UTRECHT UNIVERSITY

Assessing the response of vegetation energy fluxes to heat-waves using the H-TESSEL land surface scheme

Author:

Stamatios Petalas

Supervisor:

Bart van den Hurk

Examiners:

Bart van den Hurk

Aarnout van Delden

A thesis submitted in partial fulfilment for the
Master's degree in Meteorology, Physical Oceanography and Climate

in the
Faculty of Science
Department of Physics and Astronomy

September 2012

Dedicated to my family,
for all their efforts for me,
and the boundless love
they always embraced me with ...

Abstract

In literature it is suggested that energy fluxes over forest and grassland have a contrasting response during heat-waves. In order to investigate whether this is reproduced by the H-TESSEL land surface scheme, an 11-year offline run of the model is compared to observations.

A selection of heat-waves based on the ECA/E-Obs dataset is performed, after which the surface energy fluxes are compared for the two vegetation types. The number of heat-waves are then correlated to spring-time and summer-time soil moisture depletion. A connection is found for summer-time, but not for spring-time water depletion.

Net radiation is compared to observations, and the additional amount received during heat-waves is found to be underestimated by the model, which most probably comes from the low discernibility of the actual observations in the model forcing. Grassland in general receives the same net radiation with forest in contrast with observations, due to the similar albedo prescription in the model.

Fluxes are then compared for the two vegetation types. Forests in general compensate for the additional amount of net radiation by increasing sensible heat release, while grasslands increase latent-heat release. The observed differences in the response of forest and grassland fluxes to heat-waves are simulated by the model with a small differentiation in the partitioning of net radiation. When only heat-waves with low soil moisture are selected, there seems to be an improvement to that. Short-duration heat-waves are then filtered out, which brings the results even closer to observations.

Latent heat is overestimated during heat-waves, especially for grassland, and in combination with the fact that sensible heat is not differentiated in the same amount, points to the connection with vapour pressure deficit, which should be further investigated. Ground flux in both vegetation types remains high compared to the observations. This is potentially a bias in skin temperature calculation connected to sensible and latent heat through the surface energy balance solver, which was not investigated further.

Fluxes are then examined during the evolution of heat-waves. This shows no differentiation from previous results, while extreme water depletion in long-lasting events was not observed. A CART analysis is performed on the dataset which highlights the conservative soil water consumption of forest compared to grassland, and points to a stronger regulation of the evaporation by forests, depending on soil moisture. Finally, the possibility of using artificial neural networks for reproducing statistical relations in the model is shortly discussed.

Acknowledgements

Working on a thesis research project is a long learning process, usually bringing a fair share of excitement and disappointment, with an almost constant periodicity between the two. Here, I would like to thank those who helped me go through this process by making it an enjoyable and very rewarding journey.

First and foremost, I offer my gratitude to my supervisor Bart van den Hurk, who supported me with his patience, knowledge and understanding, and his always on the point comments and suggestions, whilst allowing me the space to work on my own way. The opportunities he gave me to expand my knowledge were numerous, including the valuable experience of working at KNMI. His encouragement and effort combined with a friendly attitude, made this study an inspiring work for me.

In my everyday life in KNMI, I had the luck to work with a friendly and cheerful group of colleagues, who helped me with technical issues whenever I knocked on their door, and blessed me with their company during the rest of the time. I thank them all for their help and friendship.

As this report marks also the end of my studies in Utrecht University, I would like to take the opportunity here to thank all the teaching stuff in IMAU for imparting me with such a precious amount of knowledge, and making the experience of studying abroad less strenuous, with their support and understanding. I would especially like to mention Huib de Swart, who stood to me like a father during a very difficult period. I thank them all from my heart.

Beyond university life, Alex, Mirena and Francesco have been companionable house-mates and good friends, throughout my studies. I am grateful for their company and friendship, which together with our long talks helped me to learn more about myself, and walk past any obstacles that came into my way.

Finally, Stavroula helped me throughout this effort in her own special way. This whole studying and thesis experience would have been very different if it wasn't for her love, support and understanding.

Contents

Bibliography **51**

List of Figures

ix

Abbreviations

GCM	**G**lobal **C**irculation **M**odel
HWD	**H**eat–**W**ave **D**ay
ECMWF	**E**uropean **C**enter for **M**edium-range **W**eather **F**orecasts
H-TESSEL	**H**ydrology–**T**iled **E**CMWF **S**cheme for **S**urface **E**xchanges over **L**and
WMO	**W**orld **M**eteorological **O**rganization
KNMI	**K**oninklijk **N**ederlands **M**eteorologisch **I**nstituute
	(Royal Netherlands Meteorological Institute)
KMI	**K**oninklijk **M**eteorologisch **I**nstituute van België
	(Royal Meteorological Institute of Belgium)
NOAA	**N**ational **O**ceanographic and **A**tmospheric **A**dministration
ABM	**A**ustralian **B**ureau of **M**eteorology
NASA	**N**ational **A**eronautics and **S**pace **A**dministration
MODIS	**MOD**erate resolution **I**maging **S**pectroradiometer
ASTER	**A**dvanced **S**paceborne **T**hermal **E**mission and reflection **R**adiometer
LSS	**L**and **S**urface **S**cheme
DGVM	**D**ynamic **G**lobal **V**egetation **M**odel
NCEP	**N**ational **C**enters for **E**nvironmental **P**redictions
NCAR	**N**ational **C**enter for **A**tmospheric **R**esearch
ECA	**E**uropean **C**limate **A**ssessment
CART	**C**lassification **A**nd **R**egression **T**rees
ANN	**A**rtificial **N**eural **N**etwork

"Every blade in the field - Every leaf in the forest - lays down its life in its season as beautifully as it was taken up."

Henry David Thoreau

Chapter 1

Introduction

Earth's climate is a complex dynamical system which, while experiencing random fluctuations around a steady state on short time periods, has been constantly changing throughout its history. The global rise in temperature during the last hundred years though, has justifiably drawn the attention of human societies. The alarm of potentially unbalanced conditions in the climate system, was first sound by scientists, with governments and lay-people slowly understanding the need to further investigate the reasoning behind this behaviour. Although there has been an extensive debate on the actual driving mechanisms (with human-activity induced warming gradually prevailing), the fact remains that global mean temperature rose by $0.74\,°C \pm 0.18\,°C$ over the years 1906–2005, with the rate of warming over the last 50 years being almost double of that over the last 100 years [1]. Moreover this trend is evidently larger than at any 100-year period in the last 2000 years, as constructed from climate proxies.

A slightly higher global mean temperature alone might not heavily affect life on the planet, but usually it is not the averages of weather phenomena that cause disastrous effects to societies, but their extreme counterparts. It is found that extreme weather events, like heat waves and heavy precipitation, are already observed to increase in frequency and intensity. In addition, it is considered very likely (more than 90% probability) that this trend will continue in the future [1]. In order to understand the drivers of climatic variability and change, one must first understand the different components of the system, and the interactions between them.

Meteorology and climatology in their short lifespan have managed to evolve the human understanding of weather phenomena, from a state of quasi-clairvoyance to a rational science. Knowledge acquired throughout the years is now being applied extensively in numerical models, for short term weather forecast as well as climate projections. These models are continuously growing in complexity, by gradually incorporating additional physical mechanisms, feedbacks and interactions in the earth system. This, in conjunction with observations, provides the opportunity to test the theory, increase the accuracy and expand the knowledge.

The surface of the earth (i.e. land and ocean), acts as the lower boundary of the atmosphere, where all the exchange of energy, water and chemical compounds takes place. The connection between surface–atmosphere interactions and climatic variability has been widely reported (e.g. [2] for land and [3] for ocean), making evident the need for an accurate implementation in GCMs. Studies like this show that feedbacks between vegetation, ground moisture and atmospheric conditions can have a pronounced impact on the intensity and/or the probability of occurrence, of extreme events like heat waves.

In the present study, land–surface energy fluxes during heat waves, are examined for two types of vegetation cover, namely forest (or high–vegetation) and grassland (or low–vegetation). This is done by analyzing data coming from an offline GCM experiment [1], and using an observational study as a guide, to assess the performance of the model's land–surface scheme during heat wave conditions.

It is considered helpful, to continue with an overview of what is a heat wave and how it is defined, as well as the atmospheric conditions under which such incidents occur. Following that, is a brief introduction on the elements of vegetation physiology affecting and controlling energy/water fluxes, and its response to heat waves. A review of a selected observational study of land–surface energy fluxes during HWDs is then presented, concluding with an overview of the representation of land–atmosphere interaction in models. The intention for this chapter is to serve as both an introduction and a rationale for the study.

The second chapter is covering the materials and methods used, starting with a description of the H-TESSEL land–surface scheme. Then the datasets that where used are presented, together with a description of the steps to bring them in a more usable format. Finally some additional tools of the analysis procedure are discussed. The third chapter is about the process of selecting HWDs out of the 11 years of available data; how and on which basis this was done, and what was the result of this selection. In the fourth chapter the results of the analysis for the selected HWDs are presented, ending with the conclusion chapter, which wraps up the main points of this study.

The overall performance of H-TESSEL has been investigated in depth in [4]. The main purpose of the present research is to examine whether the energy fluxes modelled during extreme events like heat-waves correspond to the observations, and whether fluxes of forest and grassland show the same differentiation during these events, as in reality. Moreover, as the behaviour of these two vegetation types are defined by parameterisation in the energy balance and hydrology part of the model, it is important to point out the parts of this parameterisation that are working in a correct way, and those that need to be changed, or further examined. Although H-TESSEL demonstrates a good

---

[1]Meaning a GCM run forced by atmospheric conditions, but not coupled to the atmosphere. This is usually done to isolate the model part under examination (in this case the land–surface scheme), for having a better overview of its functionality.

overall performance, it has not been examined under extreme temperature conditions, where the energy flux response of vegetation shows a deviation from the norm. It is possible that the model will not be able to reproduce accurately the observations for such periods. Discovering whether this is the case and the reasoning behind it, acts as the main motivation for this study.

1.1 Definition and occurrence of HWDs

Heat waves are events of prolonged, excessively high temperatures (compared to the average of the period they occur in). In the common sense they are associated with an increased level of heat–discomfort, combined with drought and environmental hazards, so the definition from place to place depends on the norm for that region. In the scientific community, different meteorological institutes and organizations either adopt criteria that best fit their area of operation, or use the more generalized criterion of WMO for a heat wave period. Bellow are some examples of HWD definition for five different institutes/agencies:

- **WMO**: A period of more than 5 consecutive days, where the maximum daily temperature exceeds the average maximum temperature by $5\,^{\circ}\mathrm{C}$, the normal period being 1961–1990.

- **KNMI/KMI**: A period of at least 5 consecutive days in which the maximum temperature in De Bilt, The Netherlands exceeds $25\,^{\circ}\mathrm{C}$, provided that on at least 3 days in this period the maximum temperature in De Bilt exceeds $30\,^{\circ}\mathrm{C}$.

- **NOAA**: A period of three consecutive days where the temperature reaches or exceeds $32 - -38\,^{\circ}\mathrm{C}$ (in the USA depending on the region), with a combination of humidity threshold, to form a heat index.

- **ABM**: A period of 5 consecutive days at or above $35\,^{\circ}\mathrm{C}$, or three consecutive days at or above $40\,^{\circ}\mathrm{C}$.

The definitions of HWDs, emerge from the need to determine and eventually predict a period of potentially catastrophic events due to high temperatures. In the present day European climate, a heat wave in these terms is much more probable in summer or spring time [2]. Indeed during the first decade of the 21st century, three major, large scale incidents took place in Europe. The summer of 2003 was the warmest on record in Europe according to the WMO, with the consequent heat wave affecting most central and western European countries, and especially France (figure 1.1). As an example,

---

[2]Although winter–time high temperature anomalies can occur, they are not commonly referred to as heat waves, as their effects are muffled by the generally lower temperatures. Cold wave is a term used to define low temperature anomalies in winter, when hazards due to freezing conditions are possible

near the area of Auxerre, central France, 7 days of temperatures higher than 40 °C where recorded during that summer. The human casualties due to this heat wave were
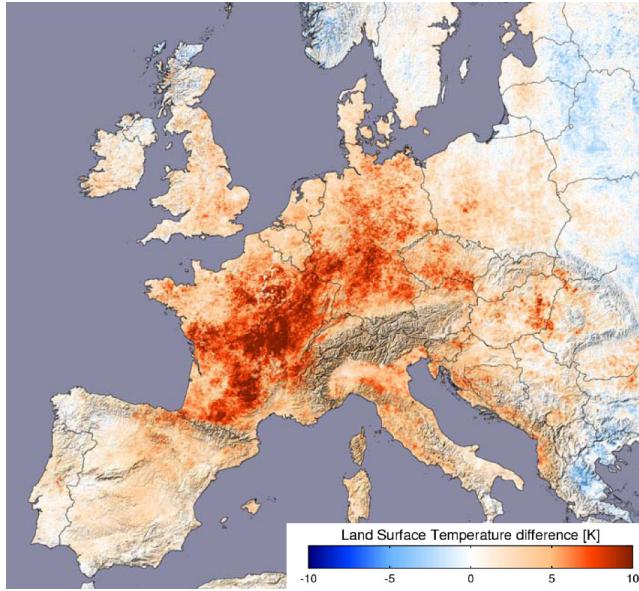
FIGURE 1.1: *Difference in land surface temperature, calculated by subtracting the average of all cloud free data during 2000, 2001, 2002 and 2004 from the ones measured in 2003, covering the date range of July 20–August 20 [Image by Reto Stöckli, Robert Simmon and David Herring, NASA Earth Observatory, based on data from the MODIS land team].*

officially over 35,000, while peer–reviewed analysis estimates this number up to 70,000 [5]. In addition, great economical losses because of massive crop failure, wildfires, and power outrages from electricity consumption spikes due to increased air conditioning use, round up the image of destruction during this event. The European heat wave of 2006 affected mostly northern and central countries, again with extraordinary temperatures and extensive casualties and damages, while the Russian heatwave of 2010, had the additional characteristic of a huge spatial extent.

Heat waves usually occur due to a combination of reasons. Initially, it is atmospheric circulation anomalies that provide the ground for such events. A persistent anti–cyclonic anomaly (or blocking), being almost stationary over a region, creates continuous subsidence, leading to clear skies and subsequent surface warming from solar radiation and warm air advection [6]. However, studies have shown that coexisting land–surface/hydrology anomalies, form the amplification factor that ultimately drives the temperatures so high. Low ground–water levels due to a precipitation deficit occurring in the winter/spring preceding the heat wave, allow for soil moisture depletion during the warm spell. This leads to reduced latent cooling, and a positive feedback loop [7]. Figure 1.2 shows results of an observational study for four summers, demonstrating the connection between spring–time precipitation anomalies and temperature anomalies in the consequent summer.
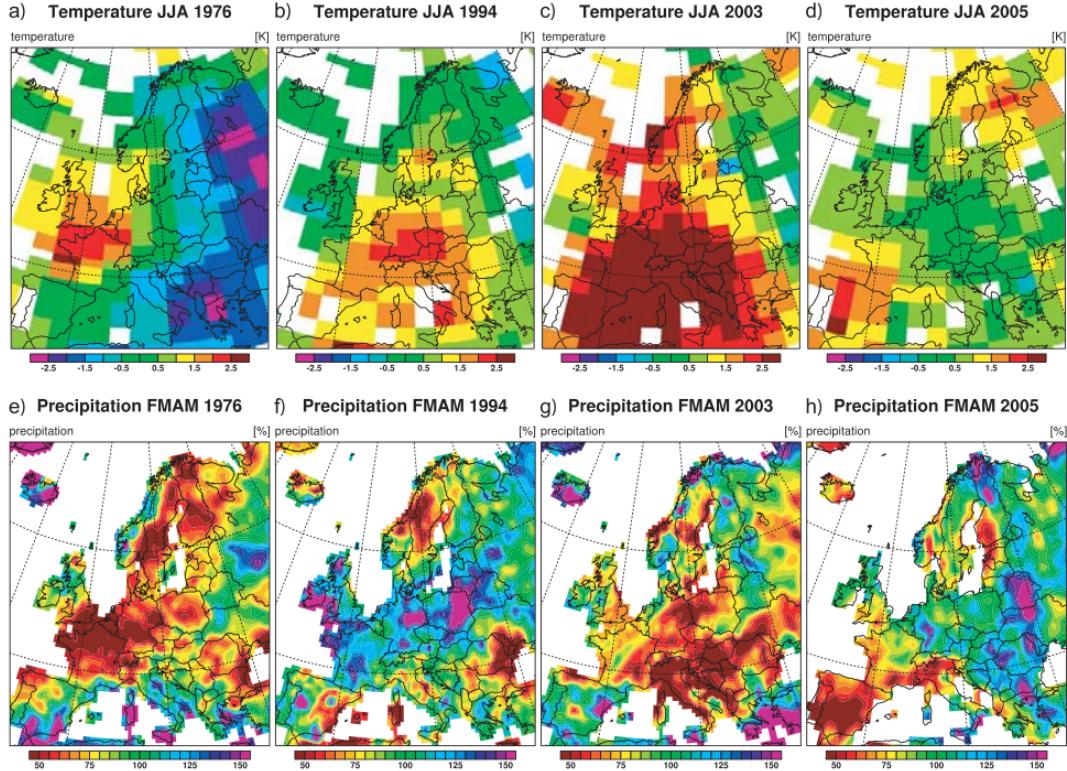
FIGURE 1.2: *Observed summer mean temperature anomalies, and precipitation percentages of the averaged over February–May, with respect to the climatological mean of 1970–2000 (Image from [8]).*

1.2 Physiological response of vegetation to heat waves

Latent heat release through water evaporation, has an important contribution to the land–surface's energy balance. This is crucial during a heat wave, when depletion of soil moisture due to heat and increased evaporation, can add a considerable feedback to the high temperatures. Plants can control the rate of their water uptake and release, a fact that differentiates vegetation canopies from passively–evaporating surfaces like bare ground or open water. This is achieved by regulating the opening and closing of the stomata plants have on their leafs, which act as portals to the outer environment. It is through these pores that chemical elements (like CO_2 and O_2) are exchanged, and water is evaporated, so a wider opening favours an increased rate for these procedures.

The propagation of water through a plant, starts at the root level, where it is absorbed from the soil, but the ultimate driver of this flow is evaporation in the stomata. This increases the solute concentration inside the plant, and the osmotic pressure gradient forces water to flow in at the roots. The main environmental factors that control stomatal aperture (or conductance), can be found in any introductory book on plant physiology, and are, in brief:

- **Radiation** Light induces photosynthesis, which boosts the nutritional cycle and in general increases stomatal conductance.

- **Temperature** Higher temperatures cause an increase in evaporation, and elevate the demand on water, so for most plants a higher temperature is connected with higher stomatal conductance.

- **Relative humidity** The rate of evaporation from a surface is depending, among other, on the relative humidity of the air above it. Usually in plants, larger relative humidity in the air close to the stomata leads to a decrease in evaporation, and stomatal aperture.

- **Wind** Higher winds induce turbulence, which in most cases mixes in dryer air and lowers the water vapour pressure close to the stomata, decreasing their conductance.

- **Soil water** Stomatal conductance is usually not depended on soil moisture when the later is ample. Bellow a certain threshold though (which depends on the plant type/species), stomatal aperture decreases with soil moisture in order to preserve water [9].

- **Nutrition** Internal and external concentration of elements needed in the plants nutritional cycle, like CO_2, can have an effect on stomatal conductance.

The exact interactions between these factors and plant biochemical mechanisms (that affect the guard cells in the stomata), is still an active field of research (e.g. [10], [11]). Coming to the case of stress conditions by elevated temperatures, photosynthesis is reduced, but the water loss via transpiration may still be high, as plants tend to trade-off conservation of water for leaf cooling, by increasing stomatal conductance [12]. The exact procedure for this behaviour depends on the plant, and different types of vegetation (e.g. grassland and forest), may have adopted different strategies for making this trade-off [13]. In a recent study on the European heat wave of 2003, it was suggested that grassland and forest responded differently during such events (figure 1.3).

In this figure the contrast between forest and crop-land is obvious. During a regular summer the vitality of vegetation (chlorophyll levels as red) appears normal in both forest and crop-land, but during the 2003 heat wave crop-land seems to have wilted (low chlorophyll levels as azure), while forest is closer to the regular for the season. The same contrast can be seen in the temperature maps. Soil moisture depletion and a feedback due to decreased latent cooling during the heat wave, is most probably the main cause of this behaviour. While evaporative cooling is usually greater over grasslands than over forests in normal conditions [15], it is likely that the reverse is happening during HWDs. In order to investigate this, an observational study on European forest and grassland sites was conducted in [16], outlined in the following.
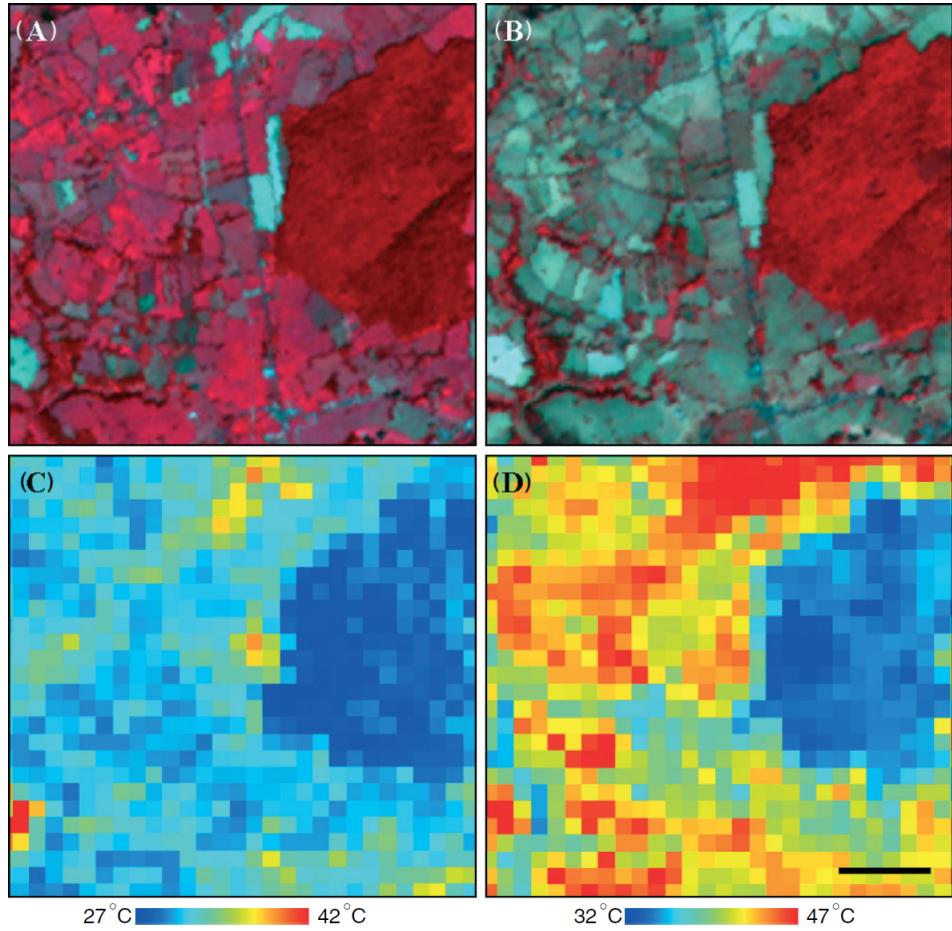
FIGURE 1.3: *Satellite images from ASTER. Atmospherically corrected surface infra-red reflectance of an area in central France (upper), and surface temperature for the same scenes (bottom), for 1 August 2000 (left) and 10 August 2003 (right). Vegetation in the upper pictures appears red, because of the high reflectance of chlorophyll in the infra-red. The red patch on the right of the pictures is a forested area, while the rest is mainly crop-lands. Scale bar indicates 500 m for all pictures (from [14]).*

1.3   A selected observational study on HWDs

An analysis of the surface energy flux partitioning in central-western Europe, was performed in [16], using observations from the FLUXNET network of eddy covariance flux towers. The goal was to compare surface energy exchange during heat wave conditions, for two surface types; forest and grassland. For this reason, tower sites were selected on the criterion of their location being either close to a grassland, or a forest region. The reference point would be the climatology of summers in the period 1997–2008, and the HWDs were defined according to the WMO criterion (but for the climatology in hand).

The results demonstrated a contrasting response of forest and grassland energy fluxes to heat waves. In more detail, the climatology shows large differences in reflected short-wave and net radiation, between forest and grassland, with forest absorbing more incoming short-wave radiation, because of its lower albedo (figure 1.4).
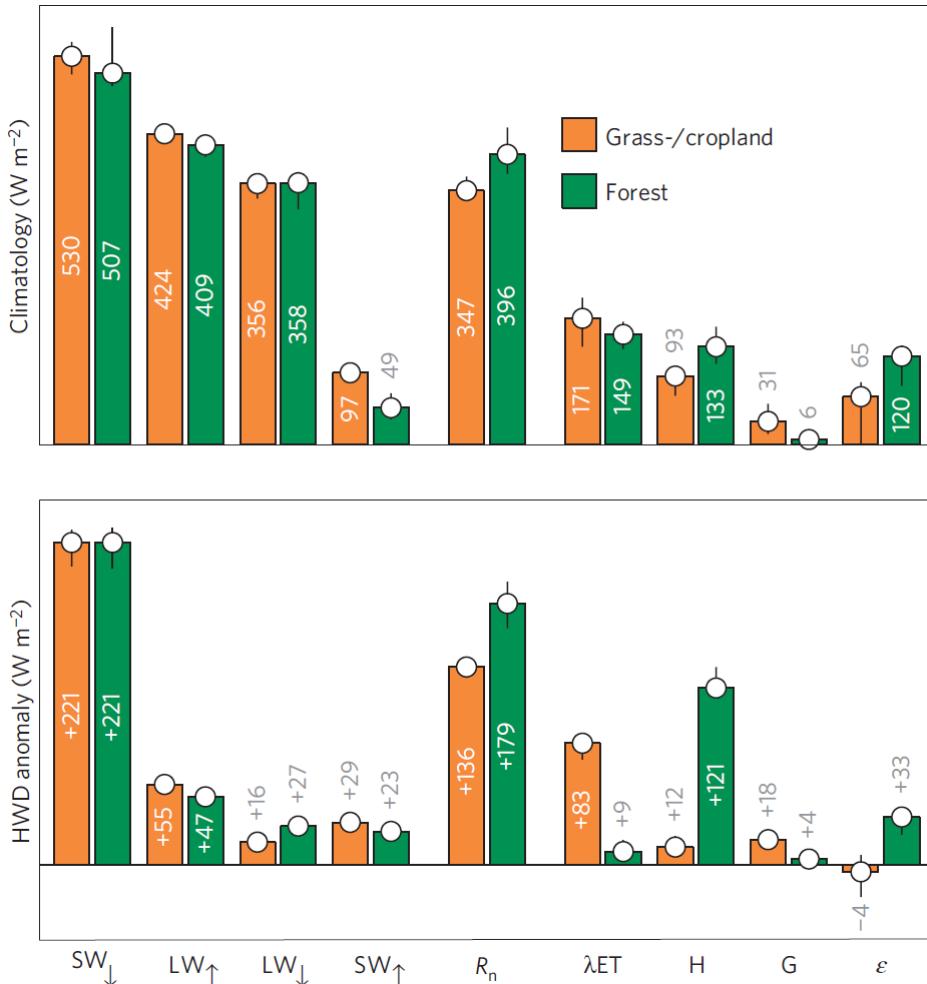
FIGURE 1.4: *Radiation and energy exchange over forest and grassland, for the climatology of summers in 1997–2008 (top), and the anomalies during heat waves in the same period (bottom). Fluxes are, incoming and outgoing short wave (SW) and long wave (LW) radiation, net radiation (R_n), latent (λET), sensible (H) and ground heat flux (G), and the residual (ϵ) (Image from [16]).*

The partitioning of net radiation on a regular summer day, is in favour of latent heat than of sensible heat for both vegetation types. Grassland though, displays a smaller Bowen ratio than forest; in other words, grassland translates a larger portion of energy into latent heat, compared to sensible heat.

During HWDs, there is a considerable anomaly in the amount of incoming solar radiation that both forest and grassland sites receive, because of the clear skies that usually accompany these events, as described previously. This extra energy is then partitioned in a totally different manner by forest and grassland (figure 1.4). Forest seems to balance out most of the additional energy through increased sensible heat release, while grassland uses the additional energy for water evaporation, increasing latent heat. This behaviour can potentially deplete soil moisture during long-lasting high temperatures, amplifying the effects of a heat wave in grassland sites, in contrast with forest sites. However during the observed HWDs, the net radiation anomaly over

forest was substantially larger than over grassland, and combined with the differences in Bowen ratio, the effect can add up to four times larger heating over forests, making them the warm spots during a heat wave. Whether or not this image can turn around due to soil moisture depletion, as models predict, is still to be investigated.

The transition from wet to dry soil conditions can be divided into three stages of drying, according to this study. In stage I drying, the evapotranspiration is independent of soil moisture, in stage II drying, evapotranspiration becomes self-limiting, and in stage III drying evapotranspiration becomes negligible. Figure 1.5 shows these stages as predicted by a conceptual model. Observations on longer lasting heat wave events are needed in order to evaluate these predictions.
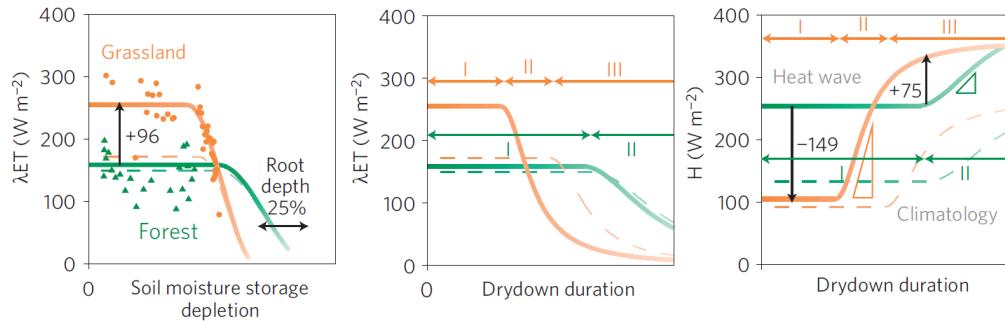
FIGURE 1.5: *Conceptual model of the evolution of fluxes over grassland and forest. In the right panel, the relation between latent heat and soil moisture. In the other panels, the temporal evolution of latent and sensible heat fluxes during the three stages of drying (Image from [16]).*

This study has been a guideline to the present thesis research, as it provided the motivation to compare the H-TESSEL modelled energy fluxes during heat waves, with the observed ones reported in it. For this reason, references on this study will be done, where applicable in the following.

1.4 Modelling Land-Atmosphere Interactions in GCM's

The land–surface component of general circulation and climate models, is usually referred to as the land surface scheme. The main purpose of it, is to model the energy balance, the water balance, the momentum exchange and –where applied– the chemical element exchange of the surface, and to propagate these information back to the atmospheric circulation part of the model. The energy balance of the surface is driven by the amount of net radiation:

$$R_n = SW \downarrow -SW \uparrow +LW \downarrow -LW \uparrow = SW \downarrow \cdot (1 - \alpha) + LW \downarrow -LW \uparrow, \qquad (1.1)$$

where SW and LW the short wave and long wave radiation (incoming and outgoing indicated by arrows), and α the surface albedo. The amount of net radiation a certain

place receives, must be balanced on the surface by the latent heat flux λET, sensible heat flux H, ground heat flux G and the chemical energy F stored by photosynthesis and released by respiration (the latter term is often omitted in climate models, as it accounts for only about 1% of absorbed insolation) [17]:

$$R_n = \lambda ET + H + G + F \tag{1.2}$$

Usually the partitioning of sensible heat and latent heat, in conjunction with atmospheric conditions, is used to calculate changes in the momentum exchange on the surface. A detailed study on feedbacks of the surface with the atmospheric boundary layer, can be found in [18].

The water cycle part in land surface schemes, is calculated by balancing out the individual terms of precipitation P, evapotranspiration E, surface/ground runoff R_s/R_g, ground water reservoir and interception reservoir change, again in connection with atmospheric conditions:

$$\frac{\partial W}{\partial t} = P - E - R_s - R_g, \tag{1.3}$$

with W being the water reservoir. In some climate models, terms of the carbon and nitrogen cycle are calculated as well in their surface schemes, to incorporate possible changes (like land cover) in the terrestrial biosphere within a changing climate.

It comes naturally that land surface schemes (LSS) evolved through time, in order to better represent the aforementioned processes. Still in this evolution, a certain amount of complexity distinguishes between different types –or generations– of models.

First generation models

The first generation land models (introduced by Manabe in 1969), used a simple energy balance equation, ignoring heat conduction into the soil. A globally constant soil depth and water-holding capacity were set, and evaporation was limited by soil water content below a threshold; if the soil moisture exceeded a prescribed limit, then further precipitation generated runoff. This parameterization of hydrology is commonly called the "Manabe bucket model" (figure 1.6).

The major conceptual limitation of first-generation schemes is the common simplification used to simulate evaporation:

$$\lambda ET = \beta \left(\frac{e^*(T_s) - e_r}{r_a} \right) \frac{\rho c_p}{\gamma}, \tag{1.4}$$

where $e^*(T_s)$ the saturated vapour pressure at surface temperature T_s, e_r the vapour pressure at a reference height, ρ the density of air, c_p its specific heat, γ the psychrometric constant, β a variable describing water availability, and r_a denoting the aerodynamic resistance. If the full complexity of the diurnal variability in λET is required from an
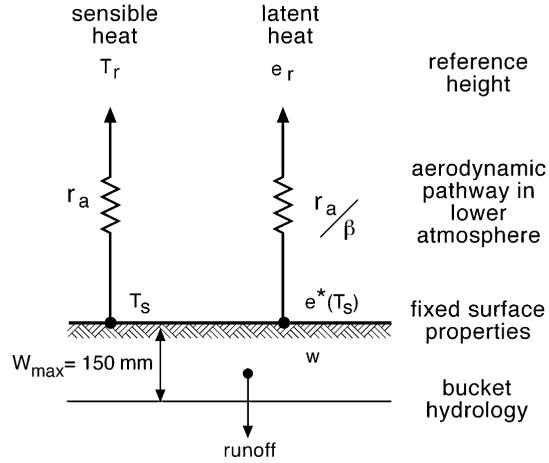
FIGURE 1.6: *Conceptual design of the first–generation LSS (Image from [17]).*

LSS, or if the function of plants is modelled explicitly, this equation fails to capture the complexity of the real system [19].

Second generation models

In the second generation land–surface models (introduced by Deardorff in 1978), a major improvement came with the simulation of soil temperature and moisture in two layers, while vegetation was perceived as a single bulk layer. The Penman-Monteith equation was used to calculate evapotranspiration, and land was perceived as partially covered by vegetation, still using a single energy balance equation. Other improvements were the introduction of an interception layer for precipitation and snow (i.e. by the leaves), the introduction of the "effective surface resistance" which simulates the resistance in transporting water vapour from the surface to air, and the "force-restore" method (by Bhumralkar in 1975 and Blackadar in 1976) which allowed for a prognostic equation of soil temperature and moisture.

Second generation models improved the representation of land–atmosphere interactions on the time scale of days, and the forecasting accuracy of precipitation [20] and winter–time surface temperatures [21], while emphasising the importance of land–surface processes in GCMs. Eventually tiled schemes were applied, with each individual grid in the model being comprised of different tiles to represent the various surface types. This will be further presented for the case of H-TESSEL in chapter 2.

Third generation models

The need to make land–schemes even more descriptive of surface processes, third generation land–surface models where developed after 1990. One of the major advances of these models, is their connection of the leaf stomatal conductance and carbon assimilation by
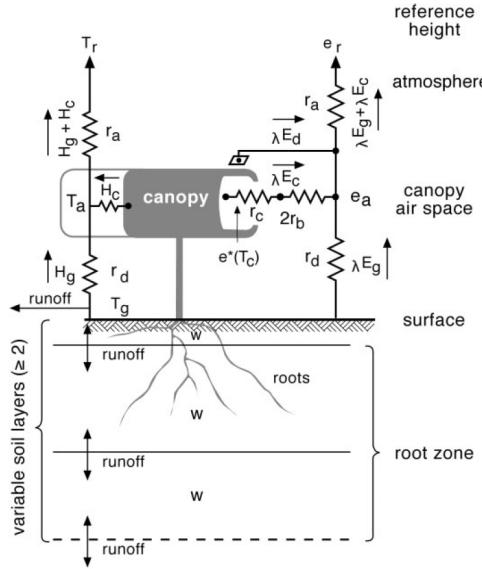
FIGURE 1.7: *Conceptual design of the second–generation LSS. Additional (Image from [17]).*

leaves. Conductance is estimated by semi-mechanistic models of leaf photosynthesis, based on how stomata are believed to function (evidence suggest that stomata try to maximize the efficiency of plant water use, and that there are limitations on leaf assimilation of carbon dioxide). These models use the same approach of second generation models to calculate certain variables (e.g. soil temperature, hydrology, runoff, etc.), while incorporating methods to model the carbon dioxide budget as well. This permits for dynamic vegetation climate models to emerge, which take into account changes in the global CO_2 levels, or changes in land use, thus making them able to simulate land–surface within a changing climate.
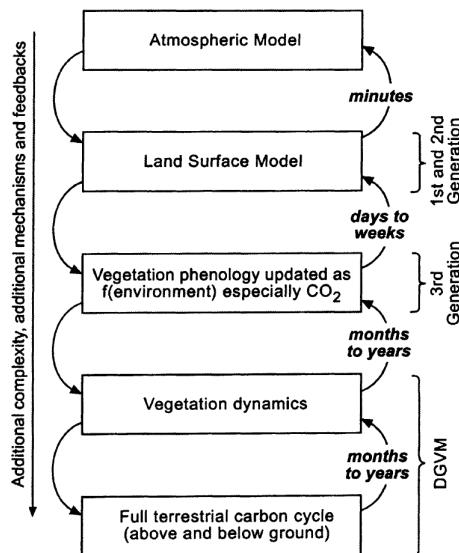
FIGURE 1.8: *Schematic of the increasing levels of detail being added into surface modelling approaches. Each level of models is introducing new aspects, while still depending on the ones that previous models already established (Image from [17]).*

Chapter 2

Materials and methods

The basic materials and methods used in this study are going to be presented next. An outline of the H-TESSEL land–surface scheme and its functionality, is considered essential at this point, as it is the core element under examination. Further on, the meteorological data that comprised the forcing for the offline run of H-TESSEL, the output data that came from the runs, as well as additional datasets used throughout this research, are going to be presented, together with the ways of transforming and manipulating them to facilitate the analysis. It should be noted here that the simulation output data were already available, as the offline runs were not part of the present study. Finally, classification and regression trees and artificial neural networks, were used as a complementary approach in the analysis, and a small introduction covering these processes is presented.

2.1 The H-TESSEL scheme

The H-TESSEL scheme belongs theoretically in the second generation of land-surface schemes, discussed in section 1.4. It is currently used in the operational model of ECMWF, for the description of processes in soil, vegetation, and snow covering the continents. The need for an accurate simulation of surface water and energy fluxes in global circulation models has already been stressed, and H-TESSEL is devised to cover that part. It is a revised version of the former TESSEL scheme, with an additional attention on hydrology; in particular surface runoff, infiltration, and soil texture differentiation. This outline will be focused on the parts of H-TESSEL that are more relevant to this study, and summer conditions (cite for more details cf. Balsamo 2009 and Hurk 2000).

A tiled approach

Numerical spatial models in earth sciences operate by dividing the area under consideration into a raster or array, and calculating the individual terms within each gridbox. In

the case of a land surface scheme, the gridbox size is usually quite large to assume a certain uniformity of land cover type. In reality forests, crop-lands, bare ground and open water, while partitioning energy in a different way, can all exist within the same gridbox, making it difficult to accurately model the surface. For this reason in H-TESSEL each surface type is represented by a different fraction or tile, within each gridbox. There are eight tiles in total, six over land (bare ground, low and high vegetation, intercepted water, shaded and exposed snow), and two over water (open and frozen water) (figure 2.1).
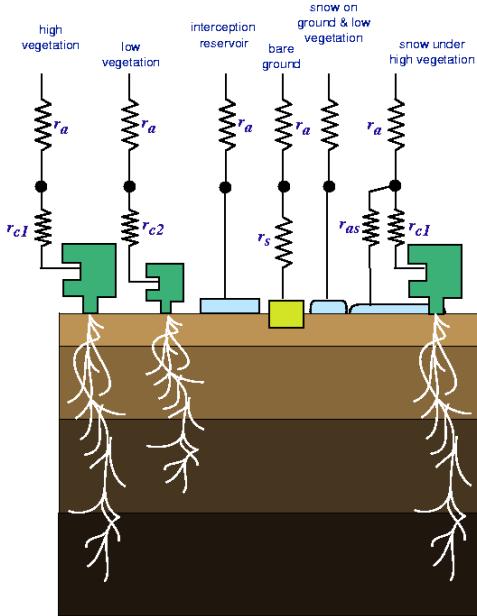
FIGURE 2.1: *Schematic representation of the structure for the six tiles over land in H-TESSEL [22].*

In this manner, each tile's energy fluxes can be treated according to the properties of the surface it represents. The total fluxes in the gridbox, are then calculated by averaging the fluxes of all tiles, weighted over the surface proportion of each tile. Surface albedo α is similar for all tiles except for snow, and long-wave emissivity ε is set to 0.99 for all the tiles. A skin layer, represents a buffering layer of air, forming the interface between the soil and the atmosphere, with temperature T_{sk}, while a skin conductivity Λ_{sk}, acts as the parameter for thermal connection to the soil. Especially for high vegetation, the height of tree trunks in a canopy create a layer that is large enough to be considered stably or unstably stratified. For this reason Λ_{sk} differs for high vegetation, to represent either an effective convective transport for unstable conditions, or a limited turbulent exchange for stable conditions.

Vertically, the soil is divided into four layers, which can be covered by a layer of snow. The depths of these layers increase almost in a geometric relation from the surface downwards, reaching a total of about 2.9 meters [22]. The energy equation is solved

following a Fourier diffusion law (taking account of water freezing/melting), with a net ground heat flux as the top boundary condition and a zero flux at the bottom. Regarding water balance, when the interception layer is saturated, the remaining precipitation water is partitioned between surface runoff and infiltration. The boundary conditions are infiltration plus surface evaporation at the top, while free drainage is taking place at the bottom, with each layer having its own sink of water in the form of root extraction (for the tiles that contain vegetation). The water removed in this way, is weighted over a function that includes root density and has a threshold for water levels bellow the wilting point.

The two tiles for high and low vegetation are characterised (apart from the area fraction), by a series of parameters [4]:

- A minimum canopy resistance $r_{s,min}$

- A vegetation coverage c_{veg}

- A leaf area index (LAI)

- A sensitivity coefficient g_d for the dependence of canopy resistance r_c on water vapour pressure deficit

- The root distribution over soil layers, specified by an exponential profile

- A vegetation roughness length z_{0v}, extracted from a global climate database

Energy balance

As mentioned before, the energy balance for the surface is solved for each of the eight tiles, assuming full coverage of that tile. Skin temperature T_{sk}, is also calculated for each tile in a gridbox. The results from each tile, can then be passed on (after weighting) to the atmospheric part of the model, making up the gridbox's energy balance. The input parameters are the same for all tiles, and consist of downward short-wave and long-wave radiation (R_s and R_T), reference atmospheric temperature (T_a, usually taken at 2 meters from the surface), specific humidity (q_a), and wind speed (U_a). The energy balance equation is then:

$$(1 - \alpha_i)(1 - f_{Rs,i})R_s + \varepsilon(R_T - \sigma T_{sk,i}^4) + H_i + L_{v,s}E_i = \Lambda_{sk,i}(T_{sk,i} - T_1), \qquad (2.1)$$

where i denotes the tile index, and $1 - f_{Rs,i}$ constitutes a partial absorption of short-wave radiation by the skin layer. T_1 is the temperature of the upper soil layer, σ is the Stefan-Bolzman constant, H_i is the sensible heat flux, L_v the latent heat of evaporation, and $L_{v,s}E_i$ the latent heat flux. Sensible and latent heat fluxes, for the high and low

vegetation tiles, are calculated through:

$$H_i = \rho_a c_p \mid U_a \mid C_{H,i}(T_a + gz - T_{sk,i}), \tag{2.2}$$

$$L_v E_i = \frac{L_v \rho_a}{r_c + r_a}(q_a - q_{sat}(T_{sk,i})) \tag{2.3}$$

Here, $r_a \equiv (\mid U_a \mid C_{H,i})^{-1}$, ρ_a is the air density, c_p the heat capacity of dry air, g the acceleration of gravity, z height of the lowest model level above the surface, $q_{sat}(T_{sk,i})$ the saturation humidity at temperature $T_{sk,i}$, and $C_{H,i}$ the turbulent exchange coefficient. The resistance r_c, is a function of downward short-wave radiation, leaf area index, average soil moisture content, atmospheric water vapour deficit and $r_{s,min}$.

The coupling to the ground is working in a similar manner; the net flux into the soil is a weighted average of fluxes from each tile, and it drives the temperature changes within the soil. The heat diffusion to the ground is controlled by the right hand side of equation 2.1, while the residual of net short-wave radiation that is not absorbed by the skin layer, is also going into the soil. The net heat flux into the soil for each grid (neglecting snow tiles), can then be expressed as:

$$G_{grid} = \sum_i C_i \{\Lambda_{sk,i}(T_{sk,i} - T_1) + f_{Rs,i}(1 - \alpha_i)R_s\}, \tag{2.4}$$

where the summation is done over all tiles (except snow), with C_i the coverage of each tile.

Finally, soil moisture transport is controlled by infiltration, surface runoff, gradient diffusion, gravity and deep water percolation. Plant transpiration is extracting water from the soil layers using a weighing function that depends on the root density. Moreover a threshold is set, so that extraction of water by plants stops, if soil moisture is bellow the wilting point of the specific vegetation type.

2.2   The Datasets and general setup

Princeton forcing dataset

This study made use of already available data from an 11-year offline run of the H-TESSEL scheme, for the years 1996–2006. The run was forced by a global meteorological forcing dataset, put together by the land surface hydrology group of Princeton University [23]. This dataset provides near-surface meteorological data for driving land surface models and other terrestrial modelling systems. It blends reanalysis data with observations and disaggregates in time and space. It was available at 1.0 degree, 3-hourly resolution globally for 1948–2008, and was constructed by combining a suite of global

observation-based datasets with the NCEP/NCAR reanalysis. For the purposes of this
simulation values from The forcing variables are:

- Atmospheric humidity

- Incoming long-wave radiation

- Incoming short-wave radiation

- Precipitation amount

- Surface air temperature (at 2 meters height)

- Surface pressure

- Surface wind

The variables that were available for the analysis in this report, was surface air temper-
ature (T2), atmospheric humidity (Q) and short-wave radiation (SW). The dataset is
going to be referred to, as 'princeton' or 'forcing' in the following.

The only change applied on this dataset, was a transposition in time. The incoming
short-wave radiation curve, can be expressed theoretically in a time-radiation plot as the
positive part of a sinusoidal function, with a maximum around noon (due to maximum
insolation), and the negative part set to zero (expressing night-time insolation). The
dataset has a 3-hourly resolution, or 8 values per day, the first being at midnight and the
last one being at 21.00 hours. These values, are in fact the average of each 3-hour interval
the day is divided in. So, the first value does not represent the beginning of an interval,
but rather the middle of it which is 1.5 hours later. For this reason, all values in the
dataset were transposed by +1.5 hours. Although this is a minor adjustment, it positions
curves in the correct place with a maximum around noon, helping the visualisation
process of daily values (figure 2.2).

Next, the area under consideration (i.e. Europe) was distinguished from the global
dataset. This resulted in a 50×50 gridbox raster, ranging in latitude from $30°N$ to
$80°N$, and in longitude from $10°W$ to $40°E$. The selection was made to match the
raster of the output values, which was already ranging over the aforementioned area.
Moreover, values of gridboxes representing sea were masked-out in order to have a better
overview when plotting.

The climatology and anomalies were then calculated for each gridbox. Climatology
refers to the period in hand (i.e. 1 January 1996 – 31 December 2006), and was deduced
by averaging every 3-hourly value with its equivalents in the same day and month,
over the 11 year period. Anomalies were calculated by subtracting every value from
its equivalent climatology in the same day and month. These newly created values are
going to be referred to as 'princeton climatology' and 'princeton anomalies' hereafter.
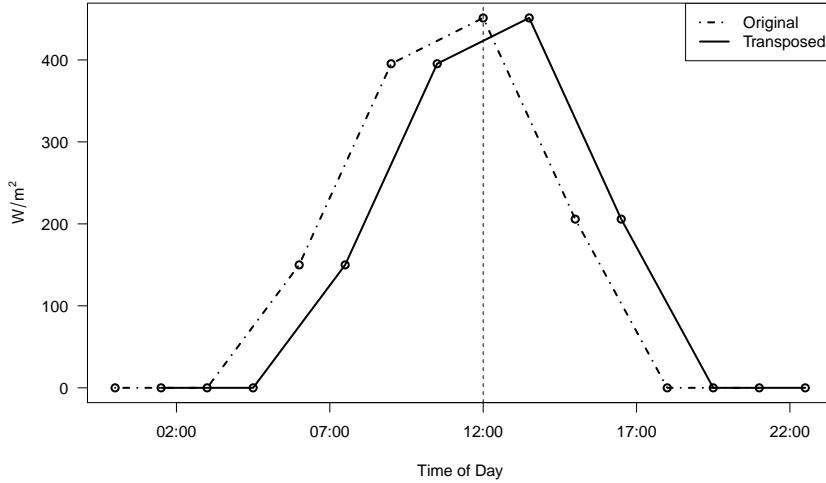
FIGURE 2.2: *Example of the incoming short-wave radiation transposition by +1.5 hours. Here only one daily cycle is shown; the same was done for air temperature and specific humidity.*

Finally, for facilitating the process of HWD selection, the maximum temperature (T2max) and maximum short-wave radiation (SWmax) were deduced from the 8 values of each day, for each single gridbox. The corresponding climatology and anomalies were then calculated for the 11-year period. The same procedure was done for calculating the mean diurnal values of temperature, short-wave radiation and specific humidity (T2mdv, SWmdv and Qmdv). Here, mean diurnal value refers to the mean of the interval between 09.00 hours and 15.00 hours for each day.

Output from H-TESSEL

Output data from the H-TESSEL offline run, were variables available in two temporal resolutions for the period of 1996–2006, namely daily and 1/2 hourly resolution, and a spatial resolution of 1 degree, with some of the variables being available in tiled format (table 2.1). As already discussed, the focus of the analysis will be on the two tiles of high and low vegetation, which may also be referred to in the following, as forest and grassland respectively. The variables though, are not sufficient for closing the surface energy balance for the tiles. Nonetheless, all the additional variables required can be extracted from the existing ones.

The incoming and outgoing short-wave and long-wave radiation were not explicitly analysed. Instead the net radiation was used in all cases, as it provides both for energy balance closure and ease of comparison with the observed one, shown in section 1.3. The incoming long-wave radiation (LWdown) for each gridbox, was simply deduced by subtracting LWup from LWnet. At this point, net radiation for forest and grassland was

TABLE 2.1: *Available variables from the H-TESSEL offline run output.*

Abbreviation	Variable	Temporal–Spatial Resolution
SManom	Soil Moisture Anomaly	
AvgSurfT	Average Surface Temperature	
CanopInt	Canopy Interception	Daily–Gridbox (1°)
SoilTemp	Soil Temperature	
SoilMoist	Soil Moisture	
—	Ice/snow variables	
SWnet	Net Short-wave Radiation	
LWnet	Net Long-wave Radiation	
Qle	Latent Heat Flux	1/2 Hourly–Gridbox (1°)
Qh	Sensible Heat Flux	
Qg	Ground Heat Flux	
LWup	Up-going Long-wave Radiation	
TileFrac	Tile Fraction	
Tskin	Skin Temperature	1/2 Hourly–Each Tile
Qle	Latent Heat Flux	
Qh	Sensible Heat Flux	

calculated by:

$$Rnet_f = SWnet + LWdown - \varepsilon \cdot \sigma \cdot Tskin_f^4,$$
$$Rnet_g = SWnet + LWdown - \varepsilon \cdot \sigma \cdot Tskin_g^4,$$

(2.5)

since SWnet and LWdown are the same for all tiles. Values for this tiled Rnet, were further confirmed by deducing the grid Rnet, weighing $Rnet_i$ (with i the tile) over each tile fraction and comparing to the expected Rnet. With $Rnet_f$ and $Rnet_g$ available, the only variable needed to close the forest and grassland energy balance is ground flux, which was calculated through:

$$G_i = Rnet_i - Qle_i - Qh_i$$

(2.6)

Having all the energy fluxes needed, the same procedure was followed, as with the princeton dataset. Climatology and anomalies were calculated in the same manner, the only difference being 48 values per day for each variable, instead of 8 for princeton (because of the 1/2 hourly resolution). In this case as mentioned before, Europe was already defined, and the sea-gridboxes were already masked-out. Any further minor treatment done on the output data will be mentioned, when needed, in the results (chapter 4).

Finally, daily values of soil moisture were available for each of the 4 levels. The rooting system of vegetation is able to extract water from all 4 layers for forest, and the top 3 layers for grassland, consequently, soil moisture was distinguished for high and low vegetation, by averaging the top 4 and 3 layers respectively. Since mostly anomalies are used for soil moisture in the following, the simplification of direct averaging is considered

safe (the potential of extracting water is not equal for each layer). The same procedure
as before was used for making climatology and anomalies.

The ECA E-Obs dataset

The E-Obs dataset was used in order to facilitate HWD selection in Europe (cf. chapter
3). ECA provides datasets derived from observational station data across Europe (figure
2.3), which are available in two formats; raw station data (after a series of tests performed
to identify obvious problems and remove suspicious values), and grided data [24]. The
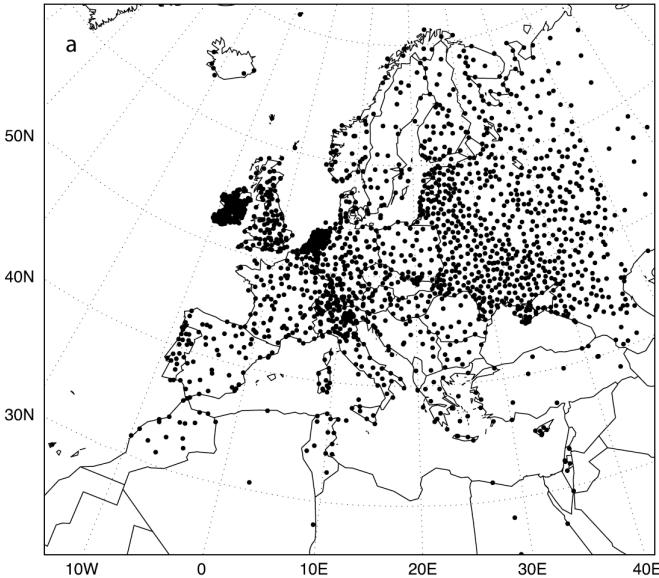
FIGURE 2.3: *Meteorological observation stations around Europe (as dots), used for
compiling the ECA E-Obs version 1.0 dataset in 2008. Station density has increased in
later versions (Image from [24]).*

grided format was constructed using a three step interpolation method, for minimising
the effect of data inhomogeneity. In the final product, the dataset is available in regular
grids of 0.25 and 0.50 degrees resolution, covering the area $25°N - 75°N \times 40°W -
75°E$. Variables included are daily mean temperature, daily minimum temperature,
daily maximum temperature, daily precipitation and daily averaged sea level pressure.

Here, the dataset for daily maximum temperature was used, in 0.25 degrees reso-
lution, and was truncated to fit the same spatial extent as the output data (except for
the northern limit at 75 instead of $80°N$). This choice was made because the available
raw data from stations, are temporally sparse in a lot of cases in the period of 1996–
2006, leaving only 174 out of around 5000 stations with fully usable data. Since these
gaps were covered by interpolation and statistical methods by the grided dataset, it was
considered more appropriate to use the latter.

The raster was then converted to 1 degree resolution, by averaging the 16 gridboxes
of 0.25 degrees (4×4) that correspond to each gridbox of 1 degree. In figure 2.4 it

is clear that although the conversion is not adequate for capturing small spatial scale variations, in the case of heat-waves where the temperature varies over a larger spatial extent, the conversion seems to give satisfactory results.

Regarding the temporal extent of the E-Obs dataset, the same period was used as the one of the output data (i.e. 1996–2006). Climatology was again calculated by averaging the maximum temperature of each day over the same day in the 11-year period, and anomalies were deduced by subtracting climatology from each day's value.
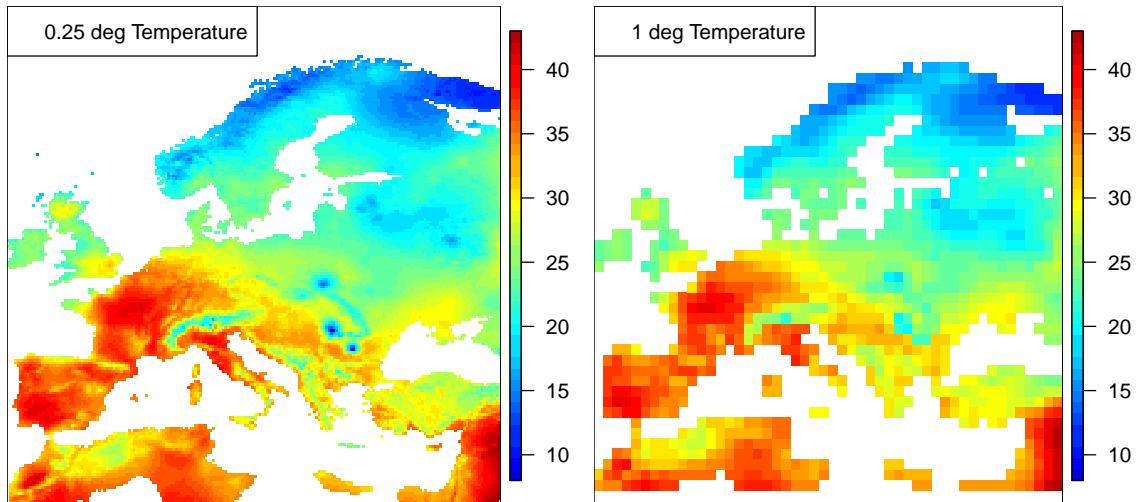
FIGURE 2.4: *Example of maximum daily temperature in $^\circ C$ for the 7th of August 2003, for the E-Obs 0.25 degree resolution and the converted 1 degree resolution.*

2.3 Classification and regression trees

In the case of a large dataset with interacting variables, it is often difficult to draw connections between them, by simple linear regression. CART analysis is a non-parametric decision tree learning technique, that produces either classification or regression trees, depending on whether the dependent variable is categorical or numeric, respectively.

The procedure of making these decision trees, is by forming a collection of rules based on the dataset in question. These rules are formed by statistical algorithms that try to find the best split, in order to differentiate certain dependent variables' values in correlation to independent variables. Once a rule splits a node into two, the same process is applied to each 'child' node, until no further gain can be made, or some pre-set stopping rules are met. If the resulting tree is very large to have a good overview, a procedure called pruning is applied, that reduces the size to a more comprehensible one.

In this way, starting from the main 'trunk' of the tree, one can move foreword from branch to branch, making a decision at each node, which depends on the value of the specific dependent variable of the node. The ending nodes or 'leaves' represent either a categorical decision in the case of categorical variables, or the expected value of the

variable in question that corresponds to the specific route taken, in the case of numeric variables (usually the mean of all possible values of the variable that fall into that route).

CART is a method with a wide range of applications in medicine, bioinformatics, climate science, policy and decision making, and wherever large datasets need to be analysed. An example of CART application in medicine and hospital decision making, is shown in figure 2.5, while application of CARTs in climate science can be found in numerous articles (e.g. [25], [26]).
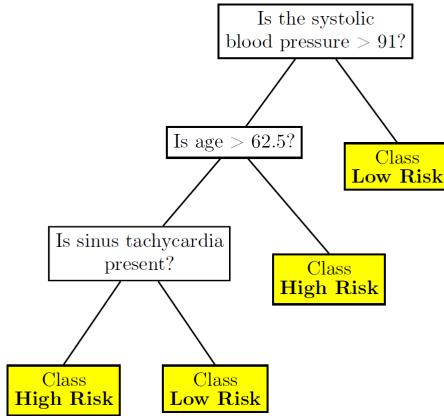
FIGURE 2.5: *A simple classification tree, used in the San Diego Medical Centre, USA, for assessing the risk level of patients. Left branches in each node indicate a positive answer to the node question, and yellow boxes indicate the terminal nodes, or 'leaves'.*

The process of building a CART algorithm from scratch is an elaborate one, so several already made materials exist to apply decision trees in datasets. Here, the 'tree' package of the R-cran statistical computing environment was used [27], which will be further discussed in chapter 4. A very helpful introduction on the use of CART in conjunction with R can be found in [28].

2.4   Artificial neural networks

Artificial neural networks are a branch of machine learning, which falls within the field of artificial intelligence. Inspired by the structure and functional aspects of biological neural networks, ANNs (or NNs) are non-linear statistical data modelling tools. Computations are structured in terms of an interconnected group of artificial neurons. A neuron can be imagined as a computation entity (figure 2.6) that receives values into one or more inputs and provides one or more output values, defined by an activation (or transfer) function. The weights applied in each input, can be adjusted in order to receive the desired output value. They are the basic feature that makes the learning process possible for a NN.

Several neurons combined make a neural network, which may have several input, output and intermediate (or hidden) neurons, and they come in a variety of different levels of complexity and functional aspects. The NN architecture chosen for this study,
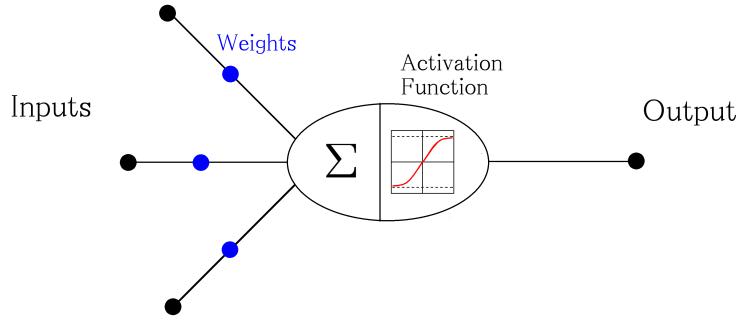
FIGURE 2.6: *Example of an artificial neuron with three inputs, the values of which after weighing, are added by the summing function and then passed on to a sigmoid activation function, giving the value of the output. Activation functions in most applications are either linear, threshold-like, or sigmoid.*

is one of the most common ones; a layered feed-forward network with a back-propagating supervised-learning algorithm, so this type will be further explained.

A basic form of such a network can be seen in figure 2.7. Values are fed into the input layer of neurons, which have outputs to communicate with each neuron of the next layer. In turn the hidden layer neurons, receive these outputs as weighted inputs, and propagate the information forward, until the output layer is reached, where the output value can be obtained.
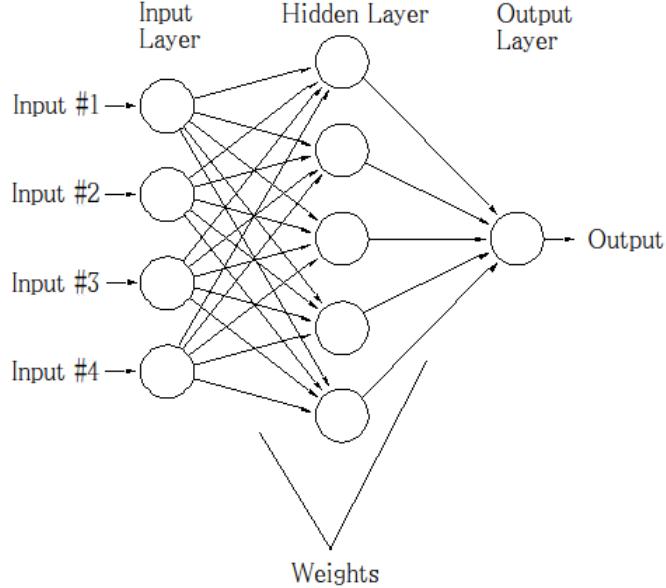
FIGURE 2.7: *Example of a neural network with four inputs, one hidden layer of 5 neurons, and a single output.*

Coming to the application of such a method in a physical system, if in a dataset some variables are known to be dependent on others, but the exact connections are either not know, or non-linear, a neural network can be trained to model any statistical relations that might exist. The procedure is as follows. First the variables that are

considered as dependent, are set as the output of a NN, and the variables they are depending on, are set as the input. After defining the number of hidden layers and the number of neurons within each hidden layer [1], the neural network is trained with a back-propagating algorithm. This involves providing the NN with a large enough dataset of input variables and expected outputs [2]. The NN then goes through the training process in steps, by using one set of inputs and adjusting the values of weights in each step, in order to produce the expected outputs. If the supervised learning is performed correctly, the neural network can then receive input values that were not used in the training procedure, and produce the correct output. Due to the nature of this method, it is clear that the selection of dependent and non-dependent variables is crucial, as the NN will learn and find interconnections even if variables have no relevance. In this case though, the results will be disappointing.

In general, neural networks of this type work at their best when input values are in a range between 0 and 1, or -1 and 1, so in this case the range of flux values had to be normalised for better performance. Output values also come out in the same range, and the training process (where output has to be prescribed in normalised intervals) forms the convention by which non-training-mode values can be converted back to regular ones.

Neural networks have multiple applications in weather forecasting, climate and environmental sciences. Several such studies have emerged lately (e.g. [30], [31], [32]), showing interesting results. In the present research the 'neuralnet' algorithm of R-cran statistical computing environment was used, which will be seen in chapter 4.

---

[1]This can be done either by applying relevant methods developed in A.I. science, or by a trial-error approach in which optimisation is achieved by minimising the error of the network. In the second approach, some rules of thumb can be found on the optimal size of a network, for a given number of inputs and outputs [29].

[2]The larger the training dataset, the larger the statistical relevance and accuracy that can be achieved.

Chapter 3

Selection of Heat-Wave Days

In order to make an analysis of energy fluxes during heat-wave days, a subset of the output dataset must be sorted out that consists of fluxes corresponding to those days. As discussed in section 1.1 there is not a universal definition of HWDs, a fact that makes the selection not a straight forward procedure. In the following, the process of making this subset is going to be presented, preluded by a demonstration of the options available and the reasoning of the final choices made.

3.1 The selection options

The period in which heat-waves are expected, must be defined to begin with, so in this study only the spring and summer periods are taken into account, neglecting days that might fit HWD criteria in autumn and winter-time as non-relevant. Then, the criterion for selecting HWDs together with the dataset on which to apply it, are needed. As the study area is covering the whole Europe, the criterion should be as general as possible, and for that reason the one of WMO was selected as the most adequate (cf. section 1.1).

The forcing dataset of the model run, containing values of atmospheric temperature at 2 meters from the surface could not be used for the selection. The main reason lies in the way it is produced (being a merged product of reanalysis data and observations), which muffles the exact observations, making temperature bursts (or peak values) less distinguishable. An attempt to characterise HWDs on this dataset applying the WMO criterion, provided poor results; the percentage of days in the 11-year period where the temperature exceeds the climatology by more than 5 degrees, was up to 30% in some areas of central urope –or an average of 60 days in every spring-summer cycle, which is considered an unacceptable overestimation. A second approach by incorporating a conjoined highest temperature – highest short-wave radiation criterion on the forcing dataset proved fruitless, as it inserted an additional uncertainty; no reliable way could be found for assigning thresholds for both temperature and SW radiation that would guarantee proper HWD selection and compatibility with the study in [16].

Since the forcing dataset is made from reanalysis and observational data, any dataset that provides pure observations of maximum daily temperature with an adequate coverage and spatial resolution for the same period, would be compatible for the selection process. The E-Obs dataset proved to be a good candidate for this purpose.

3.2   Using the E-Obs dataset

Although a sorting of HWDs can now be applied through the E-Obs dataset, the days selected as heat-waves must be also perceived as so by the H-TESSEL scheme through its forcing. This is crucial because the analysis of H-TESSEL energy fluxes would have no significance in this study, if performed for days that do not correspond to the atmospheric conditions of heat-waves. For this reason the WMO criterion was applied in three steps, described in the following.

First, by calculating the number of days in E-Obs where the maximum temperature anomaly exceeds that of climatology by more than 5 degrees, a percentage [1] of 'very warm days' is determined (figure 3.1).

In the second step this percentage was used on the forcing dataset, to distinguish the highest temperatures in the same period. The lowest one in this set, is the threshold of temperature over which a day might be considered a HWD for each gridbox [2]. In figure 3.2 this highest temperature threshold is shown for Europe. It is overall a low threshold compared to what would be expected for each area, because the calculation was based on the maximum temperature of the princeton dataset. The maximum temperature for each day in princeton is usually lower than the true one, due to the dataset's low temporal resolution.

Under the condition that the maximum temperature of a day, is both above this threshold and part of a series of at least 5 consecutive days that do so as well, it is classified as a HWD. The latter forms the third step in the application of the WMO criterion.

Once the selection of heat-wave days is complete, it is possible to proceed with the calculation of energy fluxes for the HWD conditions. In figure 3.3, the final percentage of HWDs calculated by the above process is shown. Exceptionally high percentages of heat-wave days, found in northern Scandinavia and Russia, are probably related to warm summer periods when the ice coverage is highly reduced compared to the normal, leading to high close-surface temperatures. An example of the selected HWDs for the spring-summer of 2003, is shown in figure 3.4, for a gridbox in central France.

---

[1] Number of very warm days, over total number of days in the spring-summer of the 11-year period.

[2] Highest temperatures where used instead of highest temperature anomalies, as the latter produced on average 1 or 2 5-day-lasting HW events during the 11 years, and quite often in spring. Although the highest temperature anomalies would be the optimal choice, it was discarded in exchange of better statistics and summer positioning of HWDs
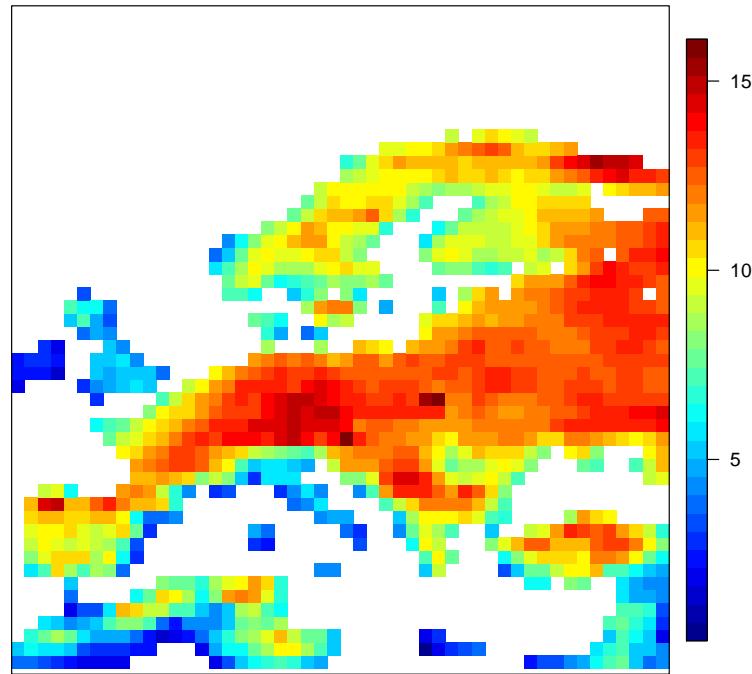
FIGURE 3.1: *Percentage of 'very warm days' in Europe for the period of March–August in 1996–2006, deduced by the E-Obs dataset.*
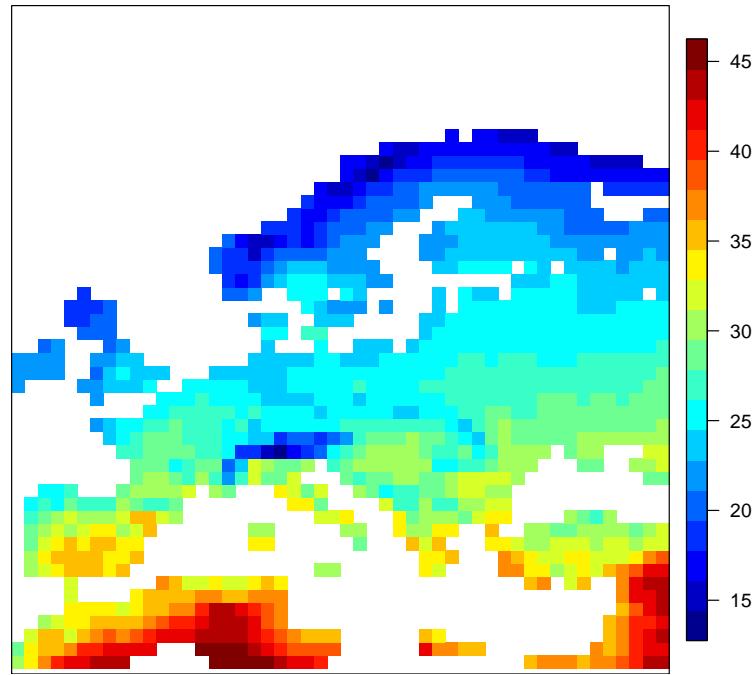
FIGURE 3.2: *The temperature threshold in °C, from the H-TESSEL forcing. Days with maximum temperature above this limit, are qualified as 'very warm days'.*
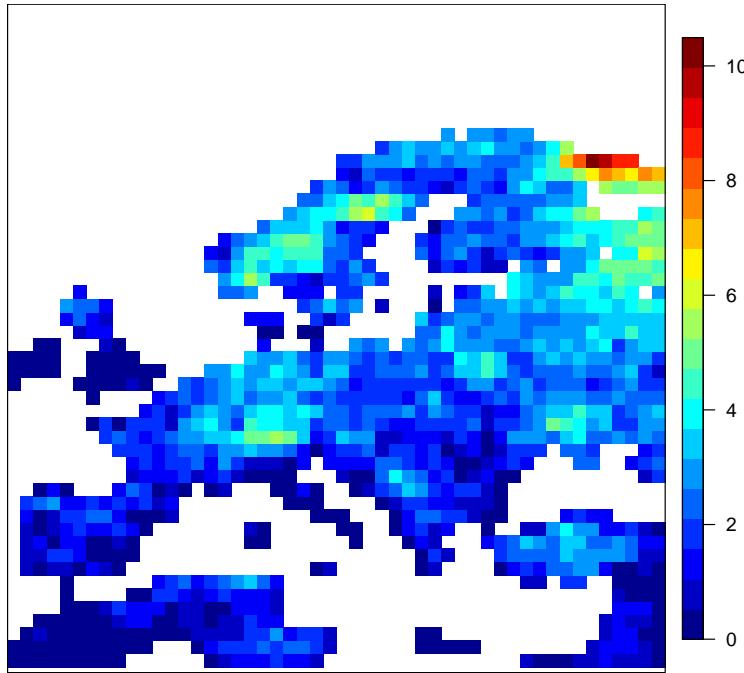
FIGURE 3.3: *Final percentage of HWDs in H-TESSEL for the period of March–August in 1996–2006, after all the criteria have been applied.*
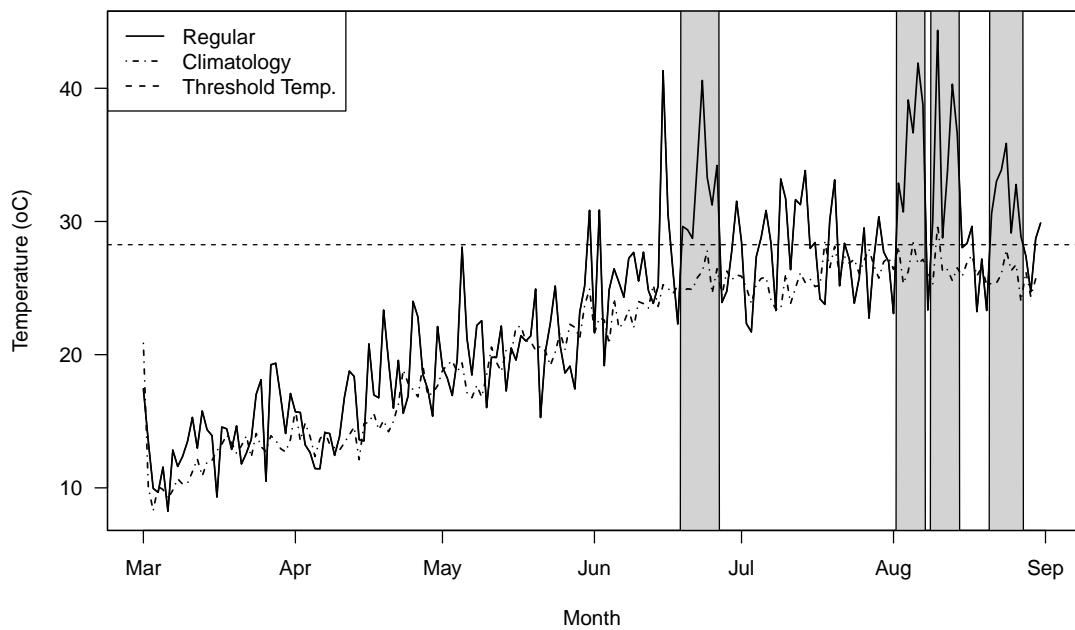
FIGURE 3.4: *Maximum temperature and climatology in the period March-August 2003, for a gridbox in central France. The periods (shaded with gray) where the temperature is above the gridbox's threshold for more than 5 consecutive days, are classified as HWDs.*

Chapter 4

Results and discussion

In this chapter, the results of the analysis are presented. Starting with the presentation of the study area, the bulk energy fluxes during HWDs and normal conditions are shown and compared to the ones calculated in [16]. The role of soil moisture (being either at ample or depleted levels), and the effect it has on energy fluxes is then discussed, followed by a case study, which demonstrates the need for discrimination of heat-wave duration, which leads to the analysis of the evolution of fluxes during heat-waves, in the form of time series of mean diurnal values, and for different groups of heat-wave duration. The chapter finishes with a CART analysis that demonstrates the dependence of latent and sensible heat, on specific humidity and soil moisture.

The calculations done on energy fluxes in the following, are spatial averages confined only in a certain area of Europe, in order to reduce the spatial variability of forcings and consequent outputs. The area was selected on certain criteria; the coverage of grassland and forest should be in a rather equal proportion, and it would be an advantage if the area included parts of Europe where heat-wave events are known to have occurred. For this reason the area shown in figure 4.1 was chosen. In addition, almost all the FLUXNET towers used for the observational study, fall within this area, which makes the comparison more relevant.

Major events, like the 2003 heat-wave in central-south France and the 2006 heat-wave in central-north Europe are captured. The few gridboxes over the alpine region in Switzerland were excluded from the averaging, as they contained tiles with high percentage of a non-relevant surface type (i.e. snow), and different atmospheric conditions which would bias the results.

The procedure of spatial averaging comes last in any other averaging performed. This means that energy fluxes for the climatology and HWDs are first averaged for each gridbox, and then averaged over the area, which will be referred to as west-central Europe in the following (WCEurope).
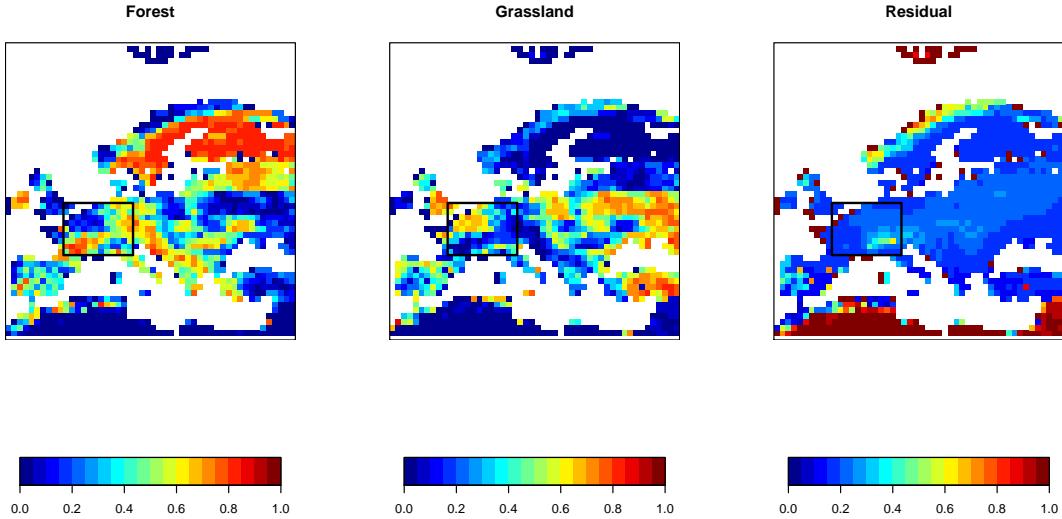
FIGURE 4.1: *Average fraction of tile coverage for forest, grassland and the residual tiles over the 11-year period. The black rectangular marks the selected area.*

4.1 The energy fluxes

The bulk energy fluxes are calculated as HWDs anomaly and climatology in WCEurope, from mean diurnal values of fluxes. Results are shown in figure 4.2. The first noticeable feature is the almost equal amount of net radiation for forest and grassland, in both climatology and HWDs. This comes from the similar albedo of the two surface types in the model, and is most probably the main driver of absolute value differentiation, from the observational results in [16]. In (1.4), due to the lower albedo forest has a higher net radiation input than grassland. Also, ground fluxes in all cases seem to be overestimated in the model for both forest and grassland, compared to observations.

In the climatology, the relative values of fluxes between forest and grassland seem to be in accordance with observations, with grassland fluxes being somewhat overestimated in that respect, except for sensible heat which is underestimated. The higher net radiation that grassland receives relatively to observations has probably a share on that, but the exception of sensible heat mentioned before, suggests that there might also be a differentiation in the partitioning of fluxes, between the model and observations.

Net radiation anomalies are almost equal for forest and grassland, while relative values of latent and sensible heat anomalies between forest and grassland are in accordance with the observed ones. An other noticeable feature is that ground flux anomaly is greater for forest than for grassland (opposed to observations), and that sensible heat flux anomaly for grassland is negative. Grassland compensates all the extra net radiation it receives, through latent heat.

The hypothesis that the above behaviour was partly due to the existence of HWDs with very low net radiation anomaly, was disproved. A differentiation between HWDs
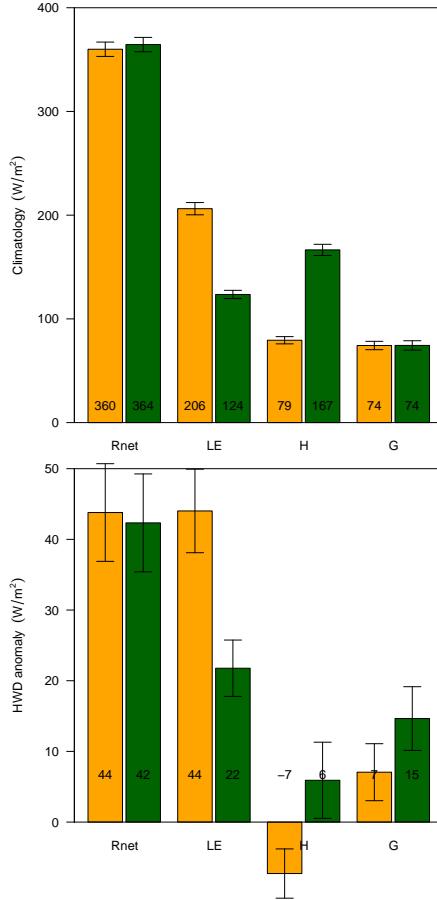
FIGURE 4.2: *Climatology and HWDs anomaly of net radiation, latent, sensible and ground heat fluxes for forest (green) and grassland (orange), for the period March-August 1996–2006, together with confidence intervals of spatial averaging. An average of 62 HWDs exist in this period for each gridbox.*

with high and low net radiation anomaly, although influencing the absolute values, did not show significant differences from the flux partitioning presented above. This is partly due to the fact that high net radiation anomalies alone, are not enough to produce effects that resemble heat-waves in short-term periods, if other conditions are not simultaneously present (like soil moisture depletion). The partitioning of net radiation among latent and sensible heat in normal conditions, shows a clear preference towards latent heat for grassland, and a slight preference towards sensible heat for forest (figure 4.3). In heat-wave conditions grassland tends on average to increase latent heat and slightly decrease sensible heat, and forest –although increasing sensible heat in half of the cases– it tends on average to partition more of the extra energy on latent than sensible heat.

Table 4.1 shows a comparison of these results with the observations (i.e. combining figures 4.2 and 1.4). Summarising the findings for H-TESSEL:

1. Latent heat in the climatology, is overestimated for grassland and underestimated for forest –the opposite holds for sensible heat. However, the differences are within
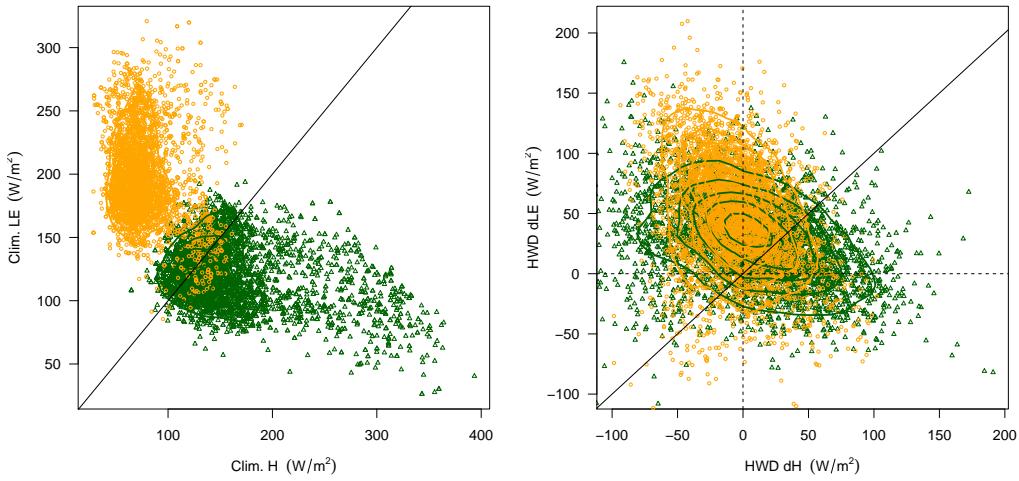
FIGURE 4.3: *Partitioning of energy between latent and sensible flux for climatology and HWD anomalies, with density contours, and one to one line. Each point represents mean diurnal values for one HWD in one gridbox (forest with green and grassland with orange).*

a reasonable level of $\sim 10\%$.

2. Ground flux in the climatology, is overestimated for both forest and grassland by $\sim 15\%$.

3. Latent heat during HWDs, is heavily overestimated in both forest and grassland by $\sim 40\%$. On the other hand, sensible heat is heavily underestimated for both forest and grassland.

4. Ground flux during HWDs, seems to be in reasonable levels for grassland, while it is overestimated for forest.

5. The residual ε in the observations, is not to be disregarded as it has a major contribution to the final partitioning.

6. The absolute values of modelled net radiation are similar with the model, while being heavily underestimated during HWDs.

The aforementioned show that although grassland receives more net radiation in H-TESSEL due to the effect of similar albedo, this is not enough to explain why H-TESSEL demonstrates differences in the partitioning of net radiation to heat fluxes, compared to observations. At this point, a further examination of the output is needed, to rule out possible biases and understand whether this behaviour is originating from the nature of H-TESSEL itself, the current HWD selection, or both.

TABLE 4.1: Comparison of the energy flux partitioning for the observations and H-TESSEL relatively to the net radiation, for climatology and anomalies. Net radiation is set to 100% in both cases to have a better overview of the partitioning. The differences are those of H-TESSEL compared to observations. ε denotes the residual in the observations, due to biases and missing balance terms. In the last row, the absolute differences of net radiation between HWDs and climatology is shown.

Variables	Grassland			Forest		
	Observed	Modeled	Diff.	Observed	Modeled	Diff.
Rnet	100	100	–	100	100	–
LE	49	57	8 %	38	34	-4 %
H	27	22	-5 %	34	46	12 %
G	9	21	12 %	2	20	18 %
ε	19	0	-19 %	30	0	-30 %
Rnet anom.	100	100	–	100	100	–
LE anom.	61	100	39 %	5	51	46 %
H anom.	9	-17	-26 %	68	14	-54 %
G anom.	13	17	4	2	35	33 %
ε	-3	0	-3 %	18	0	-18 %
Additional Rnet in HWDs	39 %	12 %	-27 %	45 %	12 %	-33 %

4.2 The role of soil moisture

Soil moisture anomalies seem to have an important role in both HWD occurrence and duration, as discussed in chapter 1. In order to investigate the degree to which this is influencing H-TESSEL results, soil moisture anomalies are calculated for each gridbox and year, first as an average of the winter/spring period (February–April), and then as an average of the consequent summer HWDs. Values of soil moisture anomalies are then correlated with the number of HWDs that occurred in the same gridbox and year (figure 4.4).

There seems to be no correlation between spring/winter anomaly and the number of summer HWDs. An attempt to correlate winter/spring soil moisture anomalies with maximum heat-wave temperature anomalies in each gridbox, did not demonstrate any connection as well. However, there is a strong connection between summers where a lot of HWDs occurred, and low soil moisture levels during those heat-wave events. When less than 10 HWDs are present in a summer (which means at most 2 heat-wave events of 5 days), the respective soil moisture anomaly can range from highly negative to highly positive, with no apparent pattern. This, in combination with the fact that conditions during heat waves are reported to include low levels of soil moisture, suggests that there needs to be a discrimination between high and low soil moisture levels, which could be a factor influencing the partitioning of fluxes seen in the previous section.
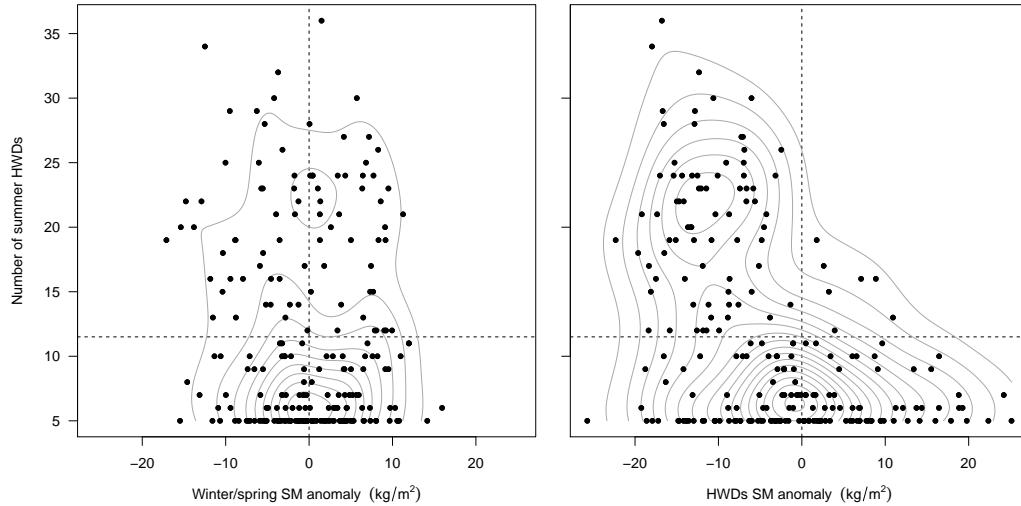
FIGURE 4.4: *Total number of summer HWDs as a function of the preceding winter/spring mean soil moisture anomaly (left), and the mean HWDs soil moisture anomaly of the same summer (right). Each point represents one gridbox in one year. Contours are density lines, and the horizontal doted line is the mean number of HWDs for all summers and gridboxes. Soil moisture here is calculated over the 3 top soil layers.*

To investigate that, fluxes where calculated again by dividing the range of soil moisture anomaly in two, and discriminating HWDs belonging to the part bellow the median of soil moisture anomalies, and HWDs belonging to the higher part of soil moisture anomalies. Results are shown in figure 4.5. This narrowing of the dataset, shows large improvements for the case of lower soil moisture anomalies, compared to observations. The climatology flux partitioning did not improve in most cases, except for the ground heat flux of grassland. Partitioning during HWDs improved largely for all energy fluxes. The differences from observations are shown in table 4.2, in comparison with the ones of table 4.1.

The investigation of the role of soil moisture in this section, suggests for a further inquiry on the role of HWD duration. The results of figure 4.4, show that there are gridboxes that may contain a single 5-day heat-wave event, some of which in addition include highly positive soil moisture anomalies –a situation that is not typical for heat-waves. This is further investigated in the following by looking at a specific short lasting heat-wave event.

4.3   A single case example

Here, an example of a 6-day heat wave event in the summer of 2003 is presented, for a gridbox in central France. The specific incident is selected to be shown, for some reasons. It occurred in an area which is known for having pronounced heat-waves in the same period. In addition, it includes negative soil moisture anomalies, and also rain events that appear as a sudden increase in soil moisture levels.
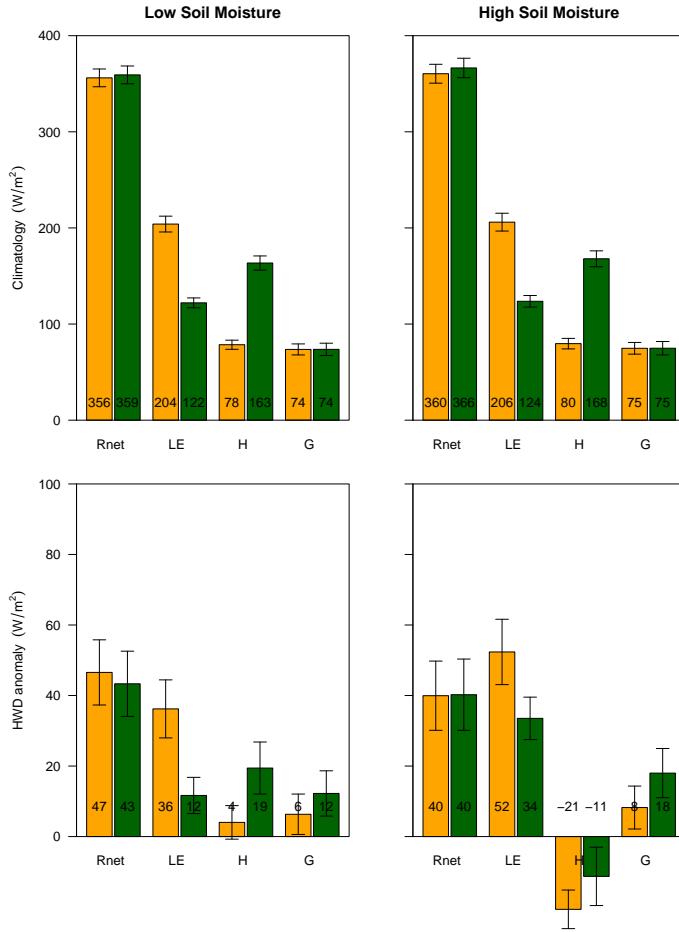
FIGURE 4.5: *Same as in figure 4.2, with the dataset divided in two; lower soil moisture (left), and higher soil moisture (right).*

TABLE 4.2: Same as in table 4.1, but for differences in energy partitioning from the observations, for the regular selection, and the additional low soil moisture selection. Bold values indicate improvements.

Variables	Grassland	Grassland (low SM)	Forest	Forest (low SM)
Rnet	–	–	–	–
LE	8 %	8 %	-4 %	-4 %
H	-5 %	-5 %	12 %	12 %
G	21 %	**12 %**	18 %	19 %
e	-19 %	-19 %	-30 %	-30 %
Rnet anom.	–	–	–	–
LE anom.	39 %	**16 %**	46 %	**22 %**
H anom.	-26 %	$\sin 0$ %	-54 %	**-23 %**
G anom.	4 %	$\sin 0$ %	33 %	**26 %**
e	3 %	3 %	-18	-18 %
Additional Rnet in HWDs	-27 %	-26 %	-33 %	-33 %
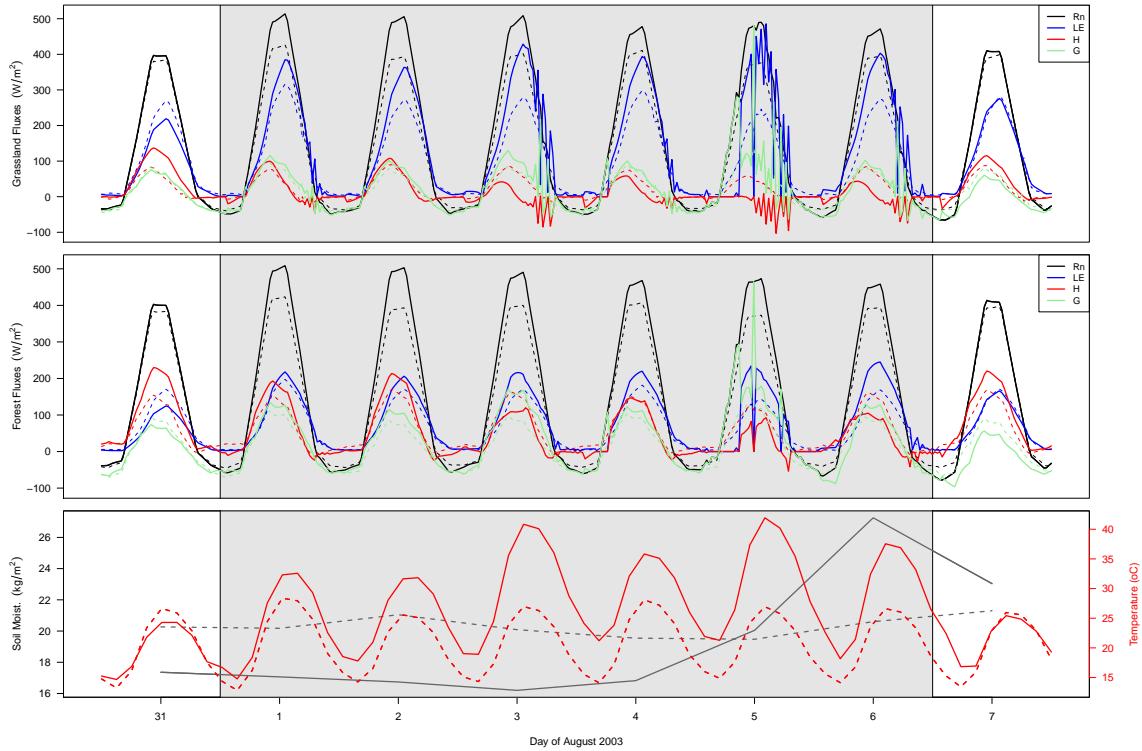
FIGURE 4.6: *Evolution of fluxes in a six-day heat-wave event between 1–6 August 2003 in a gridbox in central France.*

Figure 4.6, shows the evolution of this event from one day before the heat-wave, until one day after its end (31 July – 7 August 2003). The diurnal evolution of energy fluxes and climatology is shown, together with 2 m air temperature and average soil moisture for the top two layers.

One day before the heat-wave, soil moisture levels are already bellow the climatology, while temperature is almost normal for the day. The stress from low water availability, makes both forest and grassland to increase sensible heat while decreasing latent heat, compared to climatology. However, grassland partitions a greater part of the net radiation to latent heat, and forest a greater part in sensible heat.

During the first two days of the heat-wave, a positive net radiation anomaly drives the temperature higher, while soil moisture continues to deplete. Forest and grassland increase both latent and sensible heat, with grassland having a strong preference towards latent heat to compensate for the extra amount of forcing. This is also coming from the fact that grassland has a shorter roughness length, which does not allow for the same levels of sensible heat release as forest. In its beginning, this heat-wave is indeed showing HWD-like features.

The situation changes in the third day, as there is a distinct rise in temperature, and sensible heat drops bellow the normal levels for forest and grassland, and latent heat increases for both tiles. Grassland manages the extra radiation and heat, by increasing

latent heat in a great extent, while forest is able to increase ground flux effectively (skin conductivity can change in forest according to air stability). The jigged lines in the end of the day together with the increase in soil moisture the day after, suggest a rain event taking place that afternoon/night. In the next 3 days, soil moisture continues to rise due to rain events, especially the one on the fifth day. The incoming short-wave and net radiation during these days show no significant changes. This means that the rain events where either sudden and short-lasting, with cloud coverage not being captured by the 3-hourly resolution of radiation, or that this is again a case of low discernibility of values due to the merging with reanalysis data. The jigged lines are probably due to numerical fluctuations after the sudden changes in precipitation.

The heat-wave event ends after the sixth day –the noticeable feature being that the daily maximum anomaly was on average $+9.7°C$ during the six days, with a minimum anomaly of $+4.2°C$ and a maximum anomaly of $+15°C$. The range of maximum temperature anomaly, easily classifies these days as HWDs, but in reality they form a bias in the calculation of HWD fluxes. Although the forcing net radiation has a positive anomaly during these days and shows no fluctuations, the energy fluxes are highly affected by the rain events (through changes in the interception layer, and top soil layer). In addition, rain events are not supposed to be part of HWDs, so a further discrimination depending on HWD duration is needed.

4.4 HWD duration and energy fluxes

With the current selection of heat waves, longer lasting events are likely to be closer to the actual conditions of HWDs. It is easier to find short periods that display some HWD-like features (and are thus classified as heat-waves), without displaying the respective conditions to support it. If on the other hand an event is within the criteria set for HWDs, and holds this for an extended period of time, it shows that at least there is a consistent driving for the anomalous situation.

In order to investigate the evolution in time, HWDs have been grouped and averaged according to their duration. Here, the criterion for low soil moisture is again used, but this time all the days belonging to a single event must be bellow the median of soil moisture anomaly. Judging from figure 4.4, this is expected to exclude mostly short lasting events (indeed the number of events between 5 and 8 days was reduced by 50 % due to this criterion, while the exclusion of longer lasting events was between 0 and 15 %).

Anomalies in fluxes for the resulting heat-waves are averaged over the same day position within the heat-wave, and also calculated as a mean of all days in each specific duration. Results are shown for 5-day, 9-day and 13-day lasting events, with 77, 17, and
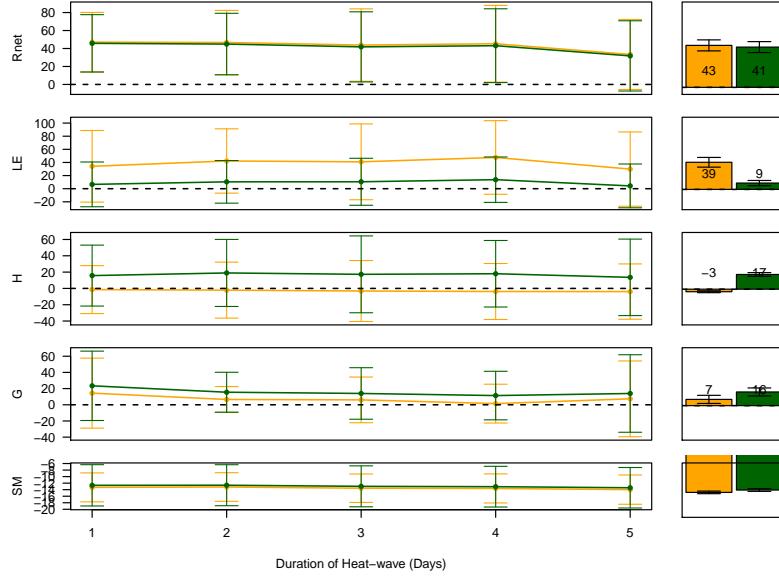
FIGURE 4.7: *Evolution of averaged daily flux anomalies for 77 events of 5-day lasting heat-waves, together with the mean for the whole event. Error bars indicate the standard deviation*
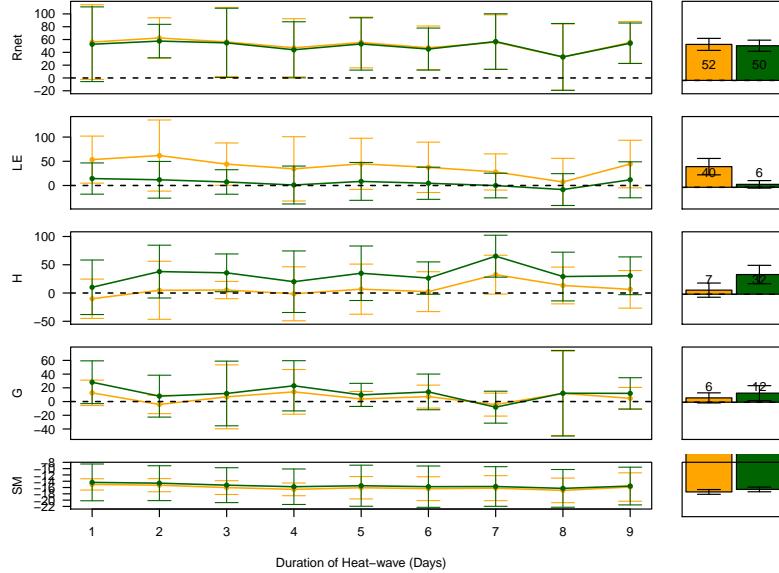
FIGURE 4.8: *Evolution of averaged daily flux anomalies for 17 events of 9-day lasting heat-waves, together with the mean for the whole event.*

5 events respectively in each category (figures 4.7, 4.8, and 4.9). Heat-waves of 13 days were the longest existing.

The results for 5-day events show a rather smooth evolution, with a net radiation anomaly driving forest to compensate with an increase in sensible heat, while grassland increases latent heat and has a negative sensible heat flux throughout the event. Ground flux anomaly is larger for forest. Soil moisture is constantly negative without fluctuations.
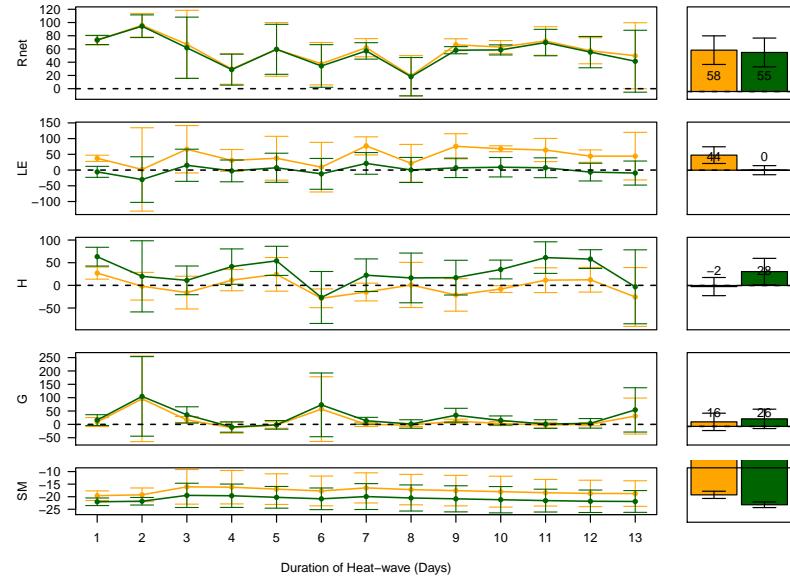
FIGURE 4.9: *Evolution of averaged daily flux anomalies for 5 events of 13-day lasting heat-waves, together with the mean for the whole event.*
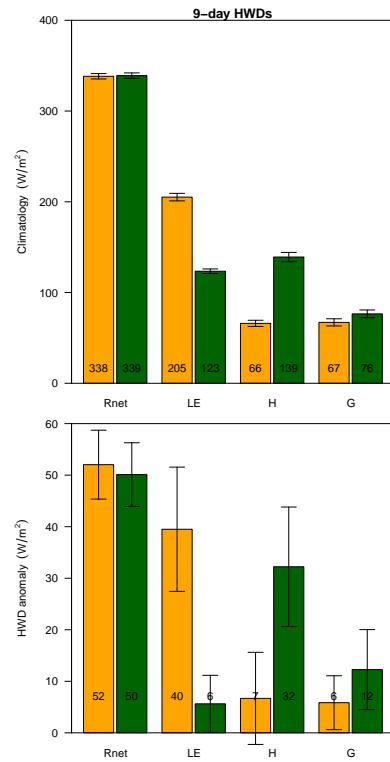

FIGURE 4.10: *Detailed fluxes, averaged over all the 9-day long heat-wave events (same as in figure 4.2).*

As the heat-wave duration increases, the evolution of flux anomalies become less gradual. Sensible and latent heat remain the main compensation factor for net radiation anomalies in forest and grassland respectively. In the 13-day heat-waves, there are a couple of interesting days, where ground flux anomaly increases to the same levels as net radiation anomaly, while latent and sensible heat anomalies experience a sudden drop. This behaviour is probably connected to rain events, with a corresponding increase in soil moisture (which was noted in the heat-wave example in the preceding section). This also disproves the hypothesis that long lasting events would not contain rainy days, and points once again to the HWD selection procedure.

Each group of heat-waves was correlated with the observations, and the one that showed the closest resemblance was the 9-day group (figure 4.10). In general it was found that the groups of 7-day to 12-day duration are the ones that contribute most to an improvement of the representation of fluxes, compared to observations. Using only these heat-waves, which account to sin 45 % of the total HWDs with low soil moisture, the improvement in the representation of fluxes is shown in table 4.3. In total, the additional criterion of heat-wave duration, showed an improvement in energy partitioning, only for the latent heat of forest, but in a large extent.

TABLE 4.3: Same as in table 4.1, but for differences in energy partitioning from the observations, for the regular selection, and the additional low soil moisture and low soil moisture–long duration selection.

Variables	Grass.	Grass. (low SM)	Grass. (low SM + 7-12 dur.)	Forest	Forest (low SM)	Forest (low SM + 7-12 dur.)
Rnet	–	–	–	–	–	–
LE	8.00 %	8.00 %	11.00 %	-4.00 %	-4.00 %	2.00 %
H	-5.00 %	-5.00 %	-7.00 %	12.00 %	12.00 %	7.00 %
G	21.00 %	12.00 %	11.00 %	18.00 %	19.00 %	21.00 %
e	-19.00 %	-19.00 %	-19.00 %	-30.00 %	-30.00 %	-30.00 %
Rnet anom.	–	–	–	–	–	–
LE anom.	39.00 %	16.00 %	15.00 %	46.00 %	22.00 %	6.00 %
H anom.	-26.00 %	0.00 %	4.00 %	-54.00 %	-23.00 %	-4.00 %
G anom.	4.00 %	0.00 %	0.00 %	33.00 %	26.00 %	22.00 %
e	3.00 %	3.00 %	3.00 %	-18	-18.00 %	-18.00 %

4.5   Analysis with CART

Low soil water levels have a pronounced effect on latent heat release, because of the stress imposed on evapotranspiration through r_c (equation 2.3), affecting the rest of the fluxes as well. Grassland and forest have a different partitioning for latent heat, mainly because of this resistance, which also depends on short-wave radiation, leaf area
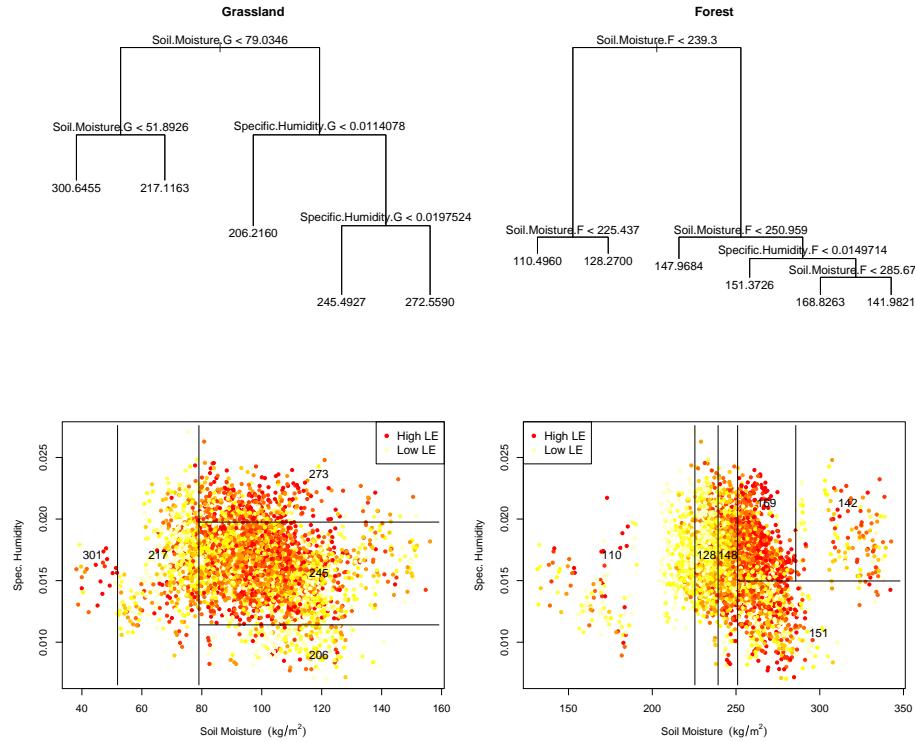
FIGURE 4.11: *CART of latent heat (mean diurnal values) dependence on soil moisture and specific humidity, for grassland and forest during HWDs. Left branches in each node represent a positive answer to the node question. The plot shows values for each day and gridbox, with red to yellow colour indicating high to low values of LE respectively.*

index, root density and water extraction capability, as well as vapour pressure deficit. In order to investigate the effect of the latter on the energy fluxes, and especially latent heat, in conjunction with soil water levels, specific humidity is used as a proxy. This correlation is of course not valid, as vapour pressure deficit depends also on atmospheric temperature, canopy temperature and pressure, but this crude approximation is used in the context that HWDs, occur within a limited range of temperature and air pressure values.

A regression tree was built with a recursive partitioning algorithm, which works by going through the dataset dividing it into smaller parts, trying to find correlations between variables within these parts. The procedure is stopped, when the error in the correlation of each division is minimised, at which point, the average of the dependent variable is calculated for each division and a construction of a decision tree is done. Here, latent heat and sensible heat during HWDs were used as the dependent variables in two distinct runs of the algorithm, while the respective specific humidity and soil moisture were the independent variables. Results are shown in figures 4.11 and 4.12.

In the case of latent heat, the behaviour of grassland and forest demonstrate some differences. Grassland evaporation, seams to be independent of soil moisture for the largest part of soil moisture range. From regions of ample soil moisture down to very
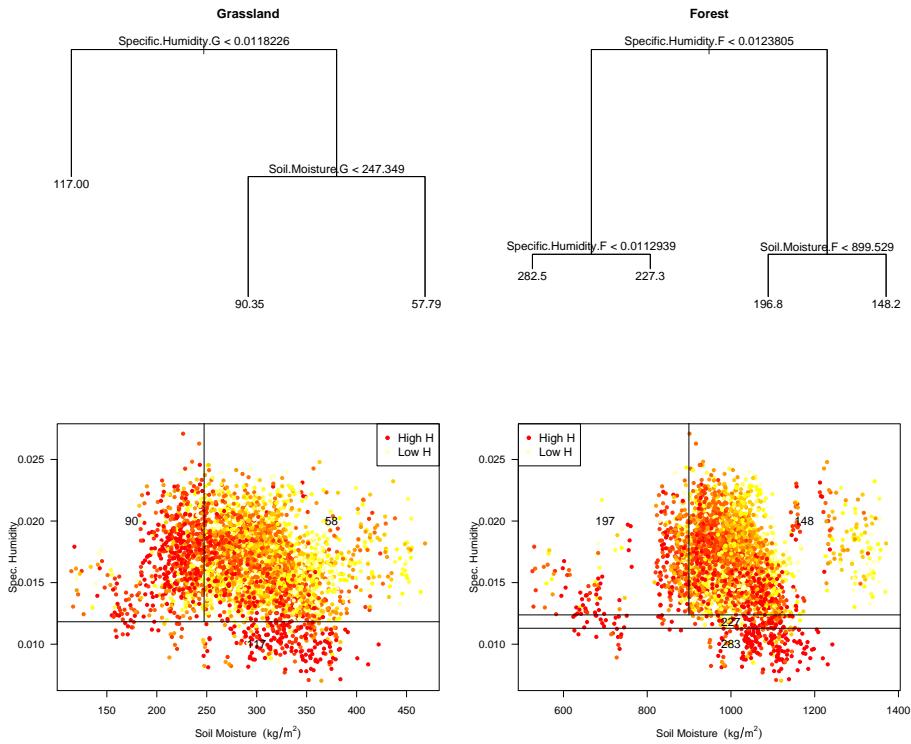
FIGURE 4.12: *Same as in figure 4.11, for sensible heat.*

low levels, grassland does not regulate latent heat depending on water availability, but instead on specific humidity. When soil moisture levels are on the very low range of their values, grassland seems to reduce latent heat and disconnect it from the dependence to specific humidity, only to increase it again up to very high levels when water is almost depleted. Most of the cases belong to the soil-water-independent range.

Forest on the other hand regulates the amount of latent heat depending on soil moisture, for the whole lower part of its values, demonstrating a dependence on specific humidity alone, only at ample moisture levels. The highest values of latent heat release occur on average in the high soil moisture – high humidity range of values, for both forest and grassland (if the burst of the values for depleted soil moisture in grassland are neglected).

In figure 4.12, the behaviour of forest and grassland seems to be very similar regarding sensible heat release, with absolute average values for forest being more than double the ones of grassland. For very low values of specific humidity, the amount of sensible heat is not depending on soil moisture, while for higher specific humidity, sensible heat is depending on soil moisture alone. The lowest values of sensible heat on average, are in the same region as the highest values of latent heat; high soil moisture and high specific humidity.

4.6   Investigating a different approach

In this final section, the possibility of using neural networks in a statistical way is described. Here, neural networks were build and trained to statistically re-produce the fluxes of H-TESSEL, depending on some variables produced by the forcing and some produced by the output of the model. The variables were net radiation, temperature, specific humidity and soil moisture as inputs, giving as output either one of latent heat, sensible heat or ground flux in each configuration.

The set-up included four input neurons, five neurons in the first hidden layer, 3 neurons in the second hidden layer, and one output neuron (figure 4.13). The procedure was to provide the network with a set of mean diurnal values of the input variables, and train it by also providing it with the expected values for the output. The neural network then, goes through cycles of weight adjustments in order to reproduce the expected output, according to the given input. This training, or control dataset included all summer days, but not HWDs. The testing dataset to check whether the network performs well, was formed only by HWDs. In this way the two datasets included different values, and the network did not have any information on the testing dataset after the training. The only way the network can produce an output in the testing dataset, is due to existing statistical relations between the input and the output, expressed by the weights from the training.

Error: 9.616896   Steps: 25227
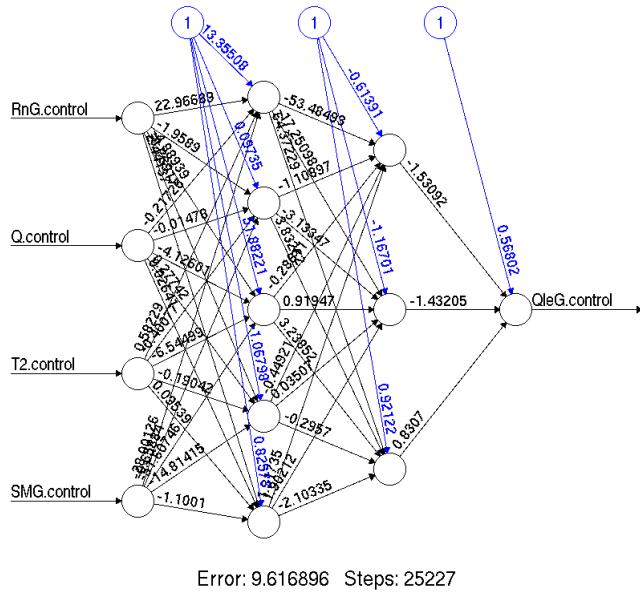
FIGURE 4.13: *The neuron configuration that was used, after the training session, for the case of latent heat in grassland. Input, hidden layers and output neurons can be seen, together with the weights assigned to them from the training.*

After training the networks for reproducing either latent heat, sensible heat or ground flux, it was found that there was a convergence of the algorithm only for the cases

of latent heat and ground flux, while the sensible heat algorithm did not converge. This means that in the dataset of H-TESSEL there is not a significant statistical connection between the input variables and sensible heat (figure 4.14). Nevertheless, sensible heat can be deduced, as it is the only unknown term for closing the energy balance.
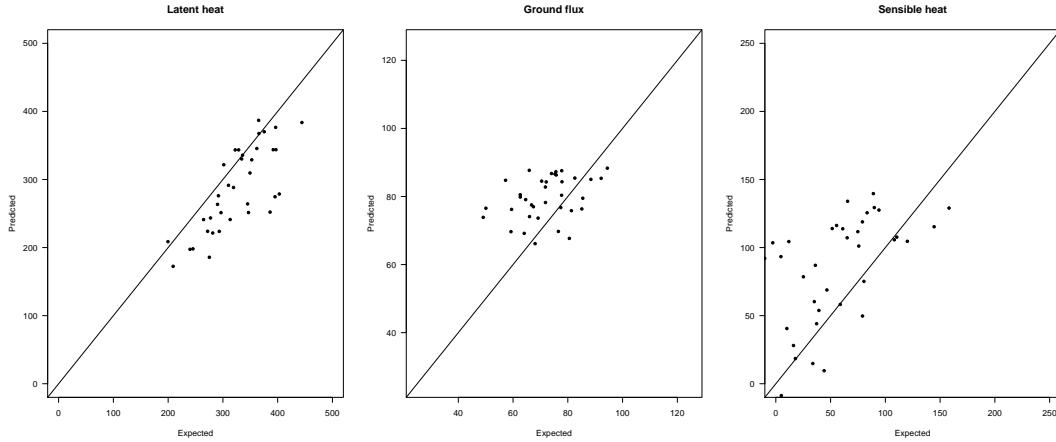
FIGURE 4.14: *Comparison of expected and predicted values of latent heat and ground flux, from the testing sessions of the networks for grassland, and sensible heat deduced by $H = R_{net} - LE - G$ (values during HWDs).*

The correlation of predicted with expected values looks promising, taking in mind the fact that the neural networks were trained by mean diurnal values. Using the full range of temporal resolution would probably result in better correlations, but the training time would increase proportionally. It should be made clear here that this is not an independent simulation; it depends on the statistical relations already existing in the H-TESSEL dataset. The main idea is that since in H-TESSEL net radiation, temperature, humidity and soil moisture can be used to statistically deduce the rest of the energy fluxes, then it might be possible to do so with observational data as well. Moreover, this approach could be used to assess the effect of changes in the model constants, without having to perform the whole offline runs each time. However this is part of future work.

Chapter 5

Conclusion

5.1 Summary of the work

Surface energy flux data of 1/2-hourly resolution from an offline H-TESSEL run output, were treated to find maximum and mean diurnal values in the period 1996–2006. The same was done for the 3-hourly resolution forcing dataset of the model, which was then used to make a selection of heat-wave days in the same period. As it was found difficult to properly select heat-wave days based on the forcing data alone, a second dataset was used to add in the procedure. Once the selection was established, an area of the output was selected in west-central Europe on the basic criterion of containing high and low vegetation in equal proportion and adequate coverage.

Energy fluxes for forest and grassland during the selected heat-wave days as well as the climatology of those days, were compared to the equivalent fluxes from observations. This was done to confirm whether the observed differences in the response of forest and grassland fluxes to heat-waves are simulated by the model. The role of soil moisture was assessed, as an aiding factor to the occurrence of heat-wave events, as well as a factor influencing the results of the model. A single event was used to investigate further the behaviour of the model on days that may have not been properly classified as HWDs, followed by an analysis of the evolution of energy fluxes depending on the duration of the events. Finally, the role of soil moisture in conjunction with air humidity is examined with a regression tree analysis.

5.2 The main points

The procedure used for HWD selection, resulted in an acceptable number of heat-waves, accounting to a fraction of around 2–3 % of the days in the period March–April 1996–2006, in most of Europe. The procedure overestimates the amount of heat-waves in areas with snow coverage for a considerable fraction of the year (i.e. alpine and scandinavic).

47

The placement of heat-waves during summer days was indeed at periods of extreme temperatures, and the summers of 2003 and 2006 contained a larger number.

The connection of low spring-time soil moisture with HWD occurrence was not found in the model data. However, there was a connection with soil moisture during these events.

Heat-waves defined by the selection procedure were found to also contain rainy days, or days that had in general high soil moisture levels, while the net radiation was not following any pattern of cloud coverage. This is probably because of a combination of low temporal resolution of incoming short-wave radiation, and of the blending process in which the dataset was produced. These days were biasing the results, and their exclusion provided a closer resemblance to the observations. A filtering of heat-waves with a duration between 7 and 12 days improved this further. This means that a HWD definition in the current level of data resolution, is much more efficient if an additional criterion for either soil moisture, or rain amount is applied.

In the climatology forest and grassland seem to behave close to the expectations, with grassland partitioning more of the energy towards latent heat, while forest towards sensible heat. Still the ground flux in both vegetation types remains high compared to the observations. This is potentially a bias in skin temperature calculation, connected to sensible and latent heat through the surface energy balance solver, which was not examined here.

The additional amount of net radiation received during HWDs, is underestimated by H-TESSEL compared to observations, which is most probably coming from the resolution of data; observations are taken on specific locations from flux towers, while the forcing of H-TESSEL is an average over an area of 10,000 km^2. This does not allow for spikes of radiation that might appear on single locations during heat-waves to be registered in a 1×1 degree resolution, which ultimately leads to a lower additional amount. This is validated by the fact that net radiation values in the climatology (which is an average in both cases), is on the same magnitude for H-TESSEL and the observations.

Grassland receives the same amount of net radiation with forest in H-TESSEL, while it receives less than forest in the observations. This is a consequence of the assignment of similar albedo for forest and grassland in the model, a fact that still cannot explain the differences in net radiation partitioning to heat fluxes, between the model and the observations.

Latent heat tends to be overestimated during HWD anomalies, especially for grassland. This, in combination with the fact that sensible heat is not differentiated in the same amount, points to the calculation of the stomatal resistance and the connection with vapour pressure deficit, which should be further investigated.

From the CART analysis, it is found that the more conservative nature of forest, concerning water use, is depicted in the dataset. While grassland seems to regulate

latent heat only at very low levels of the latter (figure 4.11), forest reduces evaporation already at median levels of soil moisture.

The differentiation in fluxes during heat-waves, due to soil dry-down was not found in any period within the dataset. Even for the longest heat-wave duration available (13 days), there was no sign of a change in the relation of latent and sensible heat between forest and grassland, as simulated by [cite Teuling 2010]. This suggests that even on the 13-day heat-waves, conditions are still on stage I drying.

Finally, latent heat seems to have a direct connection statistically, to net radiation, 2 m temperature, specific humidity and soil moisture, and probably it can be calculated directly from these variables (with a respective uncertainty). On the other hand, no statistical relevance was found between sensible heat and the aforementioned variables. This could be because of the lack of surface wind as an input, in which case it would suggest that wind affects sensible heat more strongly than latent heat.

5.3   Future work

The main limitations of the analysis here, comes from the low temporal resolution of the forcing. While it is adequate for assessing the performance of the model on the long term, events like heat-waves are defined within a low temporal resolution, and are affected by the evolution of the forcing variables. This can affect not only the selection procedure of heat-waves, but also their diurnal evolution. A single column run of the model for specific locations, forced by stationary data would probably provide more reliable results.

In such a future work, the dependence of the stomatal resistance on the combination of vapour pressure deficit and root water extraction should be examined, in search for an improvement of the calculation of latent heat.

Finally, the application of neural networks could be further examined by training networks with stationary data and different input variables, to investigate if a representation of fluxes based on statistics would give better results than the model.

5.4   Some final thoughts

This thesis work would have been much more efficient, if the amount of data was divided into smaller parts from the beginning. In that hypothetical parallel universe, the author would have started by selecting only a few stations that contain a high percentage of the respective vegetation type, and would not try to deal with the whole European continent at once. HWDs would then be determined for each station-gridpoint, which would probably make the selection faster. Fluxes could be also correlated with station data from FLUXNET, in order to determine easier which part of the model needs to

be improved. Since parallel universes are not proven to exist, this only remains as a thought for a similar situation in the future.

Bibliography

[1] IPCC, *Fourth Assessment Report: Climate Change 2007: The AR4 Synthesis Report* (Geneva: Intergovernmental Panel on Climate Change, 2007).

[2] S. I. Seneviratne and R. Stockli, "The Role of Land – Atmosphere Interactions.", Springer, 179–193 (2008).

[3] C. Wang and J. A. Carton, "A global survey of ocean-atmosphere interaction and climate variability", Geophys. Monograph, 1–19 (2004).

[4] B. van den Hurk, P. Viterbo, A. Beljaars, and A. K. Betts, *Offline validation of the ERA40 surface scheme* (ECMWF, 2000).

[5] J.-M. Robine, S. L. K. Cheung, S. Le Roy, H. Van Oyen, C. Griffiths, J.-P. Michel, and F. R. Herrmann, "Death toll exceeded 70,000 in Europe during the summer of 2003.", Comptes rendus biologies **331**, 171–8 (2007).

[6] E. Black, M. Blackburn, G. Harrison, B. Hoskins, and J. Methven, "Factors contributing to the summer 2003 European heatwave", Weather **59**, 217–223 (2004).

[7] E. M. Fischer, S. I. Seneviratne, P. L. Vidale, D. Luthi, and C. Schar, "Soil Moisture-Atmosphere Interactions during the 2003 European Summer Heat Wave", Journal of Climate **20**, 5081–5099 (2007).

[8] E. M. Fischer, S. I. Seneviratne, D. Luthi, and C. Schar, "Contribution of land-atmosphere coupling to recent European summer heat waves", Geophysical Research Letters **34**, 1–6 (2007).

[9] N. C. Turner, "Measurement and influence of environmental and plant factors on stomatal conductance in the field", Agricultural and Forest Meteorology **54**, 137–154 (1991).

[10] T. N. Buckley, "The control of stomata by water balance.", The New phytologist **168**, 275–92 (2005).

[11] E. Cominelli, M. Galbiati, and C. Tonelli, "Transcription factors controlling stomatal movements and drought tolerance", Transcription, 41–45 (2010).

[12] U. Feller, "Stomatal opening at elevated temperature: an underestimated regulatory mechanism", Gen. Appl. Plant Physiology, 19–31 (2006).

[13] A. M. Hetherington and F. I. Woodward, "Nature", **424**, 901–908 (2003).

[14] B. F. Zaitchik, A. K. Macalady, L. R. Bonneau, and R. B. Smith, "Europe's 2003 heat wave: a satellite view of impacts and land-atmosphere feedbacks", International Journal of Climatology **26**, 743–769 (2005).

[15] W. Wicke and C. Bernhofer, "Energy balance comparison of the Hartheim forest and an adjacent grassland site during the HartX experiment", Theoretical and Applied Climatology **53**, 49–58 (1996).

[16] A. J. Teuling, S. I. Seneviratne, R. Stockli, M. Reichstein, E. Moors, P. Ciais, S. Luyssaert, B. van den Hurk, C. Ammann, C. Bernhofer, E. Dellwik, D. Gianelle, B. Gielen, T. Grunwald, K. Klumpp, L. Montagnani, C. Moureaux, M. Sottocornola, and G. Wohlfahrt, "Contrasting response of European forest and grassland energy exchange to heatwaves", Nature Geoscience **3**, 722–727 (2010).

[17] J. Pitman, "The evolution of, and revolution in, land surface schemes designed for climate models", International Journal of Climatology **23**, 479–510 (2003).

[18] C. C. V. Heerwaarden, J. Vil, D. Arellano, A. F. Moene, and A. A. M. Holtslag, "Interactions between dry-air entrainment , surface evaporation and convective boundary-layer development", **1291**, 1277–1291 (2009).

[19] N. Sato, P. Sellers, D. Randall, E. Schneider, J. Shukla, J. Kinter, Y.-T. Hou, and E. Albertazzi, "Effects of Implementing the Simple Biosphere Model in a General Circulation Model.", J. Atmos. Sci., 2757–2782 (1989).

[20] A. Beljaars, P. Viterbo, and M. Miller, "The Anomalous Rainfall over the United States during July 1993: Sensitivity to Land Surface Parameterization and Soil Moisture Anomalies", Monthly Weather Review **124**, 362–383 (1995).

[21] P. Viterbo, A. Beljaars, J.-F. Mahfouf, and J. Teixeira, "The representation of soil moisture freezing and its impact on the stable boundary layer", Quarterly Journal of the Royal Meteorological Society **125**, 2401–2426 (1999).

[22] G. Balsamo, P. Viterbo, A. Beljaars, B. van den Hurk, M. Hirschi, A. K. Betts, and K. Scipal, "A Revised Hydrology for the ECMWF Model: Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast System", Journal of Hydrometeorology **124**, 623–643 (2009).

[23] J. Sheffield, G. Goteti, and E. F. Wood, "Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling", Journal of Climate **19**, 3088–3111 (2006).

[24]M. R. Haylock, N. Hofstra, a. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New, "A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006", Journal of Geophysical Research **113** (2008), `10.1029/2008JD010201`.

[25]E. Zorita and v. H. Storch, "The analog method as a simple statistical downscaling technique: comparison with more complicated methods", Journal of Climate **12**, 2474–2489 (1999).

[26]B. Orlowsky and S. I. Seneviratne, "Investigating spatial climate relations using CARTs: An application to persistent hot days in a multimodel ensemble", J. Geophys. Res. (2011), `10.1029/2010JD015188`.

[27]R. C. Team, *R: A Language and Environment for Statistical Computing*, ISBN 3-900051-07-0, R Foundation for Statistical Computing (Vienna, Austria, 2012).

[28]C. Shalizi, "Classification and regression trees.", A course on data mining (2009).

[29]J. Heaton, *Introduction to Neural Networks With Java* (Heaton Research, Incorporated, 2005).

[30]W. W. Hsieh and B. Tang, "Applying Neural Network Models to Prediction and Data Analysis in Meteorology and Oceanography", Bulletin of the American Meteorological Society **79**, 1855–1870 (1998).

[31]B. K. Hansen, *State of the Art of Neural Networks in Meteorology* (Technical University of Nova Scotia, 1997).

[32]I. Maqsood, M. Khan, and A. Abraham, "An ensemble of neural networks for weather forecasting", Neural Computing and Applications **13**, 112–122 (2004).