Department of Organic Food Quality and Food Culture
Faculty of Organic Agricultural Sciences
University Kassel
Witzenhausen

The Effect of Climate Variability on Main Components of Cow Milk in Iran

Dissertation
by
Mohammad Reza Marami Milani

September 2016

submitted to the Faculty of Organic Agricultural Sciences, Section of Organic Food Quality and Food Culture of the University of Kassel to fulfil the requirements for the degree Doktor der Agrarwissenschaften (Dr. agr.)

1. Supervisor: Prof. Dr. Angelika Ploeger
2. Supervisor: Prof. Dr. Andreas Hense

This PhD thesis is submitted to the Faculty of Organic Agricultural Sciences, Section of Organic Food Quality and Food Culture (Fachbereich 11) of the university of kassel to complete the requirements for the degree Doktor der Agrarwissenschaften **(Dr. agr.)**

It is based on three papers published to international refereed journal that are included in chapter 2, 3 and 4.

Chapter 1: gives a general introduction and short overview on methodology used in the work.

The following papers are enclosed:

chapter 2: A Pilot Investigation of the Relationship between Climate Variability and Milk Compounds under the Bootstrap Technique. Foods, 420-439 (doi:10.3390/foods4030420).

chapter 3: A Survey of the Relationship between Climatic Heat Stress Indices and Fundamental Milk Components Considering Uncertainty. Climate, 876-900 (doi:10.3390/cli3040876).

chapter 4: Applying Least Absolute Shrinkage Selection Operator and Akaike Information Criterion Analysis to Find the Best Multiple Linear Regression Models between Climate Indices and Components of Cow's Milk. Foods, 52-68 (doi:10.3390/foods5030052).

Chapter 5: is a summary of results and discussion of chapter 2, 3 and 4.

Chapter 6: introduces the used references in this study.

This dissertation is ended by acknowledgment and references.

Hereby I certify that I have done and written my thesis independently and used no other resources than are presented in the reference.

Kassel, September 2016

Tag der mündlichen Prüfung

17.11.2016

Gutachter

Prof. Dr. Angelika Ploeger
Universität Kassel

Prof. Dr. Andreas Hense
Universität Bonn

Abstract

The main purpose of this study is to assess the relationship between six bioclimatic indices for cattle (temperature humidity (*THI*), environmental stress (*ESI*), equivalent temperature (*ESI*), heat load (*HLI*), modified heat load (HLI_{new}) and respiratory rate predictor(*RRP*)) and fundamental milk components (fat, protein, and milk yield) considering uncertainty. The climate parameters used to calculate the climate indices were taken from the NASA-Modern Era Retrospective-Analysis for Research and Applications (*NASA-MERRA*) reanalysis from 2002 to 2010. Cow milk data were considered for the same period from April to September when cows use natural pasture, with possibility for cows to choose to stay in the barn or to graze on the pasture in the pasturing system. The study is based on a linear regression analysis using correlations as a summarizing diagnostic. Bootstrapping is used to represent uncertainty estimation through resampling in the confidence intervals. To find the relationships between climate indices (*THI*, *ETI*, *HLI*, HLI_{new}, *ESI* and *RRP*) and main components of cow milk (fat, protein and yield), multiple liner regression is applied. The least absolute shrinkage selection operator (*LASSO*) and the Akaike information criterion (*AIC*) techniques are applied to select the best model for milk predictands with the smallest number of climate predictors. Cross validation is used to avoid over-fitting.

Based on results of investigation the effect of heat stress indices on milk compounds separately, we suggest the use of *ESI* and *RRP* in the summer and *ESI* in the spring. *THI* and HLI_{new} are suggested for fat content and HLI_{new} also is suggested for protein content in the spring season. The best linear models are found in spring between milk yield as predictands and *THI*, *ESI*,*HLI*, *ETI* and *RRP* as predictors with p-value < 0.001 and R^2 0.50, 0.49. In summer, milk yield with independent variables of *THI*, *ETI* and *ESI* show the highest relation (p-value < 0.001) with R^2 (0.69). For fat and protein the results are only marginal.

It is strongly suggested that new and significant indices are needed to control critical heat stress conditions that consider more predictors of the effect of climate variability on animal products, such as sunshine duration, quality of pasture, the number of days of stress (*NDS*), the color of skin with attention to large black spots, and categorical predictors such as breed, welfare facility, and management system.

This methodology is suggested for studies investigating the impacts of climate variability/change on food quality/security, animal science and agriculture using short term data considering uncertainty or data collection is expensive, difficult, or data with gaps.

Zusammenfassung

Das Ziel dieser Studie ist es, den Zusammenhang zwischen sechs bioklimatischen Indizes, die in der Rinderzucht Verwendung finden, (temperature humidity (*THI*), environmental stress (*ETI*), heat load (*HLI*), modified heat load (HLI_{new}) und respiratory rate predictor(*RRP*)) und wesentlichen Milchparametern (Fett- und Proteingehalt sowie Milchertrag) unter Berücksichtigung der vorhandenen Unsicherheiten zu bewerten. Für die Berechnung der Klimaindizes wurde der Reanalyse-Datensatz der NASA-Modern Era Retrospective-Analysis for Research and Applications (*NASA-MERRA*) im Zeitraum von 2002 bis 2010 verwendet. Die Daten für die Kuhmilchproduktion liegen für die Monate April bis September (2002-2010) vor, da Kühe in diesen Monaten natürliches Futter (innerhalb des betrachteten Futtersystem) entweder im Stall oder auf der Weide nutzen. Der Zusammenhang wird mittels linearer Regression ermittelt. Die Unsicherheiten werden durch bootstrapping abgeschätzt. Um die Beziehung zwischen den Klimaindizes (*THI*, *ETI*, *HLI*, HLI_{new}, *ESI* und *RRP*) und den Milchparametern detailliert zu evaluieren werden multiple lineare Regressions Modelle benutzt. Dabei kommen die Methoden ”least absolute shrinkage selection operator (*LASSO*)” und “Akaike information criterion (*AIC*)” zur Anwendung, um das beste Prognosemodell für Vorhersagen von Milcherträgen und –qualität mit der kleinsten Anzahl von Klimaparametern auszuwählen. Zur Vermeidung von der Überanpassung wird Kreuzvalidierung (cross validation) benutzt. Die univariate Regressionsanalyse zeigt, dass die Auswirkungen der Hitzebelastung auf die einzelnen Milchparameter durch *ESI* und *RRP* im Sommer und *ESI* im Frühling am besten erfolgt. Für den Fettgehalt wird *THI* und HLI_{new} und für den Proteingehalt wird HLI_{new} im Frühling empfohlen.

Die besten multiplen linearen Regressionsmodelle werden im Frühling zwischen den Prädikanden Milchertrag und den Prädiktoren *THI*, *ESI*, *ETI*, *HLI* und *RRP* (p-value < 0.001 und R^2 (0.50, 0.49)) identifiziert. Im Sommer zeigen Milchertrag mit den unabhängigen Variablen *THI*, *ETI* und *ESI* den größten Zusammenhang (p-value < 0.001 und R^2 (0.69)). Für den Fettgehalt und Proteingehalt kann kein signifikant besseres multiple lineares Regressionsmodell gefunden werden.

Es wird empfohlen, dass neue und aussagekräftigere Indizes entwickelt werden können, um kritische Hitzebelastungszustände zu analysieren. Der Effekt von Klimavariabilität auf Milchprodukte wird in Form von Sonnenscheindauer, Qualität des Weidelands, Anzahl der Belastungstage (*NDS*), Hautfarbe mit Berücksichtigung von großen schwarzen Flecken, und kategorischen Prädiktoren wie Zucht, Versorgung und dem Managementsystem betrachtet.

Diese vorgeschlagene Methodekombination erweist sich als sinnvoll für Studien, die die Auswirkungen von Klimaveränderung und -wandel auf Nahrungsmittelqualität/sicherheit, Nutztierwissenschaften und Landwirtschaft mit Berücksichtigung der Unsicherheiten untersuchen, die bei Daten mit geringe Stichprobenumfänge auftreten, weil die Datenerhebung teuer und schwierig ist oder generell Datenlücken vorhanden sind.

dedication

I dedicate my dissertation to my family,

My loving and wonderful wife, Elham
for her enlightening and great support during my PhD.

My sweet kids, Mani and Makan
who made my student life full with hope and power.

My dear parents,
who gave me always energy with their enthusiast.

"I would appreciate my loving family that I used their time for doing my PhD"

Contents

Figures and Tables

List of Figures

List of glossary

AIC . . . . Akaike information criterion
e . . . . Vapour pressure (hPa)
e_w* . . . . Saturation vapour pressure (hPa)
ESI . . . . Environmental stress index
ETI . . . . Equivalent temperature index
HLI . . . . Heat load index
HLI_{new} . . . . Modified HLI
LASSO . . . . Least absolute shrinkage and selection operator
p . . . . See level pressure (hPa)
q . . . . Specific humidity(g/kg)
RH . . . . Relative humidity (%)
RRP . . . . Respiratory rate predictor index
SR . . . . Solar radiation (Wm^{-2})
Ta . . . . Ambient temperature ($°C$)
Td . . . . Dew point temperature ($°C$)
TD . . . . Test day, the day that cow is milked
THI . . . . Temperature humidity index
T2m . . . . 2 meter height temperature ($°C$)
v . . . . Wind speed (m/s)

Chapter 1

Introduction

It is well documented that agriculture production, farm economy, milk yield and milk components and agricultural livestock welfare are impacted by climate variability (Gauly et al. 2013). The general increase of human population and changing amount of dairy products in human diet requires on average a two percent growth of global milk production to be able to amend the increasing demand for dairy products. Furthermore, there is no doubt that cows should live in an optimum environmental condition (animal welfare aspects) to be productive in both quantitative and qualitative aspects (Darwin 2001).

According to the report of United Nation in 2013, the population in Iran will increase about 43% by 2050 up to 115 million (United nation 2012). Consumption of milk per person in Iran is 30 to 150 kg/capital/year as reported by Food and Agriculture Organization. Due to increasing human population and urbanization, the demand for food and animal production are expected to be enhanced.

However, the definitions of optimal environmental conditions, as well as the sensitivity of milk productivity by deviations from that optimal range, are still not clear. This holds especially true for developing countries and emerging economies.

In general, because animal organisms evolve together with the environmental parameters in which they live, their physiological metabolism change during time. For instance, Holstein cows are spread in the whole world, but in the tropical regions its breeds differ in temperature tolerance compare to the Holsteins bred in other regions with different climate condition (Silva and Campos Maia 2013).

The quality and quantity of milk compounds is generally a result of complex interactions of variables. These are not fixed and they can change with the time of the year, environmental conditions, and climate variability. The milk yield of Holstein cattle is found to be more sensitive to climate than those of Jersey cattle, but Jersey milk composition is more sensitive to climatic influences (Sharma et al. 1983). Coping with climate variability is easier for livestock than for crops due to their adaptation ability and resistant to climate change because of its mobility and access to feed (Reddy 2015). Aim of this research is investigating the effect of climate variability/change on cow milk components.

The previous studies in the last two decades, mostly focused on climatic parameters and indices related to temperature and humidity as more common used parameters and temperature humidity index (THI) as the most widespread indicator of heat stress (Milani et al. 2015; hammami et al. 2015). THI is a combination of temperature and humidity to build a representative index for thermal stress which is an important factor for livestock.

For including humidity and temperature as effective parameters on milk productivity, combinations such as the temperature humidity index (THI) have been suggested as an indicator for thermal stress (Gaughan et al. 1999; Thom 1959). However, THI does not consider the loss of energy by passive cooling, (e.g.), through turbulent motion in the atmosphere. It is commonly parametrized by wind speed.

Because of complexity due to animal adaptation ability and genetic variations, more predictors are required such as solar radiation in combination with wind speed, humidity and temperature also their combinations as indicators such as temperature humidity (THI) equivalent temperature (ETI), environmental stress (ESI), heat load (HLI), modified HLI (HLInew) and respiratory rate predictor index (RRP). THI, ETI, ESI, HLI, HLI_{new} and RRP indices are used in this research to consider the interactions of temperature in two meter height (T2m), humidity (RH), wind speed (v) and solar radiation (SR) on components of cow milk.

According to this motivation and answering emphatically "YES" to the question about the need of new, better and more indices, in this study some more climatic indices are calculated from several climate parameters to be able to have more combinations of the parameters which might have an effect on the quality and quantity of milk compounds.

Another purpose of this study is to compare the effect of these relatively new heat stress indices on dairy cows under different climate conditions in Iran with cold and semi-arid climate vs Caspian, mild and humid climate vs arid and semi-arid, mild and semi-warm climate conditions.

The major question to be answered is whether, indeed, large scale patterns of meteorological parameters can be identified that provide a statistically significant and physiological relevant influence upon milk productivity in terms of yield and relevant nutrient content in the milk.

Climate data, including daily averages of temperature in two meter height, wind speed, sea level pressure, specific humidity and solar radiation were taken from the Modern Era Retrospective-Analysis for Research and Applications (MERRA), which is undertaken by NASA's global modeling and assimilation office (Rienecker et al. 2011).

Milk component data in Iran is provided for fat, protein and milk yield in spring and summer. The rather short record availability (2002–2010) for climate variability purpose requires that special statistical methods are applied to avoid misinterpretations due to over-fitting. Although a nine-year period of data is not a short term data set in agricultural investigations, in climate research nine years might represent a relatively small-sized data set.

1.1 State of the art

1.1.1 Research Aims

During 8 years work experience in dairy science, milk analysis and production, I realised the varieties of milk components, accordingly affecting on the quality of products, the formulations and causing economic consequences in buying milk and products sales. The first and main question is built in my mind to find the probable reasons for that phenomenon. To find the answer, I get motivated to do this study by investigation the effect of climate variability on milk components because it was obviously happened yearly during the change of seasons. Thus, this part presents the purposes which we have considered them in this research. The research poses some questions which we would like to verify the answers in the conclusions section based on the obtained results. 1) Investigating the statistical relation between fundamental milk components and related climate parameters. It may be used to find optimum situations to minimize the negative effect of short term climate variability and longer term climate change on milk compounds. 2) Assessing the relationship between relatively new heat stress indices and milk compounds. Also comparison between the effects of these relatively new heat stress indices and the most widespread heat stress index (THI) on cow milk under different climate conditions because previous studies in the last two decades, mostly focused on THI. 3) Introducing some good methods for data preparation and statistical analysis to cope with uncertainties in time series data sets with statistical problems such as gaps in data, expensive or difficult data collection procedure and small size of data. 4) Selecting the best model with the smallest number of predictors by applying new statistical techniques and avoiding misinterpretations due to overfitting to achieve more suitable and significant model. 5) Opening some ways for enthusiast applicators to utilize more suitable conditions in designing dairy factories according to the economic justifications and feeding system to be managed with less risk of climate variability especially in developing countries.

1.1.2 Materials and Methods

1.1.2.1 Data Condition

Data selection in the study domain is done under special conditions. In this study, we used the monthly average of test-day (TD) records of milk yield (kg), with the TD representing the day that the cow was milked. The milk yield and three-times-a-week records of fat and protein contents (g/100 mL milk) were collected from almost 600 industrial Holstein herd stations. The herd sizes varied between 75 and 200 cows, and the cows had access to grazing from April to September from 2002 to 2010. The final milk monthly averaged data bank was three seperated matrices for milk yield, fat and protein with $936,227$ individuals in each matrix. Data were gathered under the condition that the cows were on days 4 to 305 of the milking time and between their third and sixth calves over their lifetime without a record of mastitis throughout the entire study period. All of the records from cows that had a dry period or mastitis illness were omitted from the primary raw data set. Finally, the individual records from all herds within one of the three climatic zones (Figure 1.1) and the selected month were averaged to enhance a potential joint signal. Bootstrapping with a resampling N = 1000 was employed to assess the sampling uncertainty of this final data preparation step. This step basically represents a resampling with replacement among the herds/stations within one of the three climatic zones in each month of the years 2002 to 2010.

Climate data (solar radiation, two meter height temperature, dew point, relative humidity, wind speed, sea level pressure and specific humidity) were taken from the NASA-Modern Era Retrospective Analysis for Research and Applications (NASA-MERRA) data set. This data set has a nominal resolution of 1/2 degrees latitude and 2/3 degrees longitude covering the same months and years as the milk data. The temperature humidity index (THI), equivalent temperature index (ETI), Environmental Stress Index (ESI), Heat Load Index (HLI), modified HLI (HLI_{new}) and Respiratory Rate Predictor (RRP) indices were used in this study.

1.1.2.2 Data and study procedure at a glance

Study domain, data and study procedure of this study are presented in Figures 1.1, 1.2 and 1.3.

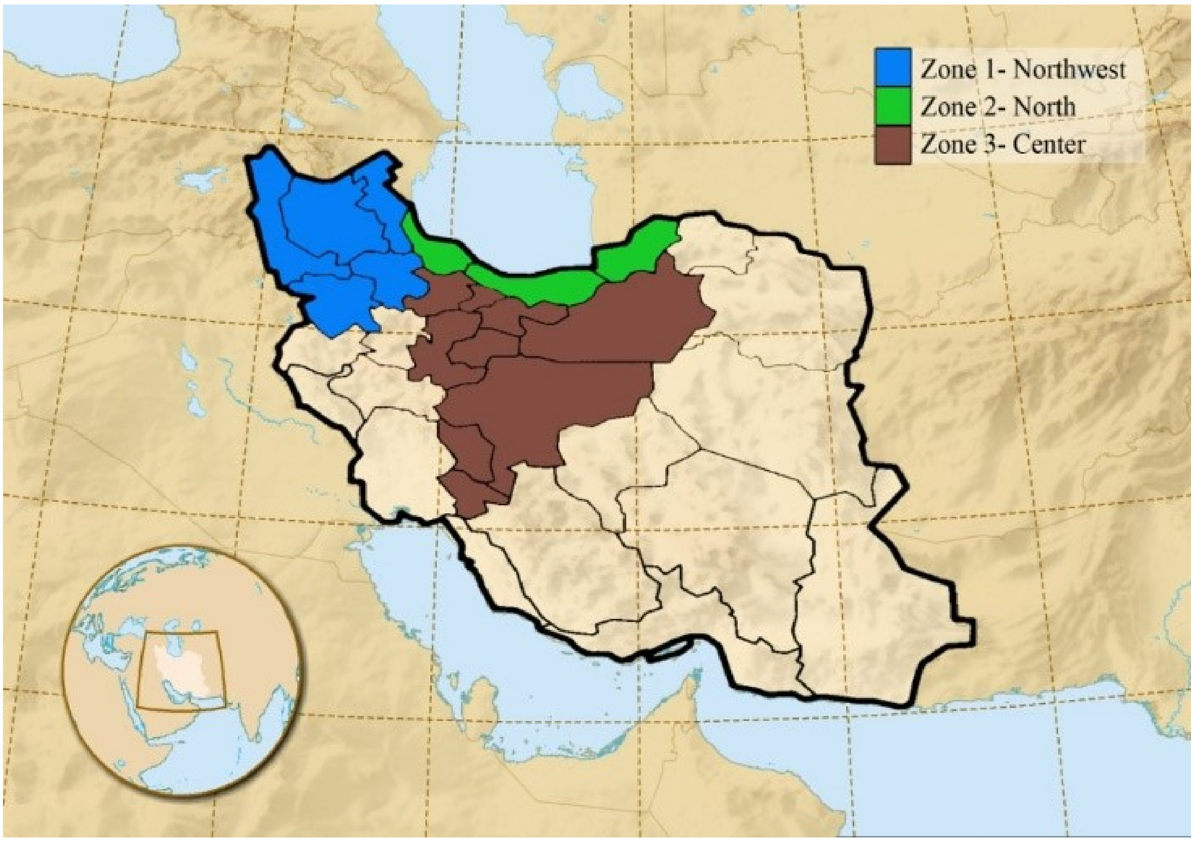


Figure 1.1: Study domain and defined zones

Data Bank: Zone {**1**, **2**, **3**}

Year	Month	All stations	Milk Data	Climate Indices
2002	{April,é ,June}	{1,2,3,é , n}	(Milk yield, Fat, Pro.) Spring	{ THI, ETI, ESI, HLI, HLI_{NEW}, RRP } Spring
2002	{July,é ,Sep.}	{1,2,3,é , n}	(Milk yield, Fat, Pro.) Summer	{ THI, ETI, ESI, HLI, HLI_{NEW}, RRP } Summer
⋮	⋮	⋮	⋮	⋮
2010	{April,é ,June}	{1,2,3,é , n}	(Milk yield, Fat, Pro.) Spring	{ THI, ETI, ESI, HLI, HLI_{NEW}, RRP } Spring
2010	{July,é ,Sep.}	{1,2,3,é , n}	(Milk yield, Fat, Pro.) Summer	{ THI, ETI, ESI, HLI, HLI_{NEW}, RRP } Summer

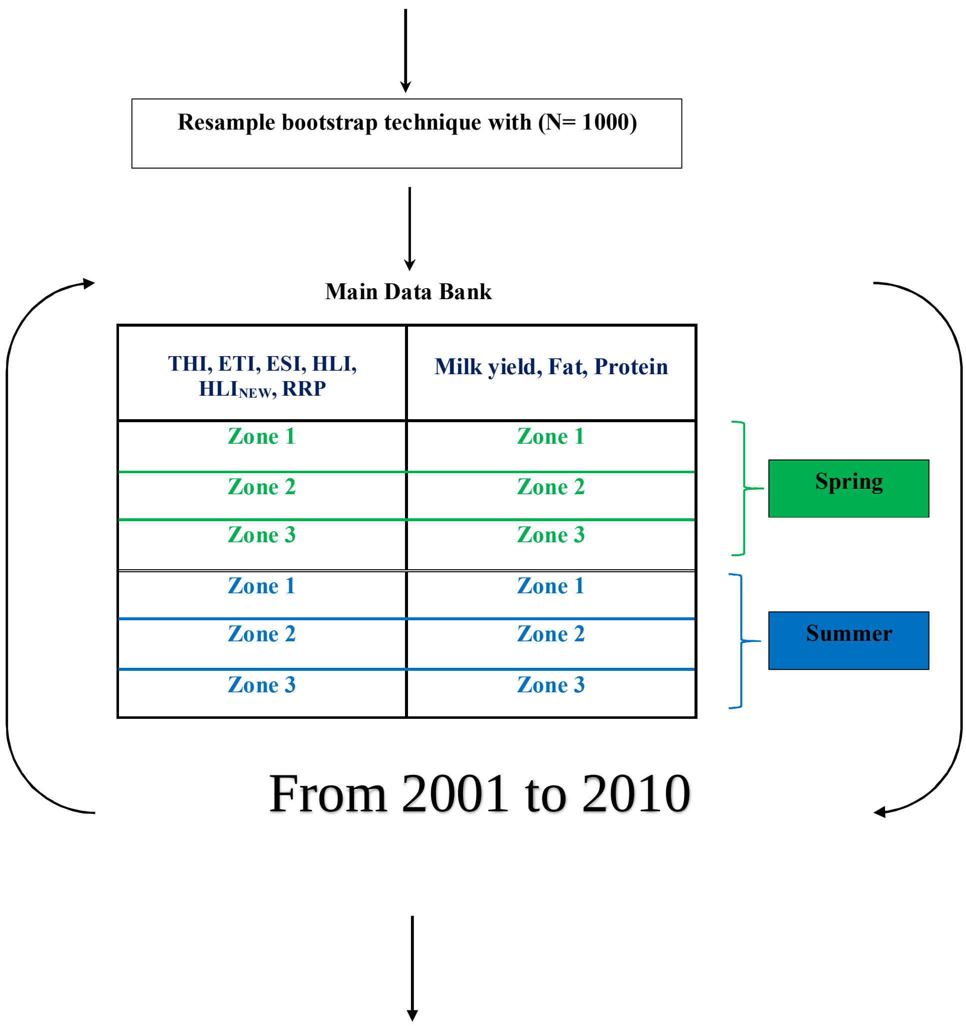


Figure 1.2: Study procedure at a glance on data structure

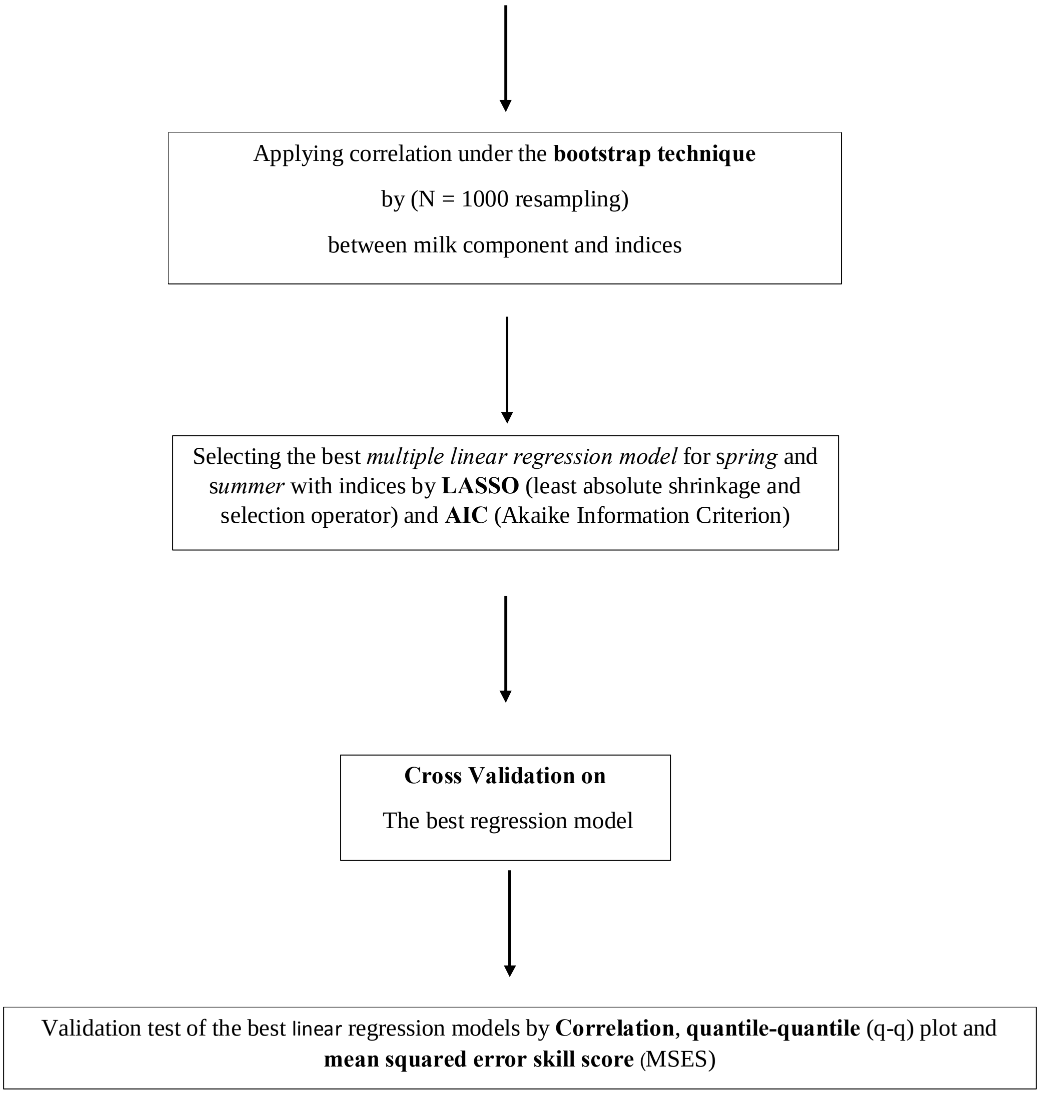


Figure 1.3: Study procedure at a glance on statistical methods

1.1.2.3 Linear relationships between climatic parameters and milk compounds under bootstrapping

The linear relations between climatic indices and milk compounds have been estimated through correlation analysis under bootstrapping. Correlation in time series data, as in our case, can be controversial. If there is no obvious relationship between parameters, it could be very useful and relative easy to apply. The inherent uncertainty was assessed by the second bootstrapping analysis. The nonparametric approach resampled among the time series from 2002 to 2010 for each month within a season and the climatic zones using the climate indices as the predictors and the milk compounds as the predictands. To account for the seasonal dependencies of the climatic input parameters and the physiological responses of cows, we analyzed two separate seasons (spring and summer) using three individual monthly averages within each season. Additionally we distinguish between the relationships between the climate data and milk data, and those describing the inter-relationship between both data sources. Correlation coefficients are averaged from the correlations across all stations within each bootstrap sample, with sample size of 1000. For correlation analysis, data is normalized to a mean of zero and standard deviation of one. We also calculated the confidence interval from inter-quartile range of the correlation coefficients across all bootstrap samples.

1.1.2.4 The best multiple linear regression model between climate indices and milk components considering uncertainty

Here, we model the observed milk components employing methods which take different models as well as parameters into account. To measure the "distance" between a specific model based on the meteorological data and the observed milk variables the "Kullback-Leibler Information" (K-L distance) together with the Akaike's Information Criterion (AIC) is used. The AIC identifies that model that minimizes K-L distance. In order to select the best model with the smallest number of predictors the LASSO and the AIC are combined. To avoid over fitting and test the validity of the obtained model, cross validation is applied on the selected model. To consider uncertainty and constructing confidence intervals bootstrap method is applied which is a field of active research in statistics, particularly for dependent data (Lahiri 2003; Efron and Tibshirani 1986). For presenting graphically the results in an appropriate statistical sense quantile – quantile plots are used. The quality of the models are assessed applying quantile-quantile (Q-Q) plot, correlation coefficient and mean squared error skill score (MSES).

1.2 Outline

This dissertation is structured as follows:

Chapter 1 describes an introduction and gives some information about the study domain, dataset and structure of dissertation.

Chapter 2, 3 and 4 explain material and methods, results and conclusions which are presented in three papers cumulatively. First two papers focused on data preparation and some statistical methods which cope with uncertainties in special data sets. Furthermore climate indices are selected from the literature, which aggregate more climatic parameters in a physiologically meaningful way. Then we looked for some basic relations between dependent (milk components) and independent(data parameters and indices) variables. In third paper, the most suitable merged statistical model between all climatic indices as predictors and milk parameters as predictands are obtained through multiple linear regression models.

Chapter 5 describes a summary of results and we conclude the study in this Chapter.

Chapter 2

A Pilot Investigation of the Relationship between Climate Variability and Milk Compounds under the Bootstrap Technique (*published*)

The purpose of this part is analyzing the linear relationship between climate variables and milk components in Iran by applying bootstrapping to include and assess the uncertainty. Bootstrap techniques are a very valuable methods to assess the influence of uncertainty within the available data sample in statistical analysis. Bootstrapping belongs to a class of resampling techniques based on the empirical probability distribution function or cumulative distribution function (cdf) of the data sample which new realizations of the sample are generated through the same distribution function.

OPEN ACCESS
foods
ISSN 2304-8158
www.mdpi.com/journal/foods

Article

A Pilot Investigation of the Relationship between Climate Variability and Milk Compounds under the Bootstrap Technique

Mohammad Reza Marami Milani [1,*], Andreas Hense [2], Elham Rahmani [2] and Angelika Ploeger [1]

[1] Department of Organic Food Quality and Food Culture, University of Kassel, Nordbahnhofstr. 1a, 37213 Witzenhausen, Germany; E-Mail: a.ploeger@uni-kassel.de
[2] Meteorological Institute, University of Bonn, Auf dem Hügel 20, 53121 Bonn, Germany; E-Mails: ahense@uni-bonn.de (A.H.); erahmani@uni-bonn.de (E.R.)

***** Author to whom correspondence should be addressed; E-Mail: marami@uni-kassel.de; Tel.: +49-228-735-101; Fax: +49-228-735-188.

Academic Editor: Christopher J. Smith

Received: 15 May 2015 / Accepted: 13 August 2015 / Published: 11 September 2015

Abstract: This study analyzes the linear relationship between climate variables and milk components in Iran by applying bootstrapping to include and assess the uncertainty. The climate parameters, Temperature Humidity Index (THI) and Equivalent Temperature Index (ETI) are computed from the NASA-Modern Era Retrospective-Analysis for Research and Applications (NASA-MERRA) reanalysis (2002–2010). Milk data for fat, protein (measured on fresh matter bases), and milk yield are taken from 936,227 milk records for the same period, using cows fed by natural pasture from April to September. Confidence intervals for the regression model are calculated using the bootstrap technique. This method is applied to the original times series, generating statistically equivalent surrogate samples. As a result, despite the short time data and the related uncertainties, an interesting behavior of the relationships between milk compound and the climate parameters is visible. During spring only, a weak dependency of milk yield and climate variations is obvious, while fat and protein concentrations show reasonable correlations. In summer, milk yield shows a similar level of relationship with ETI, but not with temperature and THI. We suggest this methodology for studies in the field of the impacts of climate change and agriculture, also environment and food with short-term data.

Keywords: climate variability; THI; ETI; milk compounds; bootstrap; uncertainty; Iran

1. Introduction

It is well documented that agriculture production, farm economy, milk yield and milk components and agricultural livestock welfare are impacted by climate variability [1].

The general increase of human population and the amount of dairy products in human diet requires on average a two percent growth of global milk production to be able to amend the increasing demand for dairy products. Furthermore, there is no doubt that cows should live in an optimum environmental condition (animal welfare aspects) to be productive in both quantitative and qualitative aspects [2].

However, the definitions of optimal environmental conditions, as well as the sensitivity of milk productivity by deviations from that optimal range, are still not clear. This holds especially true for developing countries and emerging economies. Therefore, the effects of climate variability upon milk yield, as well as amount of milk fat and milk protein (g/100 mL fresh milk), are investigated with data from Iran.

The aim of this research is investigating the effect of climate variability on three main components of milk.

According to the report of United Nation in 2013, the population in Iran will increase about 43% by 2050 up to 115 million [3]. Consumption of milk per person in Iran is 30 to 150 kg/capita/year as reported by Food and Agriculture Organization. Due to increasing human population and urbanization, the demand for food and animal production are expected to be enhanced. It shows the importance of the interdisciplinary research related to the effects of climate change on food security. Unfortunately such research has been done less frequently in the Middle East and Iran as well.

Previous and ongoing research [1,4–9] determined that tolerance of climate variability in ecosystems plays important role on fodder, rate of ability of cow physiological adaptation, ruminal fermentation, cow nutrition, husbandry systems, and has an effect on DNA integrity. The quality and quantity of milk compounds is generally a result of complex interactions of variables. These are not fixed and they can change with the time of the year, environmental conditions, and climate variability. The yield of Holstein cattle is found to be more sensitive to climate than those of Jersey cattle, but Jersey milk composition is more sensitive to climatic influences. The effect of heat stress on decreasing milk yield is more highlighted in Holstein rather than in Jersey cattle [10].

Already, three decades ago, Rodriquez [11] investigated the effect of temperature on milk composition and yield in Florida by the means of an analysis of variance for cows of Holstein and Jersey breed with monthly averages of single-day milk samples. They reported that yield increases if the maximum daily temperature increases from 8 to 29 °C. However, they also report a rapid decrease when maximum daily temperatures are higher than 29 °C. In contrast, fat and protein decline over the entire range, from 8 to 37 °C.

For including humidity and temperature as physiological derive of milk productivity index, combinations such as the temperature humidity index (THI) have been suggested as an indicator for thermal stress [12–14].

However, THI does not consider the loss of energy by passive cooling, e.g., through turbulent motion in the atmosphere. This is commonly parameterized by wind speed. Therefore, the equivalent temperature index (ETI) is used in this investigation as another climatic index, which incorporates wind speed, temperature, and humidity.

The combination of milk productivity data (yield, fat and protein concentration in milk) with physiologically relevant meteorological data and indices, taken from a third generation atmospheric reanalysis for Iran and the Middle East, is unique.

The major question to be answered is whether, indeed, large scale patterns of meteorological parameters can be identified that provide a statistically significant and physiological relevant influence upon milk productivity in terms of yield and relevant nutrient content in the milk.

The rather short record availability (2002–2010) for climate variability purpose requires that special statistical methods are applied to avoid misinterpretations due to over-fitting. Milk component data in Iran is provided for fat, protein (g/100 mL fresh milk), and milk yield, in spring and summer, from the Animal Breeding Centre of Iran. Climate data, including daily averages of temperature in two meter height (T_{2m}), wind speed, sea level pressure, and specific humidity, were taken from the Modern Era Retrospective-Analysis for Research and Applications (MERRA), which is undertaken by NASA's global modeling and assimilation office [15].

This paper focuses on data preparation and on introducing some statistical methods to cope with uncertainties in special data sets. Further climate indices are selected from the literature, which aggregate more climatic parameters in a physiologically meaningful way. Finally the paper is looking, as a pilot study, for some basic relations between dependent and independent variables.

For investigating the statistical relation between the variables in this multidimensional context, in a first step, simple linear regression is chosen because it is easier to use and to interpret. The particular choice of linear regression and correlation analysis is motivated by future investigations on the multivariate regression relations between climate variables and milk productions.

2. Materials and Methods

2.1. General Information on Study Domain

Iran is placed in Southwestern Asia, approximately at latitudes 25°00′ N–38°39′ N and longitudes 44°00′ E–63°25′ E, and covers about 1,648,000 square kilometers with a population of more than 77 million.

Topography in Iran plays a very important role on climate conditions. It is bounded on the north by the Caspian Sea, and, on south, by the Persian Gulf and the Gulf of Oman. There are two main mountain chains in Iran, named the Alborz and the Zagros, which dominate Iran in the north and northwest-southeast, respectively. These mountains have a significant effect on distributing perceptible atmospheric water vapor. Therefore, the central and eastern parts of Iran receive less precipitation than the north and west [16,17].

According to Ahrens [18], two thirds of Iran has, mostly, two climate categories: arid and semi-arid, based on the Köppen climate classification. Previously, eight climatic zones have been categorized in Iran according to a cluster analysis of rainfall [19].

In this study, our study area consisted of three zones, according to different climate conditions and data availability (Figure 1). The first zone is in Northwest Iran, with cold and semi-arid climate conditions. The second zone is placed in the north, with a Caspian, mild and humid climate. The third zone is almost in the center of Iran with arid and semi-arid, mild and semi-warm climate conditions [20].

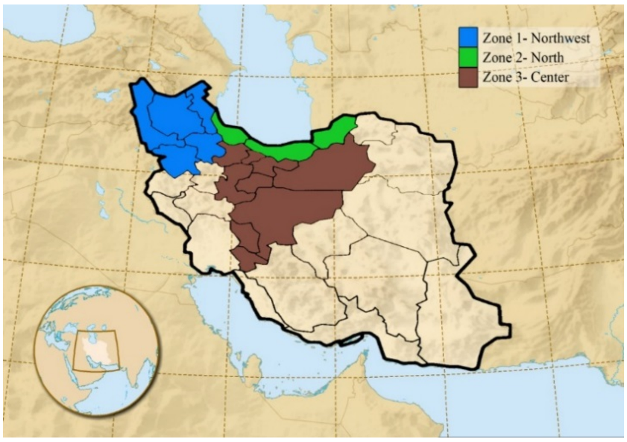


Figure 1. Classification of three study zones according to climate conditions in Iran [19,20].

In Iran there are 25,353 industrial herd stations with a capacity of about 3.3 million cows. At the moment, 18,295 industrial herd stations with about 1.3 million cows are active. Sixty-six percent of cows in Iran have a pure genetic origin (Holstein, Jersey and Brown-Swiss), 27% are hybrid and 7% are home born. About 64% of the pure cows are Holstein and 2% are Jersey and Brown-Swiss. Additionally, 19.3% of the hybrid cows have at least one Holstein parent. The total milk yield in 2013 was 3792 thousand tons. Most feed ingredients include: moisture corn, grain, alfalfa, straw, silage, wheat, soybean, cottonseed meal, and other forage [21].

2.2. Milk Data Conditions

Gathering milk data needs attention to some conditions. The most important factors in milk composition, which should be considered, are genetic, environmental conditions, management of feeding, season, cow age, rating of lactation, and pregnancy; these conditions provide a preprocessing of the milk data set of the investigation.

The yield of milk reduces when the non-lactating (dry) period is less than 40 to 45 days for first lactation and 55 to 65 days after second [22]. Milk yield also increases with increasing body weight, age, and number of lactations [23]. During the first to fifth lactation, milk fat decreases each year by about 0.2 (g/100 mL milk) and protein drops by 0.02 (g/100 mL milk) to 0.05 (g/100 mL milk) per year [24].

Mastitis, which is one of the most important and prevalent illnesses in cows, has the greatest effect on milk yield [25].

In this study, some main compounds of Holstein cow's milk in Iran were selected. Milk data is collected from 2002 to 2010, from almost 600 industrial herd stations, with about 1.2 million milk records of Holsteins cows from the seasons in which cows use natural pastures or fresh fodder. This is to say, during the months of April to September.

Industrial herd stations have been chosen because they consider the technical and health control management system, milking twice a day and a controlled nutritional feeding. They also have a regular data recording system, which was a very important aspect for this study.

All these conditions are taken into account when calculating from the raw data estimated averages of milk yield, fat, and protein concentrations of fresh milk. The milk data of the cows between their third and sixth calves were considered and the data of cows with the problem of mastitis are not considered. Additionally, we chose industrial herd stations with good health services, under controlled conditions and with veterinary care.

For the milk parameters, the yield of milk (kg/Year), fat content (g/100 mL milk), and protein content (g/100 mL milk), were used, from which the values are monthly averages in all herds and stations, respectively, in each zone. Then, the milk data bank is fixed under the mentioned conditions (lactation, health, *etc.*) with 936,227 milk records of Holstein cows from 2002 to 2010 in spring and summer.

The milk productivity data are directly calculated from the available data after correction for sampling and physiology, as described above.

2.3. Climate Data Basis and Climate Parameters

The climate database was set up for the same period as for the milk productivity data, and covers the area of Iran. As specific parameters, the daily averages of T_{2m}, sea level pressure, specific humidity, and the wind component vectors from MERRA reanalysis with a resolution of 1/2 degrees latitude and 2/3 degrees longitude are chosen. Thus, for each zone, the area, through the latitude and longitude range, and averaged the climatic parameter for all grid points in the adaptive zone, was considered. The other climate parameters, such as vapor pressure (*e*), saturation vapor pressure (e_w^*), relative humidity (*RH*), equivalent temperature index (*ETI*), and temperature humidity index (*THI*) were calculated by the following expressions.

$$e = 1.6077 \times p \times q \quad (1)$$

where, *e* is the water vapor pressure (hPa), *p* is the sea level pressure (hPa), and *q* is the specific humidity.

$$e_w^* = 6.1078 \exp[(17.1 \times T)/(235 + T)] \quad (2)$$

where e_w^* is saturation vapor pressure (hPa), which is calculated using the Magnus Equation from the two meter temperature, *T* (°C) [26].

Relative humidity (%) is calculated as:

$$RH = e / e_w^* \quad (3)$$

To assess the risk of heat stress, which is an important aspect in livestock health in humid and hot climates, we applied *THI*, which is a function of air temperature in two meter height, *T* and dew point, T_d.

Dew point is the temperature at which the actual water vapor pressure equals the saturation water vapor pressure [14,27].

T_d can be calculated by the equation below, or from the Magnus Equation, by using the saturation vapor pressure.

$$T_d = (\frac{RH}{4})^{\frac{1}{8}} \times [112 + (\frac{9T}{40})] + \frac{T}{40} - 112 \tag{4}$$

$$THI = 41.5 + T + 0.36T_d \tag{5}$$

where T_d is the dew point temperature, T is temperature in two meter height and RH is relative humidity (%).

Additionally, we calculated the *ETI*, which takes into account the loss of energy from the animal's body by conduction and turbulent transfer [28–30]. The *ETI* is calculated as follows [31]:

$$\begin{aligned} ETI = 27.88 - 0.456T + 0.010754T^2 - 0.4905RH + 0.00088RH^2 + \\ 1.1507V - 0.12645V^2 + 0.019876T \times RH - 0.046313T \times V \end{aligned} \tag{6}$$

where V is wind speed (m/s) derived for the zonal and meridional wind components. Note that the nonlinear relations between dew point temperature, water vapor pressure, and specific humidity in *THI* and between temperature, relative humidity, and wind velocity in *ETI* require daily input data to evaluate monthly means.

2.4. Bootstrap Technique

Temporally short records of data and their statistical analyses, as in the present case, require careful considerations of their inherent uncertainties. This is necessary in order to not to come up with misleading conclusions due to the small sample size. This investigation is based on a comparable large dataset of milk compound data from the Holstein breed and from various regions in Iran. However, if the intention is to find statistically meaningful connections to the inter-annual variability of near surface climate, the sample size decreases dramatically because only nine years of data (2002–2010) are available to estimate the dependency between climate variations and milk compound changes. Although we consider in this paper the results of a nine year period of climate conditions on milk compounds and yield with a high number of cows but for climate research such a period is relatively short. Considering monthly values is not helpful due to the ever-present annual cycle, which makes, for example, direct comparisons of actual spring and summer values practically meaningless ("summer is always warmer than spring").

Bootstrap techniques are a very valuable, computer-based tool to assess the influence of uncertainty within the available data sample on the results of the statistical analysis [32].

Bootstrapping belongs to a class of resampling techniques. Based on the empirical probability distribution function, or cumulative distribution function (cdf), of the data sample, new realizations of the sample are generated, which share, by this construction, the same distribution function. Depending on the way in which the cdf is estimated from the original sample, one distinguishes between parametric and non-parametric bootstrapping. In the first case, parameters of an assumed cdf are estimated from the original sample (e.g., mean and variance for a Gaussian cdf). Plugging these parameter values into the cdf allows one to generate, by drawing at random from this fully defined cdf, new sample values [33]. The second way to generate the new bootstrap samples is by estimating the cdf non-parametrically through sorting of the original sample in increasing order. Sampling at random from this empirical cdf leads directly to a random sampling of the original data set with replacements such that, in the new sample, some of the original data are left out and replaced by copies of the remaining original data.

In the present case, the non-parametric bootstrap technique was applied in two different situations. As mentioned above, the milk compound data set is quite comprehensive if taken by itself. However, comparing it to climate data requires a reduction of these data to monthly averages in order to concentrate the analysis on the relevant time scales of months to years. Therefore, in a first step, we applied the bootstrap technique to data from each station (consisting of herds) in each zone, separately generating 1000 samples of monthly means of milk yield, fat, and protein concentrations for the full period of 2002–2010, which are presented in Figure 2.

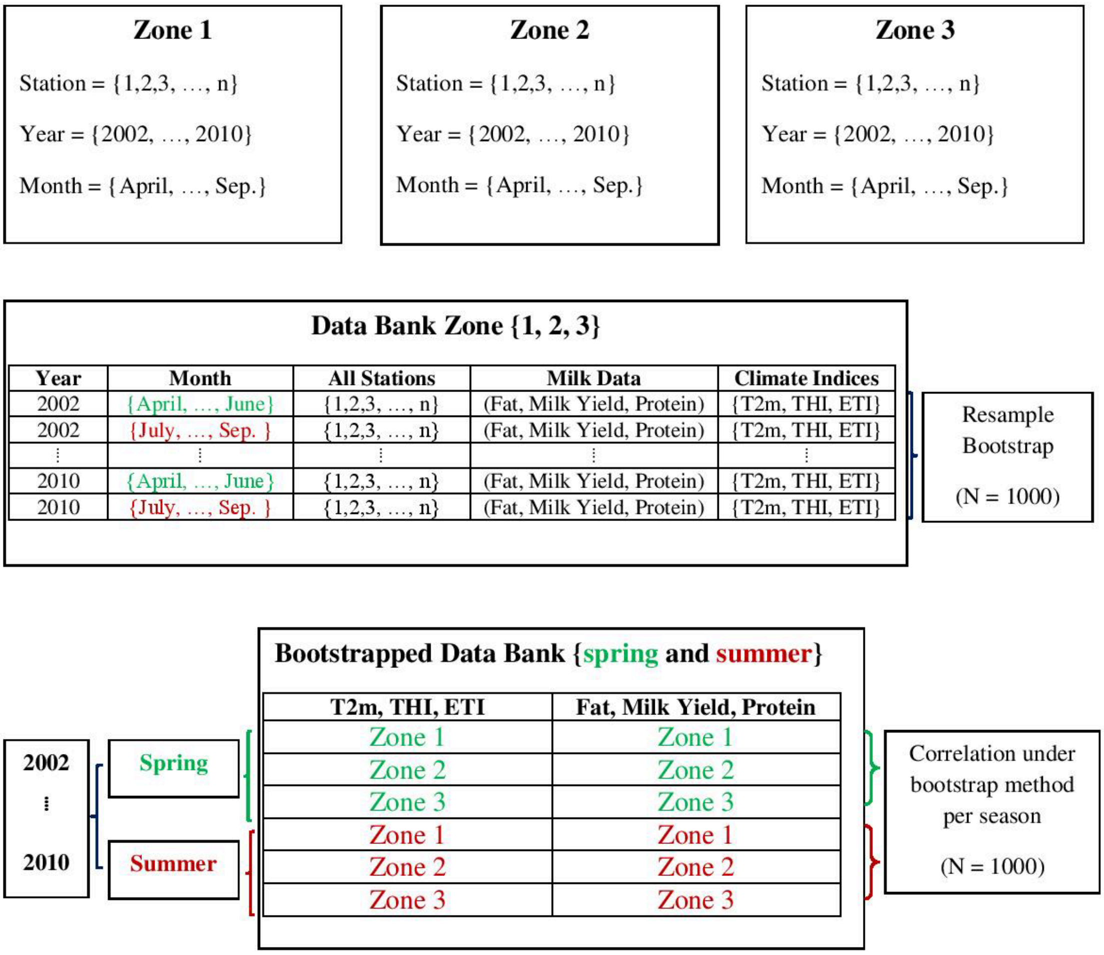


Figure 2. Flowchart of data set.

The general aim is to compare the results from the bootstrapped correlation analysis with those from the classical approach in order to assess the robustness of the linear correlation coefficients under the resampling of this rather small data set in terms of climate variability.

Technically, we implemented the analysis using R version 3.0.2, which provides various packages for statistical data analyzing, calculating, and graphical display [34].

2.5. Linear Relationships between Climatic Parameters and Milk Compounds

The linear relations between the climatic parameters or indexes (T_{2m}, THI, ETI) and the milk compounds (milk yield, protein, fat) have been estimated through correlation analysis under bootstrapping. The flowchart of the used data is shown in Figure 2.

Correlation in time series data, as in our case, can be controversial. If there is no obvious relationship between parameters, it could be very useful and relative easy to apply.

Correlation between variables x and y with Gaussian distributions and variances σ_x^2 and σ_y^2, respectively, with the range of [−1, +1] is defined by:

$$\rho_{xy} = \frac{COV(x, y)}{\sigma_x \sigma_y} \tag{7}$$

We applied bootstrap with N = 1000 resampling steps. Then, we have 1000 samples of x and y as $x = (x_1, x_2, \ldots, x_N)$, and $y = (y_1, y_2, \ldots, y_N)$. Correlation of the transformed variables of x and y is estimated as:

$$r_{xy} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\left\{\sum_{i=1}^{N}(x_i - \bar{x})^2\right\}^{0.5} \left\{\sum_{i=1}^{N}(y_i - \bar{y})^2\right\}^{0.5}} \tag{8}$$

We assessed the inherent uncertainties of bootstrapping as described above. Two separate seasons have been analyzed to account for the seasonal dependencies of the climatic input parameters and the physiological responses of cows during spring and summer. Additionally we distinguish between the relationships between the climate data and milk data, and those describing the inter-relationship between both data sources. The correlation analysis during the time period of 2002–2010 in spring and summer are presented in section 3.2. In these matrices, the respective correlation coefficients are averaged from the correlations across all stations within each bootstrap sample, with sample size of 1000. We also calculated the confidence interval from inter-quartile range of the correlation coefficients across all bootstrap samples. Rahmani *et al.* [35] also applied a bootstrapping technique for estimating the sampling uncertainty of the correlation and regression analyses.

3. Results and Discussion

3.1. Statistics of Climatic and Milk Parameters

To obtain an overview of the relevant conditions for the relationship between milk productivity (amount and nutrient content) and climate variability, the statistics of milk parameters in Figure 3 and the climatic parameters in Table 1 and Figure 4 are summarized for short-term data (2002–2010) in summer (July, August, and September) and spring (April, May, and June), separately, in the study area.

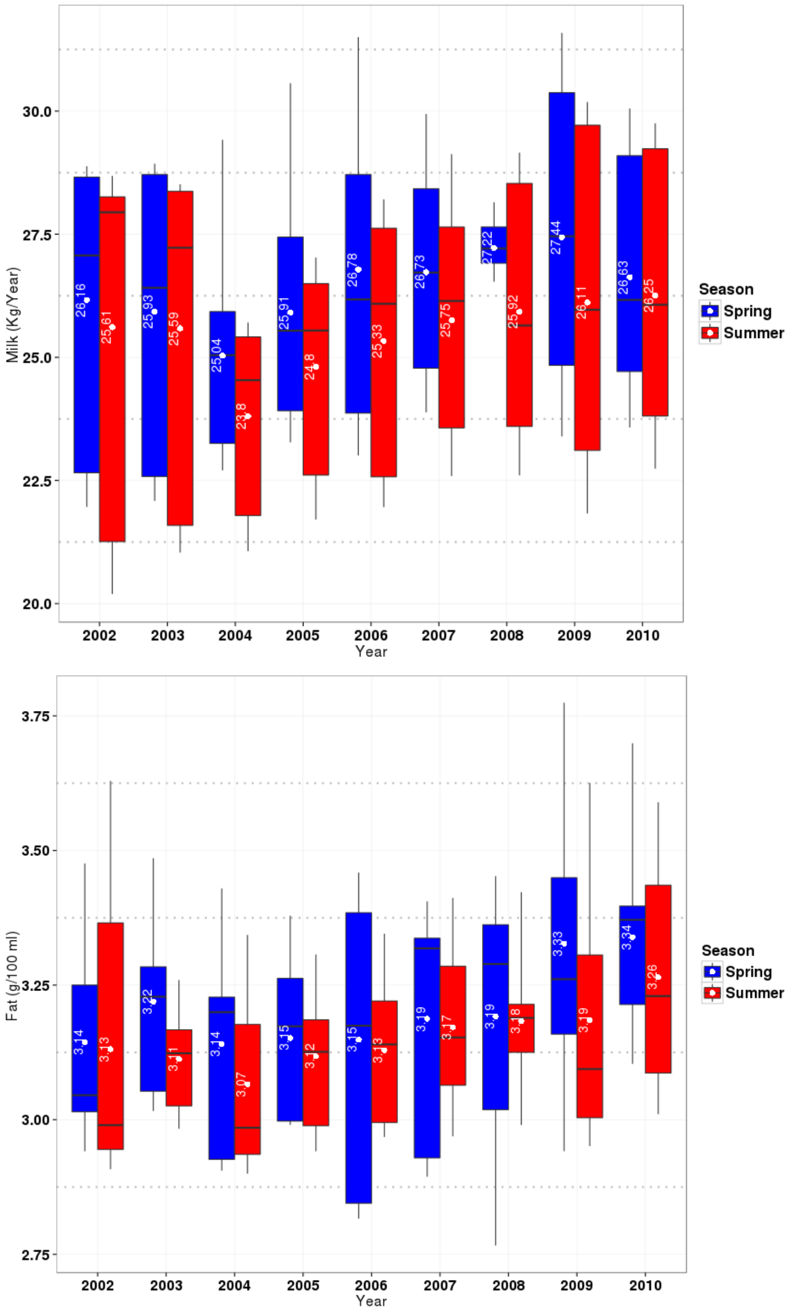


Figure 3. *Cont.*

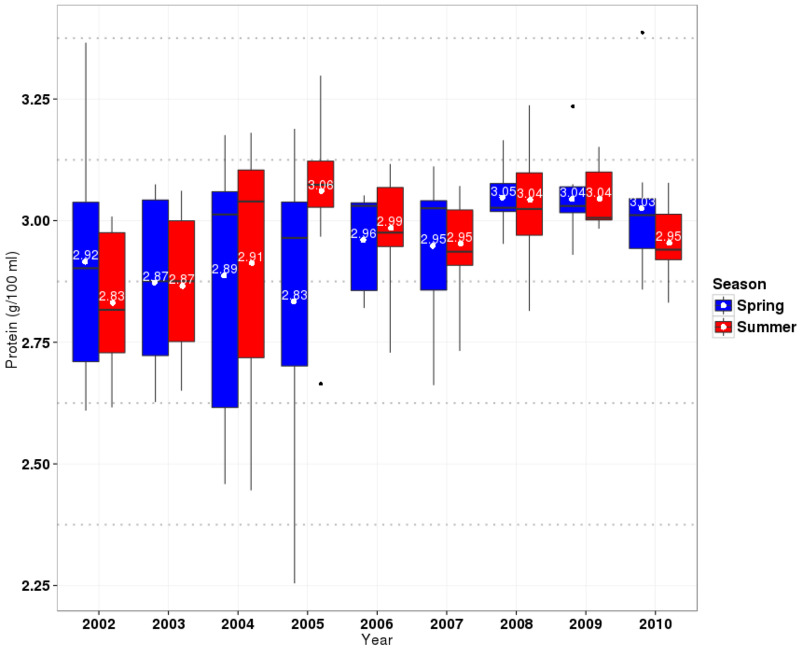
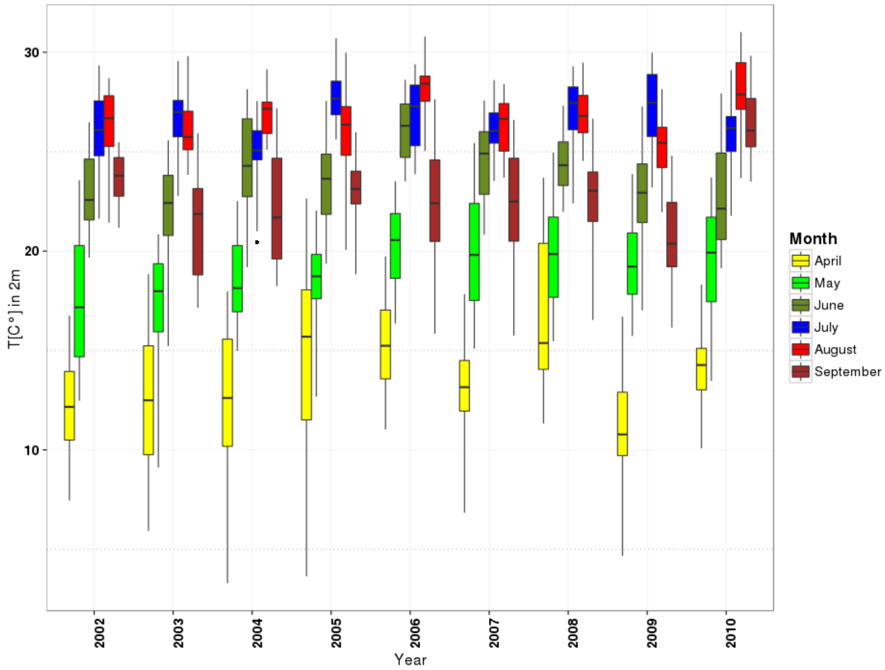


Figure 3. Statistics of milk parameters. Lines inside the box plots show the median and points show the season averages of milk compounds (Milk yield (Kg/Year), fat (g/100 mL), protein (g/100 mL) in fresh product) in spring (blue boxes) and summer (red boxes) from 2002 to 2010.

Table 1. Temperature in two meter height (T_{2m} in °C), Temperature Humidity Index (THI) and Equivalent Temperature Index (ETI) and number of days that cow is under stress (NDS), during 2002 to 2010 in spring and summer.

	Spring			Summer		
	T_{2m}	THI	ETI	T_{2m}	THI	ETI
Minimum	8.0	50.7	22.0	17.1	61.83	21.21
Median	19.7	64.3	22.9	25.0	70.41	22.50
Average	18.7	63.1	23.2	25.3	70.43	22.44
Maximum	28.6	73.1	25.6	31.0	75.20	23.18
NDS	31	6	-----	168	100	-----



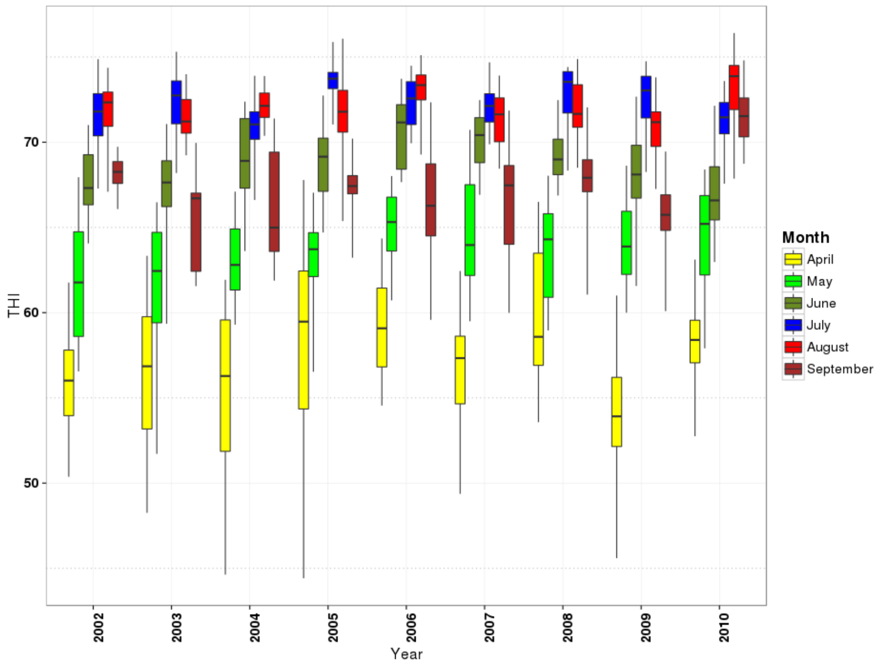


Figure 4. *Cont.*

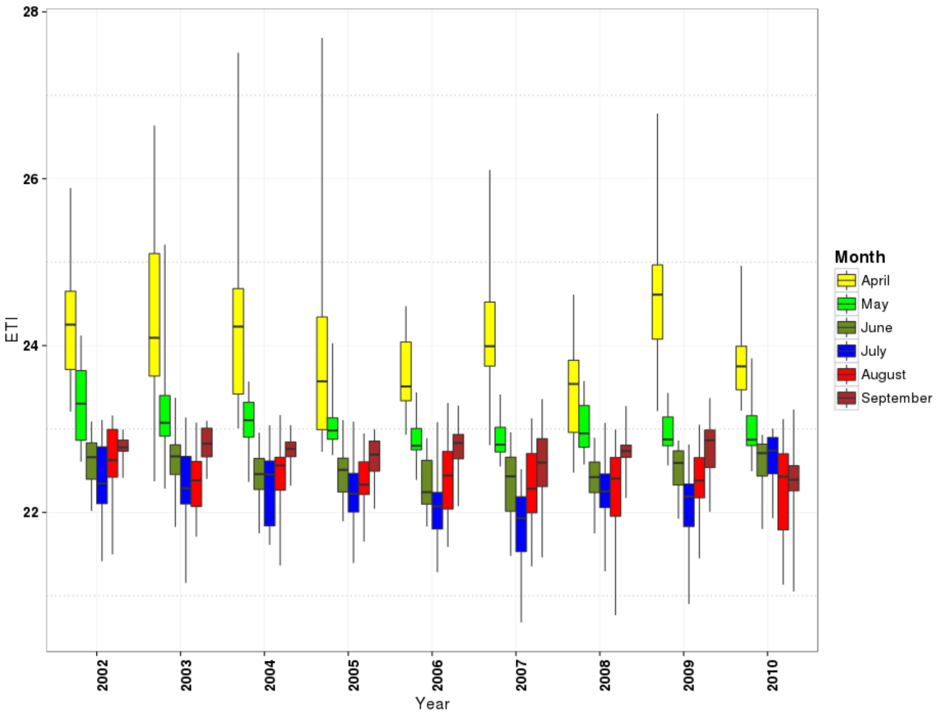


Figure 4. Statistics of climatic parameters from the box plot of T_{2m} (°C), THI, ETI, from 2002 to 2010 in spring and summer.

For milk parameters, yield of milk (kg/Year), fat, and protein (g/100 mL milk) are considered.

The box plots of milk compounds represent the scales of variability across all stations and years. There is hardly any visible seasonal signal in the three compounds of milk, which justifies using only the two seasons, spring and summer.

The results of Figure 3 indicate that from 2002 to 2010, the monthly average of milk yield decreased from spring to summer. The average of protein from spring to summer showed different trends, while the average of fat values decreased in an unremarkable way during this period. The maximum reducing rate of fat is for 2009, with 0.15 (g/100) mL milk (Figure 3).

Decreasing of milk yield is also reported by Rodriquez *et al.* and Bouraoui *et al.* [11,36]. They reported a decrease of protein and of milk yield. In this study, for protein is not big decreasing detected, which might be caused by differences in the experimental conditions.

As with the results in this investigation, Knapp and Grummer [37] and Roman-Ponce *et al.* [38] also did not find any significant relation between fat reductions and heat stress. Reduction in amounts of fat and protein from spring to summer is also reported by Bouraoui *et al.* [36].

The climate data are the directly available MERRA data (T_{2m}) or the derived compounds (THI, ETI) calculated from the respective MERRA data sets. As additional information, the average number of days of stress for cows is given.

For climate parameters, the monthly average of daily T_{2m}, minimum and maximum, average of calculated THI and ETI and their maximum and minimum values are presented in Table 1and Figure 4.

THI is determined to assess the risk of heat stress with critical values of 72, 78 and 82. When the THI rises above 72, cows are probably under heat stress. Values higher than 78 will seriously affect milk production. When THI exceeds 82, very severe heat stress on cow occurs with significant decreases in milk production [36,39].

Hence, 72 as the start point of heat stress was assumed and NDS presents the number of days with a THI above 72 in Table 1. For ETI, no critical value could be found from the literature that could serve as start point for stress. For temperature, the number of days exceeding 25 °C was considered. It is the upper critical temperature for Holstein cows that they are under stress (NDS) [40].

The monthly average values for T_{2m}, THI and ETI, respectively, change from 18.7 °C, 63.1 and 23.2 in spring to 25.3 °C, 70.43 and 22.44 in summer.

The annual cycle of the climate parameters across all stations and their short-term changes, between 2002 and 2010, can be seen in Figure 4. Here, a very prominent annual cycle is apparent for all three parameters with a strong change during spring and a plateau during the summer months. Again this special configuration is the second justification for selecting the two seasons.

3.2. Results for Correlation between Climatic Parameters and Milk Compounds

Figures 5 and 6 describe summarized results of the correlation analysis between T_{2m}, THI, ETI and milk yield, fat, protein, in spring and summer, respectively.

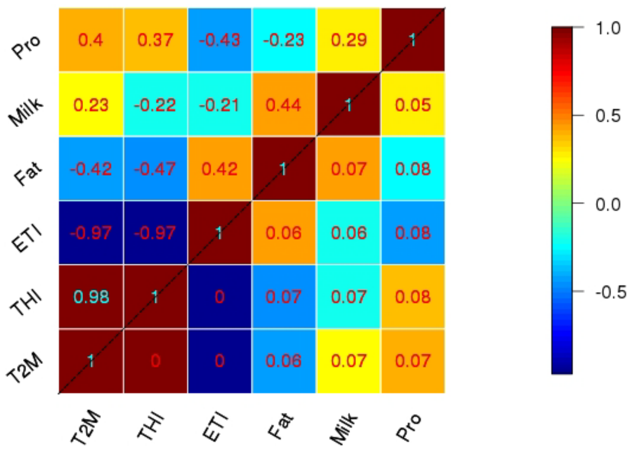


Figure 5. Correlation (above diagonal) and respective interquartile range of bootstrap sample (below diagonal) between milk compounds and climatic indices in spring.

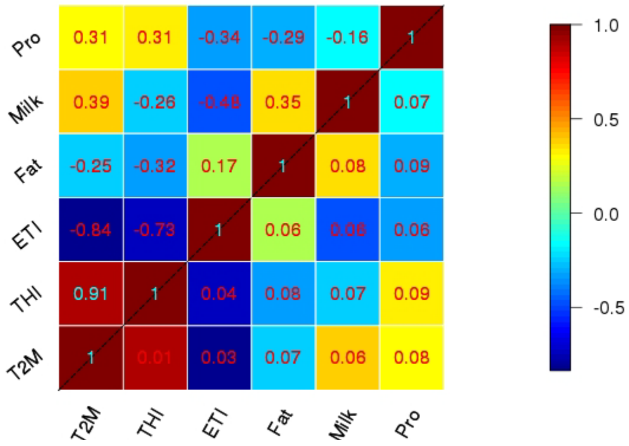


Figure 6. Correlation (above diagonal) and respective interquartile range of bootstrap sample (below diagonal) between milk compounds and climatic indices in summer.

The entries above the diagonal in these two matrices are the respective correlation coefficients as averages of the correlations across all stations during the time period of 2002–2010 within each bootstrap sample. Bootstrapping with a sample size of 1000 is applied to the total sample of size 81 (3 × 27), formed by all nine years at all three climatic zones (Figure 1). The values below the diagonal are estimated confidence intervals from the inter-quartile range of the correlation coefficients across all bootstrap samples. Inter-quartile range is the distance between 25% and 75% quantiles of the bootstrap samples.

The correlation results for spring (upper triangle in Figure 5) illustrate a positive correlation between milk yield with protein and fat content in the milk in spring, with the values of 0.29 and 0.44. This means that by increasing the milk yield, the fat and protein concentrations also increase. The correlations are weak and they explain, at most, 19% of the total variability (yield *vs.* fat with a correlation of 0.44). Correlations involving protein concentrations are also small compared to the sampling uncertainty (lower triangle in Figure 5).

In summer, milk yield still has positive correlation with fat (0.35), but, again, the correlation involving protein is small, showing that fat and protein concentrations seem to be independent of milk yield. This might be due to the rather short sample size, even when taking the different zones into account.

For both seasons, T_{2m} and THI show very high positive correlations with each other, and a high negative correlation with ETI, especially for spring. This means that, in contrast to the milk productivity data, we observe a very high dependency between the climatic variables, meaning that either variable can be chosen for analyzing the interdependency between climate variability and milk productivity. During summer, the correlations are still quite high, but the ETI has some independent information compared to T_{2m} and THI because the correlations drop to 0.7, meaning that only 50% of the total variance is common between ETI on the one hand and T_{2m}/THI on the other.

Despite the small sample size for climate investigations (nine years but a very large number of milk parameters) and the related uncertainties, an interesting behavior of the relationships between milk compounds and the climate parameters is visible. During spring only, a very weak dependency of milk yield and climate variations is obvious, while fat and protein concentrations show reasonable correlations with an explained variance of around 15% to 22%. However, for summer, milk yield shows a similar level of relationship with T_{2m} and ETI, but not with THI. Also in summer the relationships of fat concentrations with the climate parameters level off, quite in contrast to the spring results.

Figure 7 presents some additional explanations regarding regression analysis between protein concentration, as predictand, and T2M, THI, ETI as predictors in spring, before and after applying bootstrapping. The data have been normalized to zero and a standard deviation of one. To that aim, we used anomalies of the data and applied the bootstrapping technique. In this sample, the figure of the slope of best-fitting straight line between the variables presents a positive or negative correlation. The black line is the regression line without bootstrap, and consists of the predicted score on y for each possible value of x. The green lines show the range of change in the intercept and slope in resampling data during 1000 bootstraps. As is represented in Figure 5, the values of positive correlation between protein-T2M (0.4), protein-THI (0.37) show the positive slopes of the regression lines between these variables in Figure 7, but with a considerable scatter. The squared value of the correlation (0.16) allows for a quantification of linear dependency between predictand and predictor, explaining 14% of the total variability *vs.* remaining scatter, with 86% of total variability. For protein concentration *vs.* ETI, it changed to a negative correlation (-0.43), which is clear in Figure 7, as well leading to 18% for the linear dependency *vs.* 82% for unexplained scatter. In all cases, bootstrap increased the significance level of the best-fitted regression line. A question for future research is whether the amount of scatter might be significantly reduced by additional linear, or even nonlinear, predictors.

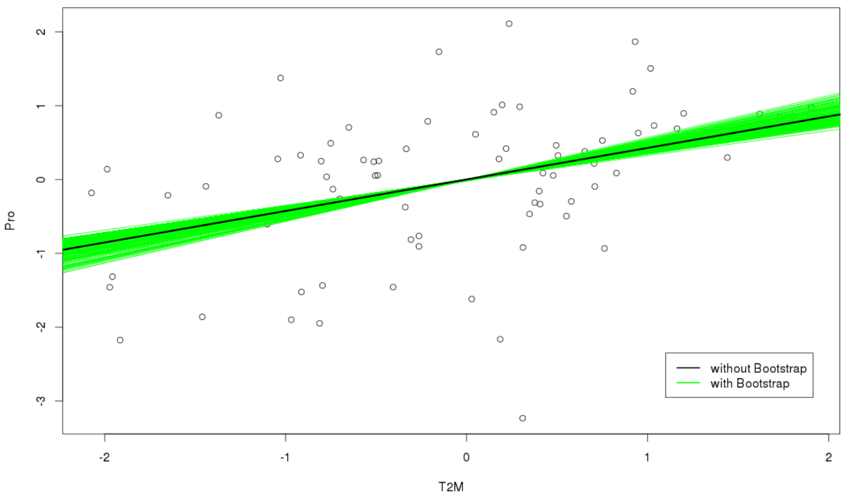


Figure 7. *Cont.*

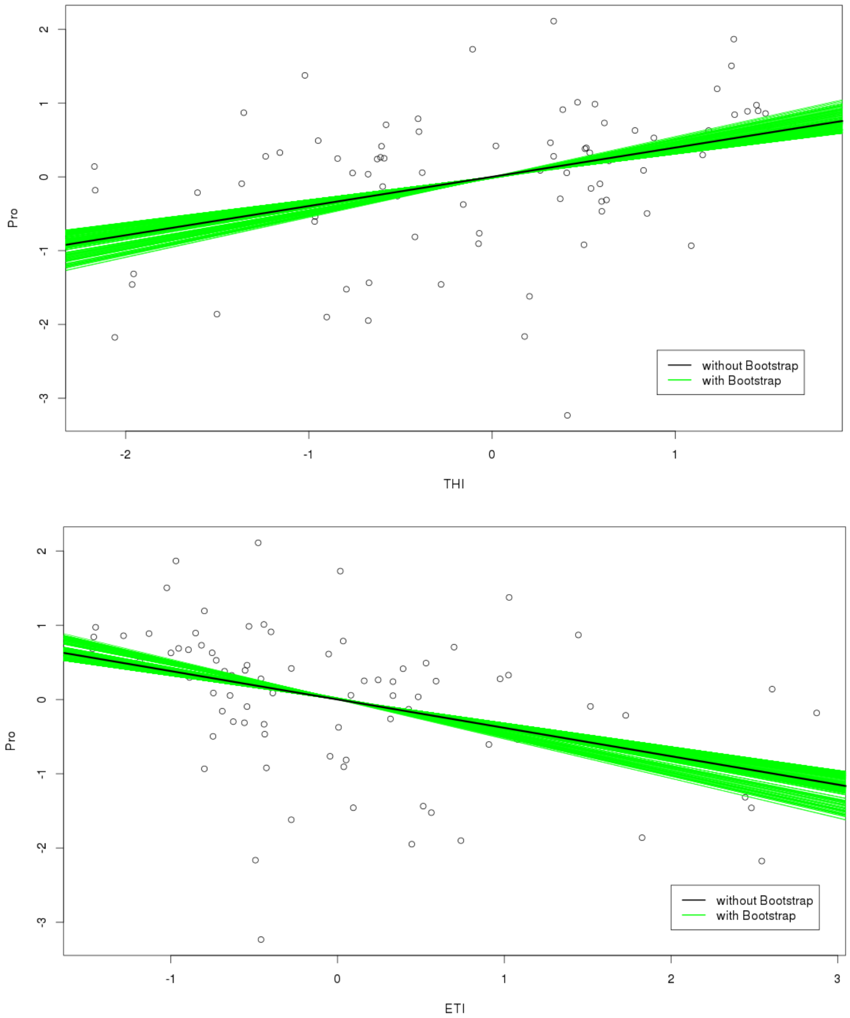


Figure 7. The best-fitted regression line before (black line) and after 1000 bootstrapping (green lines) between protein and T2M, THI and ETI indexes in spring.

4. Conclusions

As mentioned in the introduction, one of the main objectives of this research is to evaluate the Pearson correlation under a statistical bootstrapping method for comparing the influence of dairy variable time series on climate time series.

The bootstrapped linear relationships between climatic parameters and milk compounds illustrate an interesting correlation between milk compounds and the climate parameters.

Although a nine year period data for milk compounds is not short dataset in agricultural investigations but in climate research field such a period is relatively short term. This methodology can be very useful for developing countries with short-term data sets. The bootstrapping technique is confirmed to the bivariate correlation in this study, which also agreed with results reported by Lunneborg [41]. The bootstrap results expanded confidence intervals while it demonstrates against the presented results of Rasmussen [42], which explained that bootstrapping overly restricted confidence intervals.

Additionally, when the distribution of available data sets is complicated or unknown, and only a small pilot sample with a great deal of uncertainty is available, one of the best solutions to estimate the regression coefficients and confidence interval of the regression equation is the bootstrap method. Applying the bootstrap technique is also very useful and relatively simple for analyzing what data collection is expensive, difficult, or data with gaps. Fundamentally, bootstrapping can be used for finding the regression coefficients in linear or non-linear models with an increasing significance level of the test.

With reference to the study of Silva and Maia [31], the Holstein breed in different climate conditions, such as tropical, Mediterranean, arid and semi-arid, has changed its characteristics. According to the characteristic changes in cows under different conditions, the effect of climate variability on animal products, new thermal stress indices with more significant probability and interaction, between animal physiology and climate variability, are needed.

The authors strongly suggest that more research in this field is needed in order to focus on discovering the critical points of climate indices regarding each of the milk components separately. Additionally, new indices have to consider more predictors, such as solar ration and percentage of fat and protein in the diet of cows, heat tolerance, how many days did cows suffer from extremely critical heat stress days, how much of the skin surface is covered with large black spots (especially for the Holstein breed), and also the ability of herd stations to control heat stress, facility of welfare, and breed of cows. Finally, every additional predictor can change the results and help to achieve models that are closer to reality.

In addition, to increase the usability of this technique, the authors propose that it makes sense to do further simulations in other climate regions, even given the statistical problems when using short time datasets with large amount of uncertainty.

Applying statistical modeling techniques would be an advantage in industry for the improvement of the food security and preventing expensive losses due to unexpected natural issues.

Acknowledgments

The authors warmly thank from Nasim Azari at the Food Science Engineering Department of Science and Research branch of Tehran Islamic Azad University, Ahmad Moghimi Esfand Abadi and Asghar Salimi Niknam at the Animal Breeding Centre of Iran in Karaj, branch of the Iranian Ministry of Agriculture, Dr. Behnam Saremi and Dr. Ali Asadi for supporting the milk data. We are also grateful to Prof. Ali Mortazavi in Department of Food Science, Ferdowsi Mashhad University for his cooperation.

Author Contributions

A.H. and A.P. supervised the methodology of the study and the used statistical methods. E.R. contributed data and tools for analyzing. M.M. performed the statistical analysis, interpreted the results, and wrote the initial manuscript. All authors contributed to read and edit the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Gauly, M.; Bollwein, H.; Breves, G.; Brugemann, K.; Danicke, S.; Das, G.; Demeler, J.; Hansen, H.; Isselstein, J.; Konig, S.; *et al.* Future consequences and challenges for dairy cow production system arising from climate change in center Europe—A review. *Animal* **2013**, *7*, 843–859.
2. Darwin, R. *Climate Change and Food Security*; Agriculture Information Bulletin Number 765–8; United States Department of Agriculture: Washington, DC, USA, 2001.
3. United Nation, Department of Economic and Social United Affairs. World Population Prospects the 2012 Revision. Available online: http://esa.un.org/unpd/wpp/index.htm (accessed on 10 June 2014).
4. Beatty, D.T.; Barnes, A.; Taylor, E.; Pethick, D.; McCarthy, M.; Maloney, S.K. Physiological responses of Bostaurus and Bosindicus cattle to prolonged continuous heat and humidity. *J. Anim. Sci*. **2006**, *84*, 972–985.
5. Ciais, Ph.; Reichstein, M.; Viovy, N.; Granier, A.; Ogée, J.; Allard, V.; Aubinet, M.; Buchmann, N.; Bernhofer, Chr.; Carrara, A.; *et al.* Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature* **2005**, *437*, 529–533.
6. Collier, R.J.; Dahl, E.; VanBaale, M.J. Major advances associated with environmental effects on dairy cattle. *J. Dairy Sci*. **2006**, *89*, 1244–1253.
7. Myneni, R.B.; Keeling, C.D.; Tucker. C.J.; Asrar, G.; Nemani, R.R. Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* **1997**, *386*, 698–702.
8. Nikkhah, A.; Furedi, C.J.; Kennedy, A.D.; Scott, S.L.; Wittenberg, K.M.; Crow, G.H.; Plaizier, J.C. Morning *vs.* evening feed delivery for lactating dairy cows. *Can. J. Anim. Sci.* **2011**, *91*, 113–122.
9. O'Brien, M.D.; Rhoads, R.P.; Sanders, S.R.; Duff, G.C.; Baumgard, L.H. Metabolic adaptations to heat stress in growing cattle. *Domest. Anim. Endocrinol.* **2010**, *38*, 86–94.
10. Sharma, A.K.; Rodriguez, L.A.; Mekonnen, G.; Wilcox, C.J.; Bachman, K.C.; Collier, R.J. Climatological and genetic effects on milk composition and yield. *J. Dairy Sci*. **1983**, *66*, 119–26.
11. Rodriquez, L.A.; Mekonnen, G.; Wilcox, C.J.; Martin, F.G.; Krienke, W.A. Effects of Relative Humidity, Maximum and Minimum Temperature, Pregnancy, and Stage of Lactation on Milk Composition and Yield. *J. Dairy Sci*. **1985**, *68*, 973–978.
12. Gaughan, J.B.; Mader, T.L.; Holt, S.M.; Josey, M.J.; Rowan, K.J. Heat tolerance of Boran and Tuli crossbred steers. *J. Anim. Sci*. **1999**, *77*, 2398–2405.
13. Ingraham, R.H.; Stanley, R.W.; Wagner, W.C. Relationship of temperature and humidity to conception rate of Holstein cows in Hawaii. *J. Dairy Sci.* **1974**, *59*, 2086–2090.
14. Thom, E.C. The discomfort index. *Weatherwise* **1959**, *12*, 57–60.
15. Rienecker, M.M.; Suarez, M.J.; Gelaro, R.; Todling, R.; Bacmeister, J.; Liu, E.; Bosilovich, M.G.; Schubert, S.D.; Takacs, L.; Kim, G.-K.; *et al.* MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Clim.* **2011**, *24*, 3624–3648.
16. Dinpashoh, Y.; Jhajharia, D.; Fakheri-Fard, A.; Singh, V.P.; Kahya, E. Trends in reference crop evapotranspiration over Iran. *J. Hydrol.* **2011**, *399*, 422–433.
17. Kousari, M.R.; Ahani, H.; Hendi-Zadeh, R. Temporal and spatial trend detection of maximum air temperature in Iran during 1960–2005. *Glob. Planet. Chang.* **2013**, *111*, 97–110.
18. Ahrens, C. *Essentials of Meteorology: An Introduction to the Atmosphere*, 5th ed.; Thomson Learning: Belmont, CA, USA, 2008; p. 448.

19. Modarres, R.; Sarhadi, A. Statistically-based regionalization of rainfall climates of Iran. *Glob Planet. Chang.* **2011**, *75*, 67–75.
20. Rahmani, E. The Effect of Climate Variability on Wheat in Iran. Ph.D. Thesis, University of Bonn, Bonn, Germany, 2015.
21. Statistical Center of Iran. Vice-President for Strategic Planning and Supervision. Available online: http://www.amar.org.ir/Default.aspx?tabid=281 (accessed on 5 August 2014).
22. Kuhn, M.T.; Hutchison, J.L.; Norman, H.D. Dry Period Length to Maximize Production Across Adjacent Lactations and Lifetime Production. *J. Dairy Sci*. **2006**, *89*, 1713–1722.
23. Martinez, M.L.; Lee, A.J.; Lin, C.Y. Age and Zebu-Holstein Additive and heterotic Effects on Lactation Performance and Reproduction in Brazil. *J. Dairy Sci*. **1988**, *71*, 800–808.
24. Heinrichs, J.; Jones, C.; Bailey, K. MILK Components: Understanding the Causes and Importance of Milk Fat and Protein Variation in Your Dairy Herd. *Dairy Anim. Sci*. **1997**, *5*, 1e–8e.
25. Lescourret, F.; Coulon, J.B. Modeling the Impact of Mastitis on Milk Production by Dairy Cows. *J. Dairy Sci.* **1994**, *77*, 2289–2301.
26. Kraus, H. *Die Atmosphäre der Erde*, 3rd ed.; Springer: Berlin Heidelberg, Germany, 2004; pp. 67–89.
27. Bohmanova, J.; Misztal, I.; Colet, B. Temperature-Humidity Indices as Indicators of Milk Production Losses due to Heat Stress. *J. Dairy Sci*. **2007**, *90*, 1947–1956.
28. Baeta, F.C.; Meador, N.F.; Shanklin, M.D.; Johnson, H.D. Equivalent temperature index at temperatures above the thermoneutral for lactating dairy cows. In proceedings of the Summer Meeting of American Society of Agricultural Engineers (ASAE), Baltimore, MD, USA, 28 June–1 July 1987; pp. 87–4015.
29. Mukherjee, D.; Bravo-Ureta, B.E.; De Vries, A. Dairy productivity and climatic conditions: Econometric evidence from South-eastern United States. *Aust. Agric. Res. Econ.* **2012**, *57*, 123–140.
30. Gomes da Silva, R.; Morais, D.A.; Guilhermino, M.M. Evaluation of thermal stress indexes for dairy cows in tropical regions. *R. Bras. Zootec.* **2007**, *36*, 1192–1198.
31. Gomes da Silva, R.; Campos Maia, A.S. *Principles of Animal Biometeorology*; Springer Science & Business Media: Dordrecht, The Netherlands, 2013; pp. 207–227.
32. Efron, B.; Tibshirani, R. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statist. Sci*. **1986**, *1*, 54–75.
33. Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *Numerical Recipes: The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2007; pp. 340–439.
34. R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2011.
35. Rahmani, E.; Friederichs, P.; Keller, J.; Hense A. Development of an effective and potentially scalable weather generator for temperature and growing degree days. *Theor. Appl. Climatol.* **2015**, *120*, doi:10.1007/s00704-015-1477-z.
36. Bouraoui, R.; Lahmarb, M.; Majdoubc, A.; Djemalic, M.; Belyead, R. The relationship of temperature-humidity index with milk production of dairy cows in a Mediterranean climate. *Anim. Res*. **2002**, *51*, 479–491.
37. Knapp, D.M.; Grummer, R.R. Response of lactating dairy cows to fat supplementation during heat stress. *J. Dairy Sci*. **1991**, *74*, 2573–2579.

38. Roman-Ponce, H.; Thatcher, W.W.; Buffington, D.E.; Wilcox, C.J.; Van Horn, H.H. Physiological and Production Responses of Dairy Cattle to a Shade Structure in a Subtropical Environment. *J. Dairy Sci*. **1977**, *60*, 424–430.
39. Johnson, H.D. Environmental management of cattle to minimize the stress of climatic change. *Int. J. Biometeorol.* **1980**, *24*, 65–78.
40. Berman, A.; Folman, Y.; Kaim, M.; Mamen, M.; Herz, Z.; Wolfenson, D.; Arieli, A.; Graber, Y. Upper critical temperatures and forced ventilation effects for high-yielding dairy cows in a subtropical climate. *J. Dairy Sci*. **1985**, *68*, 1488–1495.
41. Lunneborg, C.E. Estimating the correlation coefficient: The bootstrap approach. *Psychol. Bull.* **1985**, *98*, 209–215.
42. Rasmussen, J.L. Estimating correlation coefficients: Bootstrap and parametric approaches. *Psychol. Bull.* **1987**, *101*, 136–139.

Chapter 3

A Survey of the Relationship between Climatic Heat Stress Indices and Fundamental Milk Components Considering Uncertainty (*published*)

The purpose of this part is to assess the relationship between four bioclimatic indices for cattle (environmental stress, heat load, modified heat load, and respiratory rate predictor indices) and three main milk components (fat, protein, and milk yield) considering uncertainty.

OPEN ACCESS
climate
ISSN 2225-1154
www.mdpi.com/journal/climate

Article

A Survey of the Relationship between Climatic Heat Stress Indices and Fundamental Milk Components Considering Uncertainty

Mohammad Reza Marami Milani [1,*], Andreas Hense [2], Elham Rahmani [2] and Angelika Ploeger [1]

[1] Department of Organic Food Quality and food Culture, University of Kassel, Nordbahnhofstr. 1a, Witzenhausen 37213, Germany; E-Mail: a.ploeger@uni-kassel.de
[2] Meteorological Institute, University of Bonn, Auf dem Hügel 20, Bonn 53121, Germany; E-Mails: ahense@uni-bonn.de (A.H.); erahmani@uni-bonn.de (E.R.)

[†] These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: marami@uni-kassel.de; Tel.: +49-228-735-101; Fax: +49-228-735-188.

Received: 21 July 2015 / Accepted: 4 November 2015 / Published: 13 November 2015

Abstract: The main purpose of this study is to assess the relationship between four bioclimatic indices for cattle (environmental stress, heat load, modified heat load, and respiratory rate predictor indices) and three main milk components (fat, protein, and milk yield) considering uncertainty. The climate parameters used to calculate the climate indices were taken from the NASA-Modern Era Retrospective-Analysis for Research and Applications (NASA-MERRA) reanalysis from 2002 to 2010. Cow milk data were considered for the same period from April to September when the cows use the natural pasture. The study is based on a linear regression analysis using correlations as a summarizing diagnostic. Bootstrapping is used to represent uncertainty information in the confidence intervals. The main results identify an interesting relationship between the milk compounds and climate indices under all climate conditions. During spring, there are reasonably high correlations between the fat and protein concentrations *vs.* the climate indices, whereas there are insignificant dependencies between the milk yield and climate indices. During summer, the correlation between the fat and protein concentrations with the climate indices decreased in comparison with the spring results, whereas the correlation for the milk yield increased. This methodology is suggested for studies investigating the impacts of climate variability/change on food and agriculture using short term data considering uncertainty.

Keywords: climate variability; climate indices; heat stress indices; milk components; bootstrap; correlation

1. Introduction

Climate variability and climate change have direct effects on livestock through stress factors during periods of high temperature as well as indirect effects on their feed and diet [1]. Coping with climate variability is easier for animals than for crops due to their ability to adapt to different conditions and feed management systems. Thus, investigating the effect of climate variability on livestock and their products is complex. This interaction needs to be investigated using different climatic indices that have considered various climatic parameters. Additionally, genetic parameters play a strategic role in heat stress tolerance by reducing the effects of heat stress on dairy cattle [2]. According to the study by Lambertz *et al.* [3] in different housing systems, heat stress resulted in decreasing milk yield, fat, and protein percentages under warm and cold climate conditions with and without access to grazing. Sharma *et al.* [4] reported that the quality and quantity of milk compounds changed over time due to environmental and climate variability. These authors showed that the milk yield of Holsteins was more sensitive to the climate than the milk yield of Jerseys, whereas the other milk composition of the Jerseys was more sensitive. The milk yield of Holsteins was also more sensitive to heat stress compared to Jersey cattle.

The strongest focus of studies in this field over the last five decades was the effect of the Temperature Humidity Index (THI), which is a heat stress index used for livestock. THI is a combination of the temperature and humidity.

The main goal of the current study was to determine which climate parameter or bioclimatic index was most influential on cow milk as a strategic product for human nutrition. To partially answer the question, we investigated the relationship between physiologically relevant climate indices of cattle (Environmental Stress Index (ESI), Heat Load Index (HLI), modified HLI (HLI_{New}), and Respiratory Rate Predictor (RRP)) and milk compound observations (milk yield, fat, and protein). Another purpose of this study was to compare the effect of these relatively new heat stress indices on dairy cows under different climate conditions.

A first explanatory analysis of our study also focused on the THI and Equivalent Temperature Index (ETI) [5]. Thus, in this study we applied relatively new bioclimatic indices such as the ESI, HLI, and RRP to evaluate the interactions of temperature, humidity, wind speed, and solar radiation with milk components. These indices were applied for the first time and had the advantage of enabling the direct use of solar radiation, relative humidity, and wind speed in the heat stress equation. A summary of the reviews concerning the climate indices used in this study is presented in Table 1.

The potential of respiratory and skin evaporation by the animal is reduced under high humidity conditions. The effect of high temperature is decreased by wind, but solar radiation increases the heat effect on metabolic processes [6]. The first effect is due to the reduction of latent heat loss in moist environments, whereas the second effect results from general increases in the bulk sensible and latent heat loss from the body due to increasing wind speed that enhances turbulence around the body. In several studies, a positive correlation was reported between the dry-bulb temperature, solar radiation, relative humidity, and respiration rate index with the physiological reactions of livestock [7–12]. For instance, Mader and Davis reported a negative correlation between wind speed and respiration [13].

Therefore, a general assessment of the relationships between physiologically motivated climate indices and milk compounds is of central interest for predictions of the effects of large-scale changes in climate variables upon cattle livestock productivity. We developed a linear regression analysis using the bootstrapping method to assess the inherent sampling uncertainty in the data. The correlation coefficient was extracted from each individual regression analysis as a summarizing diagnostic. Bootstrapping is especially important when using data with a relatively small size, an expensive or difficult data collection procedure, or data that possess basic uncertainties (e.g., due to unknown farm management practices) [5].

The results of this study may be used to find optimum situations to minimize the negative effect of short term climate variability and longer term climate change on milk compounds. Moreover, the results may allow the utilization of more suitable conditions in the design of dairy factories according to the economic justifications of dairy production and the feeding of cattle that will allow the system to be managed with less risk of climate variability, especially in developing countries.

Based on published studies from different regions of the world, there was a lack of studies in Iran. Therefore, correcting this deficit was a second purpose for this study.

Table 1. Published studies of climatic indices in different regions of the world.

Year	Indices	Abbr. of Indices	Author(s)
2001, 2003	Environmental Stress Index	ESI	Moran *et al.* [14,15]
2002	Heat Load Index	HLI	Gaughan *et al.* [16]
2003	Respiratory Rate Index	RR or RRP	Eigenberg *et al.* [17]
2008	Heat Load Index (modified)	HLI	Gaughan *et al.* [18]

This paper also focused on data preparation and on introducing statistical methods to cope with uncertainties in special data sets, which is similar to a previous study by these authors [5]. In this study, climate indices were selected from the literature to identify basic relationships between milk compound data and climate indices.

The particular choice of linear regression and correlation analysis was motivated by future investigations on the multivariate regression relationships between climate variables and milk production.

2. Materials and Methods

The main purpose of this study was to evaluate the relationship between relatively new heat stress indices and milk compounds. Therefore, additional climatic indices were calculated from several climate parameters to provide more combinations that could affect the quality and quantity of milk compounds. The Environmental Stress Index (ESI), Heat Load Index (HLI), modified HLI (HLI_{New}), and Respiratory Rate Predictor (RRP) indices were used in this study. To determine these complex indices based on fundamental physiological reasoning, climate variables such as solar radiation, two meter height (T_{2m}), dew point, relative humidity, wind speed, sea level pressure, and specific humidity were taken into account. These variables were extracted from the Modern Era Retrospective-Analysis for Research and Applications (MERRA) undertaken by NASA's global modeling and assimilation office [19]. This study is a statistical downscaling experiment of coarse climate information to local or regional information from the biosphere. Similar approaches for plant phenology have been reported (e.g., in Maak and v. Storch [20] or Matulla *et al.* [21]). These studies showed that an explicit downscaling of coarse climate information to climate information on the local scale was not necessary. Instead, a direct statistical fit between the coarse climate data and the biological variable in question provided very valuable insights. During the review process, one of the reviewers noted that humidity and wind could not be directly downscaled from the large to local scales without extended modeling. This is true, and we consider the regression model a pragmatic, readily applicable first-order statistical approximation of this model in place of resource-intensive dynamic downscaling using a regional climate model.

The milk component data consisted of the fat, protein, and milk yield in the spring and summer in Iran. A linear regression analysis was performed using the milk compound data as the predictands and the climate indices as the predictors. The bootstrap method, which represents an active field of research in statistics for dependent data [22,23], was applied to consider the uncertainty and construct confidence intervals for the regression coefficients or correlations. Lunneborg [24] reported that bootstrapping expanded the confidence intervals. In contrast, Rasmussen [25] demonstrated that bootstrapping overly limited the confidence intervals. Jhun and Jeong [26] also reported that the bootstrap method presented more reliable results.

2.1. Geography of the Study Domain

This study was performed in three selected zones in the northwest, north, and central regions of Iran according to data availability and different climate conditions (cold semi-arid, Caspian mild and humid, semi-warm, and semi-arid). Figure 1 shows the assumed geographical and climatic areas in this study [27–29].

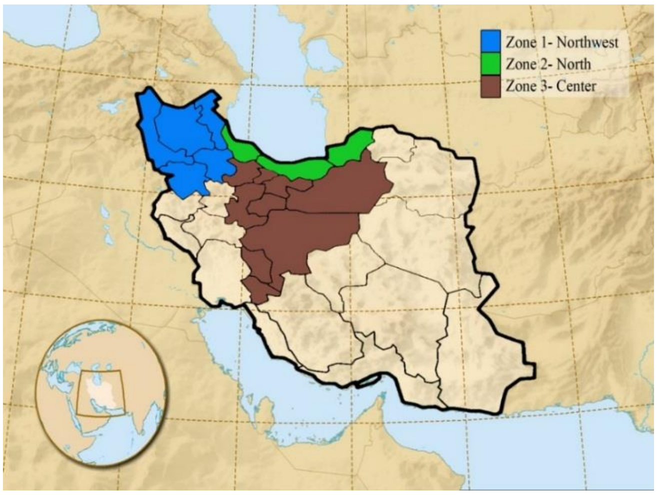


Figure 1. Classification of the considered study zones according to various climate conditions in Iran [27].

2.2. Observed and Reanalysis Data and Calculated Climate Indices

In Iran, 18,295 industrial herd stations with approximately 1.3 million cows are active from a total of 25,353 stations with the capacity for approximately 3.3 million cows. A total of 66% of the cows in Iran are pure Holstein, Jersey, and Brown Swiss, 27% are hybrids, and 7% are home-born cows. According to the published statistics, the total milk yield in Iran in 2013 was 3792 tons [30].

We considered important factors that could affect milk data gathering conditions (*i.e.*, genetics, cow age, environmental conditions, season, lactation, pregnancy, and feeding management) from industrial herd stations with good health services under controlled conditions and veterinary care.

In this study, we used the monthly average of test-day (TD) records of milk yield (kg), with the TD representing the day that the cow was milked. The milk yield and three-times-a-week records of fat and protein contents (g/100 mL milk) were collected from almost 600 industrial Holstein herd stations. The herd sizes varied between 75 and 200 cows, and the cows had access to grazing from April to September from 2002 to 2010. The final milk monthly averaged data bank was a matrix of $3 \times 936{,}227$ individuals comprised of milk yield, fat, and protein data. Data were gathered under the condition that the cows were on days four to 305 of the milking time and between their third and sixth calves over their lifetime without a record of mastitis throughout the entire study period. All of the records from cows that had a dry period or mastitis illness were omitted from the primary raw data set. Mastitis is known to be one of the most prevalent illnesses of cows and to have a major effect on milk yield [31]. Finally, the individual records from all herds within one of the three climatic zones and the selected month were averaged to enhance a potential joint signal. Bootstrapping with a resampling N = 1000 was employed to assess the sampling uncertainty of this final data preparation step. This step basically represents a resampling with replacement among the herds/stations within one of the three climatic zones in each month of the years 2002 to 2010. For further details, the reader is referred to Section 2.3.

Climate data were taken from the MERRA reanalysis [19] data set. This data set has a nominal resolution of 1/2 degrees latitude and 2/3 degrees longitude covering the same months and years as the milk data. THI is a widely used indicator of thermal conditions and the heat stress index. Marami *et al.* [5] reported a correlation of THI with fat, milk yield, and protein, with values of (−0.47, −0.22, 0.37) in spring and (−0.32, −0.26, 0.31) in summer, respectively. According to the study by West [32], variability in other climate factors, such as solar radiation, wind speed, and their interactions, affected the performance of dairy cows. We also selected relatively new indices based on these predictors. The four new climate indices are defined in the following sections.

These indices are the Environment Stress Index (ESI), Heat Load Index (HLI), modified Heat Load Index (HLI_{New}), and Respiratory Rate Predictor (RRP). The indices are calculated using the corresponding climate data parameters from the daily values at individual grid points and subsequently averaged to months and overall grid points within each of the three climatic zones. This step was performed based on previous results [20,27] that showed that, typically, the area-averaged large-scale climate information provided the most important predictor for variability in the biosphere. Due to the nonlinear dependencies between the climate indices and input climate variables (especially for HLI and ESI; see Equations 1 and 5), the results would be different when monthly or seasonally averaged climate data were used.

The structure of the data set for climatic and milk parameters is presented in Figure 2. We separately considered the parameters during the period from the years 2002 to 2010 for summer (July, August, and September) and spring (April, May, and June).

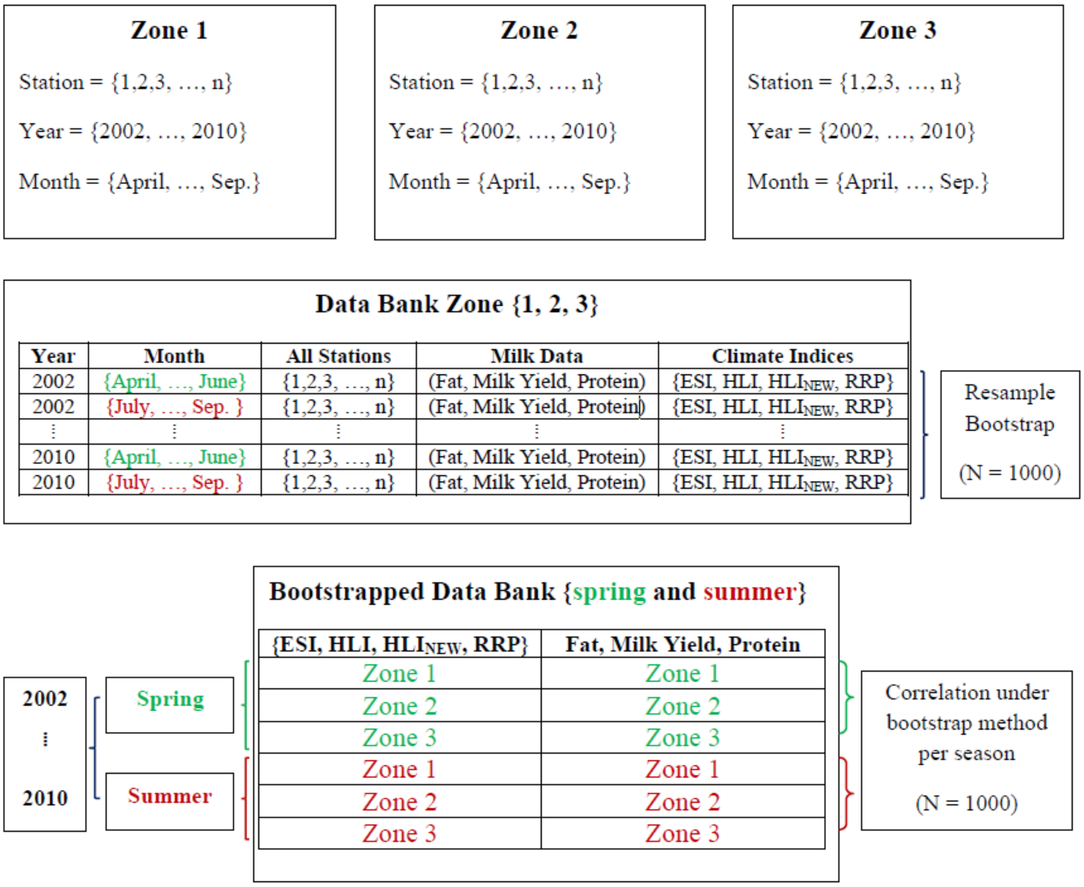


Figure 2. Structure of the climatic and milk parameter data set.

In this study, all of the analyses were implemented using R version 3.0.2, which provided various packages for statistical data analysis, calculations, and graphical display [33].

2.2.1. Environmental Stress Index (ESI)

Moran *et al.* [14] introduced an environmental stress index (ESI) as a substitute for the wet bulb globe temperature (WBGT). The ESI index considers the Ta (ambient temperature; ℃), RH (relative humidity; %), and SR (solar radiation; wm^{-2}). According to the study of Epstein and Moran [34], the WBGT index is well justified for many regions but is hard to use. The ESI is highly correlated with the WBGT ($R^2 \geq 0.981$) even for hot-wet and hot-dry climate conditions [35].

$$\mathrm{ESI} = 0.63\mathrm{Ta} - 0.03\mathrm{RH} + 0.002\mathrm{SR} + 0.0054\mathrm{Ta} \times \mathrm{RH} - 0.073(0.1 + \mathrm{SR})^{-1} \quad (1)$$

The RH is calculated by the Magnus formula [36] using other climate parameters, such as vapor pressure (e) and saturated vapor pressure (e_w^*)

$$e = 1.6077pq \tag{2}$$

where *e* is the vapor pressure (hPa), *p* is the local pressure (hPa), and *q* is the specific humidity. The difference between the local and sea surface pressure can be important in mountainous areas, where the local pressure *p* is less than the sea level pressure and leads to an overestimation of the vapor pressure.

$$e_w^* = 6.1\ \exp(\frac{17.6\,T}{T + 243.5}) \tag{3}$$

where e_w^* is the saturated vapor pressure (hPa) and is calculated by the two meter height (T_{2m} in ℃).

RH (relative humidity; %) is calculated as

$$\mathrm{RH} = \frac{e}{e_w^*} \tag{4}$$

2.2.2. Heat Load Index (HLI)

The HLI was suggested for the first time by Gaughan *et al.* [16]. The HLI is based on V (wind speed; m/s), RH (relative humidity; %), and $\mathrm{T_g^*}$ (predicted globe temperature; ℃).

$$\mathrm{HLI} = 33.2 + 0.2\mathrm{RH} + 1.2\mathrm{T_g^*} - (0.82\mathrm{V})^{0.1} - \log(0.4\mathrm{V}^2 + 0.0001) \tag{5}$$

$$\mathrm{T_g^*} = 1.33\mathrm{T} - 2.65\mathrm{T}^{\frac{1}{2}} + 3.21\log\,(\mathrm{SR} + 1) + 3.5 \tag{6}$$

Since 2008, HLI has been modified ($\mathrm{HLI}_{\mathrm{New}}$) by Gaughan *et al.* [18] to two formulas based on a globe temperature above and below 25 ℃.

When $\mathrm{T_g^*}$ < 25 ℃,

$$\mathrm{HLI}_{\mathrm{New}} = 10.66 + 0.28\mathrm{RH} + 1.3\mathrm{T_g^*} - \mathrm{V} \tag{7}$$

When $\mathrm{T_g^*}$ > 25 ℃,

$$\mathrm{HLI}_{\mathrm{New}} = 8.62 + 0.38\mathrm{RH} + 1.55\mathrm{T_g^*}\ - 0.5\mathrm{V} + \mathrm{e}^{2.4-\mathrm{V}} \tag{8}$$

Silva *et al.* [6] showed that the first version of the HLI was the best thermal stress index under tropical conditions. The classification by HLI can be divided into four categories: cool (≤70.0), warm (between 70.1 and 77.0), hot (between 77.1 and 86.0), and very hot (≥86.0) [18].

2.2.3 Respiratory Rate Predictor Index (RRP)

The RRP was suggested by Eigenberg *et al.* [17] for no-shade conditions and consisted of the wind speed (V; m/s), RH (relative humidity; %), T (predicted globe temperature; ℃), and SR (solar radiation; $\mathrm{wm^{-2}}$).

$$\mathrm{RRP} = 5.4\mathrm{T} + 0.58\mathrm{RH} - 0.63\mathrm{V} + 0.024\mathrm{SR} - 110.9 \tag{9}$$

where *V* is the wind speed (m/s) derived for the zonal and meridional wind components.

There are four categories in the RRP classification: normal (≤ 85), warning (between 85 and 110), danger (between 110 and 133), and emergency (≥ 133).

2.3. Uncertainty Consideration by Applying the Bootstrap Technique

The statistical analysis of short term data sets such as those used in this study requires careful consideration of their inherent uncertainty. This consideration is essential to avoid misleading results due to the small sample size.

A comparatively large data set of milk compounds in Iran is available. However, analyzing the statistically meaningful relationship with the inter-annual climate data results in a distinctive decrease in the sample size because only nine years of data (2002−2010) are available to estimate the dependency between near-surface climate variations and milk compound changes. Although a nine-year period of data is not a short term data set in agricultural investigations, in climate research nine years might represent a relatively small-sized data set.

Bootstrapping is a resampling technique that is a very valuable statistical method for the assessment of the influence of uncertainty within the available data sample on the results of the analysis [37] Bootstrapping is based on the empirical probability distribution function or the cumulative distribution function (cdf) of the data sample. Bootstrap methods use simulations to calculate standard errors, confidence intervals, and significance tests. They can be applied for any level of modeling and any type of statistical analysis [38]. The method used to estimate the cdf from the original sample distinguishes between parametric and non-parametric bootstrapping.

In the parametric bootstrapping method, the values of the mean and variance of an assumed Gaussian cdf are estimated from the original sample. Then, new sample values are generated by plugging these parameters and randomly drawing from this cdf [39].

The second way to generate new bootstrap samples is by estimating the cdf non-parametrically by sorting the original sample in increasing order. Sampling at random from this empirical cdf leads directly to a random sampling of the original data set with replacement such that, in the new sample, some of the original data are left out and replaced with copies of the remaining original data.

In this study, the non-parametric bootstrap technique was applied to the data and correlation analysis for the comparison of the milk and climate data. Thus, in the first step, the bootstrap technique was applied to the data from each station (*i.e.*, herds) in each zone separately by generating 1000 samples of the monthly means of the milk yield, and fat and protein concentrations for the full period (2002–2010). In the second step, a correlation analysis was performed using the non-parametric bootstrapping analysis instead of the classical linear correlation approach.

2.4. Linear Regression Analysis between Climatic Indices and Milk Compounds with Bootstrapping

The linear relationships between the indices (ESI, HLI, HLI_{New}, and RRP) and the milk compounds (milk yield, protein, and fat) were estimated via linear regression analysis with bootstrapping. The structure of the data used in this analysis is shown in Figure 2.

The inherent uncertainty was assessed by the second bootstrapping analysis. The nonparametric approach resampled among the time series from 2002 to 2010 for each month within a season and the climatic zones using the climate indices as the predictors and the milk compounds as the predictands. To account for the seasonal dependencies of the climatic input parameters and the physiological responses of cows, we analyzed two separate seasons (spring and summer) using three individual monthly averages within each season.

The linear regression analysis was performed in spring and summer during the time period from 2002 to 2010. The correlation coefficient was used as a diagnostic for each regression analysis based on N = 1000 bootstrap samples. In the resulting matrices, the respective correlation coefficients were averaged from the correlations across all climatic zones within each bootstrap sample with a sample size of 1000. Each of the N = 1000 bootstrap samples provided a linear Pearson correlation estimate coefficient across the nine years (2002 to 2010), the three individual months of each season, and the three climatic zones. For each correlation coefficient, the *p*-value for the Null hypothesis that the ensemble correlation was zero was calculated using Student's *t*-test implemented in the R function rcorr-test [33]. The interquartile range of the bootstrap sample of the correlation coefficients and the averaged bootstrap *p*-values were used to estimate the confidence intervals and the significance of the Null hypothesis of the underlying regression analysis.

3. Results and Discussion

3.1. Climatic Index and Milk Parameter Statistics

The averages of the milk parameters based on the yields of milk (TD), fat, and protein (g/100 mL of milk) of Holstein cows during spring and summer from 2002 to 2010 are presented in Table 2. On average, the seasonal mean milk yield decreased by 1 to 2 kg per day from spring into summer over the years 2002 to 2010. This decrease might be the effect of climate variability or the interaction of climate variability and the physiology of the cow and lactation. This seasonal effect was not obvious for the other two variables (fat and protein concentration).

Neither Knapp and Grummer [40] nor Roman-Ponce *et al.* [41] found any significant relationship between fat reduction and heat stress.

A decreasing milk yield was reported by Rodriquez *et al.* [42] and Bouraoui *et al.* [43], who also investigated the reduction in the fat, protein, and milk yield from spring to summer.

Table 2. Seasonal averages of fat, protein, and milk yield (Kg) of Holstein cows in spring and summer from 2002 to 2010.

	Fat (g/100 mL)		**Protein (gr/100 mL)**		**Milk Yield (kg)**	
	Spring	Summer	Spring	Summer	Spring	Summer
2002	3.14	3.13	2.92	2.83	26.16	25.61
2003	3.22	3.11	2.87	2.87	25.93	25.59
2004	3.14	3.07	2.89	2.91	25.04	23.8
2005	3.15	3.12	2.83	3.06	25.92	24.8
2006	3.15	3.13	2.96	2.98	26.78	25.33
2007	3.19	3.17	2.95	2.95	26.73	25.75
2008	3.19	3.18	3.05	3.04	27.23	25.98
2009	3.33	3.18	3.05	3.05	27.44	26.1
2010	3.34	3.27	3.03	2.95	26.64	26.26

No critical ESI value could be found from the literature to serve as the starting point of stress for the cows. The critical values of heat stress could vary under different climate conditions. Therefore, we reported this value in Table 3 and did not consider it in our analysis.

The average value of all indices increased from spring to summer. The monthly average values for the ESI, HLI, $HLI_{New,}$ and RRP indices changed from 13.32, 76.87, 67.62, and 5.85 in the spring to 16.33, 82.68, 75.07, and 31.88 in the summer, respectively.

Table 3. ESI, HLI, HLI $_{New}$, and RRP and the number of days that the cows were under stress (NDS) in the spring and summer from 2002 to 2010.

	Spring				Summer			
	ESI	HLI	HLI $_{New}$	RRP	ESI	HLI	HLI $_{New}$	RRP
Minimum	8.44	68.07	56.01	–36.19	13.40	77.18	68.17	6.64
Median	13.86	78.22	69.46	10.82	16.83	83.50	75.94	36.2
Average	13.32	76.87	67.62	5.85	16.33	82.68	75.07	31.88
Maximum	17.60	84.43	77.03	42.37	18.07	86.30	79.63	47.1
NDS	----	113	11	0	---	247	102	0

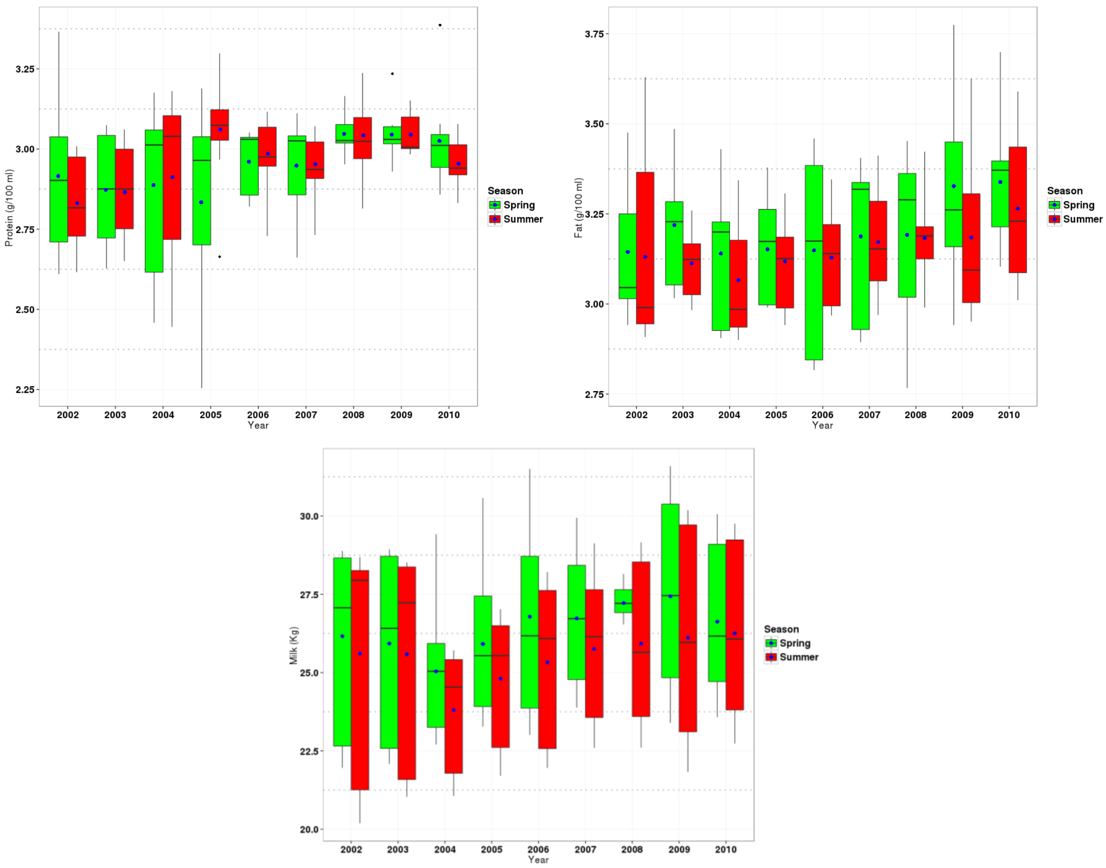


Figure 3. Lines inside the plot boxes indicate the median and the points show the season averages of the milk compounds milk yield (Kg), fat (g/100 mL milk), and protein (g/100 mL milk) in fresh products in the spring and summer from 2002 to 2010.

The monthly averages of the daily, minimum, maximum, and median of the climate indices (ESI, HLI, HLI $_{New}$, and RRP) in the spring and summer from 2002 to 2010 are shown in Table 3. Additional information on the average number of days of stress (NDS) for the cows is also provided in Table 3. The critical values of HLI and HLI $_{New}$ were assumed to be 77 for the starting point, indicating that the cows were probably subjected to heat stress. When the indices rose above 86, the cows were under extreme heat stress. The starting point of heat stress for RRP was assumed to be 110; when the value exceeded 133, the cows were in an emergency situation [44].

Figures 3 and 4 present a statistical overview of the milk yield and fat and protein content and the climate variability. The summarized statistics are separately shown in box plots for the short term data (2002–2010) in the summer (July, August, and September) and spring (April, May, and June) in the study area.

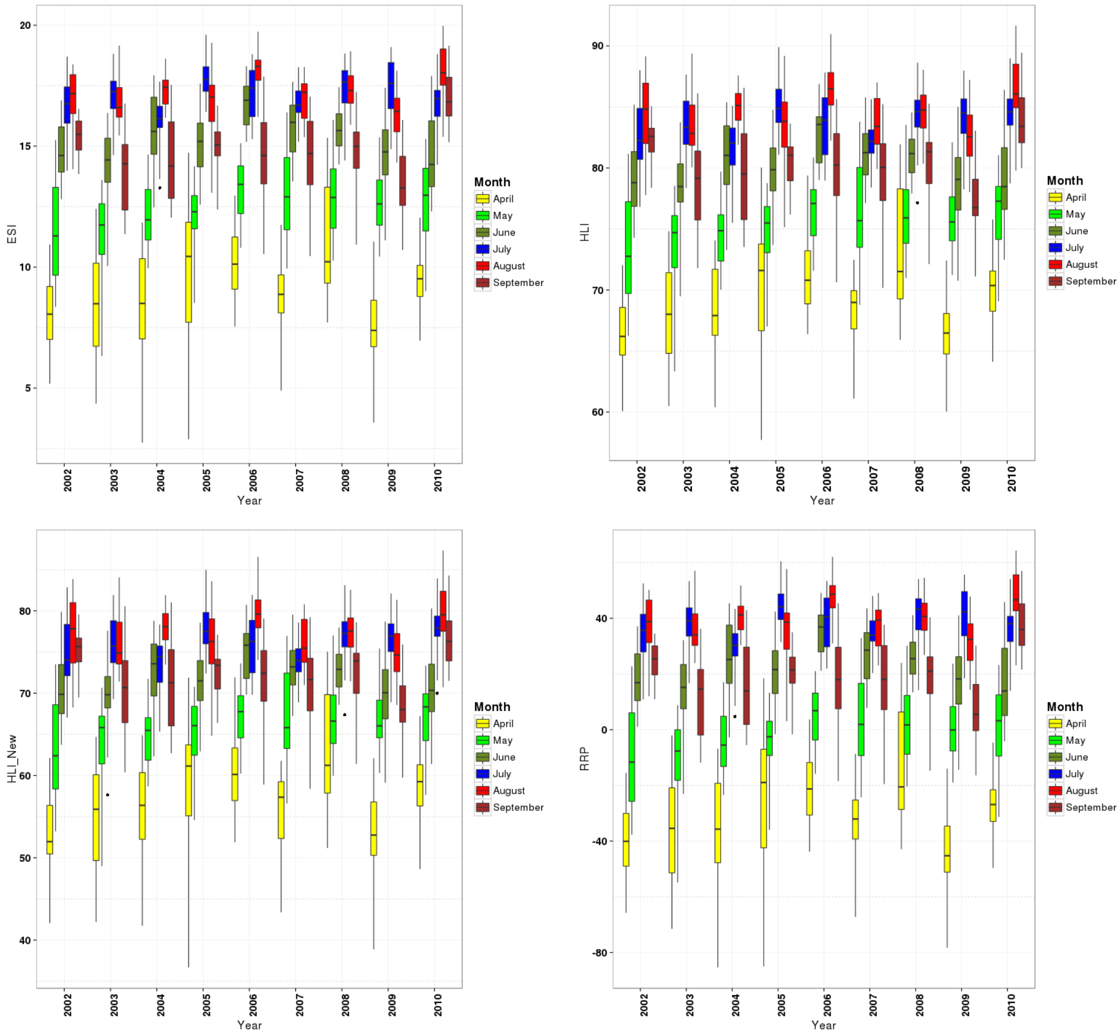


Figure 4. The box plots of ESI, HLI, HLI $_{New}$, and RRP from 2002 to 2010 in the spring and summer.

The box plots in Figure 3 represent the variability in milk compounds across all stations and years. There was hardly any visible seasonal signal in the data from the three milk compounds, which justified the use of only two seasons (spring (green boxes) and summer (red boxes)). The variability in the protein concentrations indicated by the width of the boxes changed from 2002 to the spring of 2005 to the summer of 2005 to 2010 without any clear changes in the median or mean. This lack of change is due to the shortage of available protein data from summer 2005.

Figure 4 shows the annual cycle of the climate indices across all stations and their short term changes from 2002 to 2010. A very prominent annual cycle can be observed for all climate indices, with a stronger change during the spring months compared with summer. This special configuration also justifies the selection of the two seasons.

3.2. Correlation between the Climatic Indices and Milk Parameters

Next, we investigated the relationship between the milk parameters and climate indices. Figures 5 and 6 present the correlation analysis between ESI, HLI, HLI_{New}, and RRP and the milk yield, fat, and protein in the spring and summer, respectively. The data were normalized to a mean of zero and standard deviation of one. For the correlation analysis, anomalies in the data were included and the bootstrapping technique was applied. The corrgram technique was used to visualize exploratory displays of the correlation matrix [45]. The values above the diagonal in the matrices are the respective correlation coefficients as averages of the correlations across all stations during the time period 2002–2010 within each bootstrap sample. As shown in Figure 2, bootstrapping with a random sample size of 1000 with replacement was applied to the total sample size of 81 ($9 \times 3 \times 3$) formed by all nine years in all three climatic zones in the spring (April, May, and June) and summer (July, August, and September). The values below the diagonal are the estimated confidence intervals from the inter-quartile ranges of the correlation coefficients across all bootstrapped samples. These values can be read as $\pm \Delta\rho$.

Figures 5 and 6 represent the correlation values between all variables in the spring and summer, respectively. The *p*-values for the correlation coefficients are presented in Figure 7, in which the upper diagonal shows the spring values and the lower diagonal presents the summer *p*-values.

In spring, the correlation results shown in the upper triangle in Figure 5 presented a positive correlation between all climate indices (ESI, HLI, HLI_{New}, and RRP) with the protein (almost 0.4, p-value < 0.05) and milk yield by a value of almost 0.3 (p-value$_{ESI} < 0.05$ and p-value$_{HLI, HLI\text{-}New, RRP} < 0.1$). Conversely, fat showed a negative correlation (−0.4, p-value < 0.05) with all climate indices.

Figure 5 also presented a positive correlation between the milk yield with fat (0.42, p-value < 0.05) and protein (0.28, p-value < 0.1) in the spring. This result suggested that the fat and protein concentrations increased with the increasing milk yield. Hammami *et al.* also presented a positive correlation between the milk yield with protein and fat [46]. Conversely, the negative correlation between fat and protein (−0.24, p-value < 0.1) was not significant.

As explained above, during spring, the fat and protein concentrations exhibited reasonably strong correlations (with typical correlation coefficients of 0.4 ± 0.07 and p-values < 0.05) with all climate indices. It should be kept in mind that these components themselves were strongly correlated with each other, which explained the variance of approximately 16% to 19% of the data. In contrast, the whole milk yield had a weak dependency (0.26 ± 0.06, p-values between 0.02 and 0.08) with the climate indices.

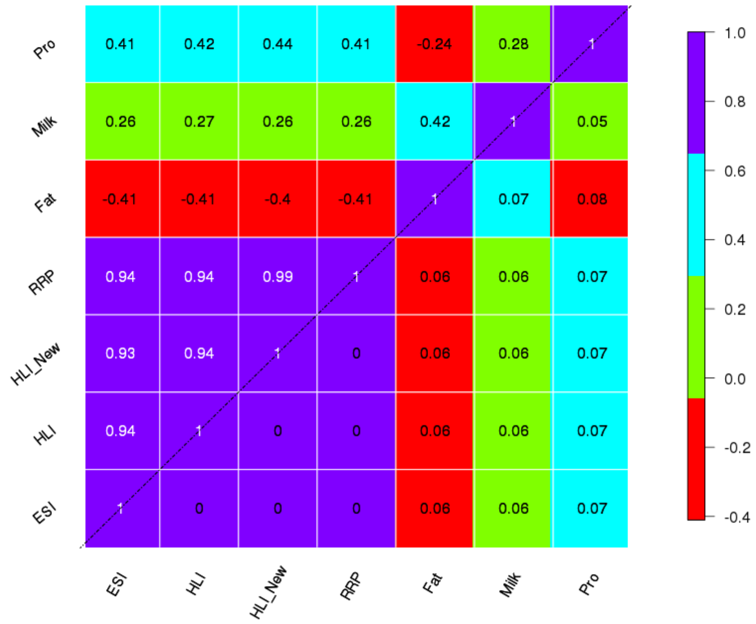


Figure 5. Correlation with bootstrapping between milk compounds and climatic indices in the spring (values under the diagonal are the estimated confidence intervals from the inter-quartile range).

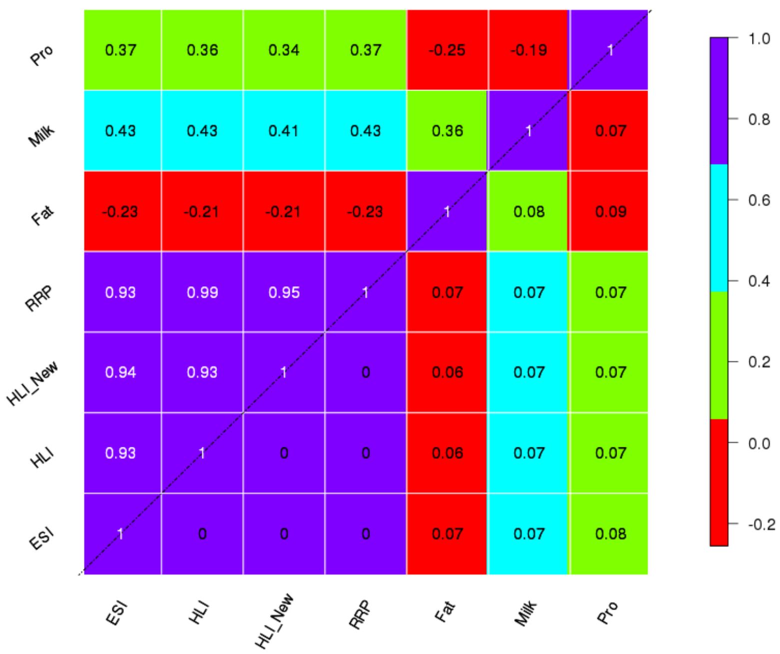


Figure 6. Correlation with bootstrapping between milk compounds and climatic indices in the summer (values under the diagonal are estimated confidence intervals from the inter-quartile range).

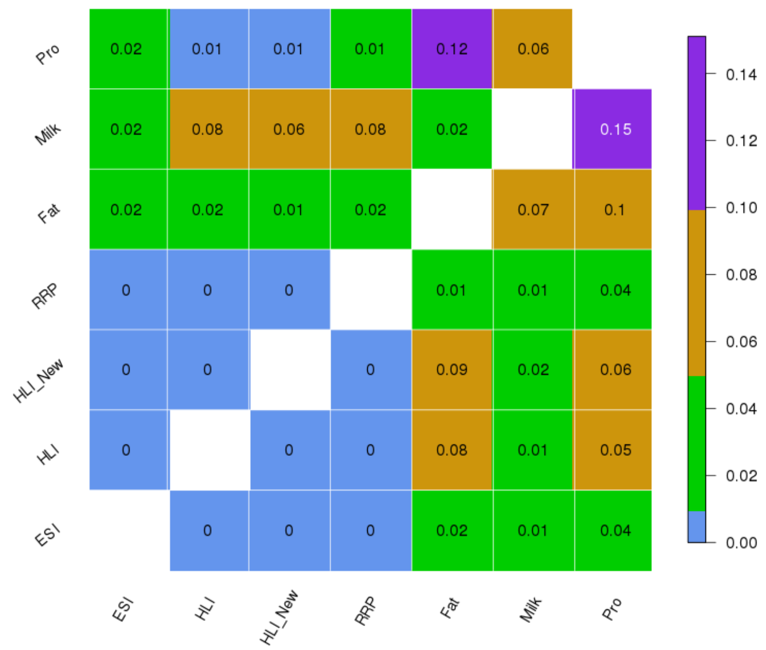


Figure 7. *P*-values of the correlations between milk compounds and climate indices. The upper diagonal indicates spring and the lower diagonal shows the summer *p*-values.

The correlation results for the summer shown in the upper triangle in Figure 6 also showed a positive correlation between milk yield and fat (0.36, *p*-value < 0.1). In contrast, a negative correlation existed between protein and fat (–0.25, *p*-value = 0.1). In contrast to spring, in summer the milk yield did not show a significant correlation with protein (*p*-value < 0.1). The squared value of the maximal positive correlation (milk yield *vs.* fat with a correlation 0.36) allowed for the quantification of the linear dependency between the predictand (or independent variable) and the predictor (or dependent variable) and explained 13% of the total variability *vs.* a remaining scatter with 87% of the total variability.

In summer, Figure 6 presented a positive correlation between all climate indices (ESI, HLI, HLI $_{\text{New}}$, and RRP) with the milk yield (almost 0.4, *p*-value < 0.05) and protein (almost 0.4, *p*-value $_{\text{ESI, RRP}}$ < 0.05 and *p*-value $_{\text{HLI, HLI-New}}$ < 0.1). In contrast, fat showed a negative correlation (−0.2, *p*-value $_{\text{ESI, RRP}}$ < 0.05 and *p*-value $_{\text{HLI, HLI-New}}$ < 0.1) with all climate indices. Fat presented less of a negative correlation with the climate indices in the summer than in the spring.

In summer, warm conditions dominate more of the climate parameters, bioclimatic indices, and cattle physiological conditions. Evidently, the dependency of the structure between milk variables and the climate indices changes between spring and summer, once again supporting the initial decision to split the data set into two seasons.

The milk yield showed the largest dependency with the climate indices (*p*-value < 0.05) compared with the spring season, whereas fat and protein showed less of a correlation compared to spring.

For both seasons, all climate indices (ESI, HLI, HLI $_{\text{New}}$, and RRP) showed a very high positive correlation with one another (*p*-value < 0.001). In contrast with the small dependency between the milk components, we observed a very high dependency between the climatic variables. Additionally, interesting relationships were present between the milk compounds and the climate indices.

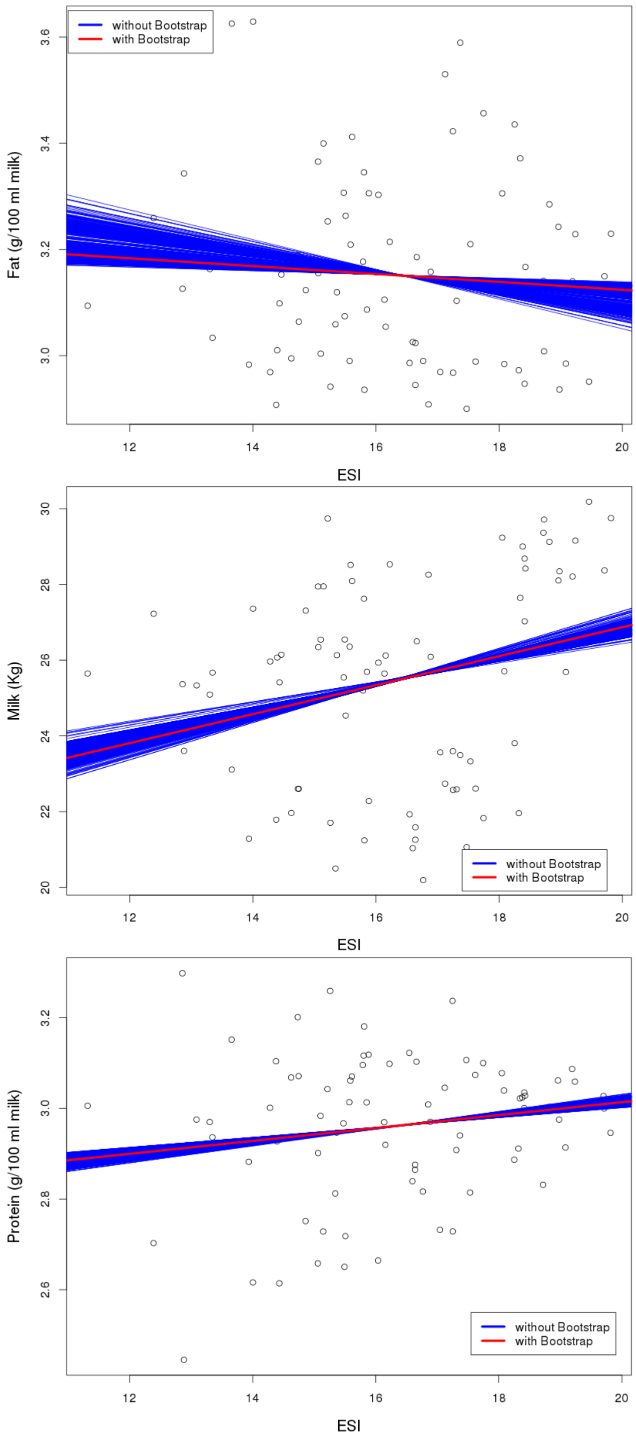


Figure 8. The best-fitted regression lines before (red line) and after 1000 bootstraps (blue lines) between milk components (fat, protein, and milk) and ESI in the summer.

According to the study by Marami *et al.* [5], THI showed less of a correlation with the milk yield and protein compared with the current indices used in this study but exhibited a slightly increased correlation (about +0.09) with fat in the spring and summer.

Figure 8 presents additional explanations regarding the regression analysis before and after bootstrapping (for more figures, please see the Appendix). For example, the milk components were considered predictands and ESI the predictor in the summer. In these samples, the slope of the best-fitting straight line between the variables presented a positive or negative correlation in the summer. The red line was the regression line without bootstrapping and consisted of the predicted score on y for each possible value of x. The blue lines showed the range of changes in the intercept and slope in the resampling data during the 1000 bootstraps. As shown in Figure 6, the values of the positive correlation between protein-ESI (0.37) and milk yield-ESI (0.43) arose from the positive slopes of the regression lines between these variables in Figure 8. These data accounted for approximately 14% and 18% of the linear dependency *vs.* 86% and 82% for the unexplained scatter. For the fat content *vs.* ESI, the regression line changed to a negative correlation (−0.23) and explained only 5% of the total variability *vs.* the remaining scatter, which explained 95% of the total variability.

4. Conclusions

This study investigates the most influential climate parameter or bioclimatic index on cow milk as a strategic product and presents a relationship between physiologically relevant climate indices for cattle (ESI, HLI, and RRP) and milk compound (milk yield, fat, protein) observations. We also compare the effects of these relatively new heat stress indices on dairy cows under different climate conditions.

The analysis was performed with the statistical bootstrapping method to compare the influence of dairy variables over a time series on climate variables over a time series while considering unavoidable uncertainty. The application of the bootstrapping method has advantages in these studies and can serve as a good tool to validate the predictive model for studies in developing countries with short term data sets.

Generally, the linear regression analysis exhibited a positive correlation between the milk yield, protein, and climate indices during the spring and summer but a negative correlation between fat and the climate indices. Protein seems to be less sensitive than fat to the negative effects of heat stress and climate variability. The results in this study agree with the results reported by other studies, such as [47–53].

Under the heat stress condition, there is a possibility for cows to choose to stay in the barn or to graze on the pasture in the pasturing system. It might be better to keep the cows indoors because shade plays an important role in management techniques to reduce the negative effect of heat stress [54].

Climate variability plays a reasonable role in cow physiological adaptation, fodder and nutrition, husbandry systems, DNA accuracy, and changes in their characteristics [44]. Based on characteristic changes and the cow's ability to adapt under different conditions, the effect of climate variability on animal products (*i.e.*, the critical point of heat stress) is also changed. Thus, cows may increase their tolerance to climate variability.

We strongly believe that new and significant indices are needed to control critical heat stress conditions that consider more predictors of the effect of climate variability on animal products, such as sunshine duration, the quality of the cow's diet, the number of days of stress (NDS), the color of skin with attention to large black spots, and categorical predictors such as breed, welfare facility, and management system.

We concluded that care must be taken in choosing the climate indices in different climate conditions to assess the relationship between milk components and climate variability. For instance, Holstein cows bred in tropical and subtropical conditions have differences in their hair coat characteristics compared with those bred in temperate regions [6,55].

Based on our results of this investigation into the effect of heat stress on milk compounds (milk yield, fat, and protein), we suggest the use of ESI and RRP in the summer and ESI in the spring. HLI_{New} is also suggested for fat and protein content in the spring. Silva *et al.* suggested HLI for tropical regions, and Moran and Epstein evaluated ESI for hot/dry and hot/wet climates [6,35].

To improve the correlations and achieve the best linear relationship, a study of the mixed effect of climate indices is strongly suggested.

Acknowledgments

The authors warmly thank Nasim Azari at the Food Science Engineering Department of the Science and Research branch of Tehran Islamic Azad University, Ahmad Moghimi Esfand Abadi and Asghar Salimi Niknam at the Animal Breeding Centre of Iran in Karaj, branch of the Iranian Ministry of Agriculture, and Behnam Saremi and Ali Asadi for supporting the milk data. We are also grateful to Ali Mortazavi in the Department of Food Science, Ferdowsi Mashhad University for his cooperation.

Author Contributions

Andreas Hense and Angelika Ploeger supervised the methodology of the study and the statistical methods used. Elham Rahmani contributed data and tools for the analysis. Mohammad Reza Marami Milani performed the statistical analysis, interpreted the results, and wrote the initial manuscript. All authors contributed by reading and editing the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Appendix

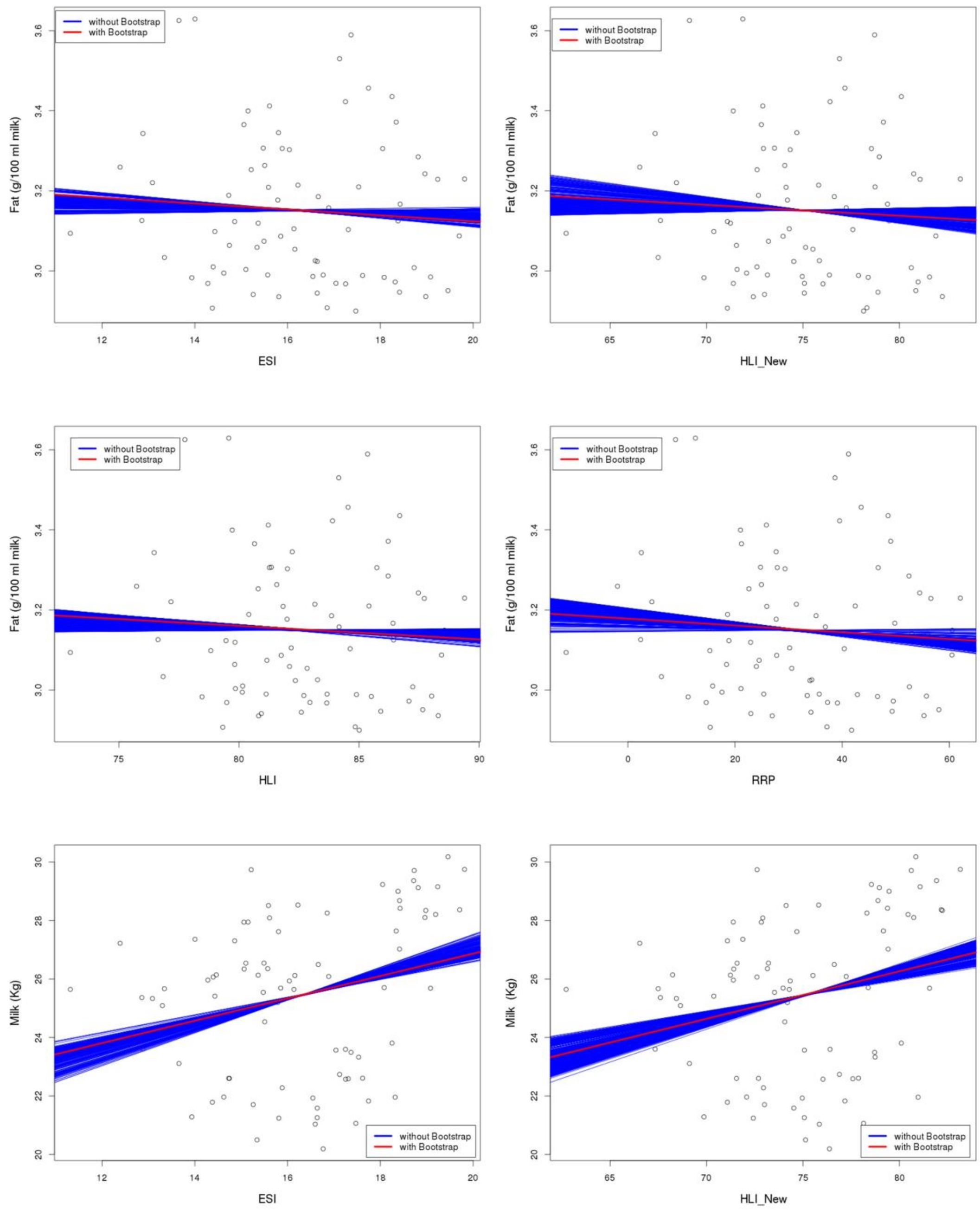


Figure A1. *Cont.*

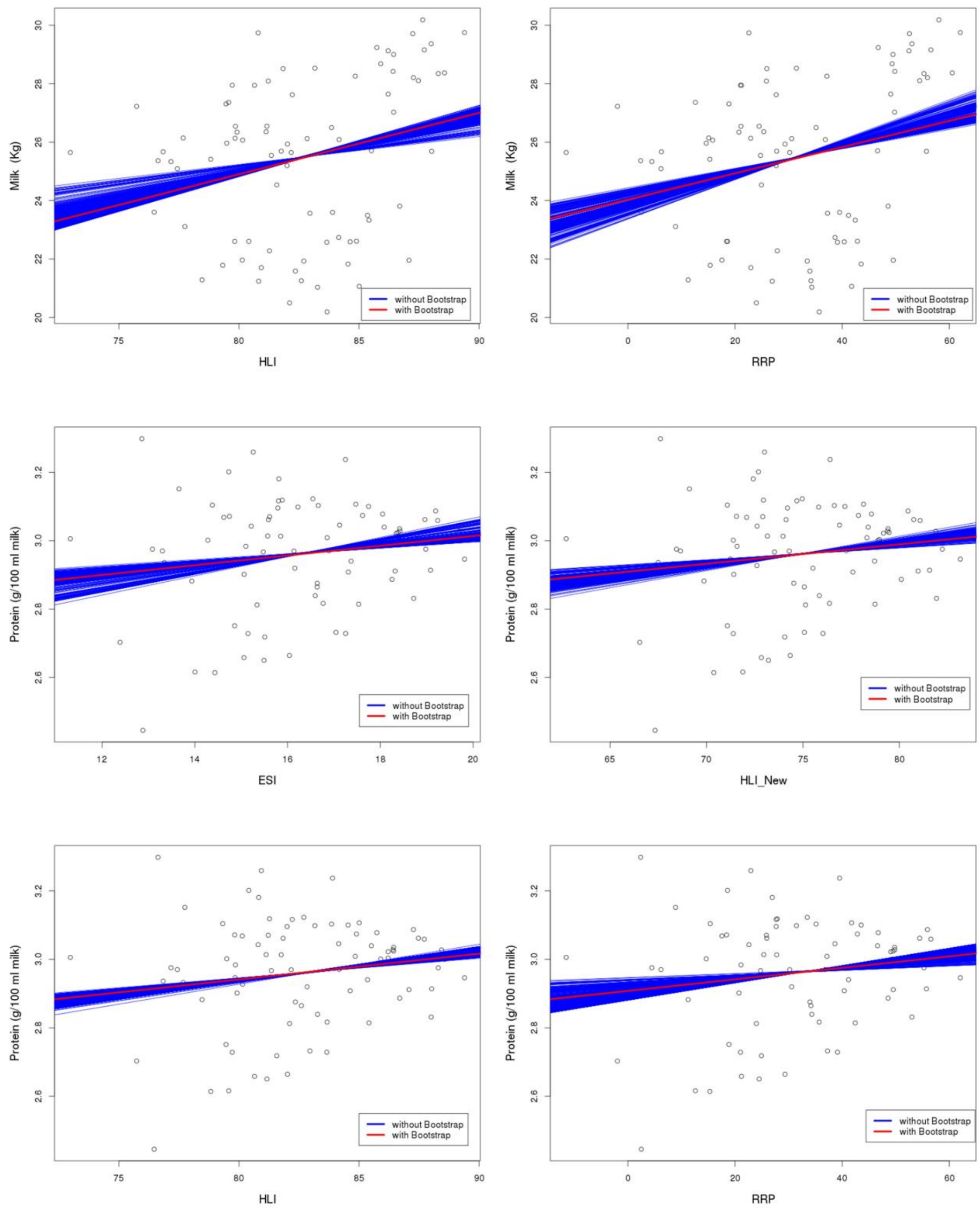


Figure A1. The best-fitted regression line before (red line) and after bootstrapping (N = 1000) (blue lines) between indices (ESI, HLI, HLI_{New}, and RRP) and milk components (fat, protein and milk yield) in the spring.

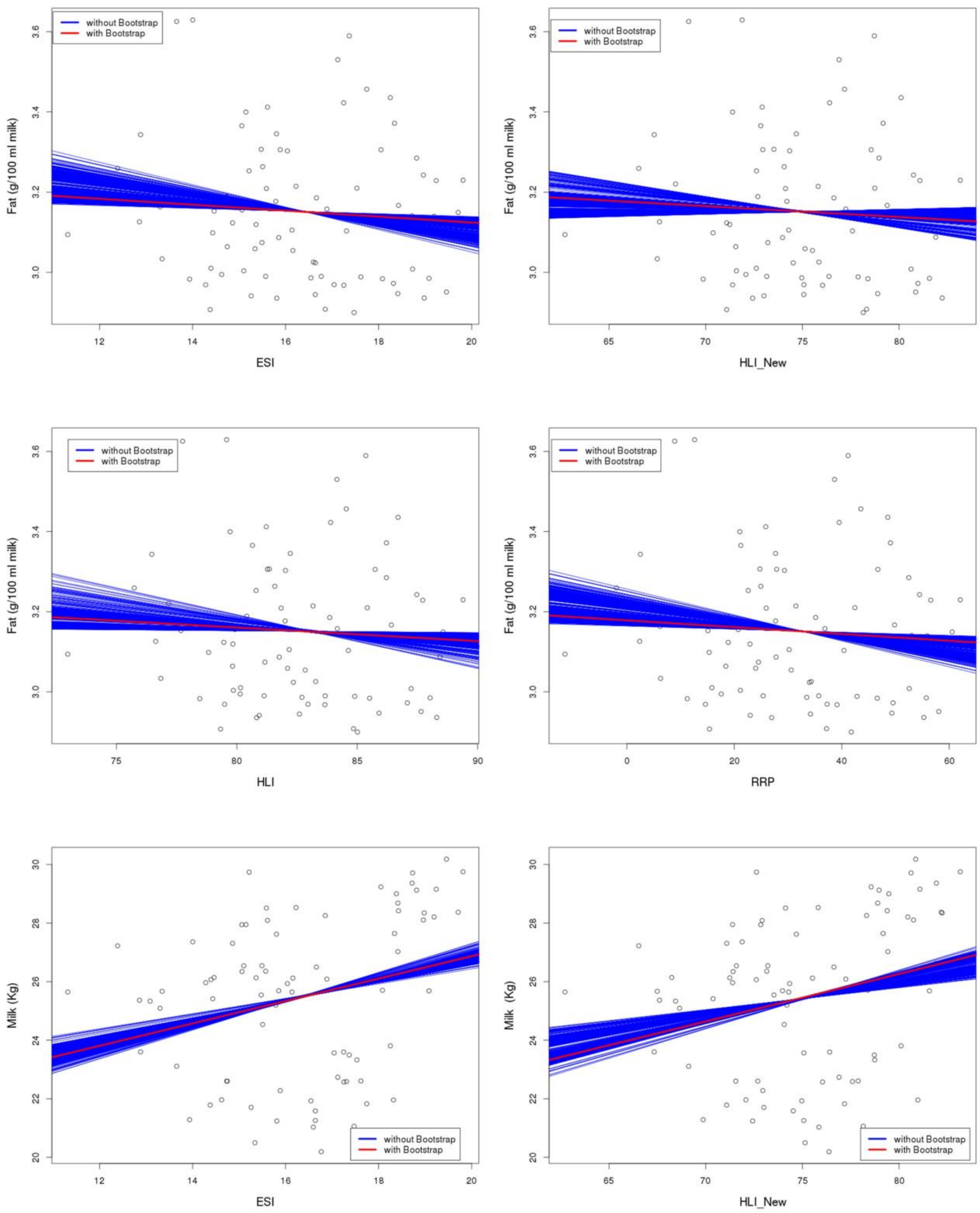


Figure A2. *Cont.*

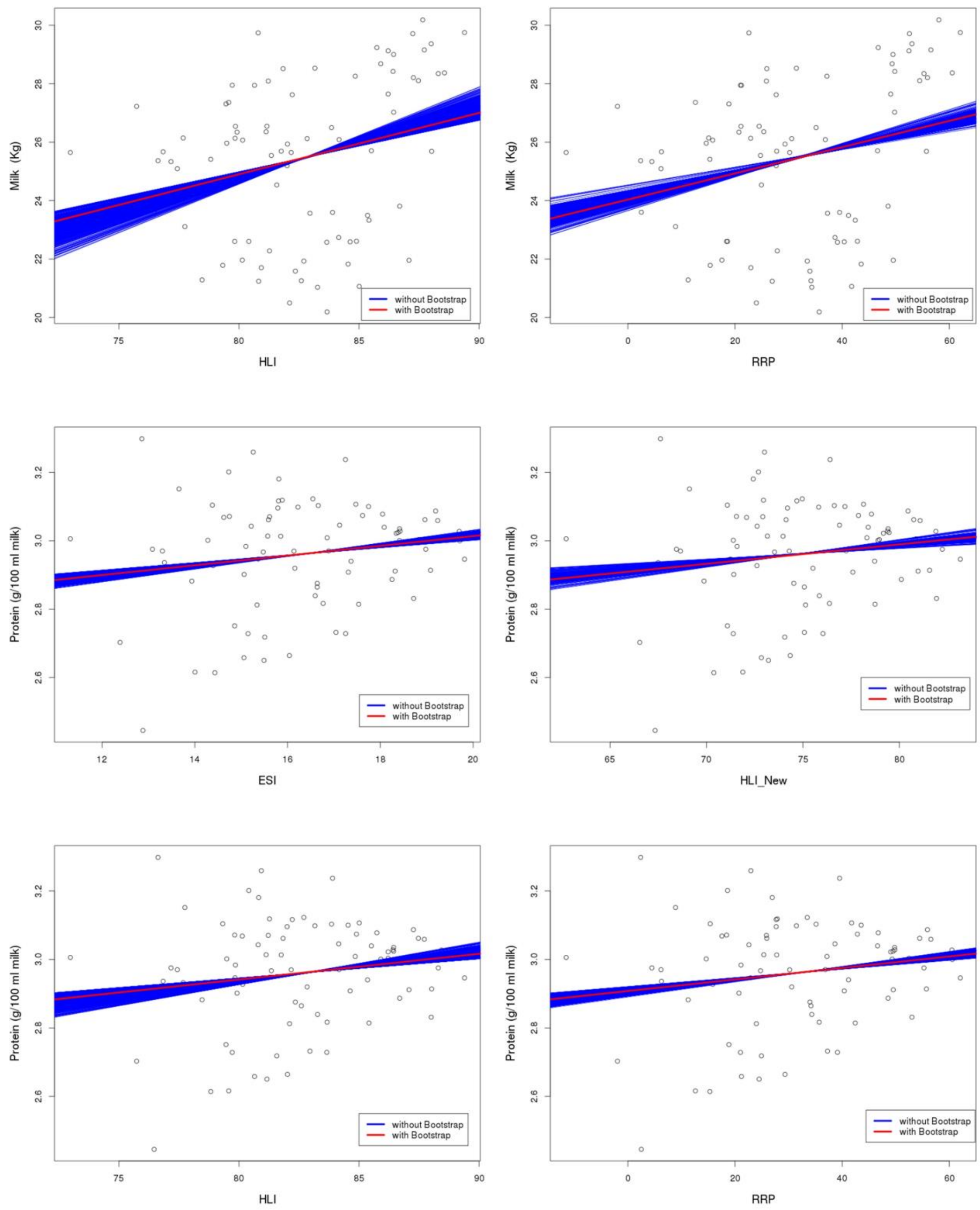


Figure A2. The best-fitted regression line before (red line) and after bootstrapping (N = 1000) (blue lines) between indices (ESI, HLI, HLI_{New}, and RRP) and milk components (fat, protein and milk yield) in the summer.

References

1. Watson, R.T.; Albritton, D.L.; Barker, T.; Bashmakov, I.A.; Canziani, O.; Christ, R.; Cubasch, U.; Davidson, O.; Gitay, H.; Griggs, D.; *et al.* Climate Chance 2001: IPCC Synthesis Report. Avalable oneline: http://ipcc.ch/meetings/session18/doc3b.pdf (accessed on 5 March 2015).
2. Renaudeau, D.; Collin, A.; Yahav, S.; De Basilio, V.; Gourdine, J.L.; Collier, R.J. Adaptation to hot climate and strategies to alleviate heat stress in livestock production. *Animal* **2012**, *6*, 707–728.
3. Lambertz, C.; Sanker, C.; Gauly, M. Climatic effects on milk production traits and somatic cell score in lactating Holstein-Friesian cows in different housing systems. *J. Dairy Sci.* **2014**, *97*, 319–329.
4. Sharma, A.K.; Rodriguez, L.A.; Mekonnen, G.; Wilcox, C.J.; Bachman, K.C.; Collier, R.J. Climatological and genetic effects on milk composition and yield. *J. Dairy Sci*. **1983**, *66*, 119–126.
5. Marami Milani, M.R.; Hense, A.; Rahmani, E.; Ploeger, A. A pilot investigation of the relationship between climate variability and milk compounds under the bootstrap technique. *Foods* **2015**, *4*, 420–439.
6. Silva, R.G.; Morais, D.A.; Guilhermino, M.M. Evaluation of thermal stress indexes for dairy cows in tropical regions. *R. Bras. Zootec.* **2007**, *36*, 1192–1198.
7. Eigenberg, R.A.; Hahn, G.L.; Nienaber, J.A.; Brown-Brandl, T.M. Development of a new respiration rate monitor for cattle. *Trans. ASAE* **2000**, *43*, 723–728.
8. Hahn, G.L.; Mader, T.L. Heat waves in relation to thermoregulation, feeding behavior and mortality of feedlot cattle. In Proceedings of the 5th International Livestock Environment Symposium, Bloomington, MN, USA, 29–31 May 1997; pp. 545–549.
9. Ingram, D.L.; Mount, L.E. Heat exchange between animal and environment. In *Man and Animals in Hot Environments*; Springer: New York, NY, USA, 1975; pp. 5–23.
10. Mader, T.L.; Dahlquist, J.M.; Hahn, G.L.; Gaughan, J.B. Shade and wind barrier effects on summertime feedlot cattle performance. *J. Anim. Sci.* **1999**, *77*, 2065–2072.
11. McLean, J.A. Loss of heat by evaporation. In *Heat Loss from Animals and Man: Assessment and Control*; Monteith, J.L, Mount, L.E., Eds.; Butterworth-Heinemann: London, UK, 1974; pp. 19–31.
12. Spain, J.N.; Spiers, D.E. Effects of supplemental shade on thermoregulatory response of calves to heat challenge in a hutch environment. *J. Dairy Sci.* **1996**, *79*, 639–646.
13. Mader, T.L.; Davis, M.S. Wind speed and solar radiation corrections for the temperature-humidity index. In Proceedings of the 15th Conference on Biometeorology and Aerobiology Joint with 16th International Congress on Biometeorology, Kansas City, MO, USA, 27 October 2002; pp. 10–28.
14. Moran, D.S.; Pandolf, K.B.; Shapiro, Y.; Heled, Y.; Shani, Y.; Matthew, W.T.; Gonzales, R.R. An environmental stress index (ESI) as a substitute for the wet bulb globe temperature (WBGT). *J. Therm. Biol.* **2001**, *26*, 427–431.
15. Moran, D.S.; Pandolf, K.B.; Shapiro, Y.; Laor, A.; Heled, Y.; Gonzalez, R.R. Evaluation of the environmental stress index for physiological variables. *J. Therm. Biol.* **2003**, *28*, 43–49.
16. Gaughan, J.G.; Goopy, L.; Spark, J. *Excessive Heat Load Index for Feedlot Cattle*. Meat and Livestock Australia: Sydney, NSW, Australia, 2002.

17. Eigenberg, R.A.; Nienaber, J.A.; Brown-Brandl, T.M. Development of a livestock safety monitor for cattle. In Proceedings of the 2003 American Society of Agricultural and Biological Engineers (ASABE) Annual Meeting, St. Joseph, MI, USA, 27–30 July 2003.
18. Gaughan, J.G.; Mader, T.L.; Holt, S.M.; Lisle, A. A new heat load index for feedlot cattle. *J. Anim. Sci*. **2008**, *86*, 226–234.
19. Rienecker, M.M.; Suarez, M.J.; Gelaro, R.; Todling, R.; Bacmeister, J.; Liu, E.; Bosilovich, M.G.; Schubert, S.D.; Takacs, L.; Kim, G.K.; *et al.* MERRA: NASA's Modern-Era retrospective analysis for research and applications. *J. Climate* **2011**, *24*, 3624–3648.
20. Maak, K.; von Storch, H. Statistical downscaling of monthly mean air temperature to the beginning of flowering of Galanthus nivalis L. in Northern Germany. *Int. J. Biometeorol.* **1997**, *41*, 5–12.
21. Matulla, C.; Scheifinger, H.; Menzel, A.; Koch, E. Exploring two methods for statistical downscaling of Central European phenological time series. *Int. J. Biometeorol.* **2003**, *48*, 56–64.
22. Lahiri, S.N. Bootstrap methods. In *Resampling Methods for Dependent Data*; Springer: New York, NY, USA, 2003; pp. 17–43.
23. Politis, D.N.; Romano, J.P.; Wolf, M. *Subsampling*; Springer: New York, NY, USA, 1999.
24. Lunneborg, C.E. Estimating the correlation coefficient: The bootstrap approach. *Psychol. Bull.* **1985**, *98*, 209–215.
25. Rasmussen, J.L. Estimating correlation coefficients: Bootstrap and parametric approaches. *Psychol. Bull.* **1987**, *101*, 136–139.
26. Jhun, M.; Jeong, H.C. Applications of bootstrap methods for categorical data analysis. *Comput. Stat. Data Anal.* **2000**, *35*, 83–91.
27. Rahmani, E. The Effect of Climate Variability on Wheat in Iran. Ph.D. Thesis, University of Bonn, Bonn, Germany, 2015.
28. Rahmani, E.; Friederichs, P.; Keller, J.; Hense A. Development of an effective and potentially scalable weather generator for temperature and growing degree days. *Theor. Appl. Climatol.* **2015**, doi:10.1007/s00704-015-1477-z.
29. Kousari, M.R.; Ahani, H.; Hendi-Zadeh, R. Temporal and spatial trend detection of maximum air temperature in Iran during 1960–2005. *Glob. Planet. Change* **2013**, *111*, 97–110.
30. Statistical Center of Iran / History. Available online: http://www.amar.org.ir/Default.aspx?tabid=281 (accessed on 05 August 2014).
31. Lescourret, F.; Coulon, J.B. Modeling the impact of mastitis on milk production by dairy cows. *J. Dairy Sci.* **1994**, *77*, 2289–2301.
32. West, J.W. Effects of heat-stress on production in dairy cattle. *J. Dairy Sci.* **2003**, *86*, 2131–2144.
33. R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2011.
34. Epstein, Y.; Moran, D.S. Thermal comfort and the heat stress indices. *Ind. Health* **2006**, *44*, 388–398.
35. Moran, D.S.; Epstein, Y. Evaluation of the environmental stress index (ESI) for hot/dry and hot/wet climates. *Ind. Health* **2006**, *3*, 399–403.
36. Kraus, H. Die Luftfeuchtigkeit. In *Die Atmosphäre der Erde*, 3rd ed.; Springer: Berlin, Germany, 2004; pp. 67–89.

37. Efron, B.; Tibshirani, R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *J. Statist. Sci*. **1986**, *1*, 54–75.
38. Davison, A.C.; Hinkley, D.V. *Bootstrap Methods and Their Application*; Cambridge University Press: New York, NY, USA, 1997.
39. Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *Numerical Recipes: The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2007.
40. Knapp, D.M.; Grummer, RR. Response of lactating dairy cows to fat supplementation during heat stress. *J. Dairy Sci*. **1991**, *74*, 2573–2579.
41. Roman-Ponce, H.; Thatcher, W.W.; Buffington, D.E.; Wilcox, C.J.; Van Horn, H.H. Physiological and production responses of dairy cattle to a shade structure in a subtropical environment. *J. Dairy Sci*. **1977**, *60*, 424–430.
42. Rodriquez, L.A. Effects of relative humidity, maximum and minimum temperature, pregnancy, and stage of lactation on milk composition and yield. *J. Dairy Sci.* **1985**, *68*, 973–978.
43. Bouraoui, R.; Lahmarb, M.; Majdoubc, A.; Djemalic, M.; Belyead, R. The relationship of temperature-humidity index with milk production of dairy cows in a Mediterranean climate. *J. Anim. Res*. **2002**, *51*, 479–491.
44. Silva, R.G.; Campos Maia, A.S. Thermal stress indexes. In *Principles of Animal Biometeorology*; Springer: Dordrecht, Netherlands, 2013; pp. 207–229.
45. Friendly, M. Corrgrams: Exploratory displays for correlation matrices. *Am. Stat.* **2002**, *56*, 316–324.
46. Hammami, H.; Vandenplas, J.; Vanrobays, M.L.; Rekik, B.; Bastin, C.; Gengler, N. Genetic analysis of heat stress effects on yield traits, udder health, and fatty acids of Walloon Holstein cows. *J. Dairy Sci.* **2015**, *98*, 4956–4968
47. Beatty, DT.; Barnes, A.; Taylor, E.; Pethick, D.; McCarthy, M.; Maloney, SK. Physiological responses of Bostaurus and Bosindicus cattle to prolonged continuous heat and humidity. *J. Animal Sci*. **2006**, *84*, 972–985.
48. Ciais, Ph.; Reichstein, M.; Viovy, N.; Granier, A.; Ogée, J.; Allard, V.; Aubinet, M.; Buchmann, N.; Bernhofer, Chr.; Carrara, A.; *et al.* Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *J. Nature* **2005**, *437*, 529–533.
49. Collier, R.J.; Dahl, E.; VanBaale, M.J. Major advances associated with environmental effects on dairy cattle. *J. Dairy Sci*. **2006**, *89*, 1244–1253.
50. Gauly, M.; Bollwein, H.; Breves, G.; Brugemann, K.; Danicke, S.; Das, G.; Demeler, J.; Hansen, H.; Isselstein, J.; Konig, S.; *et al.* Future consequences and challenges for dairy cow production system arising from climate change in center Europe—A review. *Animal* **2013**, *7*, 843–859.
51. Myneni, R.B.; Keeling, C.D.; Tucker. C.J.; Asrar, G.; Nemani, R.R. Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* **1997**, *386*, 698–702.
52. Nikkhah, A.; Furedi, C.J.; Kennedy, A.D.; Scott, S.L.; Wittenberg, K.M.; Crow, G.H.; Plaizier, J.C. Morning v.s. evening feed delivery for lactating dairy cows. *Can. J. Anim. Sci.* **2011**, *91*, 113–122.
53. O'Brien, MD. Metabolic adaptations to heat stress in growing cattle. *Domest. Anim. Endocrinol.* **2010**, *38*, 86–94.
54. Brown-Brandl, T.M.; Eigenberg, R.A.; Nienaber, J.A.; Hahn, G.L. Dynamic response indicators of heat stress in shaded and non-shaded feedlot cattle, Part 1: Analyses of indicators. *Biosyst. Eng.* **2005**, *90*, 451–462.

55. Udo, H.J.M. Hair Coat Characteristics in Friesian Heifers in the Netherlands and Kenya: Experimental Data and A Review of Literature. Ph.D. Thesis, University of Wageningen: Wageningen, The Netherlands, 1978.

Chapter 4

Applying Least Absolute Shrinkage Selection Operator and Akaike Information Criterion Analysis to Find the Best Multiple Linear Regression Models between Climate Indices and Components of Cow's Milk (*published*)

The purpose of this part is to evaluate new indices through combination of indices and climate parameters by multiple linear regression modeling. The new obtained indices might have the potential to be widely accepted and used because single heat stress index has not been universal acceptance in last 20 years (Brake and Bates 2002).

foods

MDPI

Article

Applying Least Absolute Shrinkage Selection Operator and Akaike Information Criterion Analysis to Find the Best Multiple Linear Regression Models between Climate Indices and Components of Cow's Milk

Mohammad Reza Marami Milani [1,*], Andreas Hense [2,†], Elham Rahmani [2] and Angelika Ploeger [1,†]

[1] Department of Organic Food Quality and Food Culture, University of Kassel, Nordbahnhofstr. 1a, 37213 Witzenhausen, Germany; a.ploeger@uni-kassel.de
[2] Meteorological Institute, University of Bonn, Auf dem Hügel 20, 53121 Bonn, Germany; ahense@uni-bonn.de (A.H.); erahmani@uni-bonn.de (E.R.)
* Correspondence: marami@uni-kassel.de or maramimilani@gmail.com; Tel.: +49-228-735-101
† These authors contributed equally to this work.

Academic Editor: Wijitha Senadeera
Received: 13 April 2016; Accepted: 18 July 2016; Published: 23 July 2016

Abstract: This study focuses on multiple linear regression models relating six climate indices (temperature humidity THI, environmental stress ESI, equivalent temperature index ETI, heat load HLI, modified HLI (HLI $_{new}$), and respiratory rate predictor RRP) with three main components of cow's milk (yield, fat, and protein) for cows in Iran. The least absolute shrinkage selection operator (LASSO) and the Akaike information criterion (AIC) techniques are applied to select the best model for milk predictands with the smallest number of climate predictors. Uncertainty estimation is employed by applying bootstrapping through resampling. Cross validation is used to avoid over-fitting. Climatic parameters are calculated from the NASA-MERRA global atmospheric reanalysis. Milk data for the months from April to September, 2002 to 2010 are used. The best linear regression models are found in spring between milk yield as the predictand and THI, ESI, ETI, HLI, and RRP as predictors with p-value < 0.001 and R^2 (0.50, 0.49) respectively. In summer, milk yield with independent variables of THI, ETI, and ESI show the highest relation (p-value < 0.001) with R^2 (0.69). For fat and protein the results are only marginal. This method is suggested for the impact studies of climate variability/change on agriculture and food science fields when short-time series or data with large uncertainty are available.

Keywords: AIC; LASSO; climate indices; ESI; ETI; HLI; RRP; THI; linear regression model; milk components

1. Introduction

This study is about the linear relation between some climate parameters and cow's milk compounds applied to data of milk production in Iran. Hoping to evaluate the environmental effects on animal husbandry and milk production, it would be better to consider as many environmental and physiological as well as milk compound variables as possible. In fact, animal husbandry and physiological performance is generally a result of combinations of variables affected by genetics, feed, housing/behavior as well as environmental conditions, which are very complex and potentially vary with time.

In general, because animal organisms evolve together with the environmental parameters in which they live, their physiological metabolism changes over time. For instance, Holstein cows are spread throughout the whole world, but in tropical regions they have been bred with different temperature tolerance levels compared to those Holsteins bred in regions with different climate conditions [1].

This shows that to study the interactions between environmental conditions and cattle husbandry and milk production as well as milk compounds is a complicated task and needs comprehensive information. Previous studies over the last two decades have mostly focused on climatic parameters and indices related to temperature and humidity, as more commonly used parameters, and the temperature humidity index (THI) as the most widespread indicator of heat stress [2,3]. Because of the indicated complexity due to animal adaptation ability and genetic variations, more predictors are required such as solar radiation in combination with wind speed, humidity, and temperature, and their combinations as indicators with equivalent temperature index (ETI), environmental stress index (ESI), heat load index (HLI), modified HLI (HLI $_{new}$), and respiratory rate predictor index (RRP) [4]. Lacking a detailed physiological animal model to simulate the biophysical and biochemical influences of environmental conditions on the animal, the analysis has to rely on a statistical approach. Unfortunately, the available data sample sizes do not allow the use of a wide range of environmental predictors in a classical regression approach. Therefore, the physiologically motivated indicators are combined with the recently developed method of least absolute shrinkage and selection operator (LASSO). It is a purely data driven predictor selection method. It does not share the ambiguities of other predictor selection methods like stepwise regression, which, e.g., depend on the ordering of the predictors [5].

According to this motivation and the requirement for new, better and more indices, in this study, some more climatic indices are calculated from several climate parameters to have more combinations of the parameters, which might have an effect on the quantity of milk compounds. THI, ETI, ESI, HLI, HLI $_{new}$ and RRP indices are used in this research to consider the influences of temperature, humidity, wind speed, and solar radiation on components in cow´s milk. These indices are considered as the independent variables or predictors in a set of multiple linear regression equations for milk components. The most suitable merged statistical model between all indices as predictors and milk parameters as predictands are obtained through multiple linear regression models under the LASSO constraint.

Moran and Epstein evaluated ESI as the best suitable heat stress index for hot climate in both dry and wet conditions [6]. The HLI was suggested for the first time by Gaughan et al. in 2002. Since 2008, HLI has been modified (HLI $_{new}$) by Gaughan et al. [7–9]. Silva et al. also suggested HLI for the tropical regions [8]. We decided to study both indices (HLI and HLI $_{new}$) to compare them with each other and with other indices. Marami et al. [2,4] report on correlations between different climate indices (THI, ETI, ESI, HLI, HLI $_{new,}$ and RRP) and milk compounds (fat, protein, and milk yield) using the bootstrap technique. They found that the use of a single index is valid only under certain and specific climate conditions. The results in Tables 1 and 2 summarize that each index alone does show very different correlations with milk compounds.

Table 1. Correlation between climate indices (Temperature Humidity Index (THI), Equivalent Temperature Index (ETI), Environmental Stress Index (ESI), Heat Load Index (HLI), modified HLI (HLI $_{new}$), and Respiratory Rate Predictor index (RRP)) and milk compounds (fat, milk yield, and protein) in spring (* is p-value < 0.05, § is p-value < 0.1 and $^{\#}$ is not significant). Taken from [2,4].

	THI	ETI	ESI	HLI	HLI $_{new}$	RRP
Fat	−0.47 *	0.42 *	−0.41 *	−0.41 *	−0.40 *	−0.41 *
Milk yield	−0.22 $^{\#}$	−0.21 §	0.26 *	0.27 §	0.26 §	0.26 §
Protein	0.37 §	−0.43 §	0.41 *	0.42 *	0.44 *	0.41 *

Table 2. Correlation between climate indices (THI, ETI, ESI, HLI, HLI $_{new,}$ and RRP) and milk compounds (fat, milk yield, and protein) in summer (* is p-value < 0.05, § is p-value < 0.1 and # is not significant). Taken from [2,4].

	THI	ETI	ESI	HLI	HLI $_{new}$	RRP
Fat	−0.32 *	0.17 *	−0.23 *	−0.21 §	−0.21 §	−0.23 *
Milk yield	−0.26 #	−0.43 §	0.43 *	0.43 *	0.41 *	0.43 *
Protein	0.31 §	−0.34 §	0.37 *	0.36 §	0.34 §	0.37 *

The main question to be answered is to find more accurate predictions linking several climatic indices and parameters, which characterize the quality of milk. The use of single indices has been proven to show realistic and interpretable results [2,4]. Here, we will test if the combination of certain climate indices can provide a better statistical model than those reported in Tables 1 and 2. This will analyze the information content of the set of climate indices with respect to the milk compounds beyond the simple co-linearity of the single indices.

Finding effective predictors can lead to an accurate model with the aim of studying the impact of climate change on milk components to predict for the near future.

Generally, our research follows some main aims: (1) Introducing specific methods for data preparation and statistical analysis to cope with uncertainties in time series data sets with statistical problems such as data gaps, expensive or difficult data collection procedures, and the small quantity of samples and (2) selecting the best model with the smallest number of predictors by applying recent statistical techniques and avoiding misinterpretations due to overfitting to achieve more suitable combinations of climate indices with milk compounds (milk yield, and protein) in spring and summer in terms of a better statistical model measured by higher correlation R^2 or smaller p-value.

Hypothesis of this study consisted of opening avenues for potential applications, e.g., designing dairy factories according to the economic justifications and feeding system to be managed with less risk of climate variability, especially in developing countries.

To meet the requirements for hypothesis No. (2), LASSO [5] and AIC statistics techniques are applied and cross validated to avoid errors by overfitting. To consider uncertainty and constructing confidence intervals, bootstrap method is applied which is a field of active research in statistics, particularly for dependent data [10,11].

The climate data used in this study included daily averages of solar radiation (SR), two meter height temperature (T_{2m}), dew point (T_d), relative humidity (RH), wind speed (V), sea level pressure (p), and specific humidity (q) which are taken from the NASA-Modern Era Retrospective-Analysis for Research and Applications (NASA-MERRA) reanalysis [12].

2. The Region of Study and Data Basis

The geographic region of this study is Iran, a country with variety in topography and different climate conditions. The Caspian Sea in the north, Persian Gulf and Gulf of Oman in the south, Alborz and Zagros Mountains in the north and northwest-southeast, and the two famous deserts Dasht-e-Kavir in the central plateau and Dasht-e-Lut in the southeast are the main geographical resources in Iran.

Two main data bases, climate and milk data, are taken for this study. Climate data are taken from the MERRA reanalysis dataset. MERRA provides gridded data with a resolution of 1/2 degrees latitude and 2/3 degrees longitude covering all meteorological quantities in a physical and statistical self-consistent manner. This is achieved by merging observations from satellite and meteorological stations using a meteorological forecast model [12]. This means that the meteorological variables such as temperature, wind, and humidity obey the physical relationships among each other, which are dictated by the physical laws of fluid dynamics. Additionally, the meteorological variables are combined into three-dimensional spatial temporal patterns in between the observed points, which share comparable statistics, like correlations, among the space and time points as the observations.

Climate data for the same months and years as milk data have been extracted. The four new climate indices Environment Stress Index (ESI), Heat Load Index (HLI), modified Heat Load Index (HLI_{new}), and Respiratory Rate Predictor (RRP) are calculated from the corresponding climate data parameters based on the daily meteorological grid point data.

Regarding milk data, industrial herd stations in Iran are investigated. In total 18,295 industrial herd stations are active out of 25,353 herd stations with a capacity of about 1.3 million cows. 66% of the cows in Iran are pure Holstein, Jersey and Brown Swiss, 27% are hybrids and 7% are home born cows. Corn, grain, alfalfa, wheat chaff, silage, wheat, soybean, cottonseed, and other forages are important ingredients in the cows' feed in Iran [13].

2.1. Milk Data

Milk data is preprocessed considering some important terms such as genetics, cow age, environmental conditions, season, lactation, pregnancy, and feeding management from the industrial herd stations with good health services and under controlled condition and veterinary care.

In this study, the monthly average of test-day (TD) records of milk yield (kg) are used; TD expresses the day that the cow was milked. Additionally, three times weekly records of fat and protein content (g/100 mL milk) had been collected between 2002 and 2010 from almost 600 industrial herd stations of Holstein cows, with herd sizes varying between 75 to 200 cows, with access to grazing during April to September as a base feeding. The final monthly averaged milk data are in a matrix of $3 \times 936{,}227$ individuals for milk yield, fat, and protein. Data gathering was under the conditions that cows were on days 4 to 305 milking time and between three and six calves in their lifetime without mastitis problem during the whole period of study. Records, which indicate that a cow was in a dry period or had mastitis, were omitted from the primary row data set [2,4].

2.2. Climate Indices

The equations used to compute the climate indices are presented in Table 3. The necessary input data are taken from the MERRA data set on all grid points covering the study area. The equations are the basis to investigate their relationship in a multiple linear regression model with milk compounds. The indices used are the temperature humidity index (THI), the equivalent temperature index (ETI), the environmental stress index (ESI), the heat load index (HLI), the modified HLI (HLI_{new}), and the respiratory rate predictor index (RRP). Table 4 lists the necessary input variables taken from the MERRA database of the indices presented in Table 3.

Table 3. Definition equations of the climate indices (THI, ETI, ESI, HLI, HLI new, and RRP) used for investigating the influence of environmental influences upon milk components in this study.

Climate Index	Equation
Temperature Humidity Index [14–16]	$THI = 41.5 + T + 0.36T_d \text{ where } T_d = (\frac{RH}{4})^{\frac{1}{8}} \times [112 + (\frac{9T}{40})] + \frac{T}{40} - 112,$ $RH = e/e_w^* \; e = 1.6077 \times p \times q, e_w^* = 6.1078\exp[(17.1 \times T)/(235 + T)]$
Equivalent Temperature Index [1]	$ETI = 27.88 - 0.456T + 0.010754T^2 - 0.4905RH + 0.00088RH^2 +$ $1.1507V - 0.12645V^2 + 0.019876T \times RH - 0.046313T \times V$
Environmental Stress Index [17]	$ESI = 0.63T_a - 0.03RH + 0.002SR + 0.0054T_a \times RH -$ $0.073(0.1 + SR)^{-1}$
Heat Load Index [7,8]	$HLI = 33.2 + 0.2RH + 1.2T_g^* - (0.82V)^{0.1} - \log(0.4V^2 + 0.0001) \text{ where}$ $T_g^* = 1.33T - 2.65T^{0.5} + 3.21\log(SR + 1) + 3.5$
Modified Heat Load Index [9]	$HLI_{new} = 10.66 + 0.28RH + 1.3T_g^* - V \text{ when } T_g^* < 25\ (^\circ C)$
Respiratory Rate Predictor Index [18]	$HLI_{new} = 8.62 + 0.38RH + 1.55T_g^* - 0.5V + e^{2.4-V} \text{ when } T_g^* > 25\ (^\circ C)$ $RRP = 5.4T + 0.58RH - 0.63V + 0.024SR - 110.9$

Table 4. Naming of the MERRA input variables, their definitions, and units used in the computation of the climate indices (THI, ETI, ESI, HLI, HLI $_{new}$ and RRP) in Table 3.

Parameter	Definition
T	2 m height temperature (°C)
T_d	Dew point temperature (°C)
RH	Relative humidity (%)
e	Vapour pressure (hPa)
e_w^*	Saturation vapour pressure (hPa)
p	Sea level pressure (hPa)
q	specific humidity
v	wind speed (m/s)
T_a	ambient temperature (°C)
SR	solar radiation (wm^{-2})

3. Materials and Methods

In regards to climate variability in Iran, three different climatic zones in the northwest, north, and center of Iran are selected with diverse climate conditions including cold semi-arid, Caspian mild and humid, semi warm, and semi-arid. The data of climatic and milk parameters are selected in each zone separately during 2002 to 2010 for summer and spring. Figure 1 shows the selected climatic zones in the study domain.

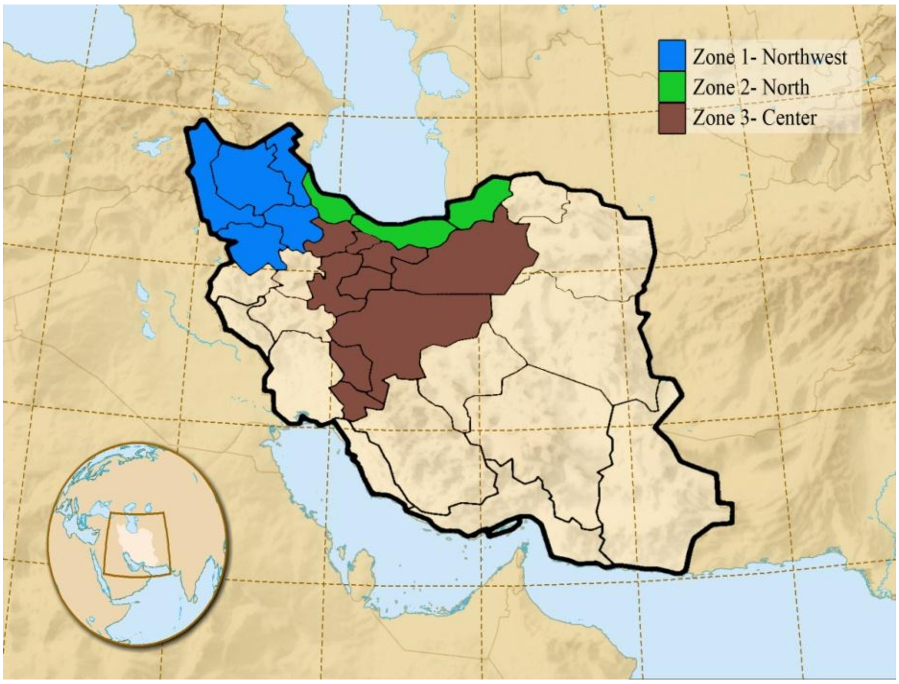


Figure 1. Classification of three study zones according to climate conditions in Iran.

To consider uncertainty, non-parametric bootstrap technique is applied to the monthly mean data of milk yield, fat, and protein concentration during 2002 and 2010 in each zone separately by generating 1000 samples which is presented in Figure 2.

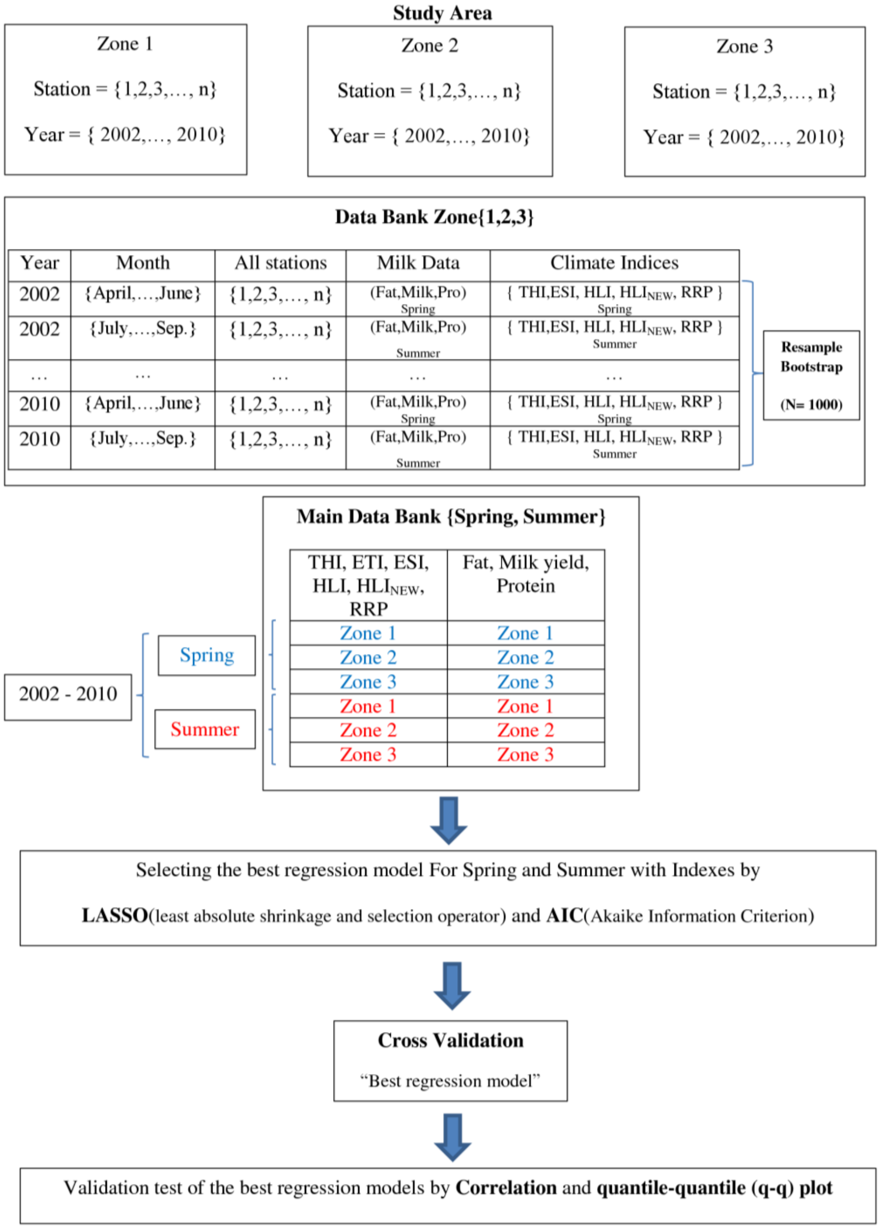


Figure 2. The flowchart of the data set of climatic and milk parameters and methodology.

For both the correlation analysis and the multiple linear regression, the data are normalized to a mean of zero and a standard deviation of one.

In this work, we model the observed milk components employing methods which take different models and parameters into account.

To measure the "distance" between a specific model based on the meteorological data and the observed milk variables, the "Kullback-Leibler Information" (K-L distance) together with the Akaike's Information Criterion (AIC) are used. The AIC identifies the model that minimizes K-L distance.

In order to select the best model with the smallest number of predictors, the LASSO and the AIC are combined. To avoid overfitting and to test the validity of the obtained model, cross validation is applied on the selected model. For presenting graphically the results in an appropriate statistical sense, quantile-quantile plots are used.

All the analyses in this study are implemented using R version 3.2.2, which provides various packages for statistical data analysis, calculation and graphical display [19]. THI, ETI, ESI, HLI, HLI_{new}, and RRP are considered as independent variables (predictors) in multiple linear regression equations, and milk components (milk yield, fat, and protein) are considered as dependent variables (predictands).

3.1. Least Absolute Shrinkage and Selection Operator Technique (LASSO)

Tibshirani [5] proposed a new method for estimation and predictor selection in linear models. The LASSO is an effective technique for shrinkage and selection method for linear regression. It minimizes the residual sum of squares like in classical linear regression to determine the unknown regression coefficients related to each predictor. As an extension to classical regression, the regression coefficients are constrained by the sum of the modulus or absolute values of the coefficients being as small as possible. The advantages of this combination of techniques is to reduce the regression coefficients as much as possible (shrinkage) and, depending on the data, setting some regression coefficients exactly to zero (selection), which is specifically achieved by the sum of the modulus of the regression coefficients.

In this study the LASSO method is used for finding the best regression models between the main compounds of milk (fat, protein, and milk yield) as predictands and environmental indices as predictors. To that aim, we applied an available package of glmnet in R [20].

LASSO presents the regression coefficients β_k (Equation (1)) as a function of a regularization parameter Lambda (λ), which determines the influence of the classical least squares contribution (first sum over n in Equation (2)) relative to the sum of modulus of the coefficients (second sum over k in Equation (2)).

$$Y \sim \sum_k \beta_k X_k + e \tag{1}$$

$$J(\beta_k) = \frac{1}{N} \sum_{n=1}^{N} \left(Y_n - \sum_k \beta_k X_{k,n}\right)^2 + \lambda \sum_k |\beta_k| \tag{2}$$

Minimization of the function $J(\beta_k)$ for varying λ values will select non zero β_k solely on the basis of the data $(Y_n,\ X_{k,n})$.

By increasing λ from zero (LASSO switched off or classical solution) to higher values (putting more weight on the absolute value constraint), glmnet sets increasingly more coefficients to zero, thereby removing them from the model. If for a range of λ values a constant number k of independent predictor variables is selected, we consider this potentially the best multiple linear regression model.

3.2. Akaike Information Criterion (AIC)

The Akaike Information Criterion (AIC) is a method for selecting a model from a set of models based on bias minimizing through Kullback-Leibler distance. Kullback and Leibler [21] measured a divergence which defines the distance measure between two probability distributions over the same event space. AIC value provides a mean which estimates the quality of each model in comparison to the other models. It selects the model with largest likelihood under the constraint of the smallest number of predictors. It was presented by Akaike in 1973 and is defined as:

$$\mathrm{AIC} = 2k - 2\ln(L) \tag{3}$$

k is the number of parameters in the model and L is the likelihood which deals with the goodness of fit. AIC is negatively oriented with smaller AIC values representing better models [22,23]. From the potentially best multiple linear regression models selected by the LASSO algorithm, the best model is defined as the one with the lowest AIC.

3.3. Cross Validating the Best Regression Model

Cross validation is a model evaluation technique which prevents overfitting by applying the model to the data that are not involved in the fitting [24,25]. In cross validation, we withhold one year (testing set) of data and estimate the regression model with the rest of data (training set) according to the rules described in the previous subsections. After selecting the best regression model by the LASSO and AIC, the withheld predictors are used to compute the milk component variables which can be compared to the withheld observations. This is done for all data sets from 2002 to 2010 for spring and summer separately. Figure 3 presents the cross validation process in this study. Afterward, the accuracy of the predicted model versus the observation is accomplished via Pearson correlation and quantile-quantile (QQ) plot [26,27].

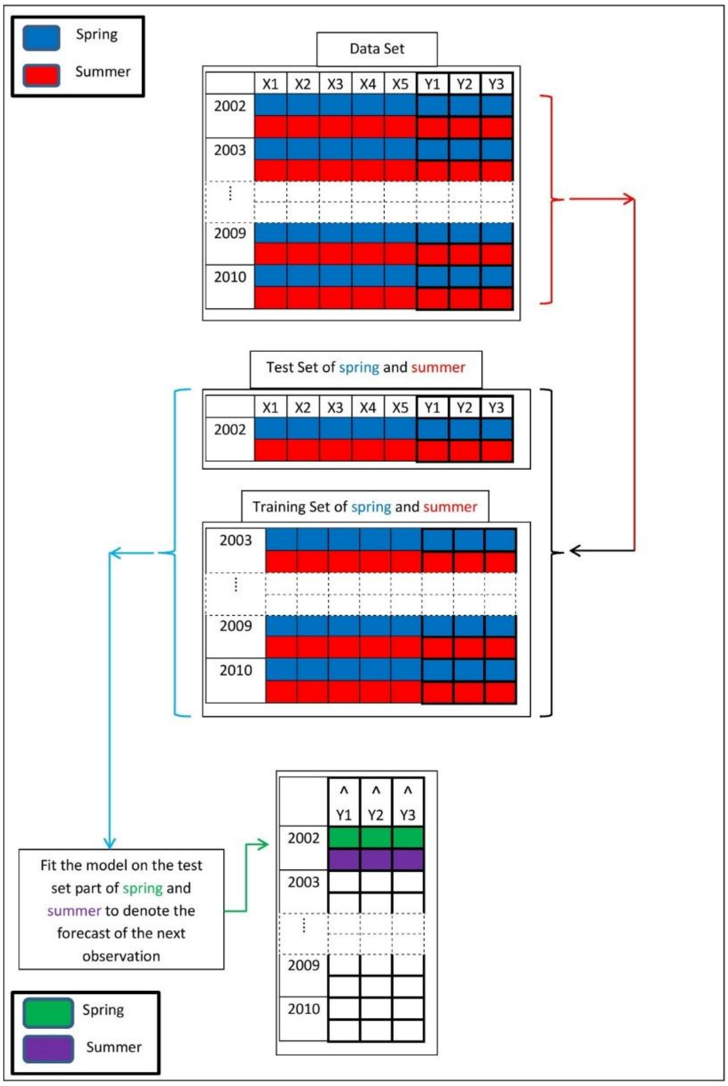


Figure 3. Cross validation procedure.

4. Results

In this section, the best models are presented through multiple linear regression analysis between climate indices and milk components in spring and summer, which are preselected through LASSO and finalized by AIC methods. Significant relationships in spring and summer could be identified between climate indices and milk yield through these methods.

The multiple linear regression models achieve higher correlation, R^2 and smaller p-value for milk yield than the simple linear regressions presented in Tables 1 and 2. For fat and protein, no significant model is suggested which performs better than what has been found with a single index by Marami et al. [2,4] in the previous studies. This is comparable to Knapp and Grummer [28] and Roman-Ponce et al. [29] who also did not find any significant relationship between fat variations and heat stress variables.

4.1. The Best Linear Regression Model between Climate Indices and Milk Components in Spring

Results of the analysis in spring are shown in Figure 4 and Table 5. Figure 4 shows the results of LASSO analysis between climate indices and milk yield (a), fat (b), and protein (c) from 2002 to 2010 in spring. On the bottom the shrinkage regularization factor is given in logarithmic units (log λ) with the smallest negative values indicating a very small weight of the LASSO regularization or an almost classical multiple linear regression with all predictors included. The axis on top of the figure indicates the numbers of LASSO selected predictors at the given log λ value.

As seen in Figure 4a, the highest weight of LASSO in the right side of the figure has selected two predictors for milk yield which are THI and HLI. Then, by moving to the left hand side of the figure, the weight of LASSO regularization decrease and selects four predictors (THI, ETI, HLI, and RRP), and then a different four predictors (THI, ETI, ESI, and RRP). With the least weight of LASSO, it selects five predictors for milk yield: THI, HLI $_{\text{New}}$, ETI, ESI, and HLI. In the next step AIC will be used to select the best multiple model within the selected models of LASSO.

Table 5 presents values of AIC, R^2 and p-values, and final selection of the best model by AIC within the preselected predictors by LASSO for the analysis of the spring data. The AIC values can be directly compared because all data (predictors as well as predictands) have been normalized to a mean of zero and a standard deviation of one.

$$\text{Milk yield Spring (Mod3)} = -3.58\ \text{THI} - 0.77\ \text{ETI} + 5.14\ \text{ESI} - 2.3\ \text{RRP} + \text{e} \tag{4}$$

$$\text{Milk yield Spring (Mod2)} = -3.485\ \text{THI} - 0.68\ \text{ETI} - 0.295\ \text{HLI} + 3.13\ \text{RRP} + \text{e} \tag{5}$$

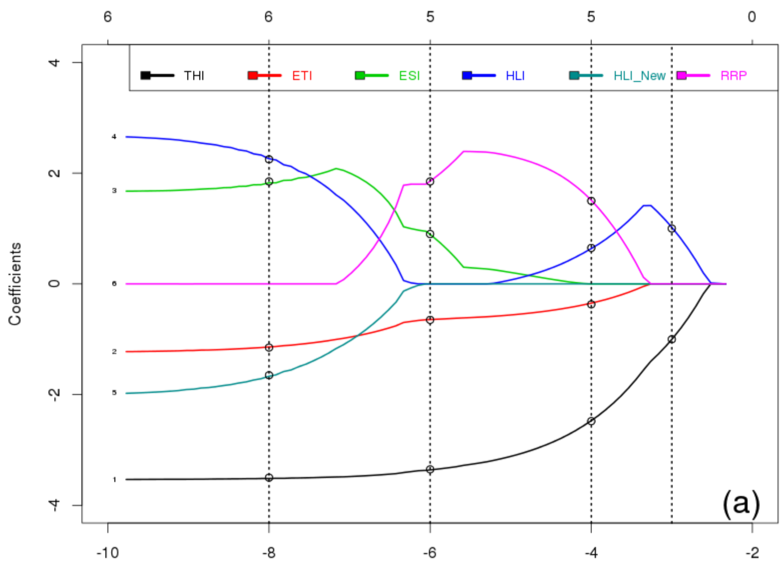


Figure 4. *Cont.*

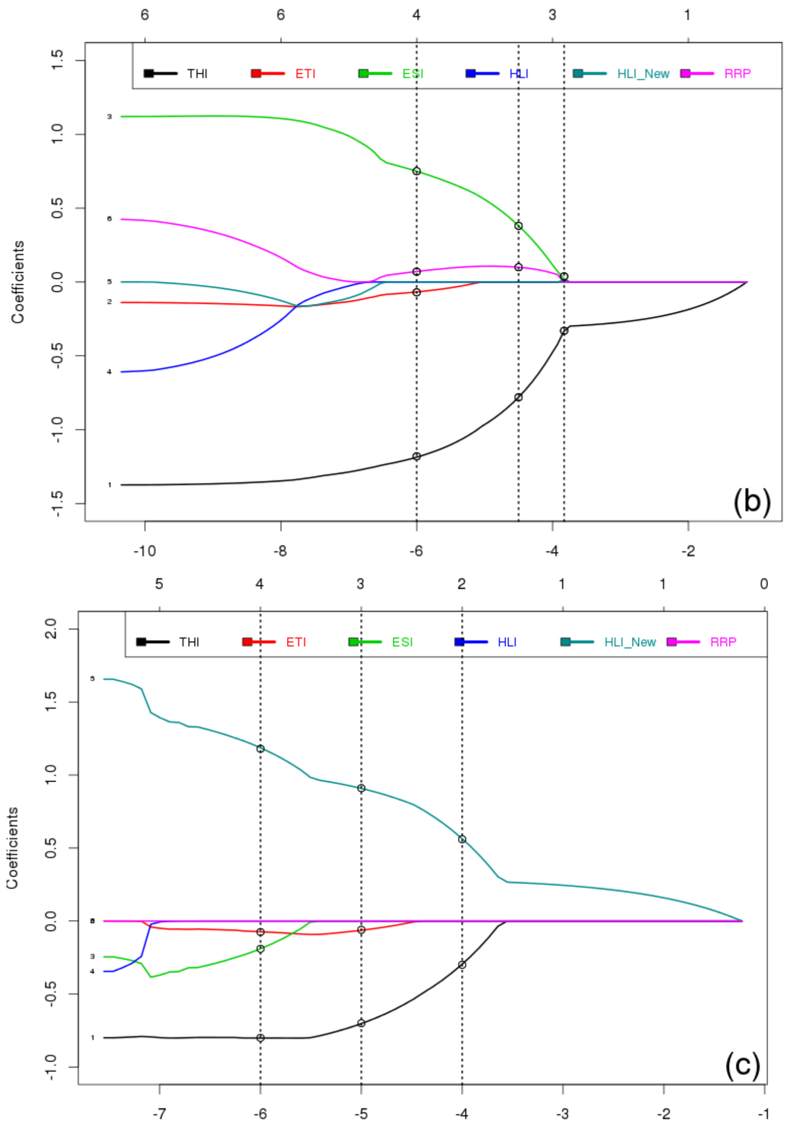


Figure 4. Least Absolute Shrinkage and Selection Operator (LASSO) analysis to find the best multiple linear regression model between climate indices and milk yield (**a**), fat (**b**), and protein (**c**), from 2002 to 2010 in spring. On the bottom *X* axis, the shrinkage regularization factor is given in logarithmic units (log λ). The number on top of the figure indicates the number of Lasso selected predictors at the log λ value.

Standard error of the estimates is calculated through Equation (6). Y is observed value, Y′ is predicted value and N is the number of pairs of scores.

$$\sigma_{est} = \sqrt{\frac{\sum (Y - Y\prime)^2}{N}} \tag{6}$$

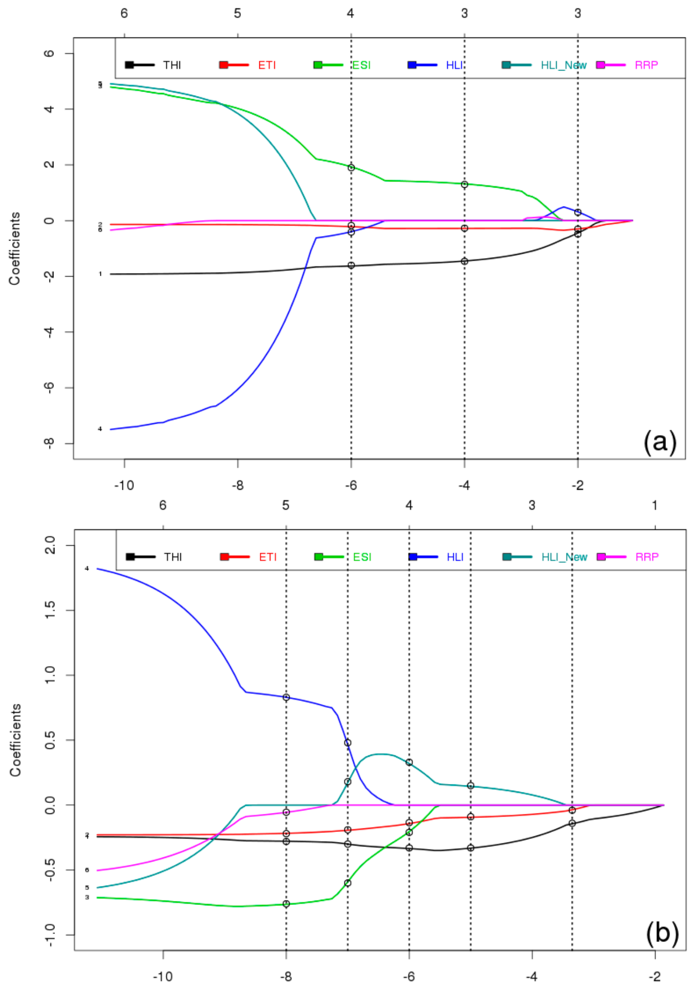
Standard error of estimate for model Equations (4) and (5) are calculated at 0.706 and 0.705 respectively. In both of the models, maximum negative effect belongs to THI with milk yield. The inclusion of the additional predictors raises the correlation from 0.22 for THI vs. yield in spring (Table 1) to 0.7, proving the added value of the other climate indices.

Table 5. Results of LASSO and Akaike Information Criterion (AIC) analysis for finding the best multiple regression model between milk parameters and climatic indices in spring. Mod1 to Mod4 in "Milk yield" refers to the dashed lines in Figure 4a; Mod1 to Mod3 in "Fat" refers to Figure 4b, and Mod1 to Mod3 in "Protein" refers to Figure 4c.

Layout Parameter	Preselected Predictors By LASSO	AIC	Δ AIC	Model Selected by AIC	R^2	*p*-Value
Milk yield	Mod 1: THI , HLI	190.90	4.51			
	Mod 2: THI, ETI, HLI, RRP	186.77	0.38	Mod 2	0.496	<0.001
	Mod 3: THI, ETI, ESI, RRP	186.39	0.00	Mod 3	0.494	<0.001
	Mod 4: THI, ETI, ESI, HLI, HLI $_{New}$	187.40	1.01			
Fat	Mod 1: THI, ESI	224.53	0.00	Mod 1	0.1249	<0.01
	Mod 2: THI, ESI, RRP	226.79	2.26			
	Mod 3: THI, ETI, ESI, RRP	228.94	4.41			
Protein	Mod 1: THI, HLI $_{New}$	225.50	0.00	Mod 1	0.1144	<0.01
	Mod 2: THI, ETI, HLI $_{New}$	227.68	2.18			
	Mod 3: THI, ESI, ETI, HLI $_{New}$	229.74	4.24			

4.2. The Best Linear Regression Model between Climate Indices and Milk Components in Summer

Similar to the spring analysis, results of the analysis for summer are shown in Figures 5 and 6. Figure 5 shows the results of LASSO analysis between climate indices and milk yield (a), fat (b), and protein (c) from 2002 to 2010 in summer.



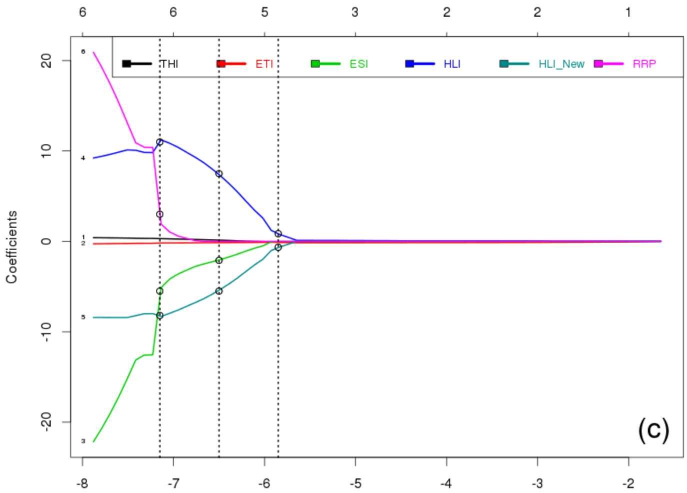


Figure 5. LASSO analysis to find the best linear model between climate indices and milk yield (**a**), fat (**b**), and protein (**c**) from 2002 to 2010 in summer. On the bottom X axis, the shrinkage regularization factor is given in logarithmic units (log λ). The number on top of the figure indicates the number of Lasso selected predictors at the log λ value.

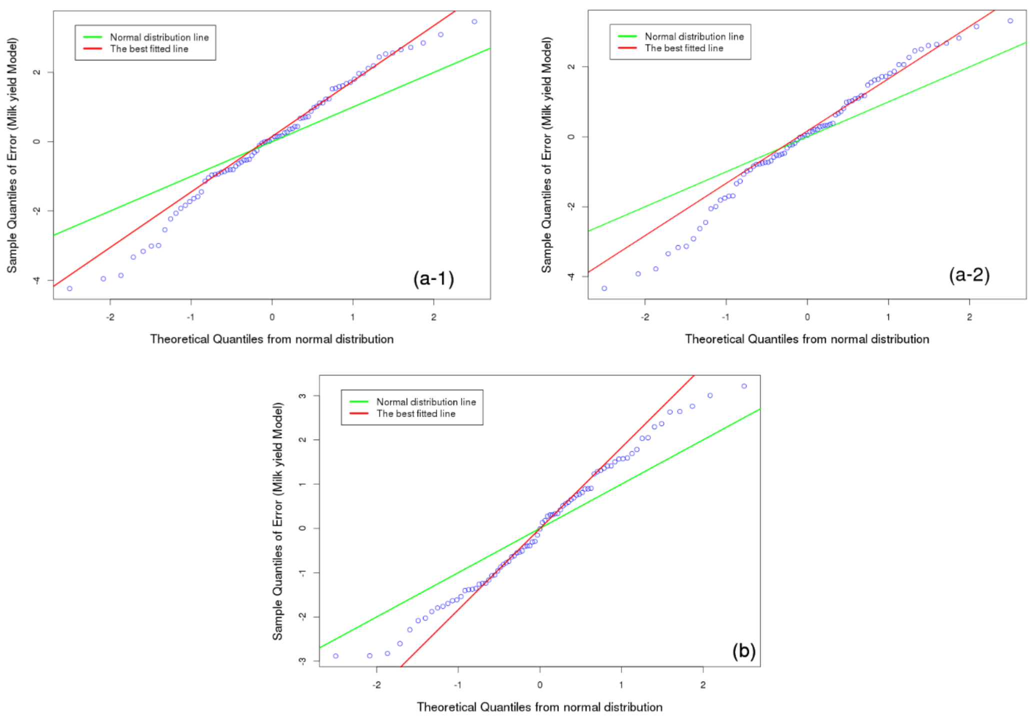


Figure 6. Quantile-Quantile (Q-Q) plot of the best models for milk yield in spring (**a**) and summer (**b**).

Referring to Figure 5a for milk yield, the highest weight of LASSO in the right side of the figure has selected three predictors for milk yield which are THI, ETI, and HLI. Then by moving to the left hand side of the figure, the weight of LASSO regularization decreases and selects another three predictors (THI, ETI, and ESI). With the least weight of LASSO, it selects four predictors (THI, ETI, ESI, and HLI). Then, in the next step, AIC will select the best multiple model within the selected models of LASSO.

Table 6 presents values of AIC, R^2 and p-values, and final selection of the best model by AIC within the preselected predictors by LASSO in summer.

Table 6. Results of LASSO and AIC analysis for finding the best multiple regression model between milk parameters and climatic indices in summer.

Layout Parameter	Preselected Predectors by LASSO	AIC	Δ AIC	Model Selected by AIC	R^2	p-Value
Milk yield	Mod 1: THI , ETI, HLI Mod 2: THI, ETI, ESI Mod 3: THI, ETI, ESI, HLI	152.40 142.70 142.54	9.86 0.16 0.00	Mod 2 Mod 3	0.697 0.707	<0.001 <0.001
Fat	Mod 1: THI, ETI Mod 2: THI, ETI, HLI $_{New}$ Mod 3: THI, ETI, ESI, HLI $_{New}$ Mod 4: THI, ETI, ESI, HLI, HLI $_{New}$ Mod 5: THI, ETI, ESI, HLI $_{New}$, RRP	234.49 235.98 237.95 240.14 240.34	0.00 1.49 3.46 5.65 5.85	Mod 1	0.0103	<0.1
Protein	Mod 1: HLI, HLI $_{New}$ Mod 2: ESI, HLI, HLI $_{New}$ Mod 3: ESI, HLI, HLI $_{New}$, RRP	231.82 229.92 225.36	6.45 4.56 0.00	Mod 3	0.1428	<0.01

The best multiple linear regression models of milk yield in summer are presented as

$$\text{Milk yield Summer (Mod3)} = -1.7\ \text{THI} - 0.126\ \text{ETI} + 2.58\ \text{ESI} - 0.91\ \text{HLI} + \text{e} \quad (7)$$

$$\text{Milk yield Summer (Mod2)} = -1.6\ \text{THI} - 0.28\ \text{ETI} + 1.47\ \text{ESI} + \text{e} \quad (8)$$

In both models, like as in spring, the highest negative relationship existed between THI with milk yield. Standard error of estimate (Equation (6)) for both models in summer is calculated at 0.546.

In contrast to the results for spring, the models (7) and (8) have a different number of predictors. In this case, one would invoke Occam's razor [30] to favor the model (8) with three predictors, all other selection criteria being equal.

4.3. Model Verification and Validation

The quality of the models are assessed by applying a quantile-quantile (Q-Q) plot, correlation coefficient, and mean squared error skill score (MSES) by Equation (9) to see how close a fitted line is to the data points. Here Y are the verifying observations, Y′ the predicted values, and $\overline{Y}$ the climatological mean as the reference prediction.

$$\text{MSES} = 1 - \left[\frac{MSE\left(Y', Y\right)}{MSE\left(Y, \overline{Y}\right)}\right] \quad (9)$$

Quantile-Quantile plots the sorted, observed quantiles of the residuals Y′-Y against the theoretically expected quantiles of a normal distribution with a mean of zero and a standard deviation of 1. If the points align along a linear increasing function, the requirement that the residuals are realizations of normally distributed random variables is well fulfilled. This checks, a-posteriori, the basic assumptions of the multiple linear regression analysis. The slope of the linear function corresponds to the standard deviation of the residuals Y′-Y.

Figure 6 (a-1, a-2) and (b) present the Q-Q plots of the residual errors Y-Y′ of the best models vs. the theoretical quantiles of a random normal variate with mean zero and variance 1 for milk yield in spring (Equations (4) and (5)) and summer (Equation (8)).

In both cases, the prerequisites of the multiple regression analysis of normally distributed residuals are well fulfilled. Such a model quality could not be achieved with a single predictor model by Marami et al. [4], nor could it be selected by prior methods.

The MSES should be equal to or less than R^2 of the model. In this study MSES of milk yield in spring (a-1, a-2) and summer (b) are 0.49, 0.49 and 0.69 respectively.

Figure 7a,b show the outcome of the best models for milk yield in spring (Equations (4) and (5) and summer (Equation (8)). They visually represent the observed and predicted values l in dimensional units. In spring, the correlation between predicted milk yield from the models and observed data is 0.7 (R^2 = 0.49, Table 5), with p-value < 0.001. For the summer, it increases to a higher value of 0.84 (R^2 = 0.7, Table 6), with p-value < 0.001. The models of milk yield in spring and summer explains almost 50% and 70% of the linear dependency vs. 50% and 30% for unexplained scatter respectively. Such a model quality could not be achieved with a single predictor model by Marami et al. [4] nor could it be selected by prior methods.

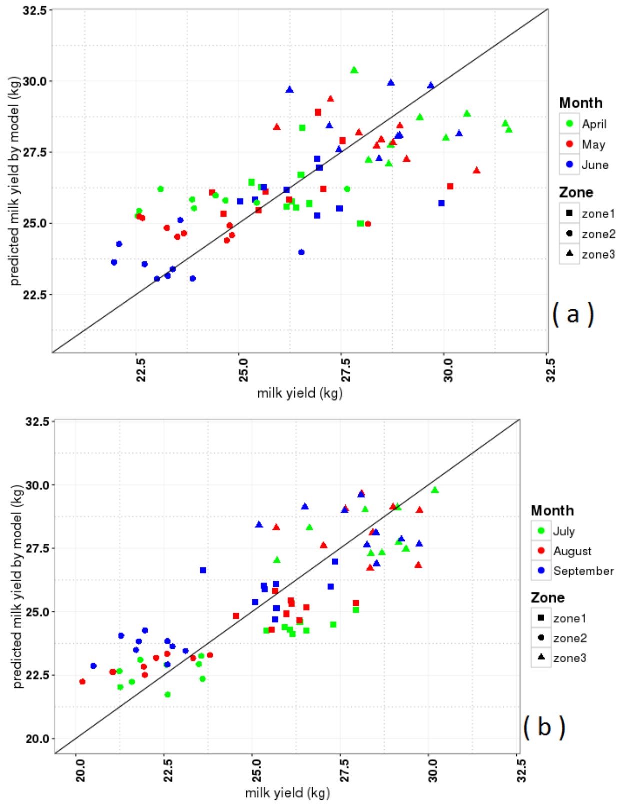


Figure 7. Layout data (observed) of milk yield vs. predicted milk yield data by models in spring (**a**) and summer (**b**). Months indicated by colors and zones (Figure 1) by different symbols.

This figure shows that the distribution of data in different climatic zones (Figure 1) and months (from 2002 to 2010) equally contributed to the overall correlation between observations and predicted data, indicating a robustness of the results.

5. Discussion and Conclusions

According to cow characteristic changes and adaptation abilities under different conditions, the effect of climate variability on animal products and quantities of nutrients and yield is very hard to work out in full detail [1].

The assumption is that the interactional effect of air temperature, humidity, wind speed, and solar radiation make a critical point of heat stress with a negative effect on cows' body metabolism. It was reported in 1988 by Dupre et al. [31] that 11% of metabolic rate is altered by changing 1 °C in body temperature. We do not consider indirect effects of climate variability, e.g., variable nutrient contents of forage crops through different growing conditions.

Heat stress has a negative effect on hypohydration, endurance and thermoregulatory control systems, and changes the blood volume and plasma osmolality of cows. As well, it affects rumen function, hormonal status, and maintenance energy requirements; it increases respiratory rate and sweating, and reduces feed intake and activity. Cows under heat stress lose large quantities of potassium which could lead to an increase on blood acidity [32–35]. To moderate the effects of heat stress, cows might be protected from direct solar radiation using a combination of fans and wetting, improving the cooling system, and developing nutritional strategies [36]. Other heat stress management possibilities for cows are to stay in the barn or to graze on pasture in the pasturing system [37].

In this study, we suggested a statistical approach to find out the best multiple linear regression model for the influence of heat stress on key components of milk. Additionally, the aim was to consider uncertainties as much as possible.

Here, the effect of heat stress on the fundamental milk compounds milk yield, fat, and protein contents were investigated. We obtained three suitable models (Equations (4), (5) and (8)) as the best multiple models between climate indices and milk yield for spring and summer.

In total, six different predictors of heat stress indicators are introduced which summarize, as much as possible, the physiological knowledge of the reaction of cattle on environmental forcing. The appropriate mix of predictors is determined by a solely data driven approach: the least absolute shrinkage and the selection operator LASSO. The result of multiple linear regression models lead to correlation and R^2 values which are highly promoted and provide more suitable and significant models than those suggested by single indices in Marami et al. [2,4] and other studies.

We found that the single use of indices is valid only under certain and specific climate conditions. The results exhibited that each index separately shows very different correlation with milk compounds although these indices have high correlation with each other (Marami et al. [2,4]). Coping with this problem, in 2002 Brake and Bates [38] developed a method of limiting metabolic rate which allows all heat stress indices to be compared with each other virtually.

In contrast with Brake and Bates, we eliminated this problem by applying multiple linear regression models to develop new indices through the combination of single heat stress indices and climate parameters together with considering uncertainties as much as possible.

For fat and protein, no significant model is better than what is suggested by Marami et al. [2,4] in the previous studies. For milk yield, a clearly superior model was identified. Note that the regression coefficients in the three models (Equations (4), (5) and (8)) can be directly compared due to the normalization of predictors as well as predictands. For both spring and summer, THI is the major driving negative component, meaning that an enhanced heat stress expressed by the temperature and humidity effects alone [31] leads to a reduction in milk yield. The difference between both seasons is that the THI is about two times more effective in yield reduction during spring than during summer. The second identified component with a negative influence upon milk yield is modeled through ETI. This index differs from THI by the inclusion of wind velocity which by the ventilation effect apparently contributes additional information beyond the co-linearity with THI. Further influences in variations of the milk yield are modeled with ESI, HLI, and RRP depending on the season, with ESI having always a positive effect upon milk yield. All three indices differ from the THI and ETI through their inclusion of solar radiation. This also gives an explanation as to why the combined indices provide a superior model in terms of explained variance of milk yield: It is the combination of temperature, humidity based heat stress with ventilation effects and the solar radiation as an additional heat source, which together serve as a driver of the metabolic processes listed above. The results of this statistical analysis can not completely rule out that the indirect effects, e.g., through nutrient content, play a role. However, the definition of the indices in use is with emphasis on the metabolic effects.

It is strongly suggested that investigating the effects of climate variability on fundamental milk components is done separately on every single parameter of milk. The critical points of heat stress indices need to be updated for different climate conditions because of characteristic changes and cows´ adaptation ability. It is very important because it is the starting point of the negative effects on milk compounds [1].

Based on the discussion above, the authors strongly recommend new indices considering different climate conditions with more predictors, such as sunshine duration, quality of cows' feed, color of skin with attention to large black spots, and some categorical predictors such as breed, welfare facility, and management systems with categorical codes to control the critical heat stress conditions.

Acknowledgments: The authors warmly thank from Nasim Azari at the Food Science Engineering Department of Science and Research branch of Tehran Islamic Azad University, Ahmad Moghimi Esfand Abadi and Asghar Salimi Niknam at the Animal Breeding Centre of Iran in Karaj, branch of the Iranian Ministry of Agriculture, Behnam Saremi and Ali Asadi for supporting the milk data. We are also grateful to Ali Mortazavi in Department of Food Science, Ferdowsi Mashhad University for his cooperation. Warmely thank to Harry Leach from the University of Liverpool for spelling check.

Author Contributions: A.H. and A.P. supervised the methodology of the study and the used statistical methods. E.R. contributed data and tools for analyzing. M.R.M.M. performed the statistical analysis, interpreted the results, and wrote the initial manuscript. All authors contributed to read and edit the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gomes da Silva, R.; Campos Maia, A.S. Thermal stress indexes. In *Principles of Animal Biometeorology*; Springer Science & Business Media: Dordrecht, The Netherlands, 2013; pp. 207–229.
2. Milani, M.R.M.; Hense, A.; Rahmani, E.; Ploeger, A. A pilot investigation of the relationship between climate variability and milk compounds under the bootstrap technique. *Foods* **2015**, *4*, 420–439. [CrossRef]
3. Hammami, H.; Vandenplas, J.; Vanrobays, M.L.; Rekik, B.; Bastin, C.; Gengler, N. Genetic analysis of heat stress effects on yield traits, udder health, and fatty acids of Walloon Holstein cows. *J. Dairy Sci.* **2015**, *98*, 4956–4568. [CrossRef] [PubMed]
4. Milani, M.R.M.; Hense, A.; Rahmani, E.; Ploeger, A. A Survey of the relationship between climatic heat stress indices and fundamental milk components considering uncertainty. *Climate* **2015**, *3*, 876–900. [CrossRef]
5. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* **1996**, *58*, 267–288.
6. Moran, D.S.; Epstein, Y. Evaluation of the environmental stress index (ESI) for hot/dry and hot/wet climates. *Ind. Health* **2006**, *3*, 399–403. [CrossRef]
7. Gaughan, J.G.; Goopy, L.; Spark, J. *Excessive Heat Load Index for Feedlot Cattle*; MLA: Sydney, Australia, 2002.
8. Silva, R.G.; Morais, D.A.; Guilhermino, M.M. Evaluation of thermal stress indexes for dairy cows in tropical regions. *Rev. Bras. Zootec.* **2007**, *36*, 1192–1198. [CrossRef]
9. Gaughan, J.G.; Mader, T.L.; Holt, S.M.; Lisle, A. A new heat load index for feedlot cattle. *J. Anim. Sci.* **2008**, *86*, 226–234. [CrossRef] [PubMed]
10. Lahiri, S.N. Bootstraps Methods. In *Resampling Methods for Dependent Data*; Springer Science & Business Media: New York, NY, USA, 2003.
11. Politis, D.N.; Romano, J.P.; Wolf, M. *Subsampling*; Springer: New York, NY, USA, 1999; pp. 7–50, 98–100.
12. Rienecker, M.M.; Suarez, M.J.; Gelaro, R.; Todling, R.; Bacmeister, J.; Liu, E.; Bosilovich, M.G.; Schubert, S.D.; Takacs, L.; Kim, G.-K.; et al. MERRA: NASA's Modern-era retrospective analysis for research and applications. *J. Clim.* **2011**, *24*, 3624–3648. [CrossRef]
13. Statistical Center of Iran. Vice-President for Strategic Planning and Supervision. Available online: http://www.amar.org.ir/Default.aspx?tabid=281 (accessed on 5 August 2014).
14. Bohmanova, J.; Misztal, I.; Colet, B. Temperature-humidity indices as indicators of milk production losses due to heat stress. *J. Dairy Sci.* **2007**, *90*, 1947–1956. [CrossRef] [PubMed]
15. Kraus, H. *Die Atmosphäre der Erde*, 3rd ed.; Springer: Berlin, Germany; Heidelberg, Germany, 2004; pp. 67–89.
16. Thom, E.C. The discomfort index. *Weatherwise* **1959**, *12*, 57–61. [CrossRef]

17. Moran, D.S.; Pandolf, K.B.; Shapiro, Y.; Heled, Y.; Shani, Y.; Matthew, W.T.; Gonzales, R.R. An environmental stress index (ESI) as a substitute for the wet bulb globe temperature (WBGT). *J. Therm. Biol.* **2001**, *26*, 427–431.
18. Eigenberg, R.A.; Nienaber, J.A.; Brown-Brandl, T.M. Development of a livestock safety monitor for cattle. In Proceedings of the American Society of Agricultural and Biological Engineers (ASABE) Annual Meeting, St. Joseph, MI, USA, 2003.
19. R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2012.
20. Lasso and Elastic-net regularized generalized linear models. R Package. Available online: https://cran.r-project.org/web/packages/glmnet/glmnet.pdf (accessed on 15 March 2016).
21. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [CrossRef]
22. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]
23. Burnham, K.P.; Anderson, D.R. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*; Springer Science & Business Media: New York, NY, USA, 2002; pp. 49–97.
24. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*, 3rd ed.; Academic Press: San Diego, CA, USA, 2011; pp. 252–254.
25. Rahmani, E.; Friederichs, P.; Keller, J.; Hense, A. Development of an effective and potentially scalable weather generator for temperature and growing degree days. *Theor. Appl. Climatol.* **2015**, *120*. [CrossRef]
26. Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection. *Stat. Surv.* **2010**, *4*, 40–79. [CrossRef]
27. Cross-validation for detecting and preventing overfitting. Available online: https://clm.utexas.edu/fietelab/QuantNeuro/readings/crossvalidation_slides_Moore_CMU.pdf (accessed on 15 October 2001).
28. Knapp, D.M.; Grummer, R.R. Response of lactating dairy cows to fat supplementation during heat stress. *J. Dairy Sci.* **1991**, *74*, 2573–2579. [CrossRef]
29. Roman-Ponce, H.; Thatcher, W.W.; Buffington, D.E.; Wilcox, C.J.; Van Horn, H.H. Physiological and production responses of dairy cattle to a shade structure in a subtropical environment. *J. Dairy Sci.* **1977**, *60*, 424–430. [CrossRef]
30. Sorkin, R. A Quantitative Occam's razor. *Int. J. Theor. Phys.* **1983**, *22*, 1091–1104. [CrossRef]
31. Dupre, R.K.; Romero, A.M.; Wood, S.C. Thermoregulation and metabolism in hypoxic animals. In *Oxygen Transfer from Atmosphere to Tissues*; Springer: New York, NY, USA, 1988; pp. 347–351.
32. Reynolds, M. Plasma and blood volume in the cow using the T-1824 hematocrit method. *Am. J. Physiol. Leg. Content* **1953**, *173*, 421–427.
33. Harrison, M.H.; Edwards, R.J.; Fennessy, P.A. Intravascular volume and tonicity as factors in the regulation of body temperature. *J. Appl. Physiol.* **1988**, *44*, 69–75.
34. Nadel, E.R.; Fortney, S.M.; Wenger, C.B. Effect of hydration state of circulatory and thermal regulations. *J. Appl. Physiol.* **1980**, *49*, 715–721. [PubMed]
35. West, J.W. Effects of heat-stress on production in dairy cattle. *J. Dairy Sci.* **2003**, *86*, 2131–2144. [CrossRef]
36. Bond, T.E.; Kelly, C.F. The globe thermometer in agricultural research. *Agric. Eng.* **1955**, *36*, 251–255.
37. Brown-Brandl, T.M.; Eigenberg, R.A.; Nienaber, J.A.; Hahn, G.L. Dynamic response indicators of heat stress in shaded and non-shaded feedlot cattle, Part 1: Analyses of indicators. *Biosyst. Eng.* **2005**, *90*, 451–462. [CrossRef]
38. Brake, R.; Bates, G. A valid method for comparing rational and empirical heat stress indices. *Ann. Occup. Hyg.* **2002**, *46*, 165–174. [CrossRef] [PubMed]

Chapter 5

Summary of Results and discussion

More than approximately 50 heat stress indices have been developed since 1900 until now. They have more focused on the effect of temperature as a main factor like as THI which is a widely used indicator of thermal conditions and heat stress index.

Heat stress indices are divided into these two main categories (Moran et al. 1998):

1) **Effective temperature** which is based on meteorological parameters only (e.g. ambient temperature, wet-bulb temperature, black-globe temperature, humidity, wind speed and solar radiation, etc.)

2) **Rational heat** which depends on the combination of environmental and physiological parameters (e.g. rectal temperature, heart rate, skin temperature, sweat rate and metabolic heat production, etc.)

In this study, we used THI, ETI, ESI, HLI, HLI_{new} and RRP indices to consider the interactions of temperature in two meter height (T2m), humidity (RH), wind speed (v) and solar radiation (SR) on milk yield, protein and fat as fundamental cow milk components.

According to the explained categories of heat stress indices, in this research we have focused directly on effective temperature and indirectly on rational heat indices. Of course interactional effect of air temperature, humidity, wind speed and solar radiation make the critical point of heat stress which has negative effect on metabolism of cow's body. It is also reported by Dupre er al. (1988) that 11% of metabolic rate is changed by changing 1 ($°C$) in body temperature.

Heat stress has a negative effect on hypohydration, endurance and thermoregulatory control systems and changes blood volume and plasma osmolality of cows. It has also effect on rumen function, hormonal status and maintenance energy requirements and also increases respiratory rate and sweating, reducing feed intake and activity. Cows under heat stress lose a lot of potassium which this deficient increases blood acidity (Reynolds 1953; Harrison et al. 1978; Nadel et al. 1980; West 2003). To moderate the stressful effects of heat stress, cows might be protected from direct solar radiation, using combination of fans and wetting cows for minimizing heat from solar radiation, improving cooling system and developing nutritional strategies (Bond and Kelly 1955).

Hertzman (1960) was also the first person who explained hypohydration (decrease in body water content) during heat stress as a "failure of the thermoregulatory system".

This study investigates the most influential climate parameter by heat stress indices on cow milk as a strategic product and presents a relationship between some important climate indices (THI, ETI,ESI, HLI, HLI_{new} and RRP) and milk fundamental compounds (milk yield, fat and

Table 5.1: Correlation between climate indices (THI, ETI, ESI, HLI, HLI_{new} and RRP) and milk compounds (fat, milk yield and protein) in spring ($\star$ is P value < 0.05, $\circ$ is P value < 0.1 and $\bullet$ is not significant).

Components	THI	ETI	ESI	HLI	HLI_{new}	RRP
Fat	-0.47$\star$	0.42$\star$	-0.41$\star$	-0.41$\star$	-0.4$\star$	-0.41$\star$
Milk yield	-0.22$\bullet$	-0.21$\circ$	0.26$\star$	0.27$\circ$	0.26$\circ$	0.26$\circ$
Protein	0.37$\circ$	-0.43$\circ$	0.41$\star$	0.42$\star$	0.44$\star$	0.41$\star$

Table 5.2: Correlation between climate indices (THI, ETI, ESI, HLI, HLI_{new} and RRP) and milk compounds (fat, milk yield and protein) in summer ($\star$ is P-value < 0.05 , $\circ$ is P-value < 0.1 and $\bullet$ is not significant).

Components	THI	ETI	ESI	HLI	HLI_{new}	RRP
Fat	-0.32$\star$	0.17$\star$	-0.23$\star$	-0.21$\circ$	-0.21$\circ$	-0.23$\star$
Milk yield	-0.26$\bullet$	-0.43$\circ$	0.43$\star$	0.43$\star$	0.41$\star$	0.43$\star$
Protein	0.31$\circ$	-0.34$\circ$	0.37$\star$	0.36$\circ$	0.34$\circ$	0.37$\star$

protein) observations. Correlation analysis between climate indices and milk components is done under bootstrap technique for considering unavoidable uncertainty.

The summery of results in spring and summer are presented in Table 5.1 and 5.2 respectively. More comprehensive explanation of these results are described in the results part of Chapter 2 and 3.

Significant relationships in spring and summer are identified in this study between climate indices and milk yield as well as protein.

But we found that the single use of indices is valid only under certain and specific climate conditions. The results in Table 5.1 and 5.2 show that each index separately shows very different correlation with milk compounds although these indices have high correlation with each other (Chapter 2 and 3, Figures 5 and 6).

Coping with this problem, Brake and Bates (2002) developed a method of limiting metabolic rate, which allows all heat stress indices to be compared with each other virtually.

Comparing with Brake and Bates (2002), we eliminated this problem by applying multiple linear regression model to develop new indices through combination of single heat stress indices and climate parameters together (Chapter 4).

We obtained three suitable models as the best linear regression models between climate indices and milk yield in spring and summer. In these models correlation and R^2 values are highly promoted and they are more suitable and significant models than the suggested alone indices in our first and second papers. For fat and protein any significant linear regression model could not be found better than single indices which are suggested in our previous studies (Marami et al [6,13]). Based on the results of effect of heat stress indices on milk compounds separately, we suggest the use of *ESI* and *RRP* in the summer and *ESI* in the spring as alone indices. *THI* and HLI_{new} are suggested for fat content and HLI_{new} also is suggested for protein content in the spring season.

Two the best multiple linear regression models for milk yield in spring are presented in Eq. (5.1) and Eq. (5.2)

$$Milk-yield_{Spring} = -3.485THI - 0.68ETI - 0.295HLI + 3.13RRP + e \tag{5.1}$$

$$Milk-yield_{Spring} = -3.58THI - 0.77ETI + 5.14ESI - 2.3RRP + e \tag{5.2}$$

The best mix linear models for milk yield in summer is found as in Eq. (5.3) as

$$Milk-yield_{Summer} = -1.6THI - 0.28ETI + 1.47ESI + e \tag{5.3}$$

According to reported critical heat stress points for THI, ETI, HLI, HLI_{new}, ESI and RRP, number of days that cows are under stress is very important because it is the start point of negative effect on milk compounds.

In regards to the obtained models, it seems that the value of critical points for starting the negative effects of heat stress on milk fundamental components are different.

As a discussion, it is strongly suggested that new and significant indices are needed to control critical heat stress conditions that consider more predictors of the effect of climate variability on animal products, such as sunshine duration, the quality of the cow's diet, the number of days of stress, the color of skin with attention to large black spots, and categorical predictors such as breed, welfare facility, and management system. For future studies It is recommended to study the variability of often used heat indices in last 50 years in different climate condition to update the critical heat stress point and investigate multicollinearity of indices.

New indices should be considered in different climate conditions because of characteristics changes and cow adaptation ability with focus on the fundamental milk components separately.

Application of the statistical methods in this study have advantages and can serve as a good tool to validate the predictive model for such studies in developing countries with short term data sets.

Last but not least in this study, statistical relationships are assessed between fundamental milk components and related climate parameters also with relatively new heat stress indices. Some good statistical methods are introduced for data preparation and analysis to cope with uncertainties in time series data sets. The best suitable and significant models with the smallest number of predictors are developed by applying new statistical techniques and avoiding misinterpretations due to overfitting.

This methodology is suggested for studies investigating the impacts of climate variability/change on food quality/security, animal science and agriculture using short term data considering uncertainty or data collection is expensive, difficult, or data with gaps.

It is hoped that these kind of studies could open some new ways for enthusiast researchers and applicators to optimize designing dairy factories in order to minimizing negative effect of climate variability on livestocks especially in developing countries.

Acknowledgment

Hereby, I warmly thank my supervisors, Prof. Dr. Angelika Ploeger and Prof. Dr. Andreas Hense.

Thanks a lot to Prof. Dr. Ploeger for including me in her research team and her financial backing and great assistance.

My special thanks to Prof. Dr. Hense who I learned many things from him and colleagues which supported me scientifically.

Deep appreciation to my loving wife, Elham for her enlightening and great supporting in the whole years of my PhD.

I am very grateful from my faculty colleagues and friends who assisted, advised and also helped me in learning Deutsch, idiom and phrase which is really my interest.

Specially, I will express my gratitude and appreciation to our good friend, PD Dr. Petra Friederichs for her kindness and splendid support.

Warmly thanks to Dr. Nasim Azari at the Food Science Engineering Department of Science and Research branch of Tehran Islamic Azad University, Ahmad Moghimi Esfand Abadi and Asghar Salimi Niknam at the Animal Breeding Centre of Iran in Karaj, branch of the Iranian Ministry of Agriculture, Dr. Behnam Saremi and Dr. Ali Asadi for supporting the milk data.

I am also grateful to Prof. Dr. Ali Mortazavi in Department of Food Science, Ferdowsi Mashhad University for his cooperation.

Chapter 6

Bibliography

This part introduces the used references in this dissertation.

Ahrens, C. *Essentials of Meteorology: An Introduction to the Atmosphere*, 5th ed.; Thomson Learning: Belmont, CA, USA, 2008; p. 448.

Akaike, H. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on* **1974**, *19*, 716-723.

Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection. *Statistics surveys* **2010**, *4*, 40-79.

Baeta, F.C.; Meador, N.F.; Shanklin, M.D.; Johnson, H.D. Equivalent temperature index at temperatures above the thermoneutral for lactating dairy cows. In proceedings of the Summer Meeting of American Society of Agricultural Engineers (ASAE), Baltimore, MD, USA, 28 June–1 July 1987; pp. 87–4015.

Beatty, D.T.; Barnes, A.; Taylor, E.; Pethick, D.; McCarthy, M.; Maloney, S.K. Physiological responses of Bostaurus and Bosindicus cattle to prolonged continuous heat and humidity. *J. Anim. Sci*. **2006**, *84*, 972–985.

Berman, A.; Folman, Y.; Kaim, M.; Mamen, M.; Herz, Z.; Wolfenson, D.; Arieli, A.; Graber, Y. Upper critical temperatures and forced ventilation effects for high-yielding dairy cows in a subtropical climate. *J. Dairy Sci*. **1985**, *68*, 1488–1495.

Bohmanova, J.; Misztal, I.; Colet, B. Temperature-Humidity Indices as Indicators of Milk Production Losses due to Heat Stress. *J. Dairy Sci*. **2007**, *90*, 1947–1956.

Bond, T. E.; Kelly, C.F. The globe thermometer in agricultural research. *Agricultural Engineering* **1955**, 251-255.

Bouraoui, R.; Lahmarb, M.; Majdoubc, A.; Djemalic, M.; Belyead, R. The relationship of temperature-humidity index with milk production of dairy cows in a Mediterranean climate. *Anim. Res*. **2002**, *51*, 479–491.

Brake, R.; Graham, B. A valid method for comparing rational and empirical heat stress indices. *Annals of occupational hygiene* **2002**, 46, 165-174.

Brown-Brandl, T. M.; Eigenberg, R. A.; Nienaber, J. A.; Hahn, G. L. Dynamic response indicators of heat stress in shaded and non-shaded feedlot cattle, Part 1: Analyses of indicators. *Biosystems engineering* **2005**, *90*, 451-462.

Burnham, K.P.; Anderson, D.R. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media: New York, NY, USA, 2002; pp. 49-97.

Ciais, Ph.; Reichstein, M.; Viovy, N.; Granier, A.; Ogée, J.; Allard, V.; Aubinet, M.; Buchmann, N.; Bernhofer, Chr.; Carrara, A.; *et al.* Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature* **2005**, *437*, 529–533.

Collier, R.J.; Dahl, E.; VanBaale, M.J. Major advances associated with environmental effects on dairy cattle. *J. Dairy Sci*. **2006**, *89*, 1244–1253.

Darwin, R. *Climate Change and Food Security*; Agriculture Information Bulletin Number 765–8; United States Department of Agriculture: Washington, DC, USA, 2001.

Davison, A.C.; Hinkley, D.V. *Bootstrap Methods and Their Application*; Cambridge University Press: New York, NY, USA, 1997.

Dinpashoh, Y.; Jhajharia, D.; Fakheri-Fard, A.; Singh, V.P.; Kahya, E. Trends in reference crop evapotranspiration over Iran. *J. Hydrol.* **2011**, *399*, 422–433.

Dupre, R. K.; Romero, A. M.; Wood, S. C. Thermoregulation and metabolism in hypoxic animals. In *Oxygen transfer from atmosphere to tissues,* Springer, US, 1988; pp. 347-351.

Efron, B.; Tibshirani, R. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statist. Sci*. **1986**, *1*, 54–75.

Eigenberg, R.A.; Hahn, G.L.; Nienaber, J.A.; Brown-Brandl, T.M. Development of a new respiration rate monitor for cattle. *Trans. ASAE* **2000**, *43*, 723–728.

Eigenberg, R.A.; Nienaber, J.A.; Brown-Brandl, T.M. Development of a livestock safety monitor for cattle. In Proceedings of the 2003 American Society of Agricultural and Biological Engineers (ASABE) Annual Meeting, St. Joseph, MI, USA, 27–30 July 2003.

Epstein, Y.; Moran, D.S. Thermal comfort and the heat stress indices. *Ind. Health* **2006**, *44*, 388–398.

Friedman, J.; Hastie, T.; Simon, N.; Tibshirani, R. glmnet : Lasso and Elastic-net regularized generalized linear models. *R package version* **2015**, 1.

Friendly, M. Corrgrams: Exploratory displays for correlation matrices. *Am. Stat.* **2002**, *56*, 316–324.

Gaughan, J.B.; Mader, T.L.; Holt, S.M.; Josey, M.J.; Rowan, K.J. Heat tolerance of Boran and Tuli crossbred steers. *J. Anim. Sci*. **1999**, *77*, 2398–2405.

Gaughan, J.G.; Goopy, L.; Spark, J. Excessive heat load index for feedlot cattle. *Sydney: MLA*. **2002**.

Gaughan, J.G.; Mader, T.L.; Holt, S.M.; Lisle, A. A new heat load index for feedlot cattle. *J. Anim. Sci*. **2008**, *86*, 226–234.

Gauly, M.; Bollwein, H.; Breves, G.; Brugemann, K.; Danicke, S.; Das, G.; Demeler, J.; Hansen, H.; Isselstein, J.; Konig, S.; *et al.* Future consequences and challenges for dairy cow production system arising from climate change in center Europe—A review. *Animal* **2013**, *7*, 843–859.

Hahn, G.L.; Mader, T.L. Heat waves in relation to thermoregulation, feeding behavior and mortality of feedlot cattle. In Proceedings of the 5th International Livestock Environment Symposium, Bloomington, MN, USA, 29–31 May 1997; pp. 545–549.

Hammami, H.; Vandenplas, J.; Vanrobays, M.L.; Rekik, B.; Bastin, C.; Gengler, N. Genetic analysis of heat stress effects on yield traits, udder health, and fatty acids of Walloon Holstein cows. *J. Dairy Sci.* **2015**, *98*, 4956–4968

Harrison, M.H.; Edwards, R.J; Fennessy, P.A. Intravascular volume and tonicity as factors in the regulation of body temperature. *Journal of Applied Physiology* **1988**, 69-75.

Heinrichs, J.; Jones, C.; Bailey, K. MILK Components: Understanding the Causes and Importance of Milk Fat and Protein Variation in Your Dairy Herd. *Dairy Anim. Sci*. **1997**, *5*, 1e–8e.

Hertzman, A. R. Failure in temperature regulation during progressive dehydration. *Journal of Occupational and Environmental Medicine* **1960**, 2, 55.

Ingraham, R.H.; Stanley, R.W.; Wagner, W.C. Relationship of temperature and humidity to conception rate of Holstein cows in Hawaii. *J. Dairy Sci.* **1974**, *59*, 2086–2090.

Ingram, D.L.; Mount, L.E. Heat exchange between animal and environment. In *Man and Animals in Hot Environments*; Springer: New York, NY, USA, 1975; pp. 5–23.

Jhun, M.; Jeong, H.C. Applications of bootstrap methods for categorical data analysis. *Comput. Stat. Data Anal.* **2000**, *35*, 83–91.

Johnson, H.D. Environmental management of cattle to minimize the stress of climatic change.
Int. J. Biometeorol. **1980**, *24*, 65–78.

Knapp, D.M.; Grummer, R.R. Response of lactating dairy cows to fat supplementation during heat stress. *J. Dairy Sci*. **1991**, *74*, 2573–2579.

Kousari, M.R.; Ahani, H.; Hendi-Zadeh, R. Temporal and spatial trend detection of maximum air temperature in Iran during 1960–2005. *Glob. Planet. Chang.* **2013**, *111*, 97–110.

Kraus, H. *Die Atmosphäre der Erde*, 3rd ed.; Springer: Berlin Heidelberg, Germany, 2004;pp. 67–89.

Kuhn, M.T.; Hutchison, J.L.; Norman, H.D. Dry Period Length to Maximize Production Across Adjacent Lactations and Lifetime Production. *J. Dairy Sci*. **2006**, *89*, 1713–1722.

Kullback, S.; Leibler, R.A. On information and sufficiency. *The annals of mathematical statistics* **1951**, 79-86.

Lahiri, S.N. Bootstrap methods. In *Resampling Methods for Dependent Data*; Springer: New York, NY, USA, 2003; pp. 17–43.

Lambertz, C.; Sanker, C.; Gauly, M. Climatic effects on milk production traits and somatic cell score in lactating Holstein-Friesian cows in different housing systems. *J. Dairy Sci.* **2014**, *97*, 319–329.

Lescourret, F.; Coulon, J.B. Modeling the Impact of Mastitis on Milk Production by Dairy Cows. *J. Dairy Sci.* **1994**, *77*, 2289–2301.

Lunneborg, C.E. Estimating the correlation coefficient: The bootstrap approach. *Psychol. Bull.* **1985**, *98*, 209–215.

Maak, K.; von Storch, H. Statistical downscaling of monthly mean air temperature to the beginning of flowering of Galanthus nivalis L. in Northern Germany. *Int. J. Biometeorol.* **1997**, *41*, 5–12.

Mader, T.L.; Dahlquist, J.M.; Hahn, G.L.; Gaughan, J.B. Shade and wind barrier effects on summertime feedlot cattle performance. *J. Anim. Sci.* **1999**, *77*, 2065–2072.

Mader, T.L.; Davis, M.S. Wind speed and solar radiation corrections for the temperature-humidity index. In Proceedings of the 15th Conference on Biometeorology and Aerobiology Joint with 16th International Congress on Biometeorology, Kansas City, MO, USA, 27 October 2002; pp. 10–28.

Martinez, M.L.; Lee, A.J.; Lin, C.Y. Age and Zebu-Holstein Additive and heterotic Effects on Lactation Performance and Reproduction in Brazil. *J. Dairy Sci*. **1988**, *71*, 800–808.

Matulla, C.; Scheifinger, H.; Menzel, A.; Koch, E. Exploring two methods for statistical downscaling of Central European phenological time series. *Int. J. Biometeorol.* **2003**, *48*, 56–64.

McLean, J.A. Loss of heat by evaporation. In *Heat Loss from Animals and Man: Assessment and Control*; Monteith, J.L, Mount, L.E., Eds.; Butterworth-Heinemann: London, UK, 1974; pp. 19–31.

Milani, M.R.M.; Hense, A.; Rahmani, E.; Ploeger, A. A Survey of the Relationship between Climatic Heat Stress Indices and Fundamental Milk Components Considering Uncertainty. *Climate* **2015**, *3*, 876-900.

Milani, M.R.M; Hense, A.; Rahmani, E.; Ploeger, A. A pilot investigation of the relationship between climate variability and milk compounds under the bootstrap technique. *Foods* **2015**, *4*, 420–439.

Modarres, R.; Sarhadi, A. Statistically-based regionalization of rainfall climates of Iran. *Glob Planet. Chang.* **2011**, *75*, 67–75.

Moore, A.W. Cross-validation for detecting and preventing overfitting. *School of Computer Science Carneigie Mellon University* **2001**.

Moran, D. S.; Shitzer, A.; & Pandolf, K. B. A physiological strain index to evaluate heat stress. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **1998**, 275, 129-134.

Moran, D.S.; Epstein, Y. Evaluation of the environmental stress index (ESI) for hot/dry and hot/wet climates. *Ind. Health* **2006**, *3*, 399–403.

Moran, D.S.; Pandolf, K.B.; Shapiro, Y.; Heled, Y.; Shani, Y.; Matthew, W.T.; Gonzales, R.R. An environmental stress index (ESI) as a substitute for the wet bulb globe temperature (WBGT). *J. Therm. Biol.* **2001**, *26*, 427–431.

Moran, D.S.; Pandolf, K.B.; Shapiro, Y.; Laor, A.; Heled, Y.; Gonzalez, R.R. Evaluation of the environmental stress index for physiological variables. *J. Therm. Biol.* **2003**, *28*, 43–49.

Mukherjee, D.; Bravo-Ureta, B.E.; De Vries, A. Dairy productivity and climatic conditions: Econometric evidence from South-eastern United States. *Aust. Agric. Res. Econ.* **2012**, *57*, 123–140.

Myneni, R.B.; Keeling, C.D.; Tucker. C.J.; Asrar, G.; Nemani, R.R. Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* **1997**, *386*, 698–702.

Nadel, Ethan R.; Fortney, S.M; Wenger, C.B. Effect of hydration state of circulatory and thermal regulations. *Journal of Applied Physiology* **1980**, 715-721.

Nikkhah, A.; Furedi, C.J.; Kennedy, A.D.; Scott, S.L.; Wittenberg, K.M.; Crow, G.H.; Plaizier, J.C. Morning *vs.* evening feed delivery for lactating dairy cows. *Can. J. Anim. Sci.* **2011**, *91*, 113–122.

O'Brien, M.D.; Rhoads, R.P.; Sanders, S.R.; Duff, G.C.; Baumgard, L.H. Metabolic adaptations to heat stress in growing cattle. *Domest. Anim. Endocrinol.* **2010**, *38*, 86–94.

Politis, D.N.; Romano, J.P.; Wolf, M. *Subsampling*, Springer: New York, NY, USA, 1999; pp. 7-50, 98-100.

Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *Numerical Recipes: The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2007; pp. 340–439.

R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2011.

R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2012.

Rahmani, E. The Effect of Climate Variability on Wheat in Iran. Ph.D. Thesis, University of Bonn, Bonn, Germany, 2015.

Rahmani, E.; Friederichs, P.; Keller, J.; Hense A. Development of an effective and potentially scalable weather generator for temperature and growing degree days. *Theor. Appl. Climatol.* **2015**, *120*, doi:10.1007/s00704-015-1477-z.

Rasmussen, J.L. Estimating correlation coefficients: Bootstrap and parametric approaches. *Psychol. Bull.* **1987**, *101*, 136–139.

Reddy, P. Parvatha. *Climate Resilient Agriculture for Ensuring Food Security*. Springer India, 2015, pp. 73.

Renaudeau, D.; Collin, A.; Yahav, S.; De Basilio, V.; Gourdine, J.L.; Collier, R.J. Adaptation to hot climate and strategies to alleviate heat stress in livestock production. *Animal* **2012**, *6*, 707–728.

Reynolds, M. Plasma and blood volume in the cow using the T-1824 hematocrit method. *American Journal of Physiology--Legacy Content* **1953**, 173.

Rienecker, M.M.; Suarez, M.J.; Gelaro, R.; Todling, R.; Bacmeister, J.; Liu, E.; Bosilovich, M.G.; Schubert, S.D.; Takacs, L.; Kim, G.-K.; *et al.* MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Clim.* **2011**, *24*, 3624–3648.

Rodriquez, L.A.; Mekonnen, G.; Wilcox, C.J.; Martin, F.G.; Krienke, W.A. Effects of Relative Humidity, Maximum and Minimum Temperature, Pregnancy, and Stage of Lactation on Milk Composition and Yield. *J. Dairy Sci*. **1985**, *68*, 973–978.

Roman-Ponce, H.; Thatcher, W.W.; Buffington, D.E.; Wilcox, C.J.; Van Horn, H.H. Physiological and Production Responses of Dairy Cattle to a Shade Structure in a Subtropical Environment. *J. Dairy Sci*. **1977**, *60*, 424–430.

Sharma, A.K.; Rodriguez, L.A.; Mekonnen, G.; Wilcox, C.J.; Bachman, K.C.; Collier, R.J. Climatological and genetic effects on milk composition and yield. *J. Dairy Sci*. **1983**, *66*, 119–26.

Silva, R.G.D; Campos Maia, A.S. Thermal stress indexes. In *Principles of Animal Biometeorology*; Springer Science & Business Media: Dordrecht, The Netherlands, 2013; pp. 207–229.

Silva, R.G.D; Morais, D.A.; Guilhermino, M.M. Evaluation of thermal stress indexes for dairy cows in tropical regions. *Revista Brasileira de Zootecnia* **2007**, *36*, 1192-1198.

Spain, J.N.; Spiers, D.E. Effects of supplemental shade on thermoregulatory response of calves to heat challenge in a hutch environment. *J. Dairy Sci.* **1996**, *79*, 639–646.

Statistical Center of Iran / History. Available online: http://www.amar.org.ir/Default.aspx?tabid=281 (accessed on 05 August 2014).

Thom, E.C. The discomfort index. *Weatherwise* **1959**, *12*, 57–60.

Tibshirani, R. Regression shrinkage and selection via the lasso. *Royal. Statist. Soc B.* **1996**, *58*, 267-288.

Udo, H.J.M. Hair Coat Characteristics in Friesian Heifers in the Netherlands and Kenya: Experimental Data and A Review of Literature. Ph.D. Thesis, University of Wageningen: Wageningen, The Netherlands, 1978.

United Nation, Department of Economic and Social United Affairs. World Population Prospects the 2012 Revision. Available online: http://esa.un.org/unpd/wpp/index.htm (accessed on 10 June 2014).

Watson, R.T.; Albritton, D.L.; Barker, T.; Bashmakov, I.A.; Canziani, O.; Christ, R.; Cubasch, U.; Davidson, O.; Gitay, H.; Griggs, D.; *et al.* Climate Change 2001: Synthesis Report. Avalable oneline: http://ipcc.ch/meetings/session18/doc3b.pdf (accessed on 5 March 2015).

West, J.W. Effects of heat-stress on production in dairy cattle. *J. Dairy Sci.* **2003**, *86*, 2131–2144.

Wilks, Daniel S. *Statistical methods in the atmospheric sciences*, 3rd ed.; Academic press: San Diego, CA, 2011, pp. 252-254.