that daily mean THI are highly correlated with daily mean, but also minimum and maximum temperatures. This confirms that THI may be preferred over temperature because it integrates relative humidity. Indeed, high humidity worsens the effect of high/low temperatures. The THI can therefore be used as single indicator. Finally, the results show that the different values of THI averaged over up to 7 days are highly correlated to a 3-day averaged THI. Therefore, we conclude that the 3-day averaged THI can be considered as indicator in the following analyses, and that other measurements would lead to similar results.

Appendix 4: Software used in the different studies

Estimation of variance components and solutions of estimated effects at the population level

Variance components and solutions were estimated using the software WOMBAT (Meyer, 2007) for French data and in the meta-analysis at the individual level, and using the BLUPF90 software suite (Miszta et al., 2014) for the Dutch and Spanish ones.

Estimation of variance components and solutions of estimated effects at the individual level

Variance components and solutions were estimated using the software WOMBAT (Meyer, 2007) for French data and in the meta-analysis at the individual level, and using the BLUPF90 software suite (Miszta et al., 2014) for the Dutch and the Spanish ones. Reliabilities were obtained with MiXBLUP (ten Napel et al., 2021) in all countries.

De-regression of EBVs

DRPs and corresponding ERCs were obtained with the “deregession” INRAE in-house program.

GWAS

GWAS were carried out with the GCTA software (Yang et al., 2011) in a model also including a polygenic effect estimated through a genomic relationship matrix estimated with the 50k genotypes.

Appendix 5: Genetic parameters in second lactation

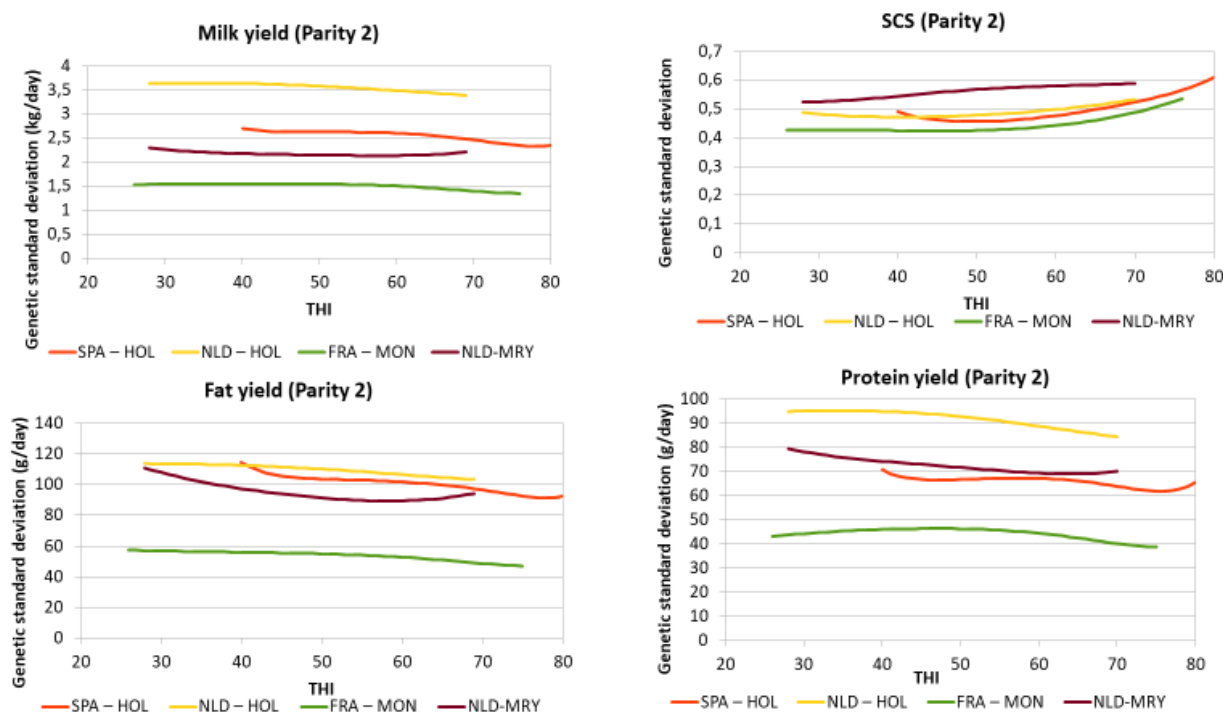


Figure 5.1. Evolution of additive genetic standard deviations for milk yield (kg/day), fat yield (g/day), protein yield (g/day), and somatic cell score (SCS) with 3-day averaged temperature-humidity index (THI) in second-parity cows, according to the country (FRA=France; NLD = the Netherlands; SPA = Spain) and to the breed (HOL=Holstein; MON=Montbéliarde, MRY=Meuse-Rhine-Yssel). Results are given at 150 days-in-milk.

Appendix 6: Comparison of estimated genetic correlations between countries (RUMIGEN vs routine Interbull international evaluations)

Table 6.1. Estimated genetic correlations between countries estimated by the RUMIGEN partners for the levels under thermo-neutral and heat stress conditions, and used by Interbull during the May 2023 routine evaluation of Holstein bulls (www.interbull.org).

	Level - Milk yield					
	Rumigen (thermo-neutral)		Rumigen (heat stress)		Interbull	
	SPA	NLD	SPA	NLD	SPA	NLD
FRA	0.96	0.94	0.92	0.89	0.92	0.89
SPA		0.92		0.86		0.9
	Level - fat yield					
	Rumigen (thermo-neutral)		Rumigen (heat stress)		Interbull	
	SPA	NLD	SPA	NLD	SPA	NLD
FRA	0.97	0.9	0.97	0.87	0.91	0.87
SPA		0.89		0.85		0.87
	Level - protein yield					
	Rumigen (thermo-neutral)		Rumigen (heat stress)		Interbull	
	SPA	NLD	SPA	NLD	SPA	NLD
FRA	0.96	0.89	0.89	0.81	0.88	0.84
SPA		0.9		0.83		0.88
	Level - somatic cell score					
	Rumigen (thermo-neutral)		Rumigen (heat stress conditions)		Interbull	
	SPA	NLD	SPA	NLD	SPA	NLD
FRA	0.95	0.88	0.96	0.88	0.96	0.94
SPA		0.89		0.88		0.91
	Level - conception rate					
	Rumigen (thermo-neutral)		Interbull			
	SPA	NLD	SPA	NLD	SPA	NLD
FRA	0.87	0.77			N/A	0.92
SPA		0.63				N/A

Appendix 7: Manhattan plots

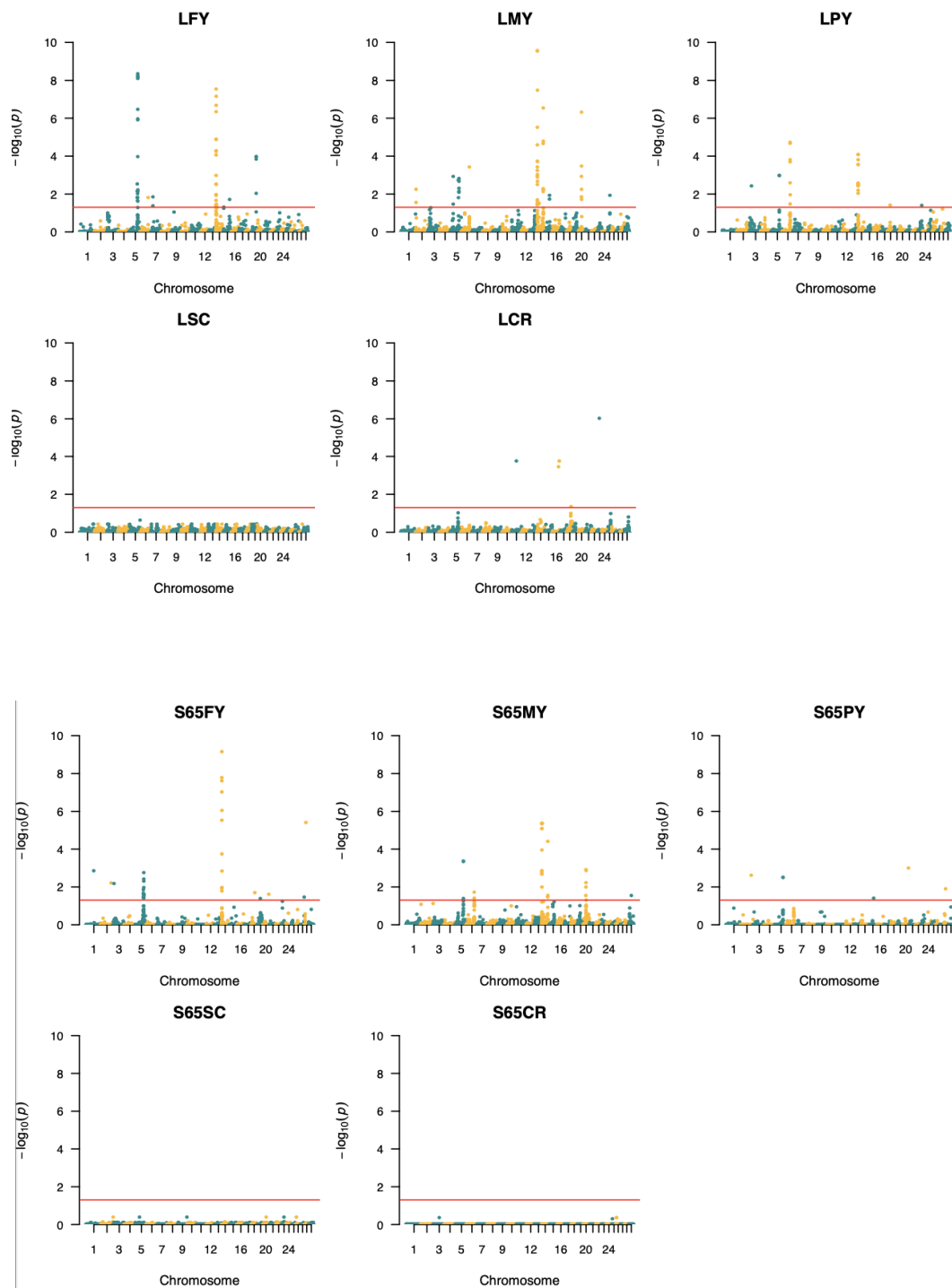


Figure 7.1. Manhattan plot with results of genome wide association study in the French population for levels and slopes of decay at heat stress ($THI=65$) of production traits fat yield (LFY/ S65FY), milk yield (LMY/ S65MY) and protein yield (LPY/S65PY), udder health trait (LSC/S65CS) and fertility trait (LCR/S65CR).

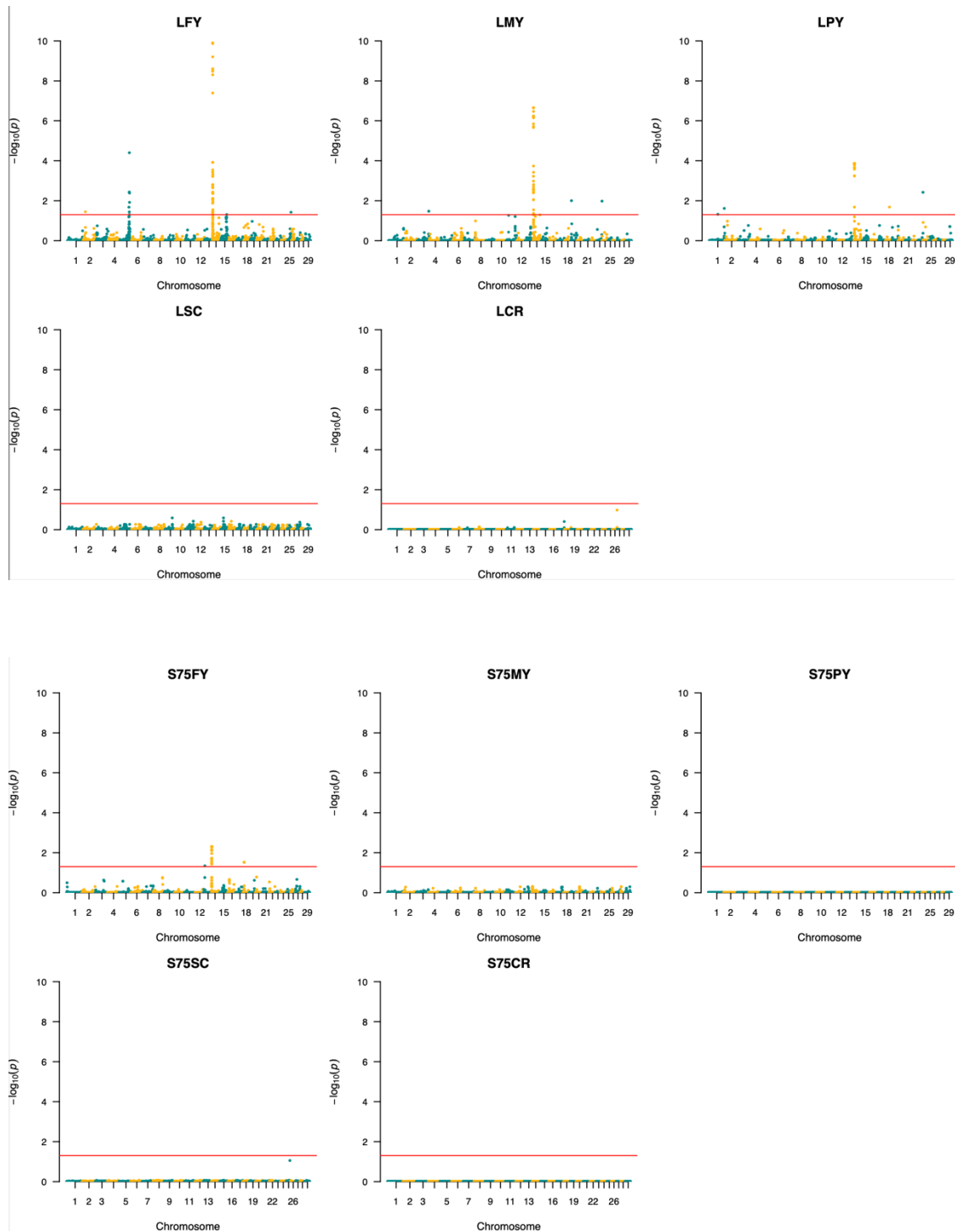


Figure 7.2. Manhattan plot with results of genome wide association study in the Spanish population for levels and slopes of decay at heat stress ($THI=75^{\circ}$) of production traits fat yield (LFY/ S75FY), milk yield (LMY/ S75MY) and protein yield (LPY/S75PY), udder health trait (LSC/S75CS) and Fertility trait (LCR/S75CR).

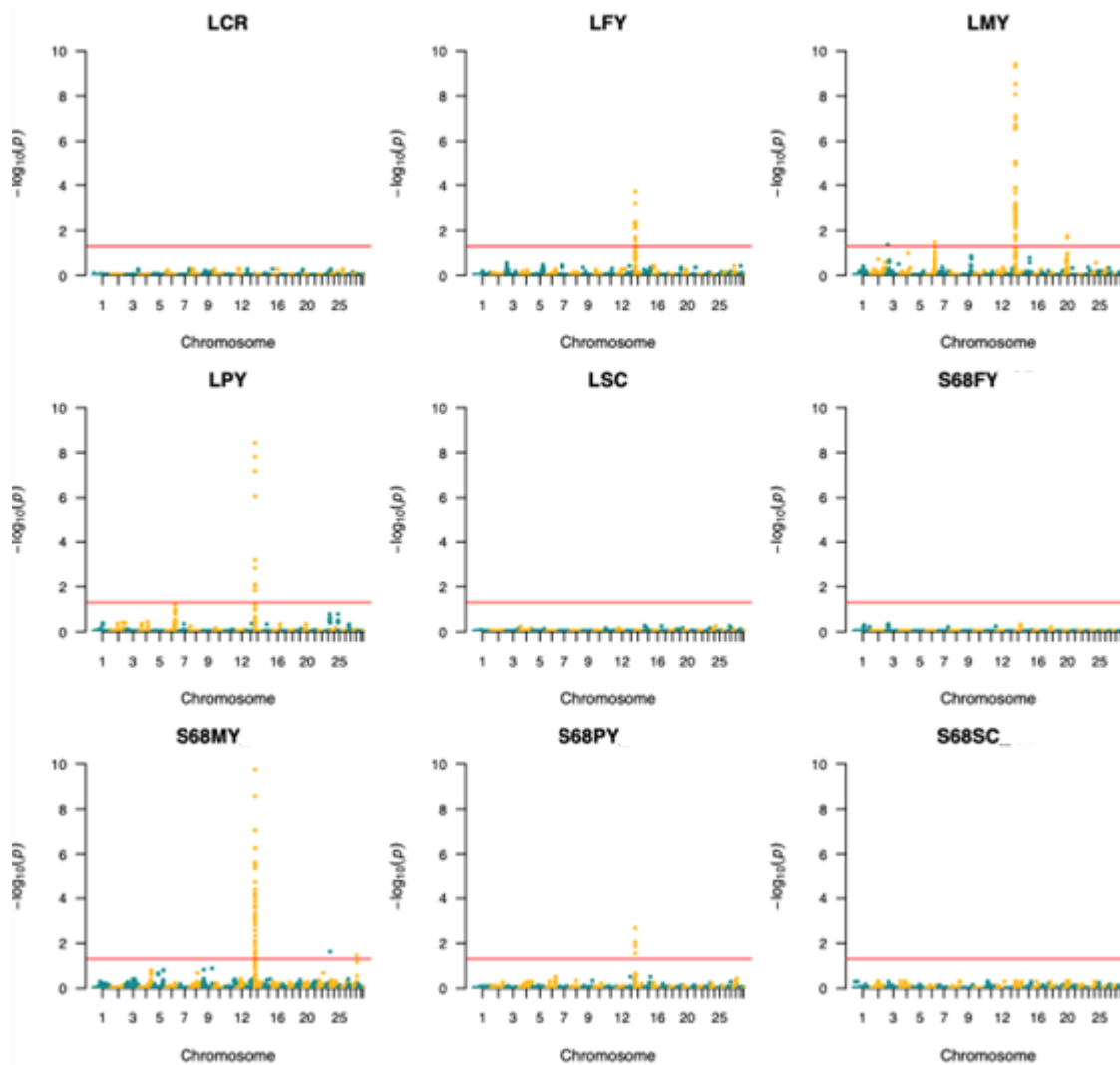


Figure 7.3.Manhattan plot with results of genome wide association study in the Dutch population for levels and slopes of decay at heat stress ($THI=68$) of production traits fat yield (LFY/ S68FY), milk yield (LMY/ S68MY) and protein yield (LPY/S68PY), udder health trait (LSC/S68CS) and Fertility trait (LCR/S68CR).

Appendix 8: Study of effect of THI at the population level in France: comparison of THI effects at the regional level

The aim of this study was to compare the effect of THI in French regions with contrasted farming systems. The method was the same as in section 4 of this document, with THI effects estimated separately in two French areas with pedoclimatic characteristics leading to different farming systems. The selected farms were located in French administrative department that are parts of two regions of Western France (see figure 1): Normandie (NOR) and Pays de la Loire (PDL). The dataset contained 399.360 and 168.713 test days of first parity Holstein cows in the regions PDL and NOR, respectively, recorded between 2015 and 2020.

In the region Normandie (NOR), most systems are based on pasture, with cows exposed to outside temperatures from early spring to autumn (except during very hot periods), while in region Pays de la Loire (PDL), most cows are kept indoor all year round and fed with a concentrate and silage based ration. Unfortunately, no data were available to better describe the farming systems and to better select the farms responding to the above description in each region.

In the region NOR, the results obtained on milk yield showed a breakpoint at THI 49-50, with a steep decline between 50 and 56. A negative slope is also observed in PDL, but much attenuated. These THI levels are not distributed regularly throughout the year: in the region NOR, 47% of THI [50-54] are concentrated between February and April, which corresponds to the period of transition between indoor and pasture.

Therefore, these results illustrate the fact that the THI effects may be affected by farming conditions. This could explain at least partially the negative effect of moderate THI observed in France. Indeed, in regions with a lot of farms with grazing systems, moderate THI are concentrated in late winter - early spring, corresponds to the period of transition between indoor and pasture (47% of THI [50-54] in NOR were observed between February and April). Therefore, the effect of moderate THI estimated could be partially confounded with a negative effect of this transition period.

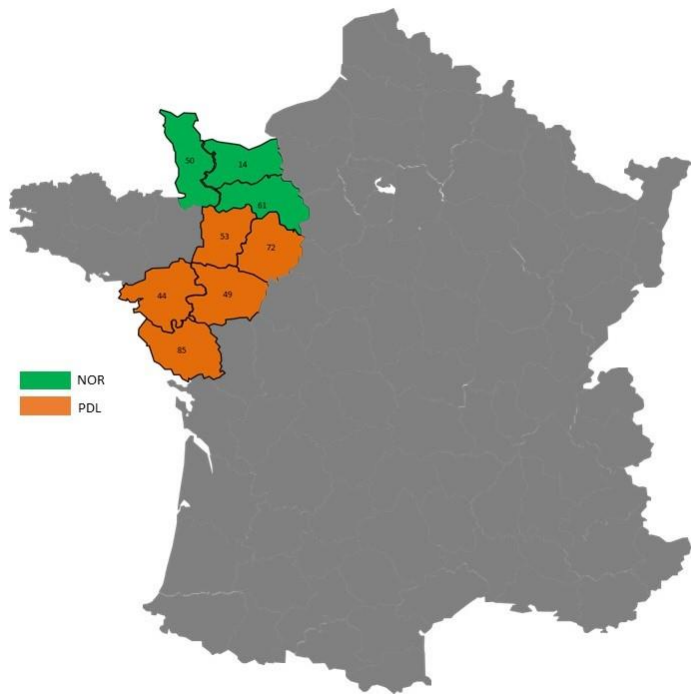


Figure 1. Regions NOR and PDL selected in the study

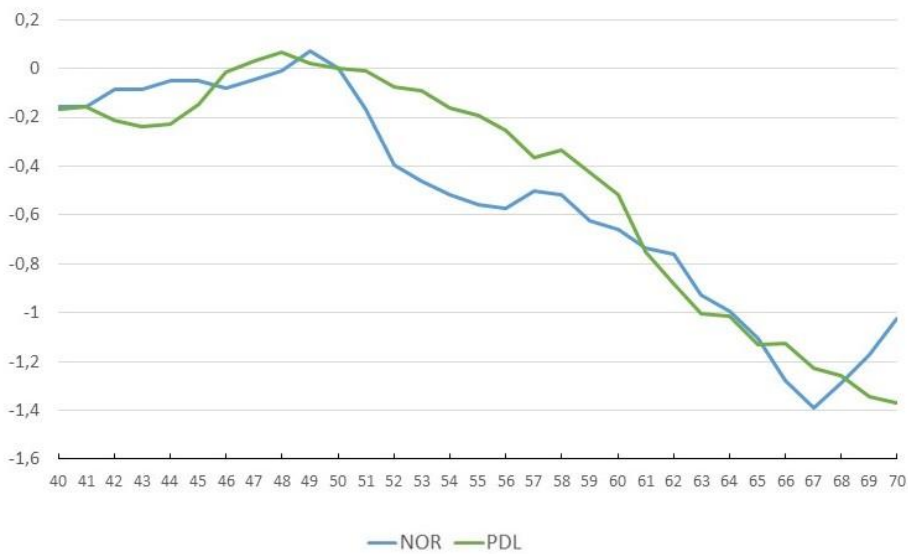


Figure 2. Effect of THI on daily Milk Yield in 2 different regions (French Holstein breed, 1st lactation)