Grant agreement number: 101000226

H2020 – Research and Innovation Action

Deliverable 3.2

Report on the trajectory of genetic and genomic correlations across heat conditions

Due date of deliverable: M36 – May 2024

Actual submission date: M36 – May 2024

WP concerned	WP3	
Lead partner	Didier BOICHARD	INRAE
Author(s)	Sophie AGUERRE, Sophie MATTALIA, Julie PROMP, Roxane VALLEE	IDELE
	Beatriz CASTRO-DIAS-CUYABANO, Aurélie VINET	INRAE
	María Jesús CARABAÑO, Clara DIAZ, Mohamed RAGAB	INIA-CSIC
	Manuel RAMON	IRIAF
	Mario CALUS, Han MULDER	WU
	Jeremie VANDENPLAS	WR
Version	V1	
Dissemination level	PU: Public	

Call: H2020-SFS-2020-2

Topic: SFS-13-2020 ı 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

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°101000226.

Table of Contents

Summary 3
1 Introduction 4
2 Description of dairy cattle and meteorological data 4
2.1 Breeds, national phenotypic data, genotypes and pedigree 4
2.1.1 Breeds 4
2.1.2 Milk production 4
2.1.3 Fertility trait 4
2.2 Meteorological information 5
3 Choice of heat stress indicators 5
4 Estimation of genetic correlations between protein yield and conception rate 6
4.1 Methods 6
4.2 Results 10
4.2.1 Genetic correlations between levels and slopes 10
4.2.2 Genetic correlations between Protein Yield and Conception Rate 12
5 Conclusion 13
6 Reference 14

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. The aim of this deliverable D3.2 is to estimate the trajectory of genetic correlations between production and fertility in the Dutch, French and Spanish Holstein cows, as well as in the French Montbéliarde cows. This work is a follow up of D3.1 and uses a similar methodology and the same sources of data.

More precisely, large-scale performance and pedigree 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. The meteorological indicators were the temperature-humidity index (THI) averaged over three days on and previous test dates for milk production, and averaged over eight days on and after the insemination date for fertility traits. These indicators were the same as in D3.1.

Because fertility has a very low heritability, accurately estimating genetic correlations between milk production and fertility requires a very large amount of data. Estimating variations in these genetic correlations with THI requires even more data, and much more than for production only as in D3.1. Ten years of data were considered in each country. The model was similar to the one used in D3.1 but was extended to consider production and fertility simultaneously. This random regression multi-trait model is computationally very challenging and several simplifications were done. More specifically, a sire model was used. In addition, in France and the Netherlands and for production, only test-day records between 80 and 200 days in milk were analyzed in order to concentrate on the part of the lactation which has the most stable genetic determinism. This restriction makes it possible to get rid of the random regression as a function of days in milk and to estimate only those that are a function of THI.

Estimates of genetic correlations between production and fertility were moderately to strongly negative according to countries. When THI increased from 50 to 70, the antagonism tended to increase slightly for MON-FR and strongly for HOL-NL. On the other hand, the antagonism remained nearly constant or decreased for the Holstein populations in France and Spain. The slopes were thought to be appealing criteria of tolerance to heat, but they have limited variability. The correlations between PY level and PY and CR slopes were negative in all populations. Correlations between slopes at THI 70 were always positive, showing that tolerance measured by productive or CR decay under heat stress tended to agree.

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 aim of this deliverable D3.2. is to estimate the trajectory of genetic correlations between production and fertility in the Dutch, French and Spanish Holstein cows, as well as in the French Montbéliarde cows. Because this work requires a lot of data to obtain meaningful estimates, it was not carried out in the MRY breed which is too small.

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 one local breed, Montbéliarde. Data for Holstein are available in France, Spain, and the Netherlands. Data for Montbéliarde are only available in France.

2.1.2 Milk production

This study focused on protein yield (g/day or kg/day) at individual test-day in first lactation. Test-day records of French, Spanish and Dutch Holstein cows, and of French Montbéliarde 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 a period of 10 years, from 2010 to 2020. Detailed description of the data in each country is presented in Table 1. Around 10, 0.7 and 0.5 million test-day records originated from 3, 0.1 and 0.3 million Holstein cows from France, Spain, and the Netherlands, respectively. A total of 2 million test-day records from 650,000 French Montbéliarde cows were available.

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 or failure (1/0) of the first artificial insemination (AI) in the first lactation. AI data were extracted from the French, Spanish, and Dutch national data bases. The CR phenotype was determined as follows: the outcome was set to 100 (successful) if the cow calved following the expected gestation length of the breed (e.g., 281 and 285 +/- 15 days in Holstein and Montbeliarde, respectively), and 0 otherwise.

All datasets covered a period of 10 years, starting from 2010. For the Holstein populations in France, Spain, and the Netherlands, the datasets included around 3.4, 0.3, and 0.8 millions of Holstein cows, respectively. Around 650,000 CR status of French Montbéliarde cows were available. More details on the data sets are shown in Table 1.

Table 1. Number of cows and records, number of sires, means and standard deviations (s.d.) of conception rates (CR) and protein yields (PY), and means and s.d. of average temperature-humidity index (THI) by trait in the five populations.

	HOL-FR	HOL-SP	HOL-NL	MON-FR
Number of cows	3,368,605	302,152	289,268	656,164
Number of CR	3,351,068	278,588	224,026	649,814
Number of PY	10,245,692	703,574	497,127	1,966,985
Number of sires	5,463	1,842	3,477	1,612
Average CR %	44.5 (49.7)	45.4 (49.8)	46.01 (49.8)	54.6 (49.8)
Average PY g/d	839 (176)	1,090 (211)	991 (187)	740 (164)
Average THIf (s.d.)	51.3 (8.9)	56.2 (8.6)	50.7 (9.5)	47.4 (10.6)
Average THIp (s.d.)	50.7 (9.1)	58.8 (9.6)	50.8 (9.9)	47.3 (10.7)

THIp: average THI in the 3 days leading up to the test day; THIf: average THI in the 7 days following insemination

2.2 Meteorological information

As for D3.1, 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):

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

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 2 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 ≥ 60 or even ≥ 72) could be found in each country. Not surprisingly, however, extreme hot conditions (THI>70) were more common in Spain but represented only 1-2% of the data in both other countries.

3 Choice of heat stress indicators

Following the conclusions of task 3.1 described in D3.1, heat stress indicators were defined by the daily THI averaged over a period of 3 days preceding a given test-day for protein yield (including the

date of the test). For conception rate, the general policy was similar but the daily THI was averaged over the period of 7 days following the insemination (including the day of insemination). These periods were found to be the most representative of the stress, according to the negative effect on the performances. These indicators were called THIp for protein yield and THIf for conception rate.

4 Estimation of genetic correlations between protein yield and conception rate

4.1 Methods

To estimate genotype-by-THI interactions within traits and the evolution of genetic correlations between traits across various THI conditions, bivariate random regression models were used.

Random regression models enable prediction of the performance of each genotype under a given environmental condition. The main objective was to measure the trend in the genetic correlation between CR and PY along a broad THI gradient. Because CR has a very low heritability, very large datasets are required to accurately estimate the varying genetic correlations over the THI gradient. Using a random regression model applied to millions of records, as needed to accurately estimate the trends in genetic correlations over the THI gradient, pose huge computational constraints when applied to an animal model. Therefore, we opted for a sire model, which made it computationally feasible to execute the study handling all the performance data spanning a wide range of THI conditions. In addition, since each cow had only one CR record, a sire model was deemed more appropriate for describing the effect of the THI gradient in the progeny group.

For test-day production records, it is generally recommended to use a random regression model to account for the varying genetic determinism of PY along the lactation. In WP3.1, we used a two-dimension random regression model with components that were dependent on both days in milk (DIM) and THI. Here, with a CR-PY bivariate model and a very large data set, we had to avoid this complexity and excluded the random regression function of DIM. In France and the Netherlands, we focused on PY records in the middle of the lactation (80-200 DIM), when genetic parameters are fairly stable whereas in Spain the full lactation was considered.

For protein yields, random regressions were considered with Legendre polynomials. For CR, two kinds of random regression were considered, either Legendre polynomials (in France) or Broken line (in Spain and the Netherlands). Legendre polynomials are widely used in the literature, they are flexible to fit any kind of trajectory, can be statistically interpreted, but they are prone to border effects, especially when few data are available. The broken line model assumes a constant genetic level up to a threshold (i.e. no G x E), and then a linear effect of the environmental factor, namely THI. Its biological interpretation is straightforward, but it is less flexible than Legendre polynomials as the general shape of the trajectory is fixed. In addition, it assumes the threshold to be known.

With Legendre polynomials, the following models were used for CR and PY in France, respectively:

$$y_{f_i} = \mathbf{x_{f_i}}\,\boldsymbol{\beta_f} + \sum_{k=0}^{2} z_{f_{jk}}\, a_{f_{s(i)k}} + e_{f_i}, \quad (1)$$

$$y_{p_{in}} = \mathbf{x_{p_{in}}}\,\boldsymbol{\beta_p} + \sum_{k=0}^{3} z_{p_{jkn}}\, a_{p_{s(i)k}} + \sum_{k=0}^{3} w_{p_{jkn}}\, p_{p_{ik}} + e_{p_{in}}, \quad (2)$$

with y_{f_i} and $y_{p_{in}}$ the CR and the *n*th PY phenotype of cow *i*, $\boldsymbol{\beta_f}$ and $\boldsymbol{\beta_p}$ the fixed effects and $\mathbf{x_{f_i}}$ and $\mathbf{x_{p_{in}}}$ their incidence vectors, $a_{f_{s(i)k}}$ and $a_{p_{s(i)k}}$ the *k*th genetic components of sire *s(i)*, $p_{p_{ik}}$ the *k*th animal component of the cow *i* (including the permanent environment effect and a part of the

genetic effect) for PY, e_{f_i} and $e_{p_{in}}$ the residuals, $z_{f_{jk}}$ the *k*th Legendre polynomial of standardized THIf *j*, $z_{p_{jkn}}$ and $w_{p_{jkn}}$ the kth Legendre polynomials of standardized THIp *j*.

The additive genetic effect $a_{f_{s(i)k}}$ was regressed using three functions: an intercept independent of THI and two THI-dependent functions (linear and quadratic functions). For CR, the vector $\mathbf{z_f}$ contains three non-zero terms, equal to a constant and the first- and second-order Legendre polynomials of standardized THIf, respectively, and $\mathbf{a_f}$ the corresponding vector of additive sire regression coefficients, contains three values per animal. For PY, the additive sire effects $\mathbf{a_{p_{s(i)}}}$ and the permanent environmental effect of the cow $\mathbf{p_{p_i}}$ were each regressed using four functions: an intercept independent of THI and three THI-dependent functions (linear, quadratic, and cubic functions). The vectors $\mathbf{z_p}$ and $\mathbf{w_p}$ contain four non-zero terms on each line, equal to a constant and the first-, second-, and third-order Legendre polynomials of standardized THIp, respectively. $\mathbf{a_{p_{s(i)}}}$ and $\mathbf{p_{p_i}}$, the corresponding vector of additive sire regression coefficients and vector of permanent environmental regression coefficients contain each four values per animal (*i.e.,* per sire *s(i)* for $\mathbf{a_{p_{s(i)}}}$ and per cow *i* for $\mathbf{p_{p_i}}$). The coefficients of the polynomials were standardized values of THI in the interval -1 to +1.

In preliminary French analyses, different orders of Legendre polynomials were tested in Montbeliarde breed. For each trait, models with different orders of Legendre polynomials (order two and three, featuring three and four polynomial coefficients respectively) were compared using the Akaike information (AIC) and the Bayesian information criteria (BIC). The results with respect to these criteria of model assessment indicated that the best model was that of a third order polynomials for PY (cubic, with four polynomial coefficients – for order 0, 1, 2, and 3), and that of a second order polynomials (quadratic, with three polynomial coefficients – for order 0, 1, and 2) for CR (Vinet et al, 2024).

Note that for PY in Spanish and Dutch Holsteins a similar model was used, albeit that in those cases the random regression were only second order, i.e. they contained up to quadratic functions of THI, but no cubic function.

For Spanish and Dutch Holsteins, the broken line model used for CR was the following:

$$y_{f_i} = \mathbf{x_{f_i}}\,\boldsymbol{\beta_f} + \sum_{k=0}^{1} z_{f_{jk}}\, a_{f_{s(i)k}} + e_{f_i},$$

with $z_{f_{j0}} = 1$ and $z_{f_{j1}} = \text{THIf} - t$ for THIf>t and 0 otherwise. The threshold t was set to 66 in Spain and 60 in the Netherlands, following conclusions of task 3.1.

The genetic parameters were then derived with the following formulae. **They are given here for the Legendre polynomials model used for the French data, those for broken line can be obtained in a similar and straightforward way.**

With 4 (for PY) + 3 (for CR) = 7 random regressions, the sire random regression variance matrix var(**a**) = **S** is a 7x7 symmetric matrix:

$$S = \begin{bmatrix} S_p & S_{pf} \\ S_{fp} & S_f \end{bmatrix}$$

$$\text{with } \mathbf{S_p} = \begin{bmatrix} \sigma^2_{s_{p_0}} & \sigma_{s_{p_0 p_1}} & \sigma_{s_{p_0 p_2}} & \sigma_{s_{p_0 p_3}} \\ \sigma_{s_{p_1 p_0}} & \sigma^2_{s_{p_1}} & \sigma_{s_{p_1 p_2}} & \sigma_{s_{p_1 p_3}} \\ \sigma_{s_{p_2 p_0}} & \sigma_{s_{p_2 p_1}} & \sigma^2_{s_{p_2}} & \sigma_{s_{p_2 p_3}} \\ \sigma_{s_{p_3 p_0}} & \sigma_{s_{p_3 p_1}} & \sigma_{s_{p_3 p_2}} & \sigma^2_{s_{p_3}} \end{bmatrix},$$

$$\mathbf{S_f} = \begin{bmatrix} \sigma^2_{s_{f_0}} & \sigma_{s_{f_0 f_1}} & \sigma_{s_{f_0 f_2}} \\ \sigma_{s_{f_1 f_0}} & \sigma^2_{s_{f_1}} & \sigma_{s_{f_1 f_2}} \\ \sigma_{s_{f_2 f_0}} & \sigma_{s_{f_2 f_1}} & \sigma^2_{s_{f_2}} \end{bmatrix}$$

$$\mathbf{S_{pf}} = \begin{bmatrix} \sigma_{s_{p_0 f_0}} & \sigma_{s_{p_0 f_1}} & \sigma_{s_{p_0 f_2}} \\ \sigma_{s_{p_1 f_0}} & \sigma_{s_{p_1 f_1}} & \sigma_{s_{p_1 f_2}} \\ \sigma_{s_{p_2 f_0}} & \sigma_{s_{p_2 f_1}} & \sigma_{s_{p_2 f_2}} \\ \sigma_{s_{p_3 f_0}} & \sigma_{s_{p_3 f_1}} & \sigma_{s_{p_3 f_2}} \end{bmatrix}$$

where $\sigma^2_{s_{p_i}}$ is the variance of the *i*th regression coefficient of the Legendre polynomial for PY, $\sigma^2_{s_{f_i}}$ is the variance of the *i*th regression coefficient of the Legendre polynomial for CR, $\sigma_{s_{p_i p_j}}$ is the covariance between coefficients *i* and *j* of the regression for PY, $\sigma_{s_{f_i f_j}}$ is the covariance between coefficients *i* and *j* of the regression for CR, and $\sigma_{s_{p_i f_j}}$ is the covariance between the coefficient *i* of the regression for PY and the coefficient *j* of the regression for CR.

The additive sire variances and covariances at each THIp or THIf were estimated by pre- and post-multiplying the sire variance matrix **S** by the corresponding THI-coefficients of the Legendre polynomials using the formulas:

$$\text{var}S_{PY_{thi_{p1}}} = \mathbf{z}_{p_{thi_{p1}}} \mathbf{S_p} \mathbf{z}'_{p_{thi_{p1}}}, \quad (3)$$

$$\text{var}S_{CR_{thi_{f1}}} = \mathbf{z}_{f_{thi_{f1}}} \mathbf{S_f} \mathbf{z}'_{f_{thi_{f1}}}, \quad (4)$$

$$\text{cov}S_{(PY_{thi_{p1}};\, PY_{thi_{p2}})} = \mathbf{z}_{p_{thi_{p1}}} \mathbf{S_p} \mathbf{z}'_{p_{thi_{p2}}}, \quad (5)$$

$$\text{cov}S_{(CR_{thi_{f1}};\, CR_{thi_{f2}})} = \mathbf{z}_{f_{thi_{f1}}} \mathbf{S_f} \mathbf{z}'_{f_{thi_{f2}}}, \quad (6)$$

$$\text{cov}S_{(PY_{thi_{p1}};\, CR_{thi_{f2}})} = \mathbf{z}_{p_{thi_{p1}}} \mathbf{S_{pf}} \mathbf{z}'_{f_{thi_{f2}}}, \quad (7)$$

The additive genetic variances $\text{var}G_{PY_{thi_{p1}}}$ and $\text{var}G_{CR_{thi_{f1}}}$ and covariances at each THIp/f, $\text{cov}G_{(PY_{thi_{p1}};\, PY_{thi_{p2}})}$, $\text{cov}G_{(CR_{thi_{f1}};\, CR_{thi_{f2}})}$, and $\text{cov}G_{(PY_{thi_{p1}};\, CR_{thi_{f1}})}$, were estimated by multiplying the corresponding sire additive variances and covariances by 4.

With $\text{var}G_{PY_{thi_{p1}}}$, the genetic variance of PY was evaluated at $THIp_1$; with $\text{var}G_{CR_{thi_{f1}}}$, the genetic variance of CR was evaluated at $THIf_1$; $\text{cov}G_{(PY_{thi_{p1}};\, PY_{thi_{p2}})}$ was the genetic covariance between PY at $THIp_1$ and PY at $THIp_2$; $\text{cov}G_{(CR_{thi_{f1}};\, CR_{thi_{f2}})}$ was the genetic covariance between CR at $THIf_1$ and CR at $THIf_2$; and $\text{cov}G_{(PY_{thi_{p1}};\, CR_{thi_{f2}})}$ was the genetic covariance between PY at $THIp_1$ and CR

at THIf$_2$. $\mathbf{z}_{p_{thi_{p1}}}$ is the vector of covariates estimated at THIp$_1$ and $\mathbf{z}_{f_{thi_{f1}}}$ is the vector of covariates estimated at THIf$_1$.

The permanent environmental random regression variance matrix **P** and the permanent environmental variances, defined only for PY, were defined similarly as follows:

$$\mathrm{varP}_{p_{thi_{p1}}} = \mathbf{z}_{p_{thi_{p1}}}\ \mathbf{P}\ \mathbf{z}'_{p_{thi_{p1}}}, \quad (8)$$

with **P** being the 4x4 symmetric permanent environmental random regression variance matrix.

Finally, $\mathrm{varR}_{PY_{thi_{p1}}}$, $\mathrm{varR}_{CR_{thi_{f2}}}$, and $\mathrm{covR}_{(PY_{thi_{p1}};\ CR_{thi_{f2}})}$ are the residual variances and covariance at THIp$_1$ and THIf$_2$, respectively.

All these variances enabled calculation of the heritability coefficients at each THIp$_i$ or THIf$_j$ as:

$$h^2_{PY_{thi_{pi}}} = \frac{4*\mathrm{varS}_{PY_{thi_{pi}}}}{\mathrm{varS}_{PY_{thi_{pi}}} + \mathrm{varP}_{PY_{thi_{pi}}} + \mathrm{varR}_{PY_{thi_{pi}}}} \quad (9)$$

$$h^2_{CR_{thi_{fj}}} = \frac{4*\mathrm{varS}_{CR_{thi_{fj}}}}{\mathrm{varS}_{CR_{thi_{fj}}} + \mathrm{varR}_{CR_{thi_{fj}}}} \quad (10)$$

and the correlations between traits at each THIp$_i$ and THIf$_j$ as:

$$r_{g(PY_{thi_{pi}};CR_{thi_{fj}})} = \frac{\mathrm{covG}_{(PY_{thi_{pi}};\ CR_{thi_{fj}})}}{\sqrt{\mathrm{varG}_{PY_{thi_{pi}}} * \mathrm{varG}_{CR_{thi_{fj}}}}}, \quad (11)$$

where $r_{g(PY_{thi_{pi}};\ CR_{thi_{fj}})}$ is the genetic correlation between PY at THIp$_i$ and CR at THIf$_j$.

The genetic correlation between the PY breeding value at THIp$_i$ and the PY slope of the regression on (standardized) THIp$_j$ was calculated as follows:

$$r_{g(PY_{thi_{pi}};\ slopePY_{thi_{pj}})} = \frac{\mathrm{covG}_{(PY_{thi_{pi}};\ slopePY_{thi_{pj}})}}{\sqrt{\mathrm{varG}_{PY_{thi_{pi}}} * \mathrm{varG}_{slopePY_{thi_{pj}}}}}, \quad (12)$$

with the variance of the PY slope at THIp$_j$ computed using the derivative formula:

$$\mathrm{varG}_{slopePY_{thi_{pj}}} = 4 * \frac{\mathbf{dz}_{p_{thi_{pj}}}}{\mathrm{dTHI}_p}\ \mathbf{S_p}\ \left(\frac{\mathbf{dz}_{p_{thi_{pj}}}}{\mathrm{dTHI}_p}\right)' \quad (13)$$

and the additive genetic covariance between the PY breeding value at THIp$_i$ and the PY slope at THIp$_j$ computed as:

$$\mathrm{covG}_{(PY_{thi_{pi}};\ slopePY_{thi_{pj}})} = 4 * \mathbf{z}_{p_{thi_{pi}}}\ \mathbf{S_p}\ \left(\frac{\mathbf{dz}_{p_{thi_{pj}}}}{\mathrm{dTHI}_p}\right)' \quad (14)$$

The same formula was used to estimate the genetic correlation between the CR breeding value and the CR slope at two different values of THIf by adapting the matrix **S**, the vector **Z**, and the derivative with regard to dTHIf.

Between traits, the correlation between the PY breeding value at $THIp_i$ and the CR slope of the regression at $THIf_j$ was calculated as:

$$r_{g\left(PY_{thi_{pi}};slopeCR_{thi_{fj}}\right)} = \frac{covG_{\left(PY_{thi_{pi}};slopeCR_{thi_{fj}}\right)}}{\sqrt{varG_{PY_{thi_{pi}}} * varG_{slopeCR_{fj}}}}, \quad (15)$$

4.2 Results

Random regression models used to investigate genotype-by-THI interactions allowed us to estimate additive genetic variances, heritabilities and genetic correlations for PY and CR across combinations of THIp and THIf. Heritabilities and genetic variances of PY and CR according to THIp and THIf were already presented in D3.1. As a quick summary: THI had a strong negative effect on all traits; with increasing THI genetic variances and heritability coefficients decreased for PY, and increased for CR; within trait, genetic correlations between different THI were high, therefore showing little GxE interaction, but these correlations were even higher for PY than for CR.

In this deliverable we present only new results: (a) genetic correlations between levels and slopes of PY and CR; (b) genetic correlations between PY and CR.

4.2.1 Genetic correlations between levels and slopes

For a given trait, the slope at a given THI is defined by the derivative of the breeding value with respect to THI. When estimated at high THI, it reflects how sensitive the breeding value is to changes in THI. It is therefore an indicator of tolerance. A positive slope reflects that the breeding value increases with THI, i.e. the raw performance is less affected by an increase in THI than the average of the population. Note that the performance very likely is affected, as the breeding value models the deviation from the overall trajectory across THI values.

The genetic standard deviations of the slopes are given in Table 2. On average, they were quite small, showing little genetic determinism of these indicators of tolerance. For PY, they were rather similar, as assessed by the ratio of the slope genetic standard deviation on the standard deviation of level at THI50, which was around 3%. For CR, they were more different, as the ratio varied from 2.7 to 11.2%, but these estimations are rather inaccurate.

The estimates of genetic correlations between levels and slopes are shown in Table 3. For PY, the correlation was negative, showing that high-yielding animals in present conditions will be less tolerant in future hotter conditions, under the current selection goals. This opposition was found to be moderate to high. This lower tolerance of high-yielding cows was also visible with the moderately negative correlation between PY level and CR slope. For CR, the correlation between level and slope was low to moderately positive showing that genetic differences in future conditions will be similar or larger than current genetic differences. Not surprisingly, the situation is very similar for CR level and PY slope with a generally positive correlation. Therefore, high fertility today is neutral or favourable to tolerance to THI, measured with PY or CR slopes.

Table 2. Genetic standard deviation of slopes at THI 70 for protein yield and conception rate in the 4 populations studied.

	HOL-FR	HOL-SP	HOL-NL	MON-FR
PY slope (g/day.°C)	1.14	1.07	1.43	1.00
PY level (g/day)	42	36	47	36
PY slope/PY level ratio (%)	2.7	3.0	3.0	2.8
CR slope (%/°C)	0.35	0.73	0.25	0.19
CR level (%)	8.8	6.2	3.7	7.1
CR slope / CR level ratio (%)	3.9	11.8	6.6	2.7

Finally, the correlation between PY and CR slopes was weak but slightly positive, showing that tolerance indicators measured by slopes are different but correlated and therefore compatible traits.

Table 3. Estimates of genetic correlations between level at THI 50 and slopes at THI 70 for protein yield and conception rate, and between slopes, in the 4 populations studied.

Traits	HOL-FR	HOL-SP	HOL-NL	MON-FR
PY - PY slope	-0.37 (0.06)	-0.32 (0.07)	-0.47 (0.14)	-0.71 (0.09)
PY – CR slope	-0.15 (0.08)	-0.18 (0.10)	-0.58 (0.29)	-0.54 (0.35)
CR – CR slope	-0.03 (0.09)	-0.10 (0.09)	-0.05 (0.27)	+0.37 (0.38)
CR – PY slope	+0.33 (0.06)	+0.23 (0.09)	-0.27 (0.17)	+0.20 (0.13)
PY slope – CR slope	+0.24 (0.14)	+0.33 (0.09)	+0.42 (0.42)	+0.27 (0.45)

4.2.2 Genetic correlations between Protein Yield and Conception Rate

In current conditions, PY and CR are antagonistic. This relationship has been known for many years and is attributed to a trade-off between both traits. Selection on yields has oriented cow metabolism for high productivity at the expense of a deterioration of her reproduction and health. The antagonism with fertility is higher with milk yield than with PY, probably because of the lower output of lactose for the same output of protein. It is also higher with total lactation yield than with yield in the first half of the lactation, because lactation yield is strongly affected by gestation and because cows with a delayed gestation produce more. However, even for PY in the first half of the lactation, the genetic correlation with CR is negative and around -0.3 to -0.4.

When this project was initially designed, it was anticipated that heat stress would increase the trade-off between PY and CR, i.e. increase the unfavourable genetic correlation. Should this hypothesis be verified, the antagonism between production and reproduction would be reinforced and make bovine production and selection even more difficult under heat stress. Therefore, the goal of this study was to address this question and provide estimates of this genetic correlation between traits under heat stress conditions that could be anticipated in the mid-term future.

Table 4. Estimates of genetic correlations between protein yield and conception rate levels at THI 50 and 70

Traits	HOL-FR	HOL-SP	HOL-NL	MON-FR
PY50 – CR50	-0.14 (0.02)	-0.14 (0.06)	0.06 (0.07)	-0.03 (0.06)
PY50 – CR70	-0.18 (0.03)	-0.16 (0.07)	-0.28 (0.13)	-0.20 (0.08)
PY70 – CR50	-0.06 (0.03)	-0.11 (0.06)	0.05 (0.07)	+0.01 (0.06)
PY70 – CR70	-0.08 (0.04)	0.02 (0.08)	-0.28 (0.13)	-0.16 (0.08)

The estimates of the genetic correlation between traits at THI 50 and 70 are shown in Table 4. These correlations, estimated on very large datasets, are relatively accurate. Most estimates are slightly negative but more moderate than expected. This can be partly explained by the choice of the trait (PY and not milk yield) and by the lactation stage studied, i.e. mid-lactation from 80 to 200 days for the French and Dutch analyses. This period does not include the beginning of the lactation when the antagonism is likely to be stronger, nor the end of the lactation affected by gestation.

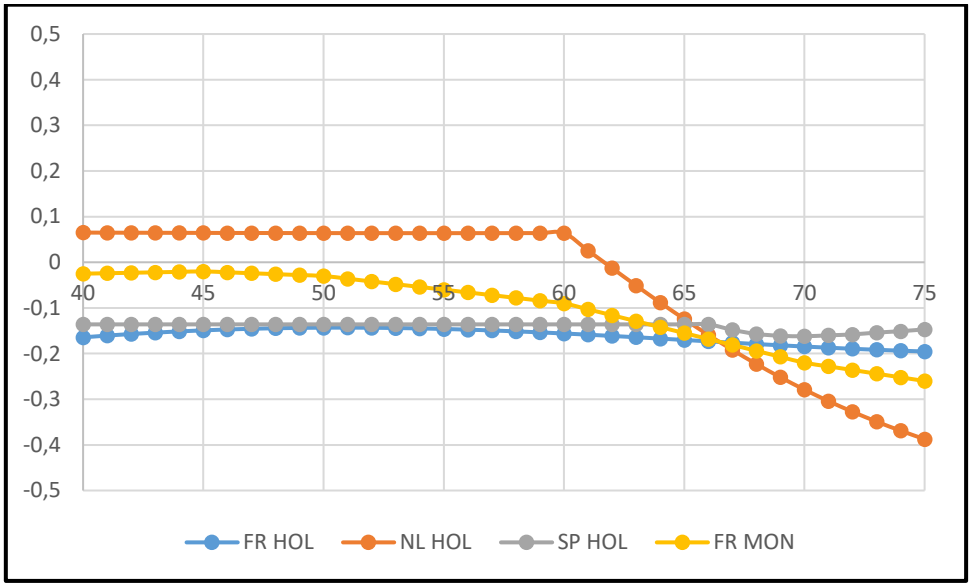
As expected, the antagonism between PY 50 and CR increases from THI 50 to 70. However, this trend is limited and likely can be overcome by selection and adaptation of the production system. Surprisingly, the corresponding correlations between PY 70 and CR 70 are not stronger.

Table 4 provides only 4 estimates per population. But estimates are available for any combination of two THIp and THIf. Figure 1 shows the trajectory of genetic correlations when THIp=THIf. And Figure 2 shows the trajectory of genetic correlations between PY at THIp=50 and CR at THIf values ranging from 40 to 75.

The results are not fully consistent across populations. The different pattern observed between the Dutch and the other populations is likely due to using the broken line model, with a discontinuity at THI 60 (Netherlands) or 66 (Spain). Between PY 50 and CR at different THI, the estimates are not very different. They are all slightly negative, although the Dutch estimates are slightly positive at THI values below 60. They tend to become more negative when THIf increases.

The discrepancies are larger for the genetic correlation between PY and CR at the same THI (Figure 2). In two populations, the correlations increase with THI whereas they decrease in the two other populations. In the French populations, they remain slightly negative. In the Spanish population, the correlation becomes increasingly positive. This was not anticipated. A possible biological interpretation is that in hot conditions, good or poor adaptation of cows leads to both good or both poor production and fertility. Another interpretation is a possible habituation of the cows. But this interpretation needs to be confirmed. In the Netherlands, the correlation becomes increasingly negative. A possible interpretation is an effect of the broken line model which may force the slope to be very negative. Nevertheless, should this strong antagonistic correlation be confirmed, keeping a good production and a good fertility would become more and more difficult in the future.

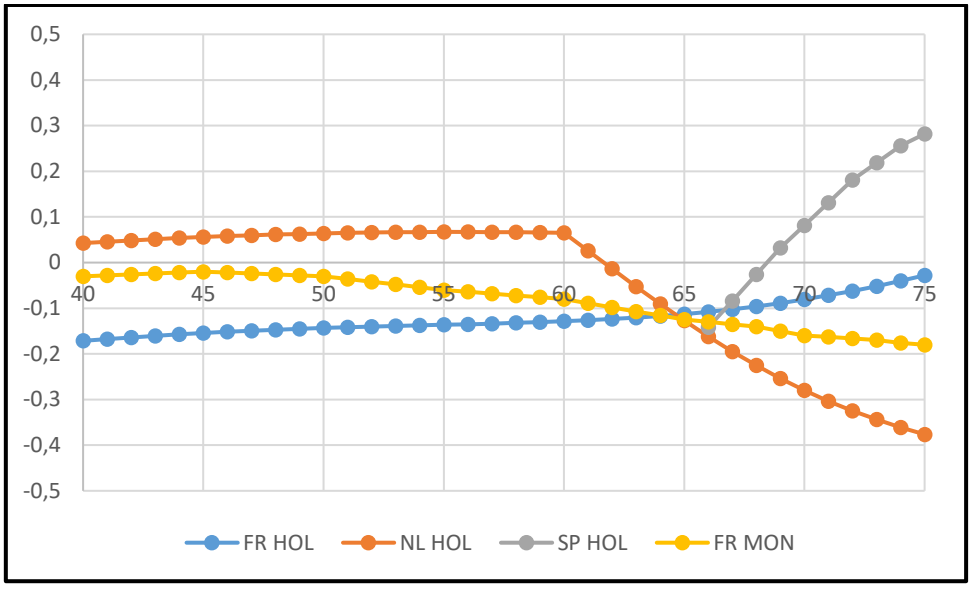
Figure 1. Trajectories of genetic correlations between PY at THI 50 and CR along the THIf, in four populations



5 Conclusion

Under thermoneutral conditions estimates of genetic correlations between production and fertility were moderately to strongly negative according to countries, but on average less antagonistic than anticipated, probably because of the PY definition, in mid-lactation. When THI increased from 50 to 70, the antagonism tended to increase slightly for MON-FR and strongly for HOL-NL. On the other hand, the antagonism remained nearly constant or decreased for the Holstein populations in France and Spain. The slopes were initially thought to be an appealing criterium of tolerance to heat. However, this conclusion is not so clear, because slopes have limited variability, poor accuracy prediction (especially for CR and high THI). As expected, the correlations between PY level and PY and CR slopes were negative in all populations. Correlations between slopes at THI 70 were always positive, showing that tolerance measured by productive or CR decay under heat stress tended to agree. The main discrepancy between populations was the genetic correlations between PY and CR at high THI and the reasons for these diverging results are not clear. They can be a consequence of the different models used (broken line vs Legendre polynomials), or different adaptation of the animals between countries, or simply the result of inaccurate estimation at high THI due to limited amount of data in very hot conditions.

Figure 2. Trajectories of genetic correlations between PY and CR at the same THI in four populations



6 Reference

Vinet A., Mattalia S., Vallée R., Bertrand C., Barbat A., Promp J., Cuyabano BCD, Boichard D. 2024. Effect of temperature-humidity index on the evolution of trade-offs between fertility and production in dairy cattle. Genet Sel Evol, 56:23. https://doi.org/10.1186/s12711-024-00889-4

.