UiO : Department of Geosciences
University of Oslo

Using NORA3 to investigate the impact of inversion trends on NO2 concentration trends in Oslo and Bergen

Katarina N. Andersen

Master's Thesis, Spring 2022

Abstract

Temperature inversions prevents vertical mixing and thereby distribution and dilution of pollutants, and can cause health hazardous concentrations of NO_2 to accumulate. Policies have been implemented to reduce the emissions, and causing the NO_2 concentration to decrease. In this study, we will use NORA3 to investigate if inversion trends have an impact on the observed trends in NO_2 concentrations in Oslo and Bergen, or if the trends in NO_2 concentrations are only due to reduction in emissions.

The statistical relationship between the observed NO_2 concentrations and inversions was assessed from 2001 to 2010 in Oslo and 2011 to 2020 in Bergen. The data was filtered to be able to discard variations in the emissions. In Oslo, the inversions, measured between 2 and 25 meter altitude at Valle Hovin, and the NO_2 concentrations, measured at Alnabrua, had a correlation of 0.7. In Bergen, the inversions, measured between 2 and 40 meters altitude at Florida/Geofysen, and NO_2 concentrations, measured at Danmarks plass, had a correlation of 0.3.

Comparison of inversions in NORA3 to the observations showed NORA3 underestimating the yearly average inversion intensity in Oslo, but the variations were similar to the observations in both Oslo and Bergen. The correlation between the modelled and observed inversions were weak for both study areas. Assessing two scenarios each in Oslo and Bergen; one where the modelled and observed inversions were similar in magnitude and one where they differed, showed for the first case that the inversions were sporadic night inversions, while for the second case, the inversions were continuous. The model surface energy balance was consistent with the 2 meter air and skin temperature. Lastly, we investigated 20 year trends in the modelled inversions and observed NO_2 concentrations from 2001 to 2020. Both study areas showed a significant negative trend in the NO_2 concentrations. There were no significant trend in modelled inversions for Oslo, but there were a significant negative trend for Bergen.

Acknowledgements

First off I would like to thank my two supervisors, Erik Berge and Terje Koren Berntsen. I have had fun working with you and learned a lot from you both, thank you for guiding me through this process. A special thank goes out to Erik for coming up with the idea for this thesis.

Thank you to NILU, MET, Susanne Lützenkirchen and Igor Esau for providing data for this thesis. I also want to thank Olvar Bergland for helping me with the statistical aspects of this thesis, Anne B. Nilsen for helping me with the topography maps, and Stefanie Falk for helping me with the technical stuff.

I have had a whole village of people helping me throughout this process that also needs some credit. Johanne and Astrid, I would never have been able to do any of this without you. Kamilla and Anna Lina, thank you both for taking your time to proof read and giving me feedback, it was very much appreciated. Mum and dad also deserves a shout out for raising and feeding me. Thank you to my fellow master students at MetOs, my ears are still ringing from the karaoke night. And no, I did not forget you, Einar. Thank you so much for your endless support and patience, you are amazing.

Contents

List of Figures

List of Tables

List of abbreviations

ABL	Atmospheric Boundary Layer
AROME	Applications of Research to Operations at Mesoscale
EPA	Environmental Protection Agency
ERA5	ECMWF reanalysis generation 5
FHI	Norwegian Institute for Public Health
HO_2	Hydrogen dioxide
HNO_3	Nitric acid
MET	Norwegian Meteorological Institute
NILU	Norwegian Institute for Air Research
NO	Nitrogen monoxide
NO_2	Nitrogen dioxide
NO_x	Nitrogen oxides
HARMONIE	HIRLAM–ALADIN Research on Mesoscale Operational NWP in Euromed
NORA3	3-km Norwegian Reanalysis
OH	Hydroxide
PAN	Peroxyacetylnitrate
SEB	Surface Energy Balance
SLR	Simple Linear Regression
SO_2	Sulfur dioxide
SSB	Statistics Norway
UCAR	University Corporation for Atmospheric Research
UNECE	United Nations Economic Commission for Europe
VOC	Volatile Organic Compound
WHO	World Health Organization

CHAPTER 1

Introduction

Studies show exposure to NO_2 may cause a variety of health issues, depending on the intensity and duration [Lowry and Schuman, 1956, Jones et al., 1973]. Shorter exposure periods to higher NO_2 concentrations may cause irritation in the air ways and aggravation of respiratory diseases, especially for children [United States Environment Protection Agency (EPA), 2021, Norwegian Institute of Public Health (FHI), 2020]. Clinical studies have shown that exposure of 300 $\mu g/m^3$ over just a few hours reduces lung function in people with asthma, and it is reason to believe that longer periods of exposure can even cause development of asthma and makes the lungs more receptive to respiratory infections.

Per 2022, the yearly average health limit defined by the Norwegian Institute of Public Health (FHI) is 30 $\mu g/m^3$, while the hourly average health limit is sat to be 100 $\mu g/m^3$. Surface near NO_2 concentrations are sourced by traffic, and according to FHI the emissions have been reduced by 50%, while the number of cars have doubled in the period from 1990 to 2020.

In 1999, Norway joined other countries in the Gothenburg protocol to reduce emissions of SO_2, NO_2, HNO_3 and volatile organic compounds (VOCs) [United Nations Economy Commission for Europe (UNECE), 1999]. As part of the protocol, Norway had to reduce the NO_2 emissions by 27% compared to the emissions in 1990 by 2010 [Norwegian Government, 2007]. Later development of the protocol stated that Norway had to reduce its NO_2 emissions by 23% compared to the emissions in 2005 by 2020 [Norwegian Petroleum Directorate, 2020]. To achieve these goals measures has been implemented, such as driving restrictions, toll increases for specific hours of the day and reducing the availability of parking spaces [Norwegian Environment Agency, 2022]. As well as the policies, the decrease in emissions has also been attributed to cleaner engines, and the increased popularity of electric cars in the last few years, especially in the larger cities.

The surface NO_2 concentrations are also highly affected by the local meteorology as horizontal and/or vertical motion contributed to the dispersion and dilution of the pollutant [Stull, 1988]. If there are little to no horizontal motion, the ability to mix is decided by the stability of the layer. In the case of a temperature inversion - where a temperature inversion is defined as a layer with potential temperature increasing with height - the layer is stable and prevents the mixing.

If the inversion persists the NO_2 emissions may accumulate, causing health hazardous concentrations. Such episodes occurs especially during winter.

In this study, we aim to understand whether the trends we observe in the NO_2 concentration can be also attributed to trends in temperature inversions, or if they only are a result of policies and emission reduction. Milionis and Davies, 2008, and Bailey et al., 2011, found that the temperature inversions are sensitive to large scale circulation, but that the sensitivity is dependent on the topography and location. Due to global warming affecting the synoptic circulation pattern, it is plausible to expect to find trends in temperature inversions. Ji et al., 2019, studied three periods (1990-2009, 2020-2039 and 2060-2079) and found there to be an increase in surface-near temperature inversion events in the southeast of Australia. Caserini et al., 2017, studied the temperature inversions in the Po Valley under RCP4.5 and RCP8.5 and found that the occurrence of temperature inversions increased for both scenarios.

The areas of interest for this study are Oslo and Bergen, due to the availability of data. Oslo and Bergen are the two most populated cities in Norway, and also the most polluted, according to FHI and the Norwegian Environment Agency. Oslo is the capital city of Norway, found in the eastern part of the country, while Bergen is located on the west coast. Both cities are valley shaped, but the terrain is more defined in Bergen than in Oslo. For Oslo, the station measuring the NO_2 concentrations is located at Alnabru, while the station measuring the meteorological conditions is placed at Valle Hovin. For Bergen, the station measuring the NO_2 concentrations is located at Danmarks plass and the two stations used to measure the meteorological conditions are placed at Florida/Geofysen.

In Oslo, Kukkonen et al., 2005 studied a period with high concentrations of PM_{10} from the 4th to the 10th of January, 2003. Their meteorological analysis, showed that the wind speed varied from weak up to 8 m/s at the period where the temperature inversion was at its strongest. Studies of temperature inversions in Bergen have previously been made by e.g. Seilen, 2018, where low-cost measurement equipment was tested from Febuary 2017 to Febuary 2018. Wolf et al., 2014, studied the vertical temperature structure in Bergen over a 2 year period showing that episodes with high pollution were associated with near surface temperature inversions. The occurrence of near surface inversion were commonly observed during the winter, even in the presence of winds at 16 m/s.

The study period of the observed inversions in both Oslo and Bergen are only ten years long each, from 2001 to 2010 in Oslo and from 2011 to 2020 in Bergen. A ten year period is too short to conclude if there are any trends, so we will therefore use model output from NORA3, a high-resolution hindcast produced by downscaling the ERA5 reanalysis in the numerical weather prediction model HARMONIE-AROME [Haakenstad et al., 2021], to compare a twenty year period trend with the trend in the NO_2 concentrations. However, to be able to use the output from the NORA3 simulations to quantify how past and future changes in meteorological conditions, in particular inversion, affect the trend in the NO_2 concentrations in Oslo and Bergen, we need to assess how robust the statistical relationship between the observed NO_2 concentrations and the meteorological variables is. This will be done by calculating the correlation between all the variables and conducting and analysing a linear regression

model using NO_2 concentrations as the dependent variable. Furthermore, we will validate the output from the NORA3 model against the observed meteorology at Valle Hovin and Florida/Geofysen. We will see how the model output correlates with the observation, investigate the magnitude and variation of the yearly average inversion intensities, and study the surface energy balance (SEB) for two periods; one where the model inversion intensity is similar to the observed inversion intensity, and one where the model underestimates/overestimates the inversion intensity compared to the observed inversion intensity. Through validating our observations and the models ability to reproduce them, we can assess if the modelled temperature inversion trends can be used to explain the trends observed for the NO_2 concentration, and if the model can be used to predict future high pollution events.

CHAPTER 2

Theory

2.1 NO_x

The term *air pollution* is according to the World Health Organization (WHO), defined as a contamination of the indoor or outdoor environment by any chemical, physical or biological agent that modifies the natural characteristics of the atmosphere, and consists of chemical compounds such as particulate matter, carbon monoxide, ozone, sulfur dioxide and nitrogen dioxide [World Health Organization (WHO), 2021]. Understanding the chemistry of NO_x, where NO_x is a collective term for chemical bindings between nitrogen and oxygen, is essential for evaluating the concentration of NO_2.

Chemistry of NO_x

At very high temperatures oxygen is thermolyzed into two single O-atoms that are highly reactant. These single O-atoms can in the presence of nitrogen produce nitrogen oxide. This reaction is given by an equilibria that shifts to the right at higher temperatures ($\sim$ 2000 K) [Jacob, 1999a].

$$O_2 \overset{\text{heat}}{\rightleftharpoons} O + O \tag{2.1}$$

$$O + N_2 \rightleftharpoons NO + N \tag{2.2}$$

$$N + O_2 \rightleftharpoons NO + O \tag{2.3}$$

It is important to note that reaction 2.2 is very slow, and only a small amount of O-atoms will follow this reaction chain. The rest of the O-atoms will react to form ozone following

$$O + O_2 + M \longrightarrow O_3 + M \tag{2.4}$$

For the NO from Equation 2.1 - 2.3 to form NO_2, a rapid null cycle takes place. Because this reaction takes place in a matter of a minute, it is common to look at both NO and NO_2 as a whole, under the term of NO_x. The null cycle is given by

$$NO + O_3 \longrightarrow NO_2 + O_2 \tag{2.5}$$

$$NO_2 + h\nu \xrightarrow{O_2} NO + O_3 \tag{2.6}$$

Where hv is used to denote photon energy from solar radiation, and it is therefore important to note that the latter part of the null cycle only takes place during daytime. At night all NO becomes NO_2, and, provided that there is sufficient O_3 present to perform the reaction, the reaction chain becomes a sink for ozone. This is also the case for wintertime in Norway, where the incoming solar radiation is more limited than during the summer.

Ozone in the troposphere acts not only as a greenhouse gas, but inhalation can also cause lung tissue damages and other cardio vascular diseases [University Corporation for Atmospheric Research (UCAR), 2014]. We will now present how NO_x can work as a catalyst for ozone in the presence of HO_x. However, note that the production of OH is dependent on the availability of solar radiation, and so the following chemical descriptions will be of little importance for wintertime NO_2. Wintertime tropospheric ozone mostly originated from the stratosphere, by transportation through the tropopause.

The chain mechanism for ozone catalyzation through NO_x and HO_x consists of the following reactions

$$\mathrm{CO} + \mathrm{OH} \xrightarrow{\mathrm{O_2}} \mathrm{CO_2} + \mathrm{HO_2} \tag{2.7}$$

$$\mathrm{HO_2} + \mathrm{NO} \longrightarrow \mathrm{OH} + \mathrm{NO_2} \tag{2.8}$$

$$\mathrm{NO_2} + hv \xrightarrow{\mathrm{O_2}} \mathrm{NO} + \mathrm{O_3} \tag{2.9}$$

With the resulting net reaction given by

$$\mathrm{CO} + 2\,\mathrm{O_2} \longrightarrow \mathrm{CO_2} + \mathrm{O_3} \tag{2.10}$$

In net reaction 2.10, neither OH, HO_2, NO nor NO_2 is consumed. Comparing the chain mechanism 2.7-2.9 to the null cycle 2.5-2.6, we find that the most notable difference is that NO reacts with HO_2 to form NO_2, and not with O_3 itself. This chain mechanism may occur repeatedly as long as all participating molecules are present, and termination occurs when NO_2 and OH react to produce HNO_3.

$$\mathrm{NO_2} + \mathrm{OH} + M \longrightarrow \mathrm{HNO_3} + M \tag{2.11}$$

M is a "third body", and can be any of the molecules in the air. At nighttime and in the winter, the oxidation of NO_x to HNO_3 happens when reacting with O_3 instead.

$$\mathrm{NO_2} + \mathrm{O_3} \longrightarrow \mathrm{NO_3} + \mathrm{O_2} \tag{2.12}$$

$$\mathrm{NO_3} + \mathrm{NO_2} + M \longrightarrow \mathrm{N_2O_5} + M \tag{2.13}$$

$$\mathrm{N_2O_5} + \mathrm{H_2O} \xrightarrow{\mathrm{aerosol}} 2\,\mathrm{HNO_3} \tag{2.14}$$

However, HNO_3 is not an effective reservoir for long transport of NO_x as it is quickly deposited. For long distance transport of NO_x, carbonyl compounds or VOCs needs to be present. By producing hydrocarbons by photochemical oxidation, they can further react with NO_x to produce Peroxyacetylnitrate ($CH_3C(O)OONO_2$), also known as PAN. An example is given by the simplest

case of acetaldehyde (CH_3CHO).

$$CH_3CHO + OH \longrightarrow CH_3CO + H_2O \tag{2.15}$$

$$CH_3CO + O_2 + M \longrightarrow CH_3C(O)OO + M \tag{2.16}$$

$$CH_3C(O)OO + NO_2 + M \longrightarrow PAN + M \tag{2.17}$$

The lifetime of PAN ranges all from one hour to several months, depending on the temperature, as the main loss of PAN is due to thermal decomposition.

$$PAN \xrightarrow{\text{heat}} CH_3C(O)OO + NO_2 \tag{2.18}$$

In the lower troposphere, PAN and NO_x are approximately in equilibrium. However, in the middle and upper troposphere, PAN can be transported and decomposed at longer distances from the source, and thus affect the tropospheric NO_x on a global scale. If NO_x is transported to areas with low NO_x concentrations, the production of ozone is enhanced as shown in Equations 2.7 and 2.9. After being thermally decomposed, NO_2 may be oxidized by OH and deposited.

Processes controlling the NO_2 concentration

In the following subsection, we will use the term NO_2 as we wish to relate the processes discussed to our study. To understand why the inversions are so important when discussing NO_2, first consider the mass balance equation for a species X within a given box

$$\frac{dm}{dt} = \sum \text{sources} - \sum \text{sinks} = F_{in} + E + P - F_{out} - L - D \tag{2.19}$$

[Jacob, 1999b], stating that the change in mass over time is equal to the sum of all sources minus the sum of all sinks, and is illustrated in Figure 2.1. The first term on the right hand side, denoted F_{in}, includes the NO_2 concentration advected into the box. The second term, E, denotes the emissions. Emissions are divided into two categories, the anthropogenic and biogenic emissions. In locations such as our study area, the anthropogenic emissions will dominate. As we are looking into the concentration of NO_2, the chemical production, P, will be the concentration of NO_2 produced by Equation 2.5 and 2.8.

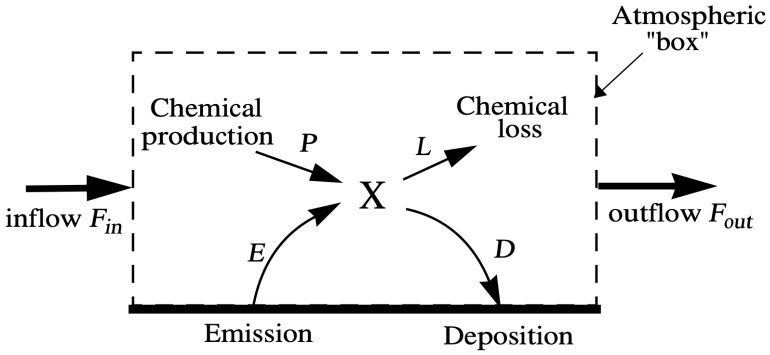


Figure 2.1: A one-box model for an atmospheric species, X, or in this case, NO_2 [Jacob, 1999c].

As for the sinks, F_{out} covers the transportation of NO_2 both vertically by turbulence and horizontally by the mean winds. In the case where the mean wind is weak, the vertical mixing through convective turbulence is important for the loss of NO_2. If the convective turbulence is suppressed by an inversion, the only sinks for NO_2 is the chemical loss, L, through the chemical reactions discussed in the previous section, and the loss through deposition, D. Note that the chemical production of NO_2 is found within the null cycles, and is therefor cancelled out by the chemical loss when NO_2 splits to NO and O_3. The removal rate of NO_2 is therefore the sum of deposition and the remaining chemical loss through production of HNO_3 and PAN. In large cities, the rate of emissions are much faster than the removal rate, leading to accumulation of NO_2 concentration.

Sources of NO_xn Norway

The contribution to emissions of nitrogen oxides made by different sectors can be found in Figure 2.2. The data was collected from Statistics Norway, and it shows that the non-road transport is, not surprisingly, the largest contributor. The oil and gas industry follows next, contributing to approximately 28% of the concentrations, and road traffic contributes to 21%.

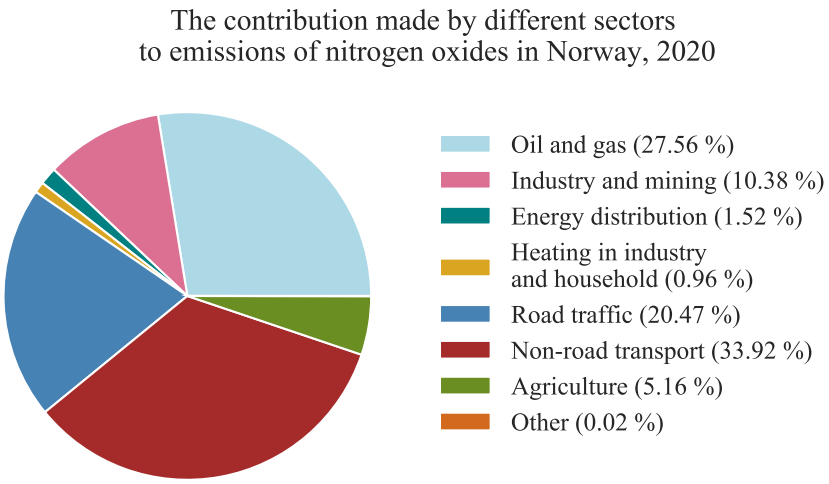


Figure 2.2: Pie chart showing the contribution of NO_x emissions by different sectors [Statistics Norway (SSB), 2021c]

Through policies trying to enhance the air quality in cities, as well as the increasing popularity of electric cars, the NO_x emissions have reduced by 25% from 1990 to 2020 [Norwegian Environment Agency, 2021]. In 2020, the emissions were at 150 000 tonnes. From 2019 to 2020, it was observed a decrease of 6.5% in emissions, through a decrease in emissions from road traffic, the oil- and gas industry as well as air travel and domestic shipping. The Covid-19 pandemic was a contributing factor, as fewer commuted to and from work, this

will especially play a role in reducing the emissions in larger cities such as Oslo and Bergen.

2.2 Atmospheric Boundary Layer

For the mass balance in Equation 2.19, we discussed the importance of mixing the pollutants horizontally and vertically through wind and turbulence. In this section, we will look at how such mixing occur, and how inversions contribute to the suppression of these processes.

Our study area lies within the Atmospheric Boundary Layer (ABL). According to Stull, 1988, the ABL is defined as *"that part of the troposphere that is directly influenced by the presence of the earth's surface, and responds to surface forcings with a timescale of about an hour or less"*. Normally, the ABL is ∼1 to 2 km thick [Wallace and Hobbs, 2006a], but has the ability to range all the way from 10 meters to 4 km. Above the ABL, we find the capping inversion, a strong stable layer, which separates the ABL from the statically stable free atmosphere, as illustrated in Figure 2.3. The ABL differs by showing diurnal temperature variations from the free atmosphere where the temperature is more stable. The diurnal temperature variation is caused by solar radiation heating the ground, and the heat transferring from the ground to air. Within the ABL, we also find surface forcings such as frictional drag, pollution and mechanical turbulence.

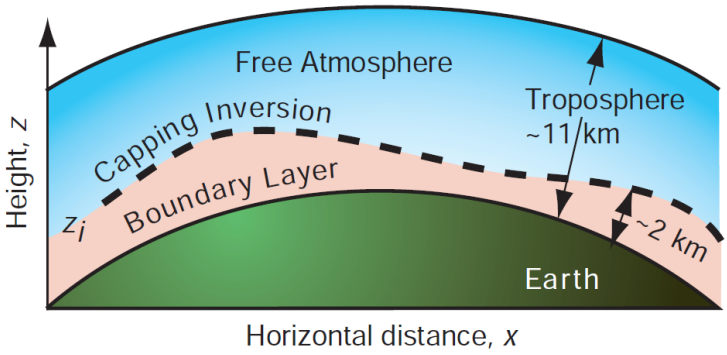


Figure 2.3: Illustration of the vertical structure of the troposphere, including the ABL, capping inversion and the free atmosphere [Wallace and Hobbs, 2006a].

Turbulence and inversion

Vertical mixing of pollution is dominated by turbulence, an irregular flow of quickly disappearing swirles called eddies. These eddies can either grow along the edge of larger eddies, or be generated mechanically through shears in the mean wind that are formed by the terrain or surrounding buildings. They can also be formed through air convection, known as convective turbulence. We mentioned the heat transfer from the surface ground to the overlaying air. Because of this heat transfer, the temperature at lower altitudes is warmer than

the temperature at higher altitudes, and buoyed forces will cause the colder air to sink and the warmer air to rise. We say that the layer is *unstable*.

Let us look further into what makes a layer stable or unstable. First, we introduce the concept of potential temperature, which is defined as *"the temperature a parcel were to have if it were to expand or compress adiabatically from its original placement to a reference pressure p_0 [Wallace and Hobbs, 2006b]"*, and is given by

$$\theta = T\Big(\frac{p_0}{p}\Big)^{R/c_p} \tag{2.20}$$

T denotes the absolute temperature of the parcel, and p is the pressure of parcel's surroundings. The pressure of the reference point, p_0, is usually 1013 hPa. R is the gas constant of air, and c_p is the specific heat capacity at constant pressure. $R/c_p \approx 0.286$.

Now consider an air parcel that is being displaced from its original placement, $[X, Z]$, to a higher altitude, $[X, Z + \delta z]$, where δz represents the change in displacement. The environment is considered to be *statically stable* if the parcel is forced back to its original location. This occurs when the potential temperature *increases* with altitude. When displaced upwards, the parcel will be colder than its surroundings, and will thus experience a negative buoyant force. However, if the parcel continues to move upwards, the environment is considered to be *statically unstable*. This occurs when the potential temperature *decreases* with altitude. When displaced upward, the parcel will be warmer than its surroundings and will experience a positive buoyant force, continuing to push it upwards. If the parcel stays at [X, Z + δz], then the parcel has been displaced to surroundings with the same temperature as the parcel, and the environment is considered to be *statically neutral*. Summarizing these conditions to:

$$\left.\begin{matrix}\text{unstable}\\ \text{neutral}\\ \text{stable}\end{matrix}\right\} \quad \text{if } \Big(\frac{d\theta}{dz}\Big)_E \quad \begin{cases} < 0 \\ = 0 \\ > 0 \end{cases} \tag{2.21}$$

For this study, we only have the observations of the absolute temperature, so we need to relate these sets of conditions to T. Taking the derivative of Equation 2.20 with respect to the height, z, finds that

$$\frac{d\theta}{dz} \simeq \frac{dT}{dz} + \Gamma_d \tag{2.22}$$

where

$$\Gamma_d = -\frac{g}{c_p} \approx 0.098\ ^\circ C/m \tag{2.23}$$

and is known as the dry adiabatic lapse rate, or *the rate the temperature of a parcel would decrease under adiabatic conditions* [Marshall and Plumb, 2008]. By solving equation 2.22 for dT/dz, the conditions of 2.21 can be rewritten as

$$\left.\begin{matrix}\text{unstable}\\ \text{neutral}\\ \text{stable}\end{matrix}\right\} \quad \text{if } \Big(\frac{dT}{dz}\Big)_E \quad \begin{cases} < -\Gamma_d \\ = -\Gamma_d \\ > -\Gamma_d \end{cases} \tag{2.24}$$

Stating that the atmosphere is statically unstable if the absolute temperature decreases with altitude faster than the dry adiabatic lapse rate. If the absolute temperature decreases slower than the dry adiabatic lapse rate, it is classified as a statically stable atmosphere. In the case where the absolute temperature changes accordingly to the dry adiabatic lapse rate with height, the atmosphere is statically neutral.

A weak, stable layer does not suppress the vertical mixing enough to define it as an inversion. We therefore need to constrain our definition of an inversion, such that weak stabilites are removed. Following Wolf et al., 2014, we will for this thesis define an inversion to be occurring when the temperature gradient of the observed air column is larger than zero,

$$\frac{dT}{dz} > 0 \tag{2.25}$$

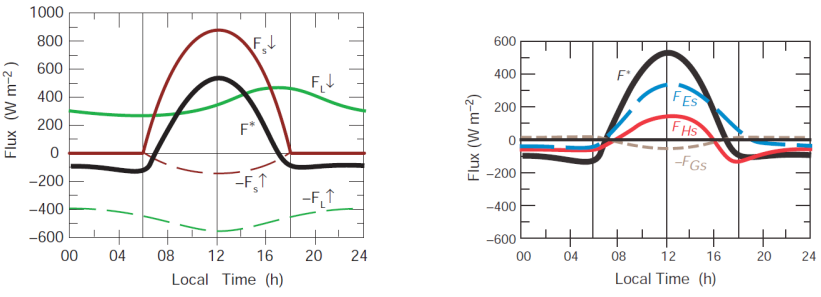
Meteorological conditions for a stable layer

Inversions can be divided in to two categories, elevated and ground inversions. We will focus on ground inversion as these are the type of inversions that trap the pollutants in the surface layer.

Consider the surface energy balance (SEB) with all terms having positive direction downwards:

$$F_S\downarrow - F_S\uparrow + F_L\downarrow - F_L\uparrow = \; F_{Hs} + F_{Es} + F_{Gs} \tag{2.26}$$

The terms $F_S\downarrow$ and $F_S\uparrow$ are the incoming and outgoing shortwave radiation fluxes, respectively, and $F_L\downarrow$ and $F_L\uparrow$ are the incoming and outgoing longwave radiation fluxes. On the right hand side, F_{Hs}, denotes the latent heat flux, and F_{Es}, denotes the sensible heat flux. The sensible heat flux heats the boundary layer directly, whereas the latent heat flux only contributes to heating and/or increase in potential energy when water vapor condensates in clouds. Lastly, F_{Gs}, is the heat conduction into the ground.



(a) (b)

Figure 2.4: The surface radiative fluxes during a cloud free day, with positive axis downwards. (a) show left hand side of Equation 2.26, (b) show right hand side of Equation 2.26

On a cloud free day with fair weather, the radiative fluxes on the left hand side of the equation depend on the incoming solar radiation, which varies with time,

latitude and season. The shortwave radiation flux down to the ground will be at its maximum when the sun is at its highest point in the sky. So will the outgoing shortwave flux, but at a much lower magnitude. The outgoing longwave radiation flux, $F_L \uparrow$ follows the Stefan Boltzmann law, and will therefore depend on the skin temperature. The downward longwave radiation flux, $F_L \downarrow$ will depend on the air temperature, which is at it highest right before sunset.

The sum of the fluxes on the right hand side of Equation 2.26 must equal that of the left hand side, but the magnitude and direction of the fluxes depend on the air temperature and relative humidity. For the same conditions previously discussed, the curves will look like the ones sketched in Figure 2.4b. The sensible heat is dependent on the temperature gradient between the air temperature and the skin temperature. When the air temperature is warmer than the skin temperature, the sensible heat flux will point upwards from the ground, whereas the flux will point downwards into the ground when the skin temperature is warmer than the air temperature. At daytime with high relative humidity, the latent heat flux will point upwards from the ground, and at nighttime with high relative humidity, the latent heat flux will point downwards to the ground.

During a ground inversion, we would expect radiative cooling at the surface. Such inversions occur frequently at cool nights with few clouds and little to no wind, when $F_L \uparrow$ cools the skin temperature faster than the air temperature. These inversions usually break up at sun rise due to heating of the ground, but during winters in higher latitudes, these types of inversions can last for many days. Ground inversions can also be enhanced due to the terrain. In valleys during cold nights with fair weather, the cold temperature can accumulate to create a cold pool. The cold pool will act as a lid, and will inhibit warm air from entering.

High pressure systems at wintertime is also associated with ground inversions. With fair weather and snow covering the ground, the skin temperature is much colder than the air above. These inversions can last for days or even weeks if the high pressure stagnates. The longer an inversion episode lasts, the more NO_2 is accumulated, and the more damage it will cause to people exposed to the pollutant.

CHAPTER 3

Study area and data

3.1 Study area and study period

In this thesis, we are studying the impacts of inversions on the NO_2 concentration in Oslo and Bergen. These are the two most populated cities in Norway and both cities have recorded high levels of NO_2 during inversion episodes. The available observational data ranges from January 2001 to December 2010 for Oslo, and from January 2011 to December 2020 for Bergen. The model data ranges from January 2001 to December 2020 for both study areas, as this allows us to study inversion trends and variability over a longer period.

Oslo

Oslo is the capital city of Norway, located on the eastern part of the country, and the Oslo municipality currently inhabits 698 660 people [Statistics Norway (SSB), 2021b]. From Table 3.1 we gather that the number of cars running on gasoline has decreased as the registered number of electric cars has increased. Registered diesel cars increased from 2008 to 2015, but decreased from 2015 to 2020. There were approximately 201 000 gas- and diesel cars registered in 2008. This number increased to ∼217 000 in 2010, and ∼241 000 in 2015, before it decreased to ∼164 000 in 2020.

	2008	**2010**	**2015**	**2020**
Gas	153 633	143 178	130 044	90 569
Diesel	47 901	74 009	111 226	73 859
Electric	311	437	11 384	58 622

Table 3.1: Registered number of cars in Oslo running on gasoline, diesel and electricity for 2008, 2010, 2015 and 2020, provided by Statistics Norway [Statistics Norway (SSB), 2021a].

The measuring stations for the meteorological data and NO_2 concentration is located at Valle Hovin and Alnabru, respectively, as shown in Figure 3.1. They are located on the east side of Oslo, in the Grorud Valley to be more specific, and are distanced approximately 2.5 kilometers from each other. The Grorud Valley spans out in the Northeast-Southwest Direction, and is at its widest

on the western side and narrows on the eastern side. The cross section of the valley between Valle Hovin and Alnabru is approximately 3.5 kilometers wide, and is approximately 2.5 kilometers at its narrowest. Årvollåsen at 360 meters above mean sea level (AMSL), and Kolsås at 250 meters AMSL is located on the northern side of the valley. On the southern side of the valley, we find Lutåsen at 360 meters AMSL. The valley itself is located approximately 130 meters AMSL, and we observe in Figure 3.1 that the stations are located at the deepest part of the valley. The area between them are slightly elevated, and the Alnabru station measuring the NO_2 concentrations is located 400 meters north of the main freeway and one of the main roads in and out of Oslo, E6. The station at Valle Hovin is relatively undisturbed from the terrain as it is located next to an open field.

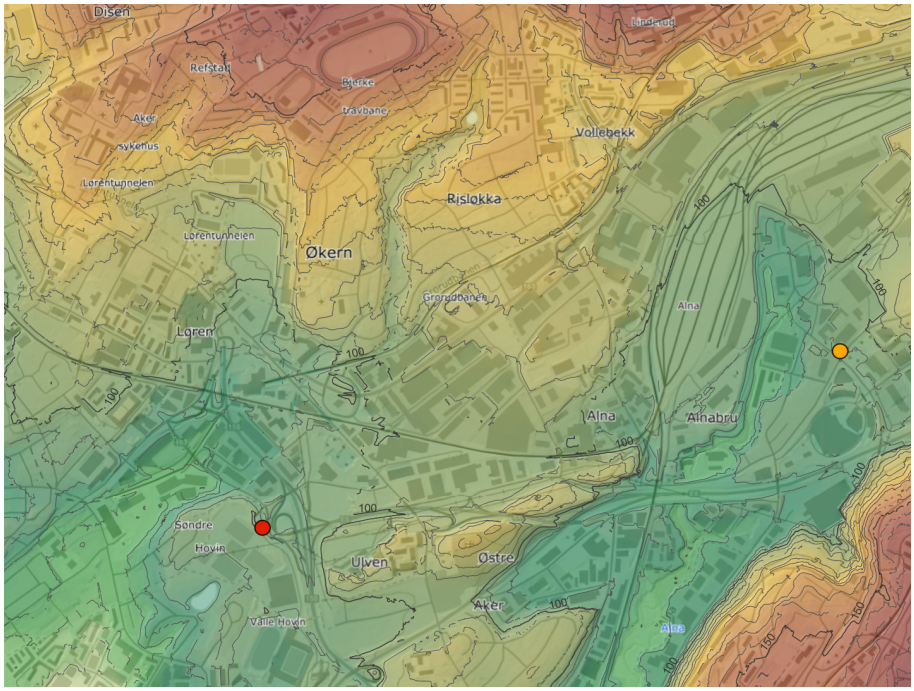


Figure 3.1: The topography surrounding the observational stations. The red dot shows the location of the station measuring the meteorological data at Valle Hovin, and the orange dot shows the station measuring the NO_2 concentration at Alnabru. The color map displays the height difference. The map was gathered from høydedata.no, and the height curves were downloaded from kartkatalog.geonorge.no with an interval of 5 meters, marked by number for every 50 meters.

The continental climate in Oslo has four defined seasons, and the relative humidity is high due to the Oslo fjord. The winters are usually cold, and with average temperatures between 0 and -5 °C, and the summers are usually warm with average temperatures between 16 to 20 °C [(MET), 2021]. For our study period covering 2001 to 2010, the highest average temperature was registered in 2006, with a value of 7.7 °C, and the lowest average temperature was measured

in 2010, with a value of 4.8 °C. The winter of 2010 deviated almost -2.5 °C from the normal 1961-1990 temperatures, and the winter of 2008 deviated 7 °C.

Bergen

Bergen is located on the western coast, and is the second largest city in Norway. Bergen municipality currently have 286 567 inhabitants, and in 2020, it was registered approximately 15 300 commuters in and out. Table 3.2 shows the registered number of cars fueled by gas, diesel and electricity, and like in Oslo, the number of cars running on gas has decreased as the number of electric cars has increased. Summing up the number of gas- and diesel cars we find that the number ranged between 102 000 and 105 000 from 2008 to 2015, before dropping down to 76 400 in 2020.

	2008	**2010**	**2015**	**2020**
Gas	76 334	69 579	58 118	39 312
Diesel	29 434	32 982	46 863	37 062
Electric	149	155	6 716	27 337

Table 3.2: Registered number of cars in Bergen running on gasoline, diesel and electricity for 2008, 2010, 2015 and 2020, provided by Statistics Norway [Statistics Norway (SSB), 2021a].

The measuring stations for Bergen are placed on Danmarks Plass and Florida/Geofysen. The station on Danmarks Plass measures the NO_2 concentration, and the two stations placed at Florida/Geofysen provides with the meteorological data. They are placed approximately 800 meters apart, on either side of Store Lungårdsvannet. The Bergen valley is trapped between Løvstakken on the southwest/west side and Ulriken on the northeast/east side. The terrain, shown in Figure 3.2 is more defined than in Oslo, and the mountains are steep with the height of Løvstakken being 477 meters AMSL, and Ulriken being 643 meters AMSL. The mountains shelter the valley from harsher winds [Wolf et al., 2014], but hinders incoming solar radiation in the winter due to the latitude of the city. On the shortest of days, Bergen only experience a couple of hours of direct sunlight. The valley runs from northwest toward southeast, and a smaller valley also runs from east toward west between Ulriken and Fløyfjellet, and ends up at Store Lungårdsvannet. The cross section between Florida/Geofysen and Danmarks plass is approximately 2 kilometers wide.

One of the main roads in and out of Bergen city passes through Danmarks plass, and the area is thus heavily polluted. The station measuring the temperature at 40 meters is placed on top of Geofysen, the building for Geophysical Institute at the University of Bergen. The station measuring the wind and temperature at 2 meters is placed on the ground next to the building. The building proximity effect on the temperature measurements are complex. It may lead to decrease in temperature through shading, but increase in the temperature at nighttime due to release of heat energy [Zhan and Lan, 2017, Ha et al., 2020]. This is something to consider when discussing the results.

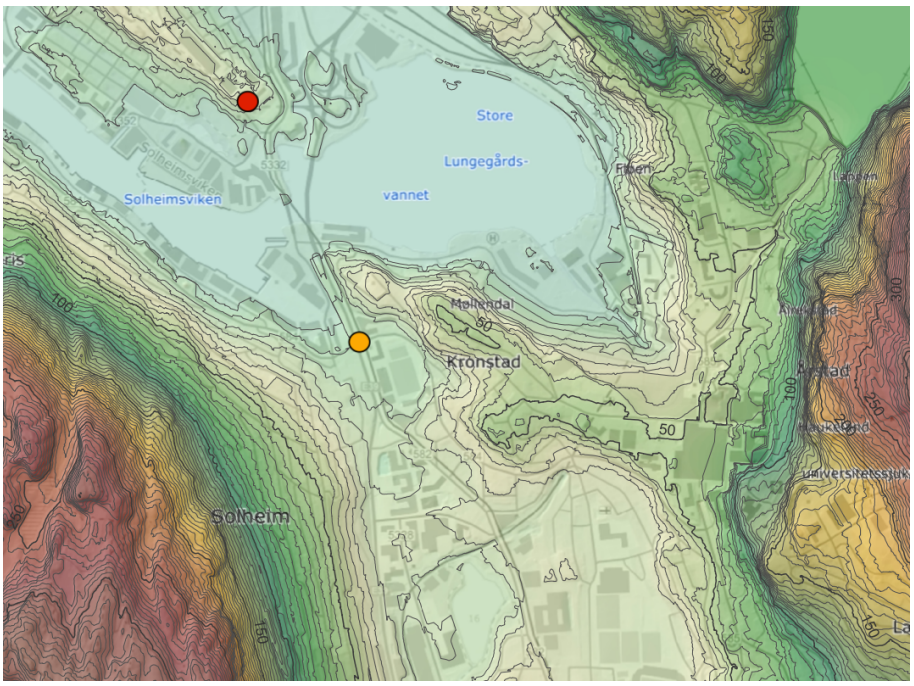


Figure 3.2: The topography surrounding the observational stations in Bergen. The red dot shows the location of the station measuring the meteorological data, and the orange dot shows the of the station measuring the NO_2 concentration. The colour map displays the height difference. The map was gathered from høydedata.no, and the height curves were downloaded from kartkatalog.geonorge.no. The height curves have an interval of 5 meters, and is marked by number for every 50 meters.

The maritime climate in Bergen provides with mild winters and high pressures in the Norwegian Sea advecting warmer southweasterlies toward the city. The warm and wet winters allows for a well mixed layer, but these conditions will be interrupted by the occurrence of cold and dry episodes. During these episodes we will experience the temperature inversions causing the accumulation of air pollutants.

3.2 Observational data

Oslo

Meteorological observations at Valle Hovin has been provided by Susanne Lützenkirchen at Oslo kommune. For the NO_2 data, we collected data from NILU's API. We chose the Alnabru station because of its proximity to Valle Hovin, and with the data availability in mind. The data availability of the observations from Valle Hovin and Alnabru is illustrated in Figure 3.3. Here, we calculated the average for each month, and the orange areas displays months that contained no data. The station at Valle Hovin has no data for Febuary, May, June, July and August in 2001, June, July and August in 2002, all months

except January and December in 2006 and in September, 2010. We lack NO_2 observations from Alnabru in June, July and August of 2002, from January to August in 2008 and in September and October in 2010.

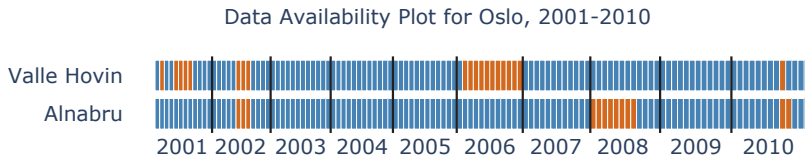


Figure 3.3: Data availability at the two observation stations in Oslo, in the period from 2001 to 2010. The blue areas indicate months containing data and the the orange areas indicate months with no data.

We treat the lack of data by discarding averages that has less than 80% data coverage. For the case of yearly averages, we will discard looking into 2001, 2002 and 2006 for the meteorological observations. For Alnabru, we will discard yearly averages for 2002 and 2008. For the seasonal average of the meteorological observations, we will discard the winter average of 2005/2006, the spring average of 2001 and the autumn average of 2010. For the observed NO_2 concentration we will discard the winter average of 2007/2008 and the autumn average of 2010.

Bergen

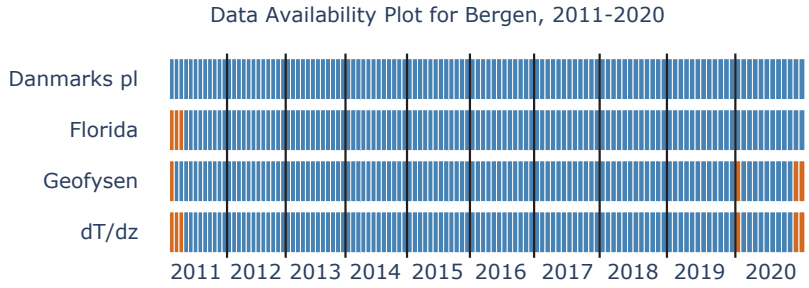


Figure 3.4: Data availability at the three observation stations in Bergen, in the period from 2011 to 2020. dT/dz is a product of data from Florida and Geofysen. The blue areas indicate months containing data and the the orange areas indicate months with no data.

The wind data and temperature measurements at 2 meters were gathered from the API of the Norwegian Meteorological Institute, Frost, and were extracted from the Florida station. The temperature at 40 meters were measured using MTP-5HE, a scanning temperature profile radiometer from ATTEX, on top of the Geofysen building. The data was provided by Igor Ezau. We gathered the data on the NO_2 concentrations from NILU's API, chosing to use the station at Danmarks plass.

The data availability plot in Figure 3.4 show that the NO_2 data is available for the entire period. The Florida station lacks data in the beginning of 2011, and the MTP-5HE sensor on top of Geofysen lack data in January 2011, and January, November and December 2020. The data coverage is better than in Oslo. However, due to our set limitation of at least 80% data coverage, we will refrain from calculate the yearly average dT/dz and thereby also inversion intensities for 2011 and 2020. The average for the winter of 2019/2020, the spring of 2011 and the autumn of 2020 will also be discarded.

3.3 Model data

The model data was collected from NORA3, a non-hydrostatic high-resolution hindcast model from the archives of the Norwegian Meteorological Institute. Its domain, displayed in Figure 3.5, includes Scandinavia, Finland, Iceland and the United Kingdom, as well as the North, Norwegian and Barents Sea, a total of 900×1500 grid points [Haakenstad et al., 2021]. The model is a product of the numerical weather prediction model HARMONIE-AROME yielding a downscaling of the ERA5 reanalysis, and has a 3 kilometer grid resolution. The vertical resolution consists of 65 levels, ranging from approximately 12 meters up to 10 hPa. This has allowed for a more improved representation of wind fields, specially in mountain areas [Solbrekke et al., 2021]. NORA3 uses a Lambert conformal conic projection with central meridian grid at 42°W and origin at 66.3°N.

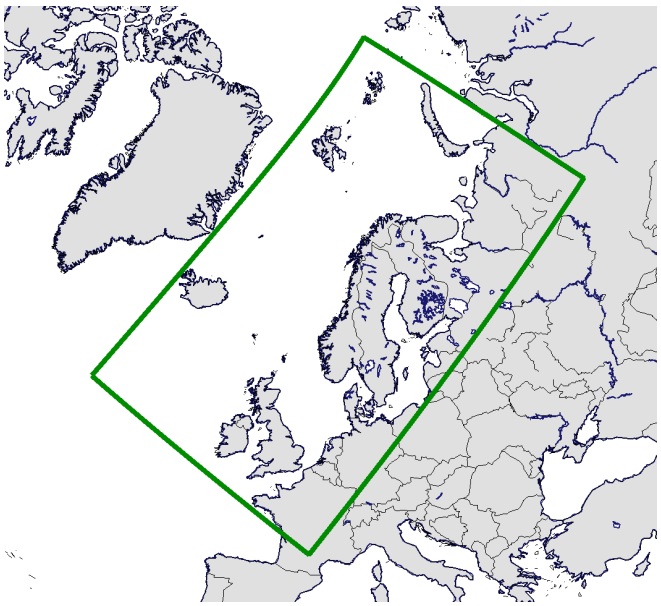


Figure 3.5: The domain of NORA3, covering Scandinavia, Finland, Iceland and the United Kingdom, as well as the North Sea, the Norwegian Sea, and the Barents Sea [Haakenstad et al., 2021]. Figure used with permission by Hilde Haakenstad.

Through a surface analysis, the surface field has been adjusted to observations.

The module used to represent the land and ocean surface platform in HARMONIE-AROME is the SURFEX model. It describes the surface fluxes of four different types of surfaces: nature, urban areas, inland water and ocean [Masson et al., 2013]. Each surface grid cell receives basic atmospheric meteorology, such as the air temperature, the horizontal wind components, the longwave radiation, the shortwave direct radiation and the diffuse radiation.

The microphysics in HARMONIE-AROME consists of cloud ice, snow and graupel and hail, and uses a run option that enhances the cloud physics in colder environments. The turbulence in HARMONIE is represented with RACMO Turbulence (HARATU), where it uses a prognostic equation that combines turbulent kinetic energy with a diagnostic length scale, determined by the Richardson's number within a single parcel [Meijgaard et al., 2012, Lenderick and Holtslag, 2004]. The rapid Radiative Transfer Model has been used for the parameterization of the longwave radiation, and the ECMWF operational shortwave scheme has been used for the parameterization of the shortwave radiation.

The resolution of NORA3 might be too coarse to identify key characteristics of the topography near our study areas, and as topography highly affect the local meteorology, the model may not be able to represent all local meteorological features found at the measuring sites. The Bergen valley at its narrowest is approximately 1 kilometer wide, and the Grorud valley near Alnabru is approximately 1.5 kilometer wide. It is also important to have a good parameterization of the urban areas, as the temperature measurements (specially in Bergen) are close to roads and buildings. Lastly, because we are studying the inversions close to ground, it is important that the model has a good vertical resolution.

The model runs are divided in to four runs for every 24 hours, starting at 00:00, 06:00, 12:00 and 18:00. Each run lasts for nine hours, where the first three hours of the run is used to stabilize the initialization. Some of the variables are only available for every three hours, such as temperatures at model levels, which (as we will show) is important in the calculation of inversions. During the work on this thesis, NORA3 data have been available from 1984 to 2021. When the data archive is finished data will cover the period 1979-2021. The variables collected for this thesis is shown in Table 3.5.

3.4 Data management

Observations, Oslo

The dataset for Valle Hovin recieved by Suzanne Lützenkirchen consisted of wind speed, wind gust, wind direction, temperature measurements from 2 and 25 meters, stability between 8 and 25 meters, relative humidity and precipitation. Relative humidity and precipitation was removed as we expect their correlation with NO_2 to be small. We also removed wind gust from the data set. This leaves us wind speed, wind direction, temperature measurements from 2 and 25 meters and stability between 8 and 25 meters. We obtain two variables for dT/dz. The first one by calculating the temperature difference between 2 and 25 meters and divide it by the height difference, and the second one by dividing

the measured stability and divide it by the height difference. The second dT/dz between 8 and 25 meters may be referred to as the *directly* measured dT/dz. As explained in Section 2.2, we define our inversion by drawing the positive values of the two dT/dz. The summary of the finale variables used for the thesis can be found in Table 3.3

To avoid unrealistic values, the dT/dz between 2 and 25 meters were constrained to range between $\pm 0.25\ ^\circ C/m$, as the temperature change with respect to height rarely exceeds this range. The wind speed was not allowed to exceed 20 m/s.

Table 3.3: The observational data in Oslo applied to this study

variable	**description**	**location**	**units**
NO2	Hourly average NO_2 concentration	Alnabru	$\mu g/m^3$
wind direction	Hourly measured wind direction	Valle Hovin	$0^\circ(360^\circ)$
wind speed	Hourly measured wind speed	Valle Hovin	m/s
temp 2m	Hourly average temperature at 2 meters	Valle Hovin	$^\circ C$
temp 25m	Hourly average temperature at 25 meters	Valle Hovin	$^\circ C$
direct dT/dz	Hourly average temperature gradient between 8 and 25 meters	Valle Hovin	$^\circ C/m$
direct dT/dz > 0	Hourly average temperature inversions between 8 and 25 meters	Valle Hovin	$^\circ C/m$
dT/dz	Hourly average temperature gradient intensity between 8 and 25 meters	Valle Hovin	$^\circ C/m$
dT/dz > 0	Hourly average temperature inversions intensity between 2 and 25 meters	Valle Hovin	$^\circ C/m$

Observations, Bergen

The data set for Bergen consists of the wind data and temperature at 2 meters from Florida, the temperature at 40 meters from Geofysen and the NO_2 concentration from Danmarks plass. We calculated the temperature gradient by dividing the temperature difference between 2 and 40 meters by the height difference. By extracting only the positive values, we also obtained the data for the inversions. We applied the same constraints on the wind data and temperature gradient/inversions as we did for Oslo. The variables can be found in Table 3.4

Table 3.4: The observational data in Bergen applied to this study

variable	description	location	units
NO2	Hourly average NO_2 concentration	Danmarks plass	$\mu g/m^3$
wind direction	Hourly measured wind direction	Florida	$0°(360°)$
wind speed	Hourly measured wind speed	Florida	m/s
temp 2m	Hourly average temperature at 2 meters	Florida	$°C$
temp 40m	Hourly average temperature at 40 meters	Geofysen	$°C$
dT/dz	Hourly average temperature gradient intensity between 2 and 40 meters	Florida, Geofysen	$°C/m$
dT/dz > 0	Hourly average temperature inversion intensity between 2 and 40 meters	Florida, Geofysen	$°C/m$

Model

The longitude and latitude of Valle Hovin are given by [10.7, 59.9] in Decimal Degrees (DD) Cooridnates, while the longitude and latitude of Florida are given by [5.3, 60.4]. Because the model uses Lambert Projection, we converted the DD coordinates to [X, Y] and chose the grid point closest to our study area. We obtained the data for the land area fraction, in order to check the land

coverage in the model. For both Oslo and Bergen, the value of the land area fraction was 1.

In order to calculate the temperature gradient and the inversions in the model, we obtained the data for temperature at 2 meters, and the air temperature for each model hybrid level. In order to compare the model results to the observations, we needed to convert the hybrid levels [Müller et al., 2017] to meters. The pressure at each model level is given by

$$p(n,k,j,i) = ap(k) + b(k) \cdot p_s(n,j,i) \tag{3.1}$$

where $ap(k)$, $b(k)$ and $p_s(n,j,i)$ are given in the NORA3 data set, n is the time index, k is the vertical index, and i and j are the horizontal indices. This further allows us to use the hypsometric equation in order to find the geopotential height difference between two model levels

$$z_2 - z_1 = \frac{R \cdot \bar{T}_v}{g} \cdot \ln\left(\frac{p_1}{p_2}\right) \tag{3.2}$$

R is the specific gas constant for dry air and is equal to 287.058 $J/(kg \cdot K)$, $\bar{T}_v$ (K) is the mean temperature calculated from the temperatures at k and $k+1$, g is the gravity constant and is equal to 9.81 m/s^2, and lastly, p_1 and p_2 are the pressure at k and $k+1$, respectively, with units in Pascal. For the initial level, k_0, the initial temperature, T_0, is given in terms of the temperature at surface, T_s and the initial pressure, p_0, is given as p_s. One thing to note is that the calculated height varies in time because $p_s(n,j,i)$ varies in time. To take care of this, we interpolated the height such that it became independent of time. Further, we chose the air temperature for 25 meters in Oslo and 40 meters for Bergen.

We want to study the energy budget in the model. The energy fluxes were recalculated from the orginal flux variables integrated over the total model run time (9 hours). The cloud area fraction was calculated from the variables *high type cloud area fraction*, *medium type cloud area fraction* and *low type cloud area fraction*, using the following equation

$$\bar{b}_k = 1 - \prod_{j=1}^{k} \frac{1 - \max(b_{j-1}, b_j)}{1 - b_{j-i}} \tag{3.3}$$

where k denotes the number of layers (in this case 3), and $\bar{b}_k$ is the total cloud area fraction [Sundqvist et al., 1989] for the column. If the cloud coverage is monotonically increasing or decreasing with height, we will find that $\bar{b}_k = \max(b_j)$. If not, then the equation will assume a random overlapping, and we will find that $\bar{b}_k > \max(b_j)$. As we have three layers of cloud area fraction, $k = 3$.

Table 3.5: Variables obtained from NORA3

variable name	units	Time resolution
Land area fraction		
Air temperature 0 m (skin temperature)	$^{\circ}C$	1 hr
Air temperature 2 m	$^{\circ}C$	1 hr
Air temperature at 25 m for Oslo and 40 m for Bergen	$^{\circ}C$	3 hr
dT/dz	$^{\circ}C$/m	3 hr
dT/dz > 0	$^{\circ}C$/m	3 hr
Surface net downward shortwave flux	W/m^2	1 hr
Surface net downward longwave flux	W/m^2	1 hr
Surface downward latent heat flux	W/m^2	1 hr
Surface downward sensible heat flux	W/m^2	1 hr
Surface downwelling shortwave flux in air	W/m^2	1 hr
Surface downwelling longwave flux in air	W/m^2	1 hr
Surface upwelling shortwave flux in air	W/m^2	1 hr
Surface upwelling longwave flux in air	W/m^2	1 hr
Wind direction	$0^{\circ}(360^{\circ})$	1 hr
Wind speed	m/s	1 hr
Cloud area fraction		1 hr

3.5 Statistical tools

This section covers the statistical tools used for the study.

Correlation

One of the most useful tools in statistics is to calculate the joint behavior of two variables. The correlation determines to which degree two variables are linearly related. For this study, we only have a sample of data to investigate the relationship between NO_2 and inversions, thus the correlation calculated will be a sample correlation [Jay L. Devore, 2012a].

The sample correlation coefficient

Consider a sample of n, given by $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. If the x's increases with the increase of the y's, then the relationship between x and y is determined to be positive. Reversely, if the x's decrease with the increase of the y's, the relationship is negative. This part of the relationship between x and y can be described in terms of S_{xy}:

$$S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \tag{3.4}$$

where $\bar{x}$ and $\bar{y}$ is the mean of x and y, respectively. Note that if the product of $(x_i - \bar{x})(y_i - \bar{y})$ is positive, then this indicates that either both x_i and y_i is larger than the mean, or smaller than the mean. In the case where one is smaller than the mean, and the other is larger than the mean, S_{xy} will be negative.

In order to describe the magnitude of the relationship however, S_{xy} must be modified, and as such, we obtain the equation for the sample correlation coefficient:

$$r = \frac{S_{xy}}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}} = \frac{S_{xy}}{\sqrt{S_{xx}}\sqrt{S_{yy}}} \tag{3.5}$$

Note that r is not affected by the units of x and y, and its value ranges from -1 to 1. If the relationship between x and y is strongly negative, then r will be close to -1, and, reversely, if the relationship between x and y is strongly positive, then r will be close to 1. In the case where there is no evident relationship between x and y, r will be close to 0. Another property of r that is important to note, is that in the case of a simple linear regression between independent variable x and dependent variable y, the r^2 tells us the proportion to which y can be described by x. This is known as the coefficient of determination.

Linear Regression

In addition to understanding the strength and direction of the relationship between variables, it is also interesting to understand how they are related. In this section, we will investigate the definition of linear regression, and further use this tool to see how NO_2 is related to inversions and other variables later on.

Simple linear regression model

We start of by defining a simple linear regression as *a statistical tool to summarize the relationship between two quantitative variables so that one variable (dependent variable, y) can be predicted from the other* (independent variable, x) [Xu, 2019a].

It is worth to note that in the making of a probabilistic relationship, there is still uncertainty to what the value of a dependent variable will be when the independent variable is fixed. Because the uncertainty is not know, the dependent variable, denoted by Y, for a fixed independent variable, x, is random [Jay L. Devore, 2012b]. This can be described as

$$Y = f(x) + \epsilon \tag{3.6}$$

where $f(x)$ is some deterministic function of x and ϵ is a random error, and incorporates all the variations that are not explained by the variation in x. In the case where $\epsilon = 0$, then y would fully be determined on x, suggesting that the variations in y is only caused by the variations in x. If $\text{var}(\epsilon)$ is positive, then the point (x, y) falls above $f(x)$, and if $\text{var}(\epsilon)$ is negative, then the point (x, y) falls below $f(x)$. We make the assumption that the mean of ϵ is zero, implying that we expect (x, y) to follow $f(x)$. This is however, rarely happens.

Now that we are familiar with the rv-term, we can determine a s*imple linear regression model* (SLR). For a SLR, the deterministic relationship between dependent variable y and independent variable x is given by

$$y = \beta_0 + \beta_1 x + \epsilon \tag{3.7}$$

where the term β_0 is the intercept coefficient, and the given value of y when $x = 0$, and β_1 is the slope coefficient, determining how much y increases in respect to one unit increase in x.

Multiple linear regression model

As we now have introduced simple linear regression, we want to build a probabilistic model that relates the dependent variable to more than only one independent variable, also known as multiple linear regression model (MLR) [Jay L. Devore, 2012c]. For two or more independent variables $x_1, x_2, \ldots, x_k$, then the MLR is given by

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_k x_k + \epsilon \tag{3.8}$$

Where again β_0 is the intercept, and $\beta_1, \beta_2, \ldots, \beta_k$ gives us the increase in y for one unit increase of $x_1, x_2, \ldots, x_k$, respectively, and ϵ is the random error term.

Applying MLR to our data

When building a linear regression model it is important to note that the independent variable can be used to describe trends and variations in the dependent variable, but not vice versa. In MLR, the independent variables should have strong correlation, but weak correlation with other independent variables, and each independent variable's relationship with the dependent variable is not affected by the value of the other independent variables [Xu, 2019b].

In our study, we have built the variable `dt/dz > 0` on `dt/dz`. Because these two variables are non-independent of each other, we need to make a conscious decision to disregard one of the them. We are mainly interested in how the *inversions* affect the NO_2, and so we disregard `dt/dz` in our MLR. Because `dt/dz > 0` does not contain any values below zero, this is important to note when studying the model's performance.

It was also important to handle all the missing data. Data from different observational stations had different time gaps, and so it was important to align these time gaps with each other to not disturb the performance of the model.

Hypothesis testing

By defining a hypothesis in a study, one makes the assumption that the study may give a specific outcome. An initial hypothesis is called a *null hypothesis* and is denoted by H_0. The alternative, or rather the opposite of the null hypothesis, the *alternative hypothesis*, is denoted by H_a.

Testing if the sample strongly contradicts the null hypothesis or not gives us an indication of whether or not the stated null hypothesis is reasonable. We reject the null hypothesis only if there is strong evidence suggesting that it is wrong. If there is not, we will continue to assume that our null hypothesis is plausible.

For this thesis, we will test our hypothesis using a t-test. Let our null hypothesis be $\mu = \mu_0$, and our alternative hypothesis $\mu \neq \mu_0$, and let n be the number of sample of observations, $x_1, x_2, \ldots, x_n$. The t-statistic is given by

$$t_c = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} \tag{3.9}$$

where $\bar{x}$ is the sample mean, μ is the estimated sample mean from our null hypothesis and the term $(S/\sqrt{n})$ is the estimated standard deviation of $\bar{x}$. The critical value, $t_{\alpha,n-1}$, can be found reading of a t-table, and by definition, our null hypothesis can be rejected if either $t_c \geq t_{\alpha,n-1}$ or $t_c \leq -t_{\alpha,n-1}$ [Jay L. Devore, 2012d].

In this thesis, we will conduct an SLR or a MLR, and test whether the linear regression is of significance. For a SLR, we will test if the slope $\beta_1 = 0$, and for a MLR, we will test if $\beta_1 = \beta_2 = \ldots = \beta_k = 0$.

CHAPTER 4

Results and discussion

In this chapter, we will present and discuss the results of the study. We will start of by presenting the observational data from Valle Hovin and Alnabru in Oslo, and from Danmarks plass and Geofysen/Florida in Bergen, and discuss the relationship between the NO_2 concentrations and the meteorological variables. Further, we will compare the model output of NORA3 to the observations to see how well the hindcast performs. We will look into periods where the model is able to yield the same intensity as the observed inversions, and other periods where the modelled inversions differs, to study the surface energy balance in the model. Lastly, we will compare the trends of the observed NO_2 concentrations to the trends of the modelled inversions.

4.1 Assessing the observational data

Oslo

Because the occurrence and disappearance of inversions happens in a matter of hours, we decided to show both the hourly and the monthly average concentrations of NO_2 in order to compare the plot against inversions later. We present the measured hourly average NO_2 and the monthly average in Figure 4.1.

Both show evidence of seasonal variations with higher concentrations during the winter months, and lower concentrations during the summer months. The hourly averaged concentrations has a maximum value of 563.8 $\mu g/m^3$, which was measured in January in 2003. The average of the maximum value for each year is 391.03 $\mu g/m^3$, and we observed that 2002 was the year with the lowest maximum value of 281.7 $\mu g/m^3$. During the summer months (June, July and August), the maximum concentration averages around 100 $\mu g/m^3$.

The plot for the monthly average show that there are higher concentrations during the winter months between December, 2002 and February, 2003, and between December, 2005 and February, 2006, compared to the other winters in the period. For these winters the maximums reaches above 75 $\mu g/m^3$. The beginning of 2010 also show higher concentrations just under 75 $\mu g/m^3$. The winter of 2004/2005 was a year with lower winter concentrations of NO_2, with a maximum of approximately 55 $\mu g/m^3$, than compared to the other years. Both the summer of 2001 and 2010 had low concentrations, with a minimum close

to 25 $\mu g/m^3$. The other years has a minimum during the summer months at around 35 $\mu g/m^3$.

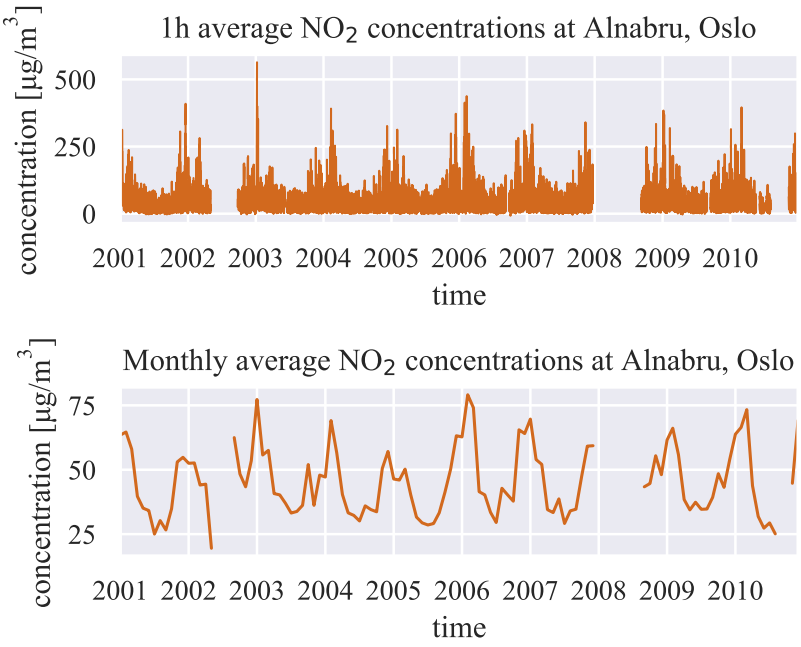


Figure 4.1: Hourly average (top) and monthly average (bottom) NO_2 concentrations measured at Alnabru, Oslo, between 2001 and 2010 with units [$\mu g/m^3$].

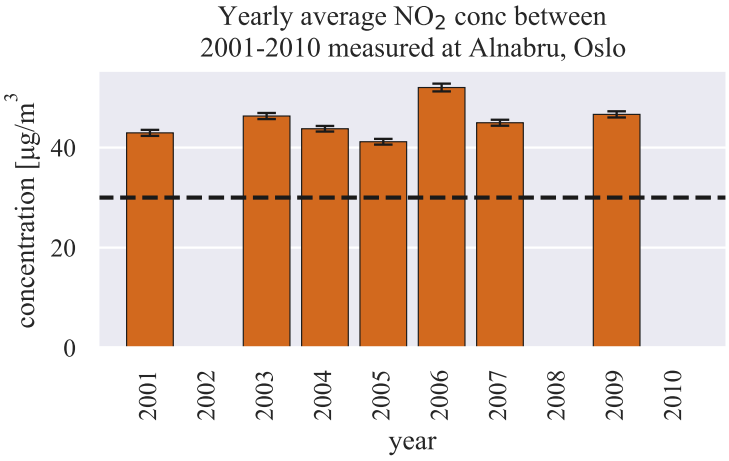


Figure 4.2: Yearly average NO_2 concentrations measured in [$\mu g/m^3$] at Alnabru, Oslo, with 95% confidence interval around the average. The dashed line indicates the annual recommended health limit [Norwegian Institute of Public Health (FHI), 2020].

Figure 4.2 displays the yearly average. From the figure, we gather that in Oslo, the yearly average surpasses the annual recommended health limit every years and was especially high during 2006, obtaining a value of over 50 $\mu g/m^3$. 2005 was the year with the lowest average concentrations, being around 40 $\mu g/m^3$. Because the number of observations are high, the uncertainty of the yearly average is small.

The seasonal variation of NO_2 is found in Figure 4.3, shown as a boxplot. We find the higher concentrations in December, January and February, while it shifts towards lower concentrations in the summer. Another thing to note is that the distribution widens during the winter months, and is much more narrower in the summer months. This suggests that the monthly average concentrations during winter varies more than it does during summer.

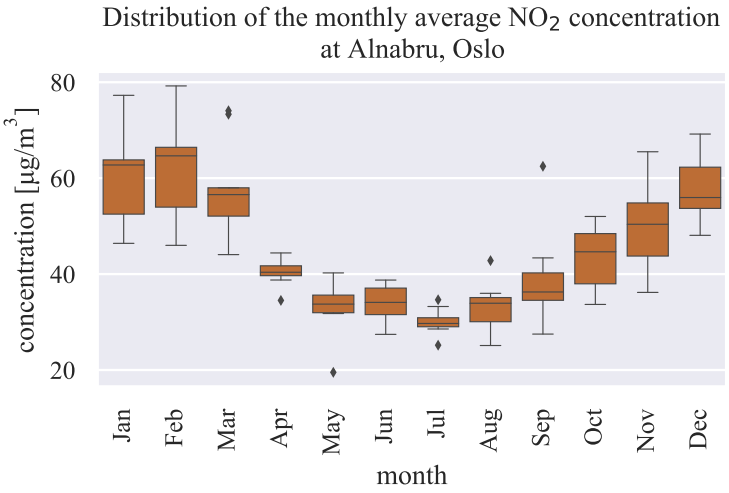


Figure 4.3: The distribution of monthly averages for each month, containing data from 2001 to 2010. The line within the box shows the median, the box' edges are determined by the 1st and 3rd quantile, and the error bars shows the maximum and minimum value. The diamonds indicates the outliers.

In the following, the inversion data for Oslo will be presented, starting with the temperature difference obtained from an altitude of 2 and 25 meters. We will further show the temperature difference measured between 8 and 25 meters in altitude, and compare the temperature gradients and inversions obtained from the two measurements.

In Figure 4.4, the plot for the hourly observed average dT/dz is very "noisy", but we can still see some seasonal variation. During summer, dT/dz is more negatively shifted than during the winter. This is a result of varying solar radiation, as in summer we have surface heating, causing more vertical mixing, and in the winter we have surface cooling. This is an important cause to why we observe the low NO_2 concentrations in Figure 4.1 and 4.3 during the summer. Because the data was restricted to range between -0.25 $°C/m$ and 0.25 $°C/m$, the maximum and minimum value ranges accordingly. The summer of 2003

shows the overall strongest negative dT/dz, but comparing this to Figure 4.1, this does not seemingly affect the NO_2 concentrations any more than any other years.

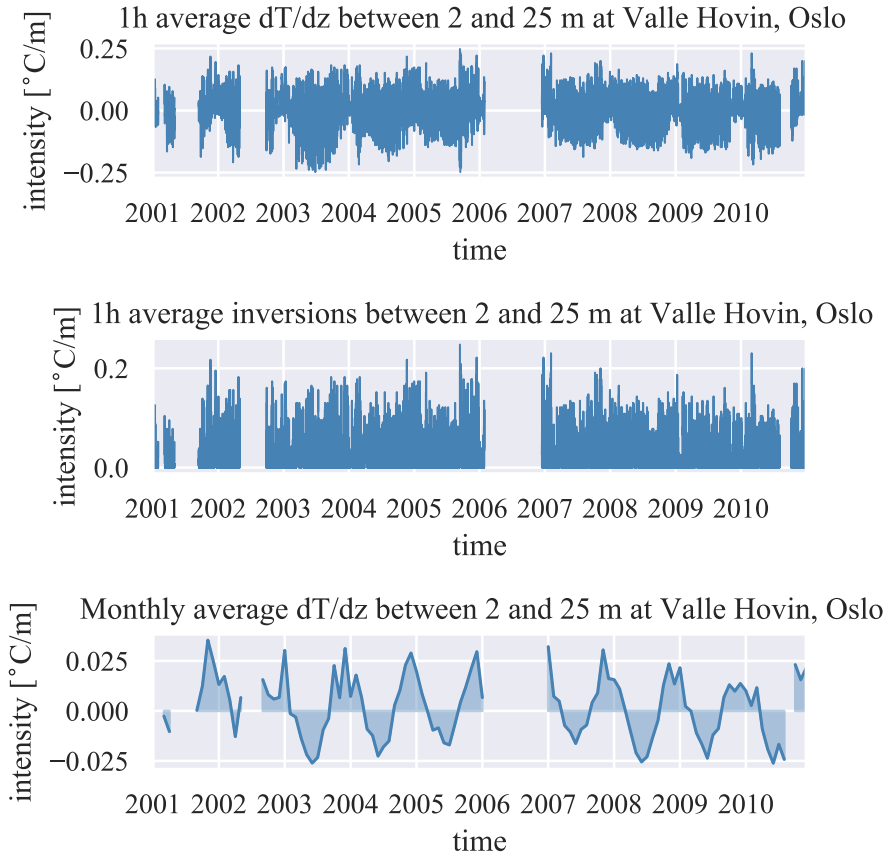


Figure 4.4: The hourly average temperature gradient, dT/dz (top), the hourly average inversion, $dT/dz > 0$ (middle), and the monthly average temperature gradient (bottom) at Valle Hovin, Oslo, between 2 meters and 25 meters, measured in intensity [°C/m].

The middle plot of Figure 4.4 shows the hourly average inversions. It is hard to decipher any seasonal variation, especially between the end of 2002 and all the way to the end of 2004. The winter of 2002/2003, 2003/2004, 2007/2008 and 2008/2009 show weaker inversion episodes compared to the other winters, while the winter between 2005 and 2006 show the strongest inversion episodes with an intensity of 0.25 $°C/m$. Interestingly, Figure 4.1 shows some strong NO_2 concentrations in the same period. But, for instance, the episode containing the high concentrations in the early 2003, is seemingly not caused by any strong inversion.

The seasonal variation is more apparent in the plot showing the monthly average dT/dz. What we find is that the peak of the winter monthly averages are shifted

towards the left of the beginning of each year, showing that the highest average inversion intensity is found during December, or even late November. Note that the three spikes in dT/dz showing in the winter of 2003/2004 might be causing the spikes in NO_2 we find for the same period in Figure 4.1 for the monthly average NO_2 concentrations. From the winter between 2007 and 2008 to the winter between 2009 to 2010, the winter monthly averaged dT/dz is seemingly decreasing. The maximum value of 0.035 $°C/m$ is found in the late 2001, and the minimum value in 2003 is found to be -0.026 $°C/m$.

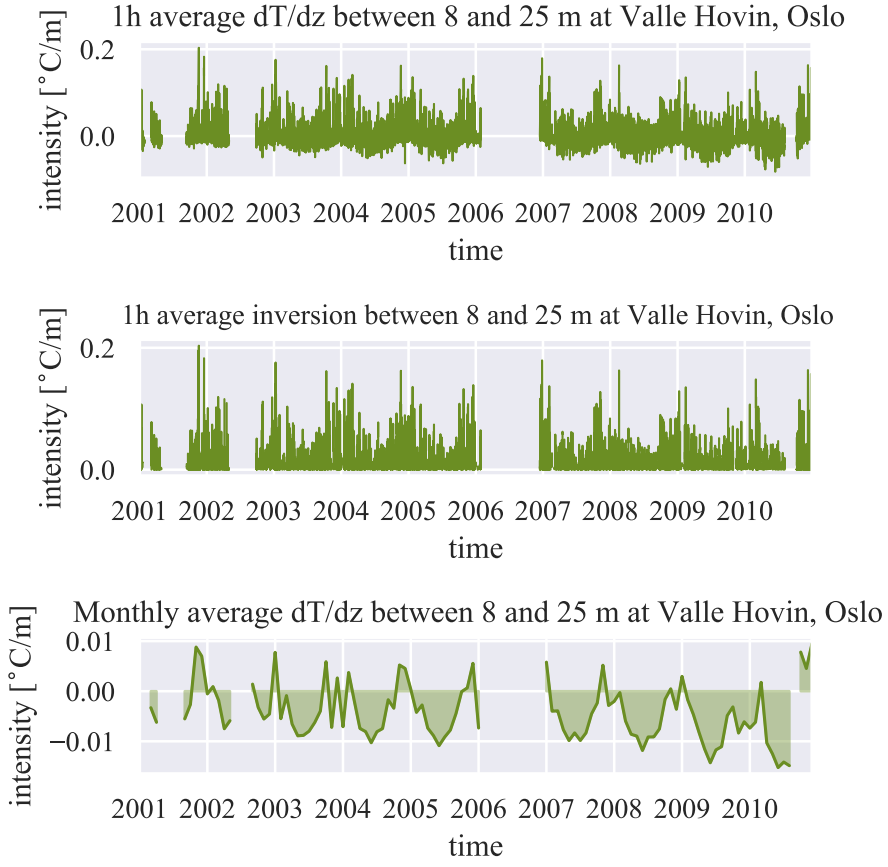


Figure 4.5: The hourly average temperature gradient, dT/dz (top), the hourly average inversion, $dT/dz > 0$ (middle), and the monthly average temperature gradient (bottom) at Valle Hovin, Oslo, between 8 meters and 25 meters, measured in intensity [°C/m].

We will now look at the measured temperature gradient and inversion from measurements between 8 and 25 meters, shown in Figure 4.5. Note that the layer we are studying in this plot are higher above the ground and also narrower than the layer studied in Figure 4.4. Here, we find that the seasonal variation is more apparent in the hourly average dT/dz, and the summer months are not as dominated by negative dT/dz as they were in the previous plot, even though there still are more negative values for dT/dz in the summer than in the winter.

Both the upper and the middle plot show a much stronger seasonal variation in the inversions compared to the temperature difference obtained from 2 and 25 meters, with stronger positive episodes occurring during winter times, because the summer inversions are much more shallow and do not reach up to higher altitudes.

With further comparison of Figure 4.5 to Figure 4.4, the strong inversion episodes taking place in the winter of 2005/2006 in Figure 4.4 are not as strong in Figure 4.5, and this shows that the inversion episodes in this period takes place closer to the ground. For the plot of the hourly average directly measured inversions between 8 and 25 meters, the strongest inversion episode occur in the winter of 2001/2002, and reaches a maximum of 0.2 $°C/m$. Another thing we note is that many of the stronger summer inversions observed in Figure 4.4 takes place below 8 meters, as we do not observe any summer inversions above 0.1 $°C/m$ between 8 and 25 meters.

In the bottom plot of Figure 4.5, showing the monthly average, we find that the monthly averaged dT/dz ranges from a maximum of 0.088 $°C/m$ in the late 2001 to a minimum of -0.015 $°C/m$ in the summer of 2010. By comparing the minimum and maximum with the ones found for dT/dz between 2 and 25 meters, we gather that this is caused by the inversions occurring closer to the ground and/or that there are cases where the inversion only take place between 2 and 8 meters, as there is a larger difference between the two maximum values than the two minimum values.

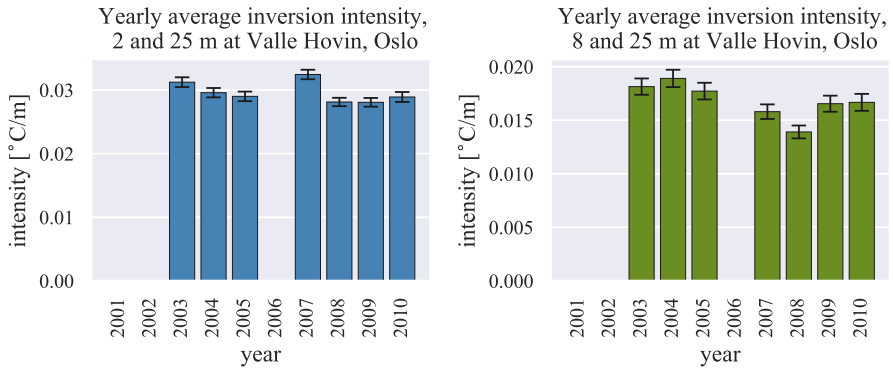
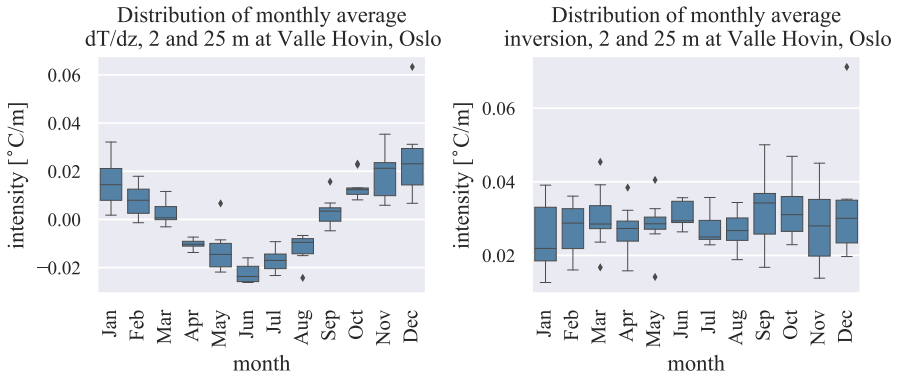


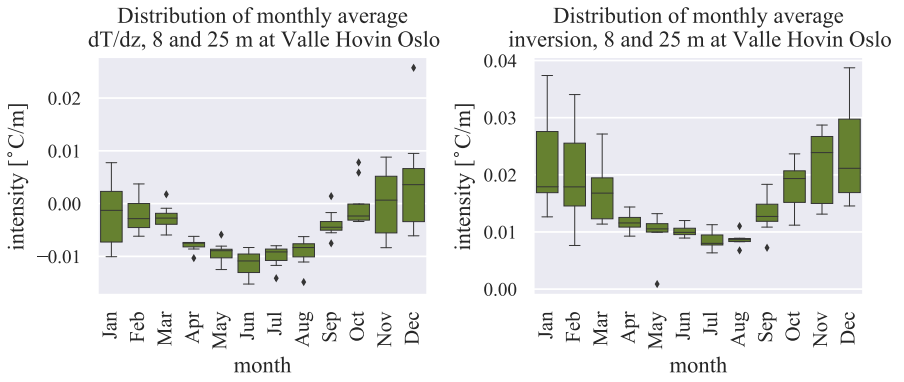
Figure 4.6: The yearly average inversion intensity measured at Valle Hovin, Oslo, between 2 and 25 meters (left) and 8 and 25 meters (right). The error bars displays the 95 % confidence interval around the average.

We will now take a look at the yearly averages for both dT/dz measurements, which is shown in Figure 4.6. For the inversions measured between 2 and 25 meters, the yearly average drops form 2003 to 2005, while peaking in 2007 with a value of above 0.032 $°C/m$. From 2008 to 2010, the yearly average increases. All of these values lie around 0.03 $°C/m$. For the yearly average inversion measured between 8 and 25 meters, shown to the right, the maximum average is found in 2004, with a value of approximately 0.019 $°C/m$. The lowest yearly average is found in 2008 with a value of approximately 0.014 $°C/m$.

Comparing the two plots in Figure 4.6, we can gather that 2004 was a year where the inversions reached higher altitudes than compared to 2007 and 2008, where the difference between the two measurements of inversions is larger.



(a)



(b)

Figure 4.7: Distribution of the monthly average temperature gradient (left) and inversion (right) for each month, measured between (a) 2 and 25 meters, and between (b) 8 and 25 meters.

By sorting the monthly average dT/dz and inversion for both measurements, a box plot was produced in order to display the seasonal variations more clearly, as shown in Figure 4.7. For both dT/dz measurements, there is a clear seasonal variation with the intensity being higher during the winter months and lower during the summer month. We also find that the spread of the monthly data is wider towards the winter months than compared to the summer months.

When only looking at the distribution for the monthly inversions, we find that the seasonal variation is lost for the measurements taken between 2 meters and 25 meters (upper right plot), but is still present for the measurements taken between 8 meters and 25 meters (lower right plot). Comparing the distribution of monthly average inversions between 2 and 25 meters to Figure 4.4, we found that the inversions taking place during summer were strong, which is why we do not observe any seasonal variation in this plot.

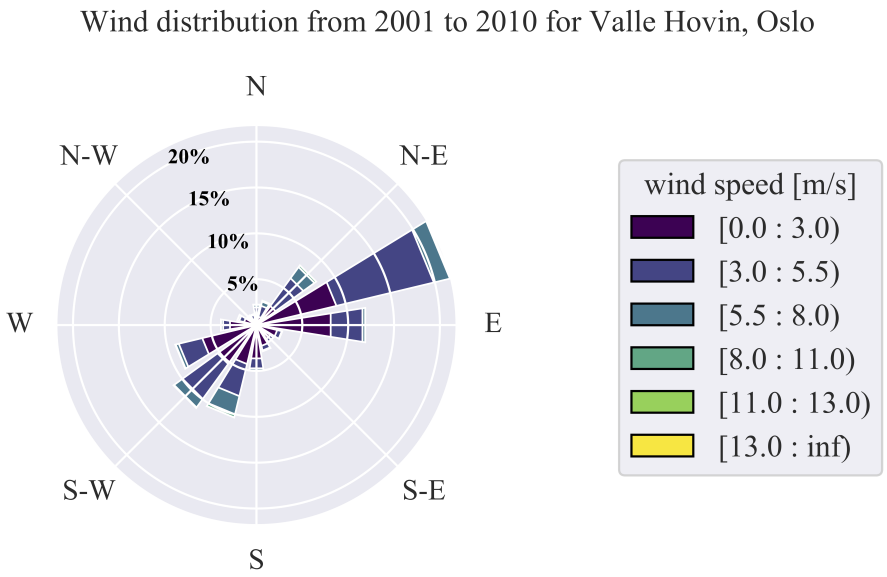


Figure 4.8: The wind distribution for the overall observational period from 2001 to 2010 measured at Valle Hovin, Oslo.

From the mass balance equation (2.19), we know that the winds have an impact on the concentrations of NO_2 as well. We therefor present the wind distribution measured at Valle Hovin for the observational period in Figure 4.8. The majority of the wind blows towards and from northeast/east and southwest/west, with over 20% arriving from northeast/east and 8% arriving from southwest/west, as it follows the topography and terrain of Grorud valley. Approximately 11% is coming from southwest, 10% is com from south/southwest, 12% is arriving from the east, and 7.5% is arriving from the northeast. There is no wind moving in the northwest/south-east direction as the topography prevents this. The wind speed arriving at Valle Hovin is usually less than 5.5 m/s.

The wind distribution for each season is displayed in Figure 4.9. For the winter months, almost 30% of the wind is arriving from the northeast/east, bringing along cold temperatures. The winds arriving from northeast, northeast/east and east are dominating. In the spring months, a larger share of the wind measurements show winds arriving from northeast, northeast/easy and east, compared to the winter months. As for the summer months, the amount of winds arriving from southwest/west, southwest and southwest/south (approximately 34%) are almost equal to the amount of winds arriving from northeast, northeast/east and east (approximately 37%). The highest wind speed recorded between 2001 and 2010 reached a speed of 13.6 m/s and was recorded on the 14th of Febuary 2004, but there is no evidence of a response in the NO_2 concentrations to this when looking back at Figure 4.1.

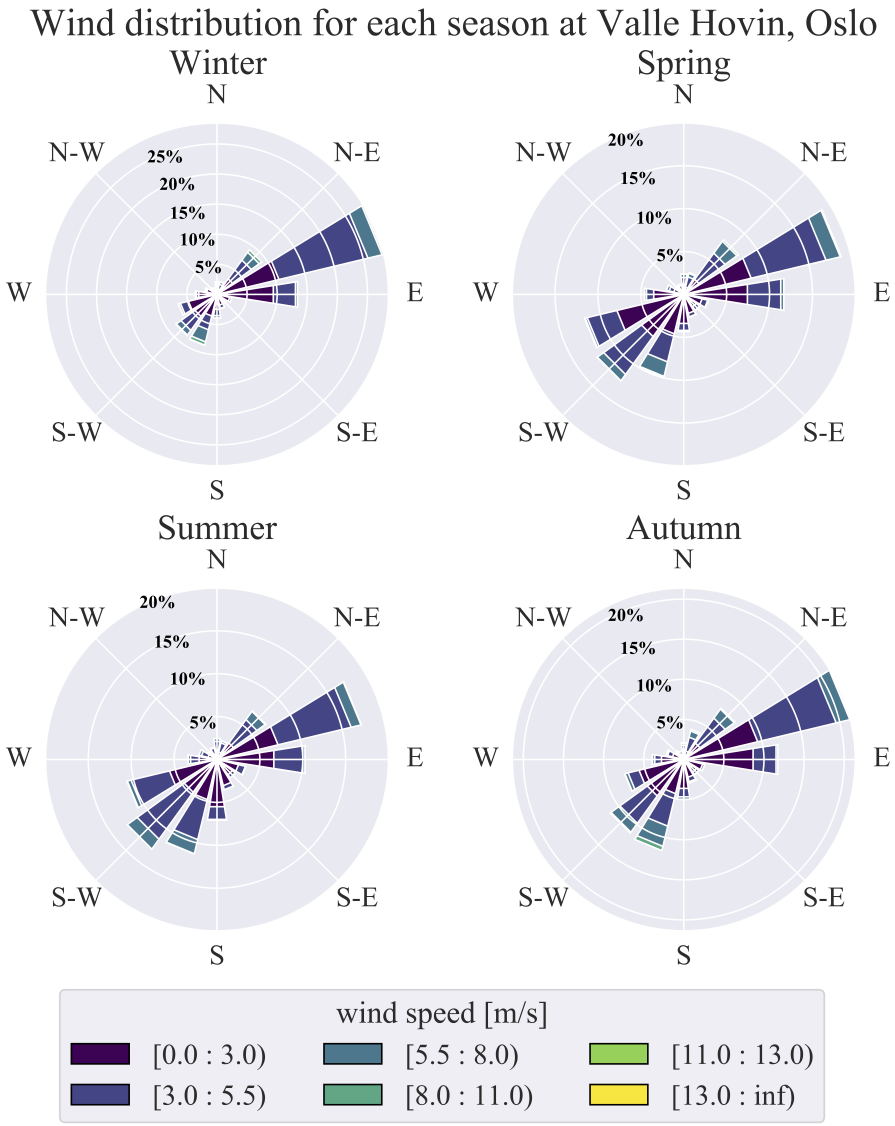


Figure 4.9: The wind distribution for each season measured between 2001-2010 at Valle Hovin, Oslo.

We have so far looked at each of the variables observed. Now, we will investigate the relationship between the NO_2 concentrations and the inversion intensity. In addition, we will also look into the relationship between NO_2 and the temperature gradient, the wind speed and the wind direction. Thus, a correlation table was made as shown in Figure 4.10, to give an overview of the the linear relation the variables have with each other.

The first thing to note in the correlation table is that the correlation between NO_2 and the inversions are statistically weak, with the correlation between NO_2 and inversion measured between 2 and 25 meters being 0.37, and the

correlation between NO_2 and inversions measured between 8 and 25 meters being 0.28. Conversely, NO_2 has a better correlation with the measured temperature gradient between 8 and 25 meters than the temperature gradient between 2 and 25 meters. The wind speed has a weak negative correlation with the NO_2 concentrations. This is expected, as winds distribute and dilute the NO_2 concentrations.

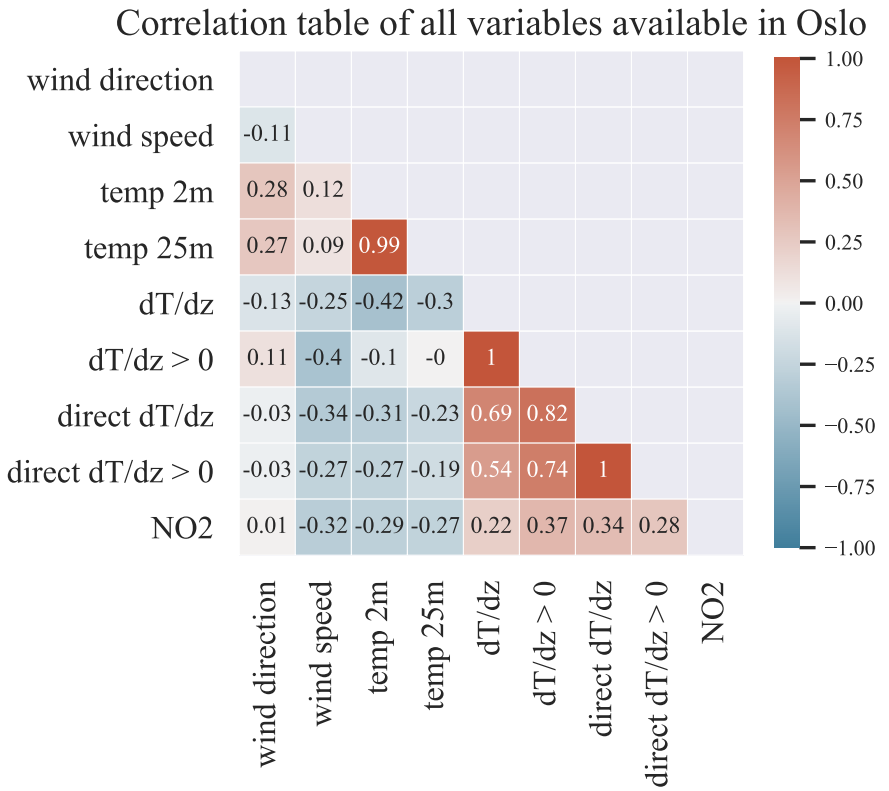


Figure 4.10: Correlation table of the observed variables at Valle Hovin and Alnabru in Oslo.

But, because the NO_2 concentrations also depend on how the emissions vary, we do not get the full picture of the relationship between NO_2 and the variables in Figure 4.10. We therefor need to filter the data such that the NO_2 concentrations is as independent of the emissions as possible. By sorting the NO_2 concentrations by the day and hour they were measured and taking the average, we gather the multiyear daily and hourly average concentrations, as shown in Figure 4.11. Here we identify that the concentrations is higher during weekdays and lower during the weekend, as one would expect, considering there are fewer commuters. We see that the multiyear hourly average is at its highest at 08:00, as commuters are driving to work, before decreasing to 13:00, increasing to 18:00 (commuters driving home from work) and decreasing again to lower concentrations through the night.

For both Monday and Thursday, the multiyear daily average is approximately

the same, and the multiyear hourly average is approximately the same at 07:00 and 08:00. This allows us to make the assumption that the concentrations will be less dependent on the emissions if we filter our data to only contain values from 07:00 and 08:00 on Mondays and Thursdays. By doing so, and calculating the correlation, we obtain the correlation table shown in Figure 4.12. We now see that the correlation between NO_2 and inversions between 2 and 25 meters increased to 0.67, suggesting that there is a linear relationship, and the correlation between NO_2 and the directly measured inversions between 8 and 25 meters increased to 0.47. Again we find that NO_2 has a slightly higher correlation with directly measured dT/dz between 8 and 25 than dT/dz between 2 and 25 meters. This suggests that the vertical mixing at higher altitudes are more important in the removal of NO_2 pollution.

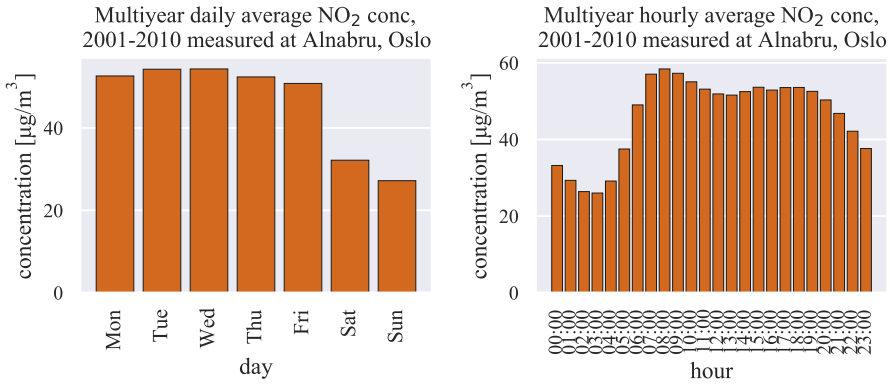


Figure 4.11: The average NO_2 concentrations for each weekday (left) and the average NO_2 concentrations for each hour (right), from 2001 to 2010.

The correlation between dT/dz between 2 and 25 meters and directly measured dT/dz between 8 and 25 meters has a correlation of 0.69 for the non-filtered data in Figure 4.10 and 0.66 for the filtered data in Figure 4.12. We would expect the correlation to be higher. This might be caused by the summer inversions we observed in Figure 4.4 not reaching all the way up to 8 meters as we saw in Figure 4.5, but it may also be caused by errors in the calibrations when measuring the temperature at 2 meters and 25 meters. The correlation between temperature at 2 meters and 25 meters shows that they have a strong linear relationship. Another thing to point out is that the correlation between NO_2 and the wind speed did not change much when filtering the data, but this is to be expected as the wind does not vary diurnally.

The correlation value gives us information of the linear relationship between values, but to investigate the relationship further, a scatter plot was made as displayed in Figure 4.13. For both measurements of the unfiltered dT/dz (left panel), we observe that for stronger negative values, the scatter point lie closer to the linear regression line, given that there is more vertical mixing for negative values dT/dz. As the value of dT/dz increases, so does the spread of the scatter points. There are scatter points showing that some of the higher concentrations of NO_2 is measured during low values of inversion intensities, and some of the

lower concentrations are measured at high values of inversion intensities. As the NO_2 concentrations in the unfiltered data is also strongly dependent of emissions, this explains why we see higher concentrations of NO_2 during weaker inversions. During nighttime, we might have strong inversions, but no source of NO_2.

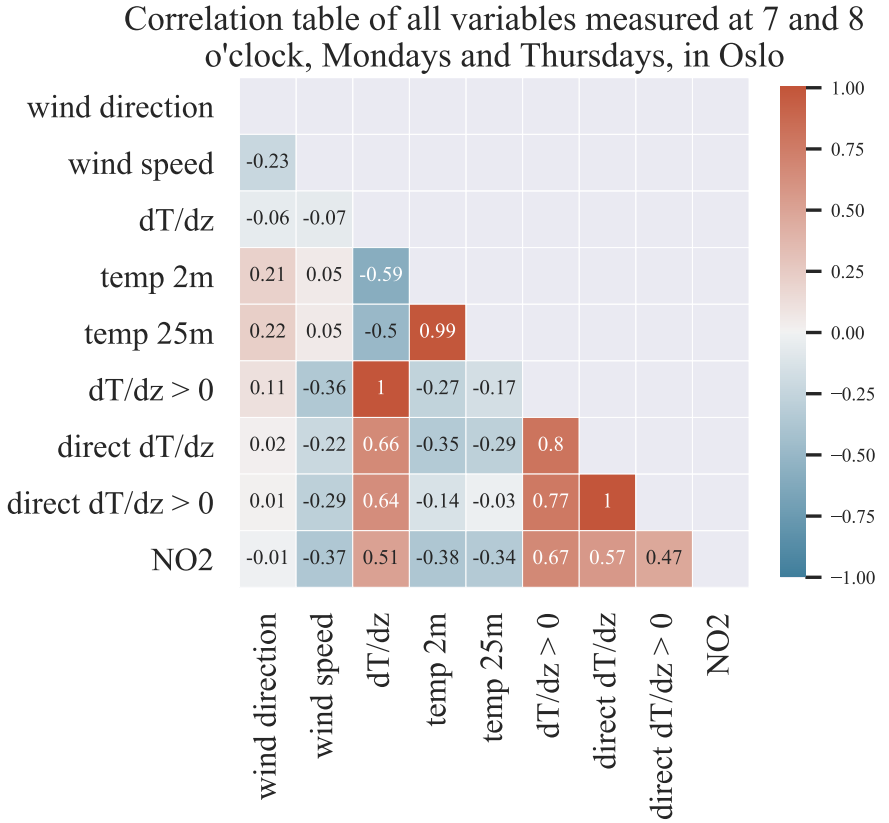
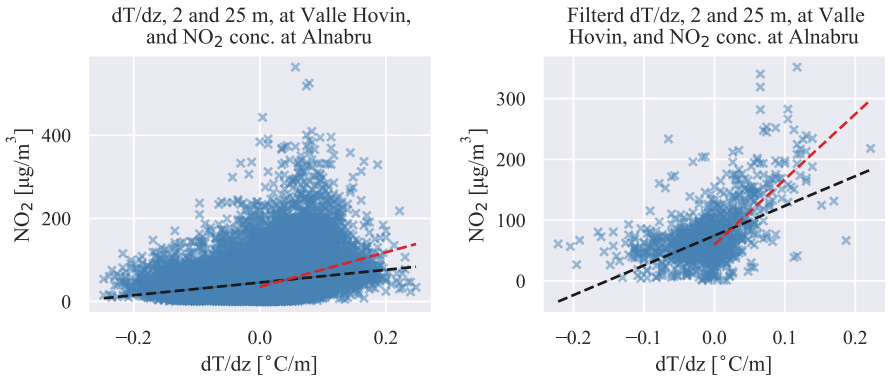
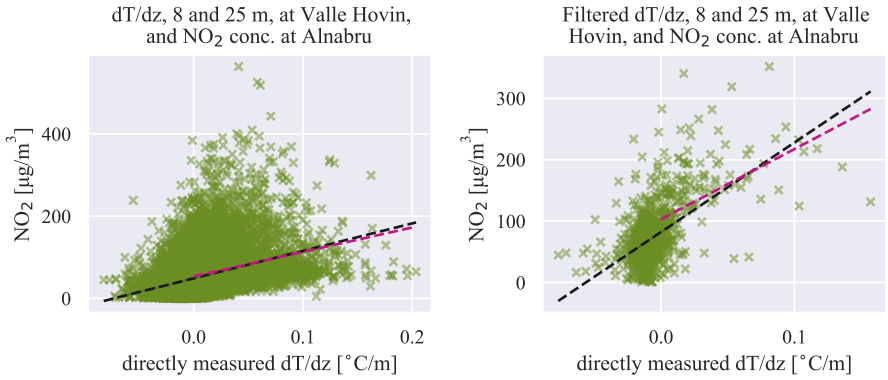


Figure 4.12: Correlation table of the observed variables filtered to only contain measurements from 07:00 and 08:00 on Mondays and Thursdays at Valle Hovin and Alnabru in Oslo.

To the right in Figure 4.13a, which shows the scatter plot of the filtered data, we observe that most of the higher concentrations at lower inversion intensities and lower concentrations at higher inversion intensities disappears. The relationship is more linear, especially for between NO_2 and the inversions. To the right in Figure 4.13b, most of the lower concentrations at higher inversion intensities disappear, but the scatter plot still shows higher concentrations for weaker inversion episodes. Because the correlation of NO_2 and inversions between 2 and 25 meters is higher than the correlation of NO_2 and inversions between 8 and 25 meters, the inversions closer to the ground is more important in determining the NO_2 concentrations. For the rest of the result chapter, we will discard the temperature gradient and inversion measured between 8 and 25 meters.



(a)



(b)

Figure 4.13: Scatter plot of NO_2 and dT/dz where (a) is measured between 2 and 25 meters and (b) is measured between 8 and 25 meters, showing both the unfiltered data (left) and the filtered data containing only measurements from 07:00 and 08:00 on Mondays and Thursdays (right). The dashed black line is the linear regression using dT/dz as explanatory variable and NO_2 as the dependent variable. The red dashed line is the linear regression using inversions as the explanatory variable and NO_2 as the dependent variable.

We will now create a multiple linear regression of the filtered data, with NO_2 as the dependent variable. Because we are mainly interested in the inversions, we chose to discard dT/dz as well as the temperatures at 2 and 25 meters because a linear regression requires that the possible explanatory variables are independent of each other. This leaves us with inversions, wind speed and wind direction as possible explanatory variable. The first step of the regression is to conduct a simple linear regression with the variable NO_2 has the strongest correlation with, in this case the inversions as we saw and discussed in Figure 4.12. Doing the calculations numerically, we find that the simple linear regression model, which we term as **model 1**, is given by

$$\texttt{NO2} = 59.2 + 1066.8 \cdot \texttt{inversion} \tag{4.1}$$

This simple linear regression is also displayed by the red dashed line to the right in Figure 4.13a. The summary table (Figure A.1) shows that the R-squared value for the model is 0.4. Further, we calculate the partial correlation between the remainder independent variable and NO_2. The partial correlation between NO_2 and the wind speed is -0.3, and the partial correlation between NO_2 and the wind direction is -0.1. Because the wind speed has the highest partial correlation with NO_2, the variable was added to find **model 2** to be given by

$$\texttt{NO2} = 83.1 + 913.4 \cdot \texttt{inversion} - 6.8 \cdot \texttt{wind speed} \tag{4.2}$$

The R-squared has now increased to 0.5, and to make sure the addition of wind speed is significant, we perform a t-test. The summary table (Figure A.2) gives us that the absolute of the statistical value of t is 9.1, and is greater than the critical t-score of 1.96, thus confirming that the addition is significant. Lastly, by adding wind direction we obtain **model 3**:

$$\texttt{NO2} = 89.1 + 926.8 \cdot \texttt{inversion} - 6.8 \cdot \texttt{wind speed} - 0.05 \cdot \texttt{wind direction} \tag{4.3}$$

The summary for the model is found in Figure A.3, and from there we find that the R-squared increased to 0.6. Even though the increase is small, the absolute of the statistical t-score is 2.8. The addition of wind direction is thereby significant. Model 3 is plotted together with the observed NO_2 concentrations in Figure 4.14

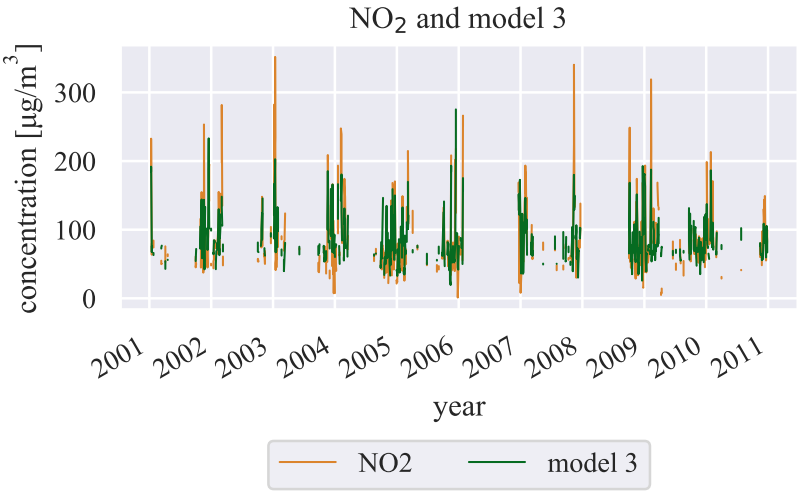


Figure 4.14: NO_2 (orange) and model 3 (green) for the study period.

Figure 4.14 shows that the model underestimates the larger concentrations values. The statistical properties are shown in Table 4.1, and it shows, together with the plot, that the model is not able to reproduce the lowest values either. Both the median and the 1st quantile is slightly higher than the actual concentrations, while the 3rd quantile is approximately similar. But overall, the model is able to yield the same concentrations within the 1st and 3rd quantile well.

	NO_2	model 3
Min.	0.9	19.4
1st Qu.	53.1	62.7
Median	70.8	74.7
Mean	84.3	84.3
3rd Qu.	99.25	98.9
Max.	351.7	275.3

Table 4.1: Summary table comparing the statistical properties of NO_2 and model 3.

Bergen

We will now look at the variables present for Bergen, in the period from 2011 to 2020. Firstly, the NO_2 concentrations are shown in Figure 4.15, containing the hourly averaged and the monthly averaged concentrations.

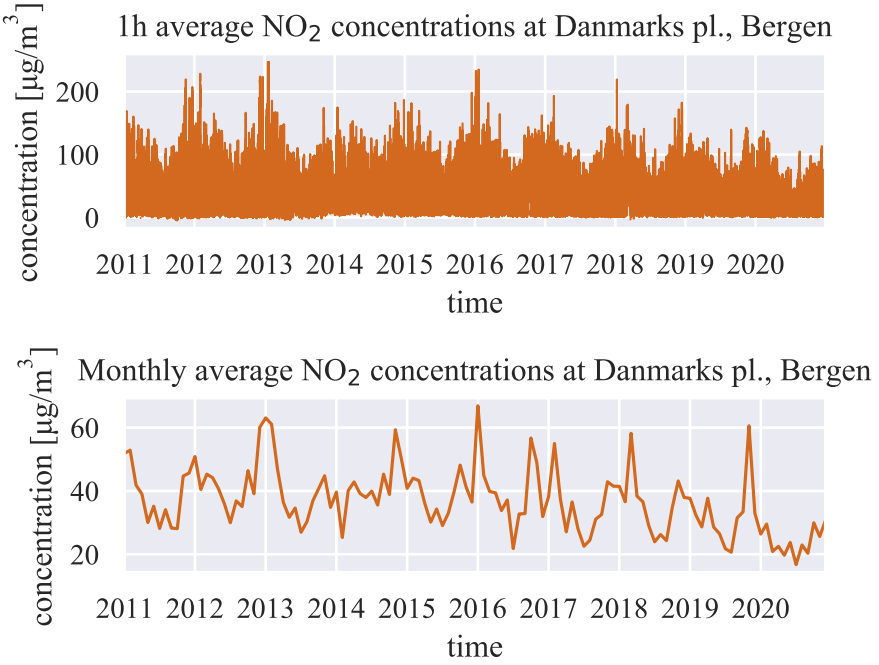


Figure 4.15: Hourly average (top) and monthly average (bottom) NO_2 concentrations measured at Danmarks plass, Bergen, between 2011 and 2020 with units $[\mu g/m^3]$.

Like in Oslo, we observe a seasonal variation in Bergen too, with higher concentrations during the winter months that shifts towards lower concentrations during the summer months. 2013 contained episodes of particularly large concentrations of NO_2, the highest value recorded to be at 247 $\mu g/m^3$ on the 22nd of January. The average maximum value for each year is 200.7 $\mu g/m^3$, and

the maximum of the summer concentrations usually lie just below 100 $\mu g/m^3$. The most important feature for this plot is that the maximum concentrations for each year declines from 2016 and all the way to the end of 2020.

From the Figure 4.15 we can also observe the monthly average NO_2 concentrations. The seasonal variations are not as prominent as they were for Oslo in Figure 4.1, but we still observe the maximums during winter and the lower concentrations appear during summer. During the winter of 2014 and 2019 there were lower winter concentrations than normal, compared to other winters during the study period. The monthly average plot reveals that there is a decline in the summer concentrations from 2018 to 2020. The monthly average is lower in Bergen than in Oslo, with a maximum of 67 $\mu g/m^3$ in 2016.

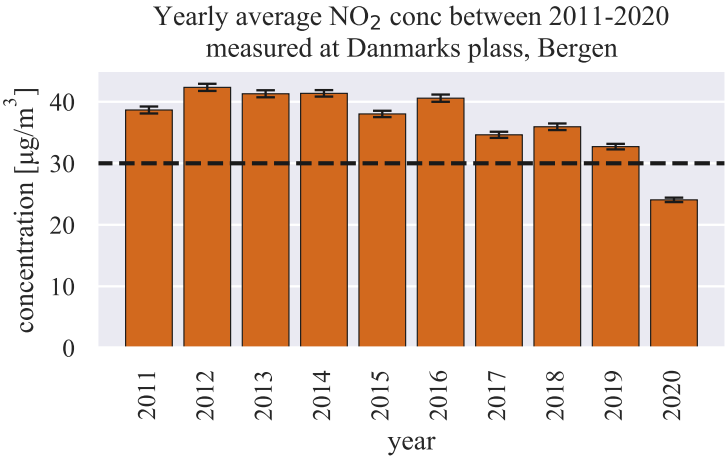


Figure 4.16: Yearly average NO_2 concentrations measured in $[\mu g/m^3]$ at Danmarks plass, Bergen, with 95% confidence interval around the average. The dashed line indicates the annual recommended health limit[Norwegian Institute of Public Health (FHI), 2020]

In Figure 4.16 we find the yearly averaged NO_2 concentrations that were measured at Danmarks plass in Bergen. We note that almost all of the years, with the exception of 2020, had averages above the yearly average health limit. 2012 was the year with the highest recorded yearly average with a value of 42 $\mu g/m^3$, and the yearly average was even up til 2016. From 2018 to 2020, we observe a steady decline in the yearly averages.

The distribution of all the monthly averages of the NO_2 concentrations in Bergen was also plotted, shown in Figure 4.17. It shows that the overall monthly average concentrations is below 40 $\mu g/m^3$ during summer, and shifts towards higher concentrations in the winter. Note that the median and the 1st and 3rd quantile in November are higher than in December, but December has a higher maximum. The spread of the distribution widens in the winter, and narrows in the summer.

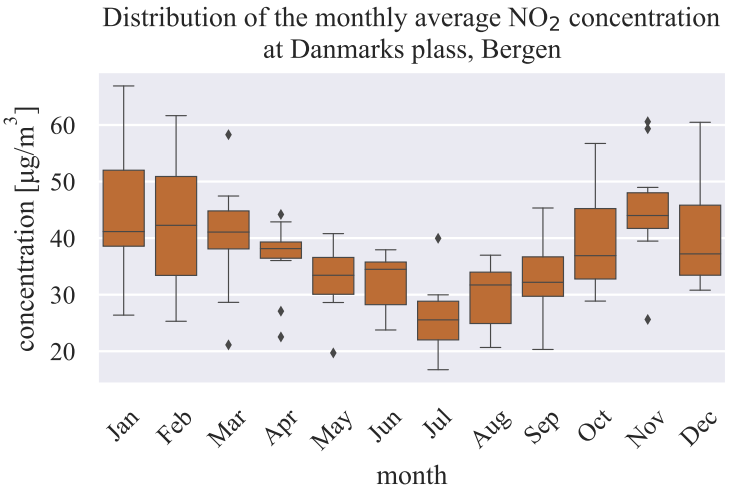


Figure 4.17: The distribution of monthly averages for each month, containing data from 2011 to 2020. The line within the box shows the median, the box' edges are determined by the 1st and 3r quantile and the error bars shows the maximum value and minimum value. The diamond displays the outliers.

We will now discuss the temperature gradient measured between the MTP-5HE sensor on top of Geofysen at 40 meters and the Florida MET station at 2 meters, as well as the inversions detected. This is illustrated in Figure 4.18.

From the upper plot of Figure 4.18, there is a large spread during the summer, with both strong negative and positive values of dT/dz. The strongest inversion events appear during summer, as we see for example in 2013, where the strongest measured inversion episode had an intensity of 0.18 $°C/m$. During winter, the inversion episodes are weaker in intensity. This is further displayed in the middle plot. The seasonal variation is not as expected, as we would assume to get the same result as Wolf et al., 2014. They showed that the ground inversions were stronger and more frequent during the winter, but here the inversions were measured from 40 meters and up to 1000 meters. One possible explanation as to why we observe this might be due to the sensor at 2 meter standing too close to a building, which might increase the measured temperature.

The monthly average dT/dz shows a seasonal variation closer to what we would expect, but the values are almost always below zero and shows that there is more vertical mixing during summer, than in winter. A reason as to why we observe weaker values in Bergen than in Figure 4.4 for Oslo is that we are looking at a wider span of altitude, ranging over 38 meters. There will therefore occur instances where the inversions are not through the entire layer.

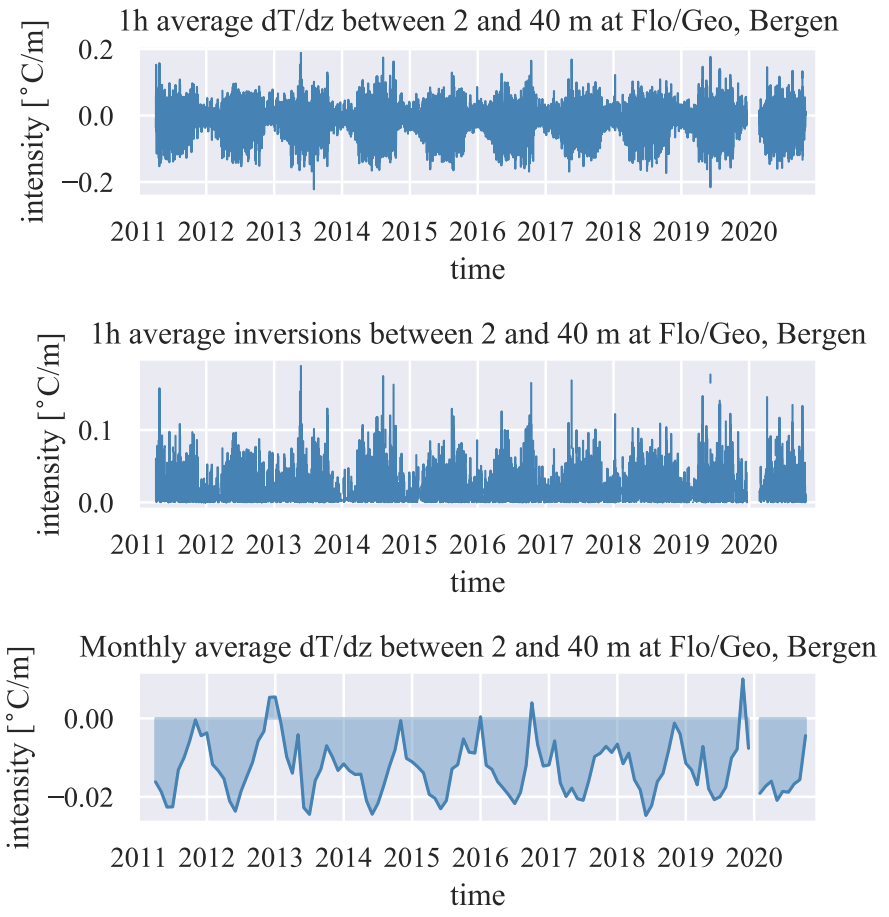


Figure 4.18: The hourly temperature gradient, dT/dz (top), the hourly average inversion, $dT/dz > 0$ (middle), and the monthly average temperature gradient (bottom) at Florida/Geofysen, Bergen, between 2 meters and 40 meters, measured in intensity $[°C/m]$.

Like for the Oslo data, a box plot of the monthly average temperature gradient and inversion was made, shown in Figure 4.19. The distribution for January, November and December contains higher values than the distribution for the other months. January is the the only month with a distribution with values above zero, and the weakest temperature gradients are measured in June and July. Just as in Oslo in Figure 4.7, the spread is wider during winter and narrows when it is shifting towards summer. Note that the seasonal variation is similar to the one observed in Figure 4.17.

To the right in Figure 4.19 which displays the distribution of the monthly average inversions, it can be observed that the seasonal variation has lost its shape, with the higher values of inversion intensity being measured during April and May. The averages stays even during June to September, before decreasing to lower values in November and December. The values are also low for January, February and March. The overall distribution of the monthly averaged inversions

are much wider than the distribution of the monthly averaged dT/dz. The magnitude of the inversions during winter are lower compared to the ones measured in Oslo, but are almost equal for April to October.

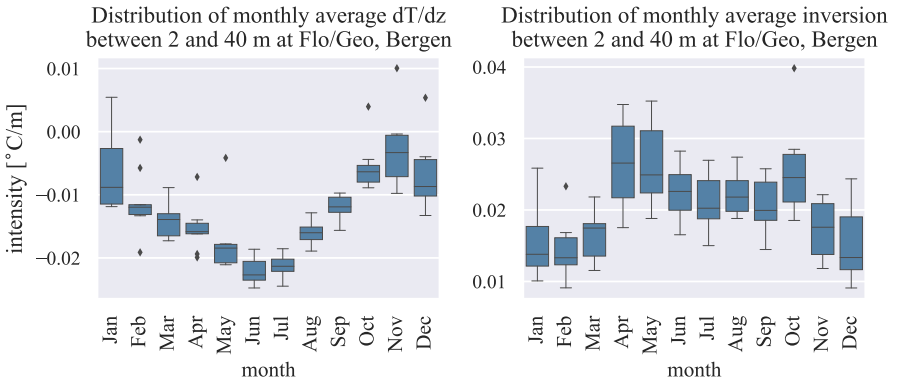


Figure 4.19: The monthly average dT/dz distribution between 2 and 40 meters to the left and the distribution of the corresponding monthly average inversions to the right. Measured at Florida/Geofysen, Bergen.

Figure 4.20 shows the yearly average inversion intensity, along with the 95% confidence interval, where the years 2011 and 2020 was discarded due to the lack of enough sufficient data. The strongest yearly average was measured in 2013 with a value of 0.024 $°C/m$, and the weakest was measured in 2017 with a value of 0.019 $°C/m$. The value decreases from 2013 to 2015, and increases from 2017 to 2019. Comparing this to Figure 4.16, there is no similarity in the variation.

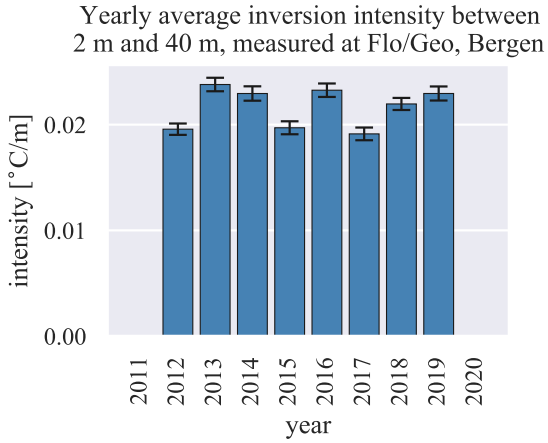


Figure 4.20: The yearly averaged inversion intensity measured at Florida/Geofysen, Bergen, between 2 and 40 meters. The error bars display the 95% confidence interval around the average.

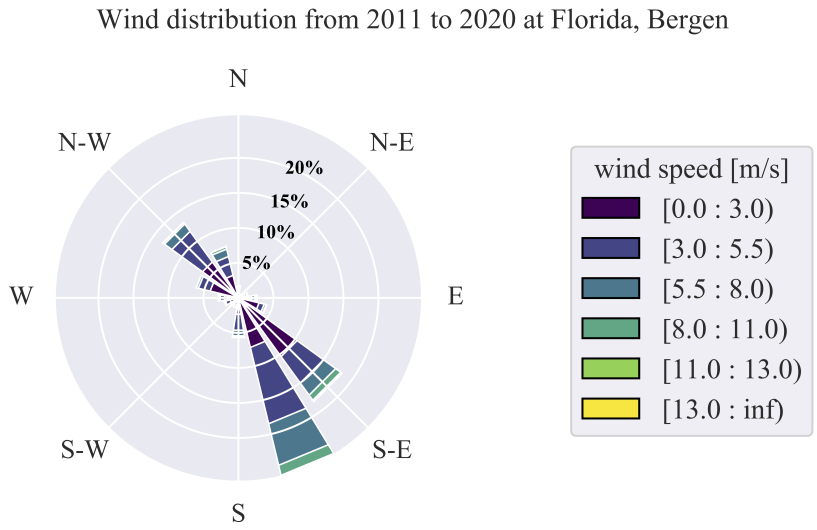


Figure 4.21: The wind distribution for the overall observational period from 2011 to 2020, observed at Florida, Bergen.

The wind distribution at the station at Florida is plotted in Figure 4.21. Approximately 47% of the wind comes either southeast/south or southeast, and 13% comes from the northeast. The terrain showed in Figure 3.2 hinders the wind from arriving in the northeast-southwest direction. The fastest wind speed to be measured during the study period was at 20 m/s on the 10th of January, 2015, but Figure 4.15 shows no response in the NO_2 concentrations.

The distribution for each season was also plotted, shown in Figure 4.22. In the winter, over 60% of the measured wind comes from either southeast/south or southeast, advecting warmer air to Bergen. This is due to the low pressure system above Iceland, part of the North Atlantic Oscillation (NAO) [Scaife et al., 2005]. Most of the wind arriving have a wind speed of less than 5.5 m/s, and a large portion has a wind speed between 8 and 11 m/s. Winds faster than 13 m/s have been recorded arriving from southeast. A small fraction comes from the region between the west and north direction. Come spring, and more wind appears from northwest. 17% of the wind comes from northwest, 10% comes from northwest/north, and 7-8% comes from the northwest/west. In the summer, there is more wind arriving from northwest, northwest/north and northwest/west than south, southeast/south and southeast. In the autumn, the winds from northwest disappears, and stronger winds from southeast appears.

The warmer winds in the winter will lessen the appearance of inversions, which would explain why we observe weaker inversion episodes in the winter. It would also contribute to the spread and dilution of NO_2 pollutants. The seasonal wind plot also shows that the winds in summer are usually weaker compared to the other seasons.

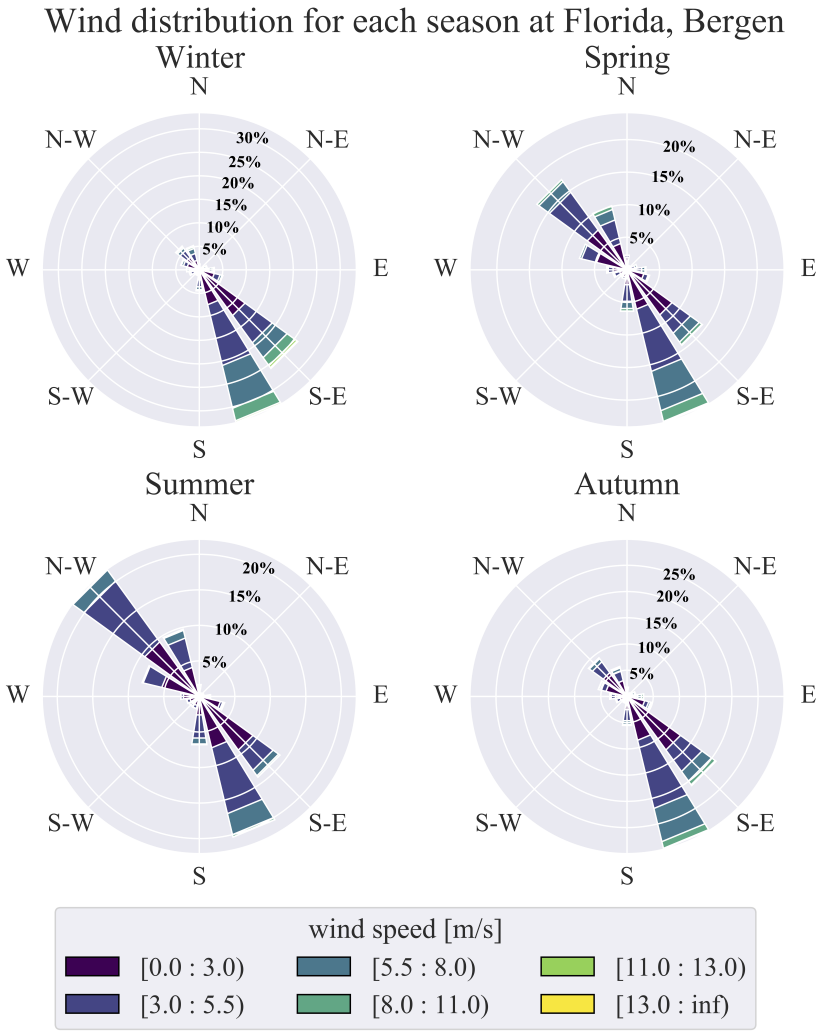


Figure 4.22: The wind distribution for each season in Bergen, showing the wind speed and wind from direction

We will now present the relationship between inversions, temperature gradient, the winds and the NO_2 concentrations. The correlation between all the observed variables for Bergen can be found in Figure 4.23. The correlation between the measured inversions and the NO_2 concentrations is weak, and gives no reason indication of a linear relationship. The correlation between the temperature and NO_2 is even weaker. The plot also reveals that the correlation between NO_2 concentrations and the wind speed is negative. This is not surprising as more winds lead to more dispersion of the pollutant and thereby lessening the NO_2 concentrations. There can be several reasons as to why we observe a weak correlation between NO_2 and the inversions. From Figure 3.2, we saw that the measuring station for NO_2concentrations and the measuring stations for the inversions are separated by Store Lungårdsvannet. Due to little friction, this allows for more surface winds to disperse the pollutant, even if there is a

measured inversion at Florida/Geofysen. It can also cause convection when warmer air is advected over the colder water. The buildings surrounding the stations at Florida/Geofysen can affect the temperature measurements, and thereby also affect the measured inversions.

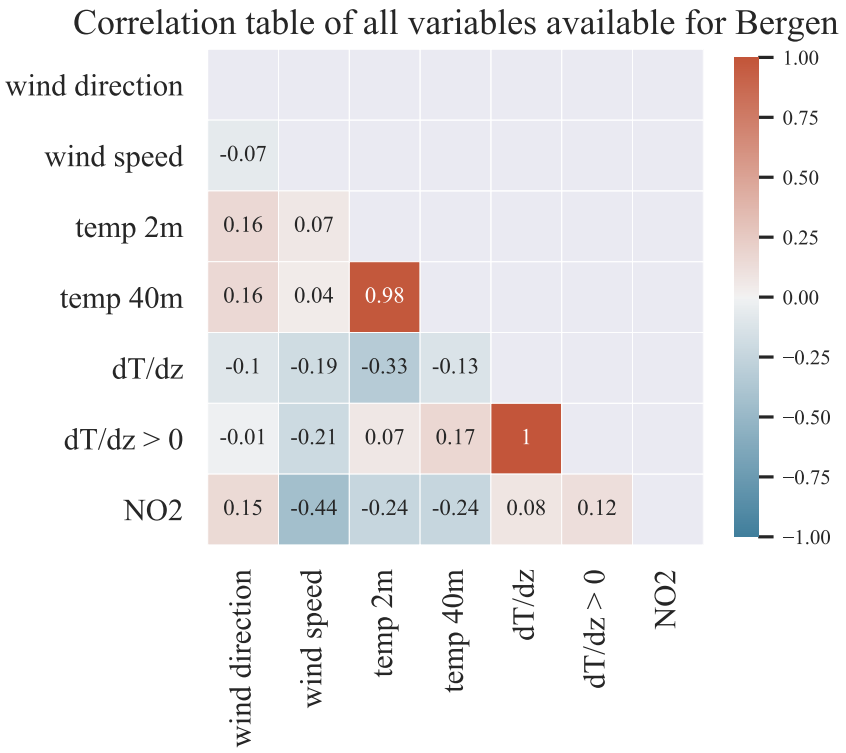


Figure 4.23: Correlation table of the observed variables at Florida/Geofysen and Danmarks plass in Bergen

The data was filtered to make the NO_2 concentrations less dependent on the emissions from car traffic, giving a better idea of the effect of the inversions. The average for each weekday was calculated, and is shown in Figure 4.24. The figure shows that the multiyear average concentrations lie around 40 $\mu g/m^3$ from Monday through Friday, before dropping down to 28 $\mu g/m^3$ on Saturday and 25 $\mu g/m^3$ on Sunday. To the right in Figure 4.24 the multiyear average concentrations for each year is displayed. It shows that the concentrations is small during the night, with the lowest average lying just below 20 $\mu g/m^3$. The multiyear average gains its peak at 07:00 with a value of 50 $\mu g/m^3$, and decreases until the commute from work to home starts at 14:00. It increases to 45 $\mu g/m^3$ and steadily decreases from there.

It was chosen for the filtered variables to contain data from 07:00 and 08:00 for all the days except Saturday and Sunday. If the inversions was constricted to only contain data for 07:00 and 08:00 on Mondays and Thursdays like we did for the Oslo data, then the amount of data would be insufficient to continue with the statistic analysis.

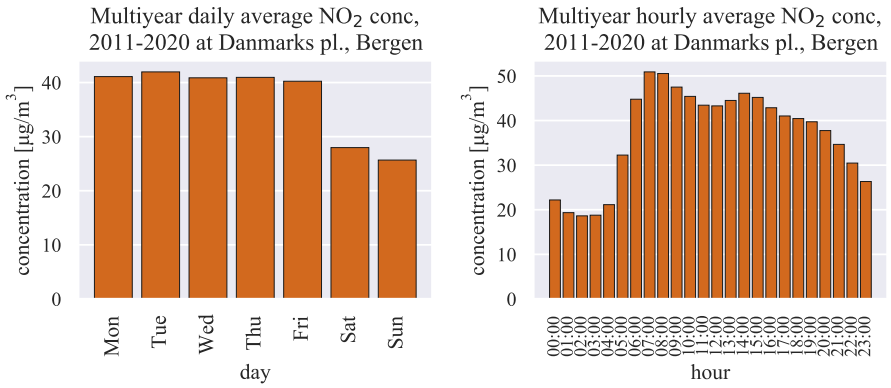


Figure 4.24: The plot on the left side shows the average NO_2 concentrations for each weekday and the plot to the right shows the average NO2 concentrations for every hour. Both averages containing data from 2011 to 2020, measured at Danmarks plass, Bergen.

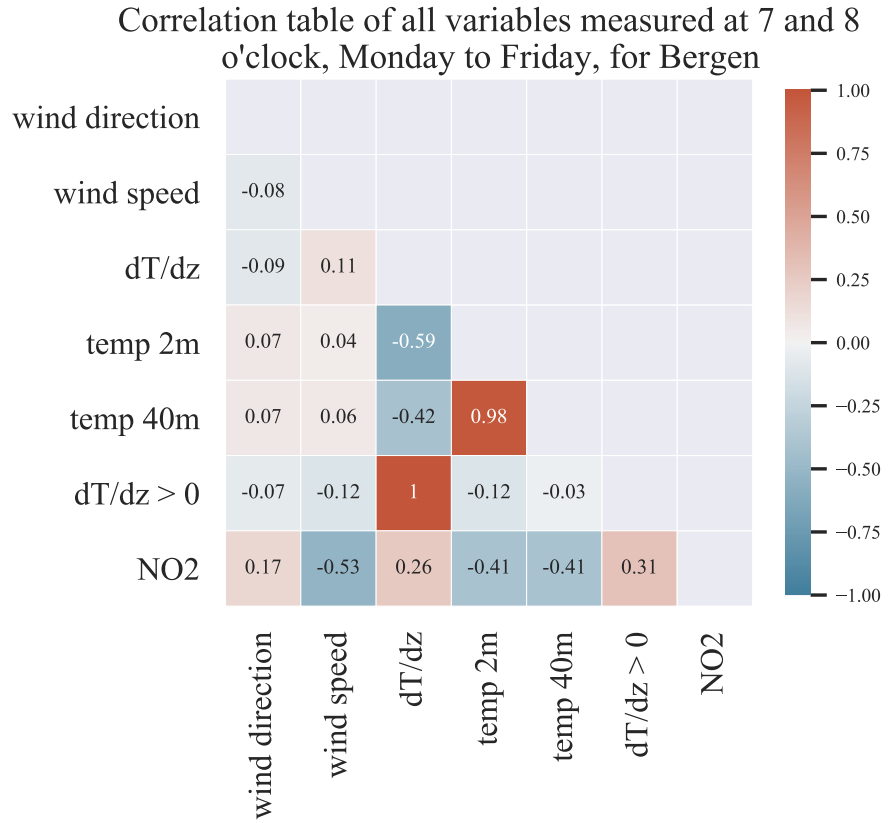


Figure 4.25: Correlation table of the observed variables filtered to only contain measurements from 07:00 and 08:00 from Monday through Friday at Florida/Geofysen and Danmarks plass in Bergen

The correlation between the variables of the filter data is shown in Figure 4.25. The correlation between NO_2 and the inversions ($dT/dz > 0$) has increased to 0.31. The correlation between NO_2 and the temperature gradient also increased to 0.26, whilst the correlation between NO_2 and the wind speed has increased in magnitude to -0.53. It is notable that the inversions in the filtered data will contain mostly shallow morning inversions. The correlation between dT/dz and the temperature at 2 meter increased in magnitude from -0.33 to -0.59, and the correlation between dT/dz and the temperature at 40 meters increased to in magnitude from -0.24 to -0.42.

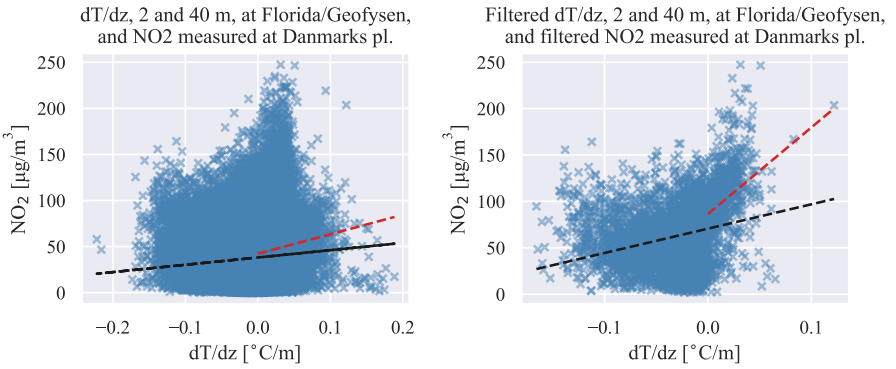


Figure 4.26: Scatter plot of NO_2 concentrations and dT/dz between 2 and 40 meters to the left and the filtered data of NO_2 concentrations and dT/dz to the right. The black dashed line shows the linear regression where dT/dz is the explanatory variable, and the red dashed line shows the linear regression where $dT/dz > 0$ is the explanatory variable

Figure 4.26 shows the scatter plot of dT/dz against the concentrations of NO_2, both for the non-filtered data (to the left) and the filtered data (to the right). The non-filtered data shows that the largest of NO_2 concentrations are measured when dT/dz is between 0 and 0.5 $°C/m$. This is likely due to high emissions of NO_2 Like in Oslo, the stronger inversions during night will weaken the correlation between NO_2 and the inversions and the correlation between NO_2 and dT/dz. The plot to the right in Figure 4.26 reveals that when the data is filtered, these strong inversions measured disappear. For both plots, the inversions have a higher correlation with the NO_2 concentrations than the temperature gradient.

Like for the Oslo data, a multiple linear regression was made for the Bergen data. For the MLR, the inversions, wind speed and wind direction was used as the explanatory variables for the NO_2 concentrations. Because the NO_2 concentrations was highest correlated with the wind speed, the first step of making an MLR was to make an SLR with NO_2 as the dependent variable and the wind speed as the independent variable. However, a scatter plot between the two variables revealed that the relationship was not linear, and so in order to continue with the statistical analysis, the SLR was made between NO_2 and the logarithm of the wind speed data. **Model 1** is given by

$$\texttt{NO2} = 113.2 - 27.8 \cdot \texttt{log(wind speed)} \tag{4.4}$$

The summary table from (Figure A.4) shows that the R^2 has a value of 0.3. Further calculations reveals that the inversions have the highest partial correlation, and the variable is therefore added to make **model 2**

$$\texttt{NO2} = 99.7 - 25.6 \cdot \texttt{log(wind speed)} + 724.9 \cdot \texttt{inversion} \tag{4.5}$$

The R^2 value increased to 0.36, and to make sure the addition of the inversions is significant, a t-test was done. The absolute of the statistical value of t is given to be 7.2 (Figure A.5), and the addition is therefore significant. Lastly, the wind direction was added to the MLR, but a t-test showed that the addition was not significant (Figure A.6). Model 2 is plotted in Figure 4.27, and its statistical properties are shown in Table 4.2. The plot and the summary table shows that Model 2 is not able to reproduce the higher and lower concentrations, e.g. the peak in the beginning of 2014 and the peak in the winter 2016/2017. Comparing the results for Bergen to the results for Oslo in Figure 4.14, the R^2 is much smaller for Bergen. This suggests that the station at Florida/Geofysen is too far away from the station at Danmarks plass to paint a good picture of how the meteorological conditions affect the NO_2 concentrations.

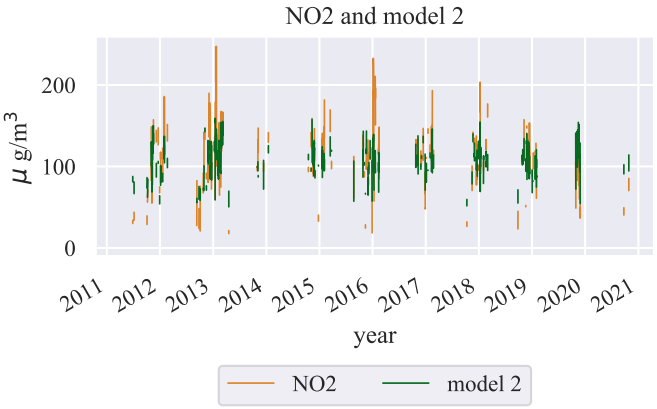


Figure 4.27: NO_2 (orange) and model 2 (green) for the study period

	NO_2	model 2
Min.	2.1	39.7
1st Qu.	79.5	86.4
Median	105.8	104.
Mean	103.3	103.3
3rd Qu.	128.1	120.7
Max.	247.3	162.1

Table 4.2: Summary of the statistical properties of NO_2 and model 2, Bergen.

4.2 Comparing NORA3 output to the observations

We will now look at how the NORA3 models the temperature gradients and wind patterns compared to the observations. We will study how well the model performs, specially with regards to the inversions, and if it is able to yield some of the seasonal variations shown in the observations.

Oslo

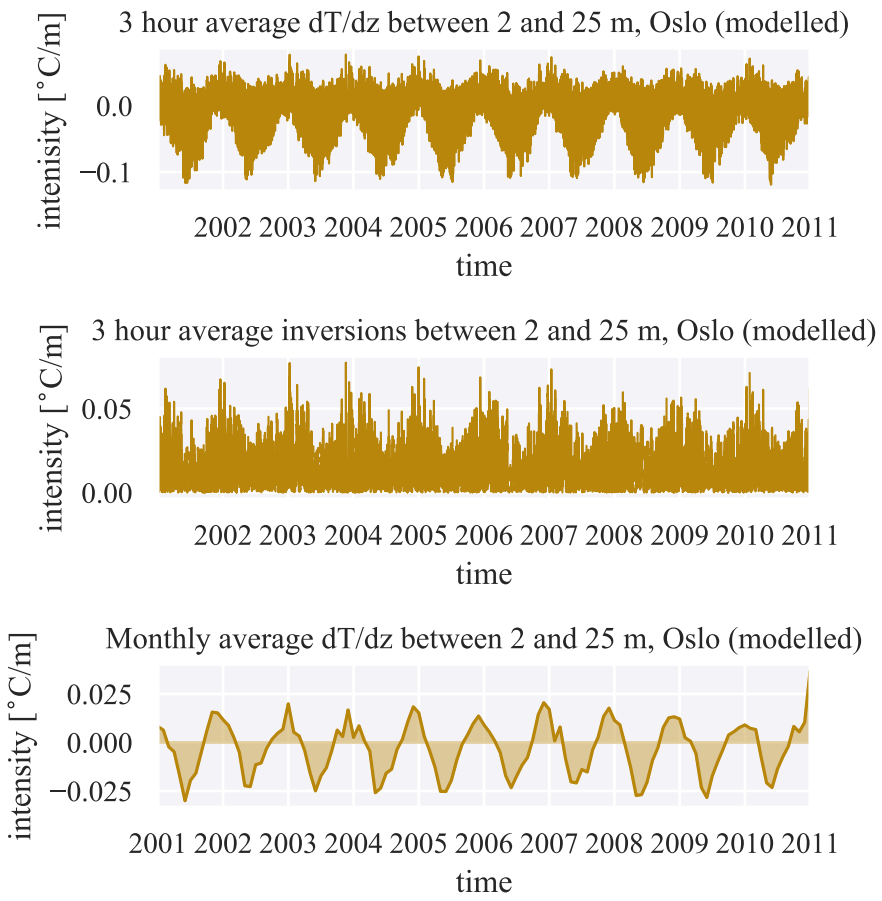


Figure 4.28: The 3 hour average temperature gradient, dT/dz (top), the 3 hour average inversion, $dT/dz > 0$ (middle), and the monthly average temperature gradient (bottom) for Oslo, between 2 meters and 25 meters, measured in intensity [$\circ$ C/m]

In Figure 4.28 the hourly average dT/dz and inversions, and the monthly average dT/dz obtained from the model is shown. Comparing the hourly average dT/dz with the observed hourly average dT/dz in Figure 4.4 shows that the model has a bias towards negative dT/dz, with the model reaching values below -0.1 $°C/m$. The highest value of dT/dz is 0.08 $°C/m$, while the lowest value is -0.12

$°C/m$. The inversion episodes are stronger during winter, whilst the layer is more unstable in the summer.

Comparing the modelled hourly average inversions to the observations finds that the model shows weaker inversion episodes. The strongest modelled inversion episode has a value of approximately 0.08 $°C/m$ at the end of 2003. During the summer, the episodes usually lie around 0.025 $°C/m$, while the observations showed that the summer episodes usually lay around 0.1 $°C/m$. It is interesting to note that the model shows weaker inversion episodes during the winter of 2007/2008 and the winter of 2008/2009. This can also be seen in the observations. However, the model shows strong episodes in the winter of 2002/2003 and the winter of 2003/2004, while the same period for the observations does not.

The monthly average dT/dz also show bias towards negative values of dT/dz. For the negative values of dT/dz, the model is able to yield the same magnitude as the observed dT/dz, both reaching a value of approximately -0.025 $°C/m$. The model manages to capture the shape of the winter of 2003/2004, and the switch from positive to negative monthly dT/dz around the same time as the observations. The peaks of the monthly average also appears at the same time as for the observations.

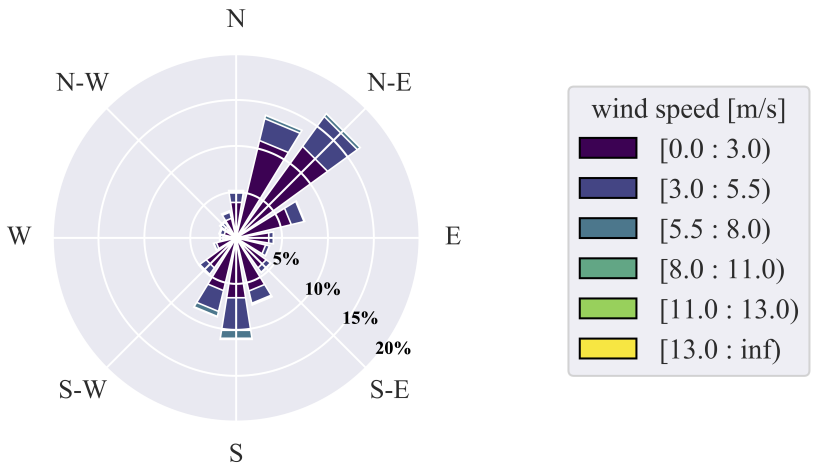


Figure 4.29: The model wind distribution for the overall observational period from 2001 to 2010.

Comparing the modelled wind distribution in Figure 4.29 to the observed wind distribution in figure 4.8, reveals that the modelled wind is more shifted towards north/south. While most of the observed wind arrives from northeast/east and east, the modelled wind arrives from northeast and northeast/north. The model has a horizontal grid resolution of 3 km, and might therefore not be able to reproduce the topography of Grorud valley, which is important in capturing the local meteorology. The modelled winds are also weaker than the observed. More winds have a wind speed of 3 m/s or slower, and fewer winds have a wind speed of 8 m/s or faster. The modelled wind speed had a maximum of 13 m/s.

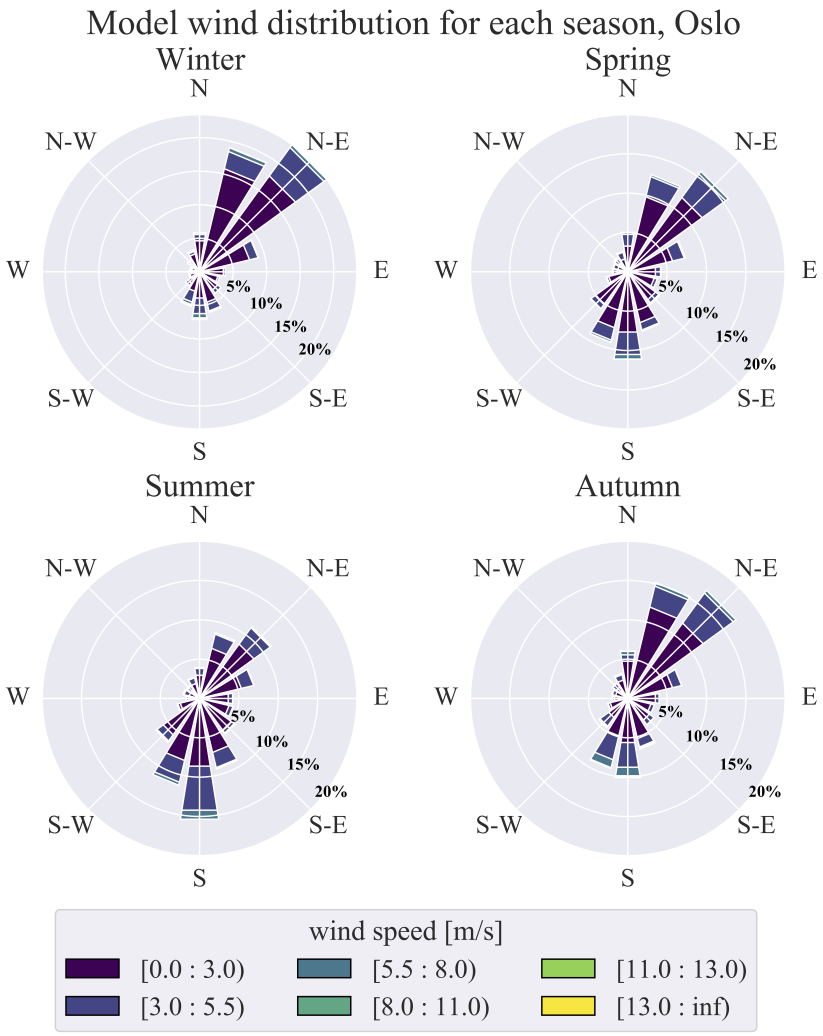


Figure 4.30: The model wind distribution for each season in Oslo

The seasonal wind plots also show that the model winds are more shifted towards north-south than the observations. The variation is captured well by the model, showing that most of the winds during the winter comes from northeast region, while in spring, the winds are more evenly distributed between the northeast and southwest region. During summer, the model displays more winds coming from the southwest region than the observations, but manages to capture the autumn distribution. This suggests that the seasonal large scale circulation is well represented in NORA3. The wind roses also reveal that the model underestimates the wind speed for all the four seasons, but particularly for winter and spring.

To compare the modelled and observed seasonal variations of the temperature gradient and the inversions, they were plotted together in Figure 4.31. The plot displaying the multiyear monthly average dT/dz displays that the model has a

negative bias, just as we saw in Figure 4.28, with the exception of April and July. The difference in modelled and observed data is most prominent for the positive values. The model performs best in replicating the monthly averages in July and August, and performs worst in November and December.

For the lower plot of Figure 4.31, showing the multiyear average, we find that the model for all seasons underestimate the observed inversion intensity average, especially during summer. The difference in modelled and observed inversions are smallest in January, February and April, and is largest in May and June. The model is not able to reproduce the shallow and short lived summer inversions, either by strength or occurrence.

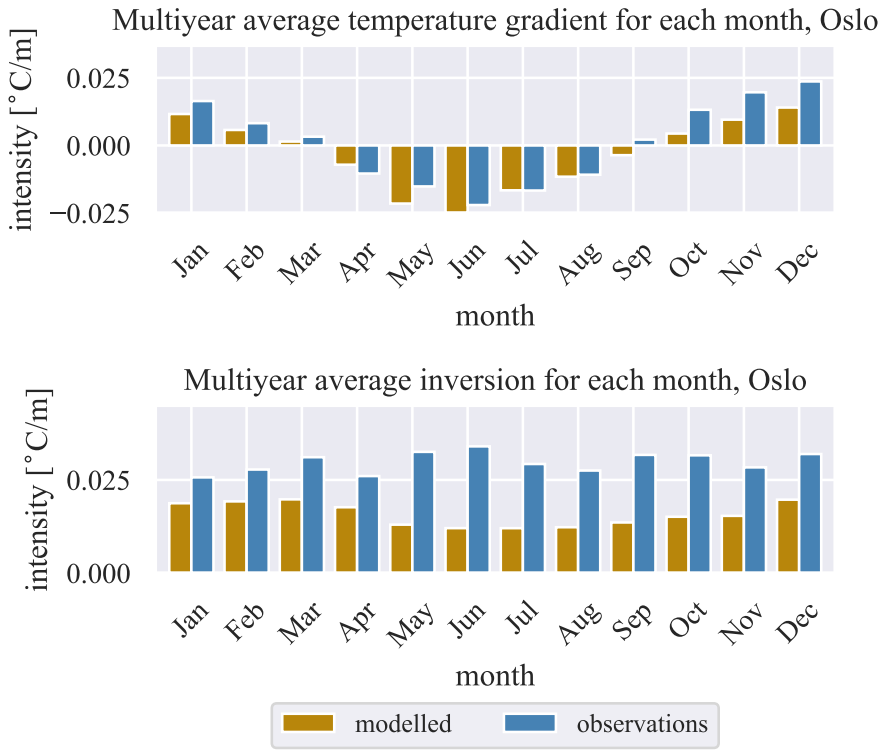


Figure 4.31: The multiyear monthly average dT/dz for each month (top), and the multiyear monthly average inversion intensity for each month (bottom) from the model output (yellow) and the observations (blue).

Scatter plot showing the correlation between the observed and modelled temperature gradient and inversions is shown in Figure 4.32. The correlation between the modelled and observed dT/dz has a value of 0.8, while the correlation between the modelled and observed inversions has a value of 0.53. This confirms what was shown in Figure 4.31; the model is better at showing the observed variations of negative value of dT/dz than the inversions,

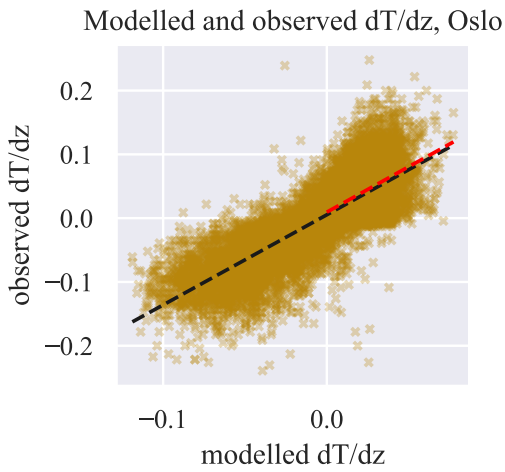


Figure 4.32: Scatter plot of the observed dT/dz and modelled dT/dz. The black line shows the linear relationship between the observed and modelled dT/dz, while the red line shows the linear relationship between observed and modelled inversions.

Looking at the yearly average, we find that the model is able to yield the same variation as the observations that are available, though again, the model shows weaker intensities. The model data shows an increase from 2001 to 2003, and both the model and the observations show a decrease from 2003 to 2005. The model average again increases from 2005 to 2007, and the model and observed average decreases from 2007 to 2008. The observed averages increases slightly from 2008 to 2009, while the modelled averages decreases slightly from 2008 to 2009, but the difference is small. From 2009 to 2010, both the observed an modelled average increase.

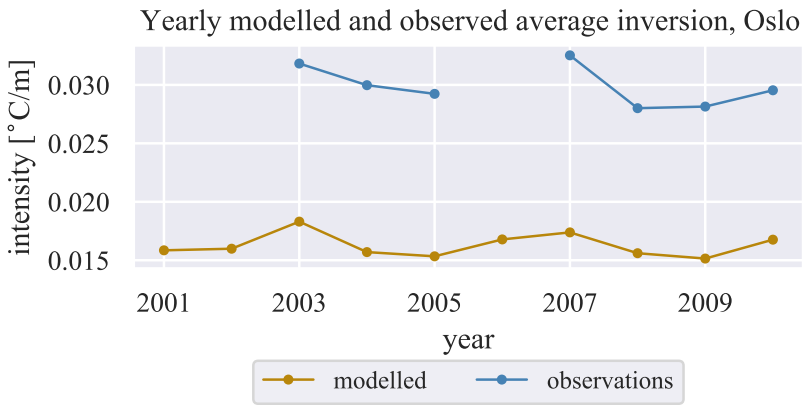


Figure 4.33: The modelled (yellow) and observed (blue) yearly average inversion intensity for Oslo.

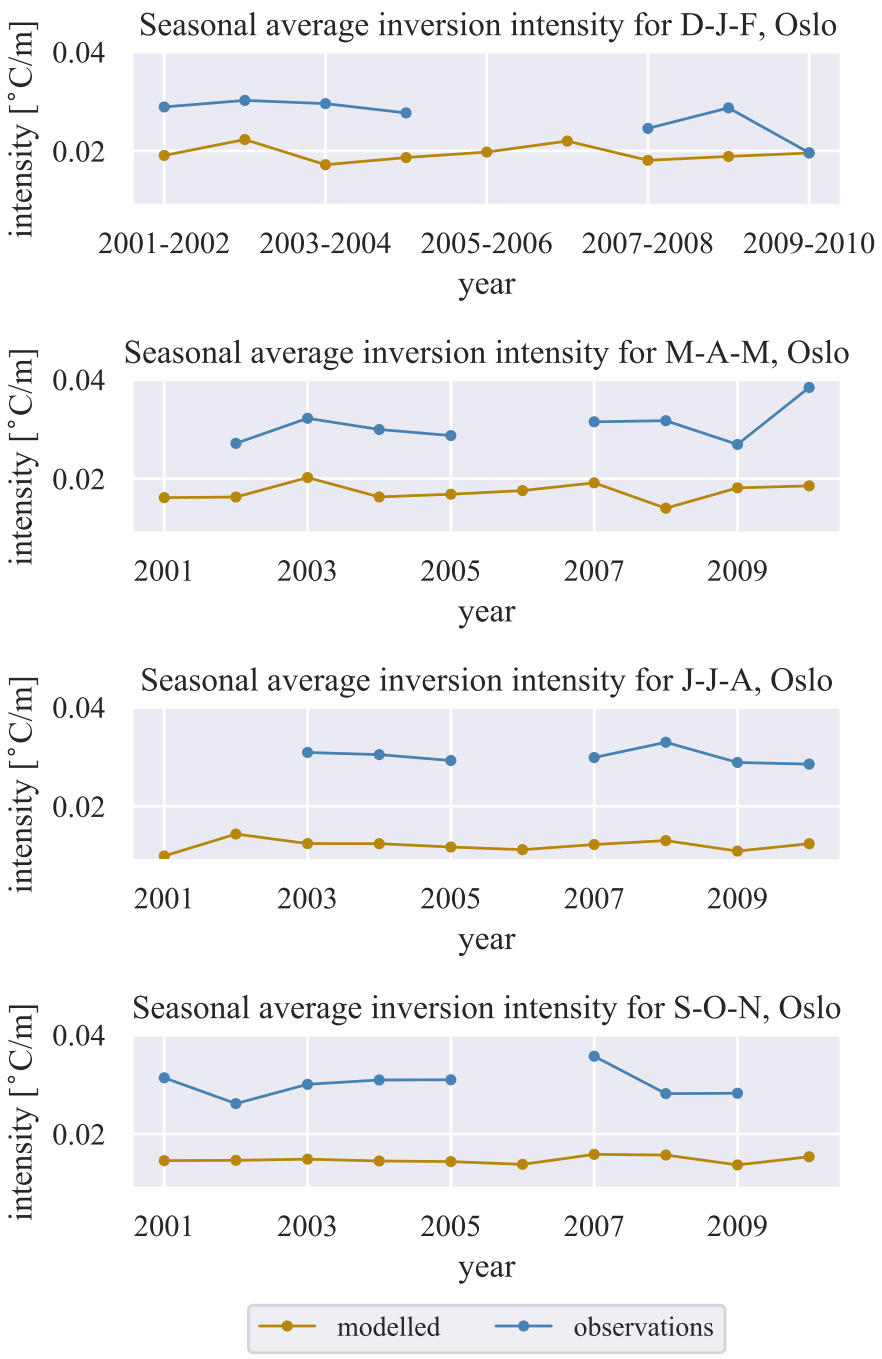


Figure 4.34: The model (yellow) and observed (blue) seasonal average inversion intensity for Bergen. From the top: winter (December, January, February), spring (March, April, May), summer (June, July, August) and autumn (September, October, November).

In Figure 4.34, we compare the variation of the modelled and the observed inversions for each of the seasons. The upper plot, displaying the variations in the winter averages, shows that the model underestimated the average intensity for the winter of 2001/2002 by almost 0.01 $°C/m$. Both the modelled and observed inversion intensity increases in the winter of 2002/2003, and decrease in the winter of 2003/2004. However, the model does not show the same variations as the observations from the 2007/2008 winter to the 2009/2010 winter.

The second plot from the top shows the variation for the spring average. Here, the magnitude of the model average is approximately 0.01 $°C/m$ below the observed average until til 2007. In 2008, the difference increased, while in 2009 it decreased. In 2010, the difference between the modelled and observed average was almost 0.02 $°C/m$. The model average shows the same variation as the observed average from 2002 to 2005, but differs from 2008 to 2010.

The third plot from the top shows the variations for the summer months. The model average is even throughout the study period, but both the model and the observations show a slight decrease from 2003 to 2005, and an increase from 2007 to 2008. From 2008 to 2009 they both show a slight decrease, but the model show an increase from 2009 to 2010, while the observations do not. For the summer averages, the model underestimates the average intensity with almost 0.02 $°C/m$.

Lastly, we compare the average for autumn for the study period, as can be shown in the lower plot in Figure 4.34. The difference between the model and the observations is approximately 0.15 $°C/m$. Whereas the model show almost no variation from 2001 to 2005, the observations had a decrease in the average in 2002 and a slight increase from 2002 to 2005. From 2007 to 2008, the observations show a steeper decrease than the model, and from 2008 to 2009 the observed average showed no variation, while the modelled average decreased.

Bergen

We will now look at how the model output performs for Bergen. The hourly average dT/dz and inversions and the monthly average of dT/dz is found in Figure 4.35. We recall from Figure 4.18 that the observations showed strong negative and positive values for dT/dz during summer, and weaker inversions during winter. The modelled dT/dz in the upper plot of Figure 4.35 aslo show strong negative values during summer, but most of the stronger inversions can be found in the late autumn or early winter. The highest value of dT/dz is 0.13 $°C/m$ and the lowest is -0.09 $°C/m$, both intensities being weaker than the observed maximum an minimum.

The middle plot, showing the model output of the hourly average inversions, show an average of the maximum summer inversions at approximately 0.05 $°C/m$, weaker than the observed. Here, the winter inversions are stronger. The observations showed winter maximums lying around 0.05 $°C/m$, while in the model the winter maximums lie around 0.1 $°C/m$.

Comparing the model monthly average to the observed monthly average reveals that the model shows a bias towards positive dT/dz. In Figure 4.35, the magnitude between positive and negative monthly average dT/dz is more evenly

distributed, whereas the observation showed most of the monthly average to be below zero. The model and the observations display similar variations, as we find the same shapes for most of the winters in the observational period.

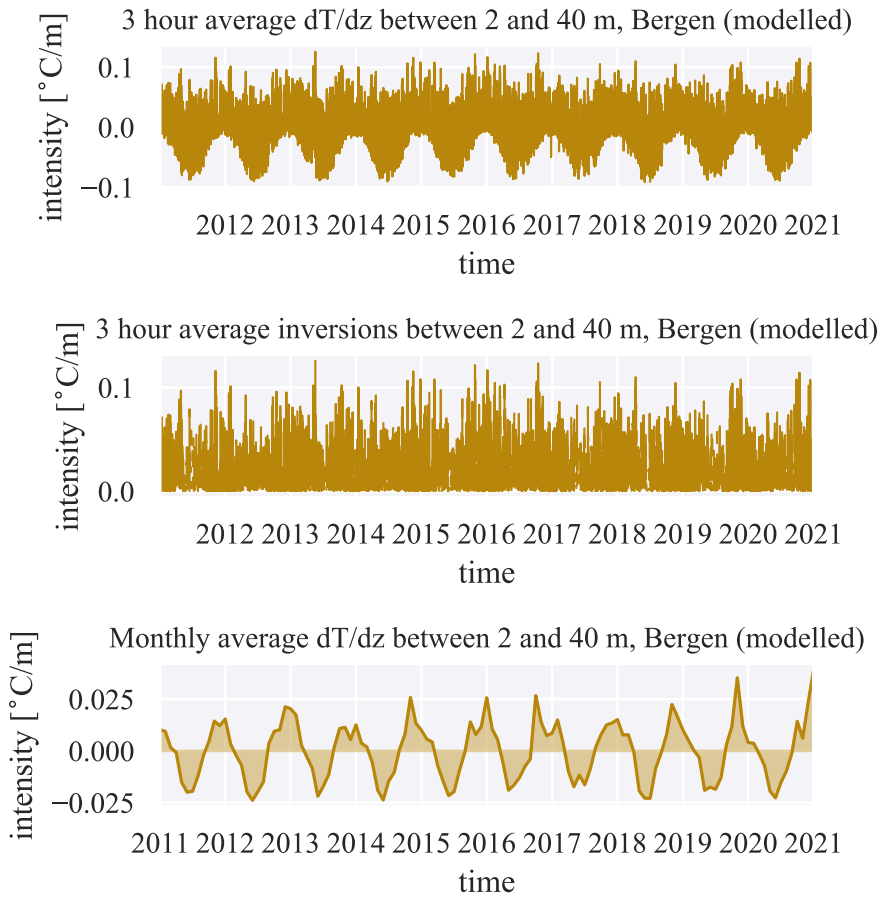


Figure 4.35: The 3 hour average temperature gradient, dT/dz (top), the 3 hour average inversion, $dT/dz > 0$ (middle), and the monthly average temperature gradient (bottom) for Bergen, between 2 meters and 40 meters, measured in intensity $[°C/m]$

The modelled wind distribution is shown in Figure 4.36. Here we find that the model is able to reproduce the wind direction well, specially compared to the modelled wind direction in Oslo. However, the modelled wind distribution shows that there are weak winds coming from northeast/east and southwest/west, which was not shown in the observations. It overestimates the abundance and strength of the winds coming from the south, and the abundance of winds coming from the east. The model also show a larger number of winds between 8 and 11 m/s coming from the southeast/east. Like the observation, more winds are arriving from the southeast region than the northwest region. One plausible explanation as to why the winds are better represented in Bergen than in Oslo

is that the Bergen valley is more defined than the Grorud valley, and therefore more easily captured in the model due to its vertical resolution.

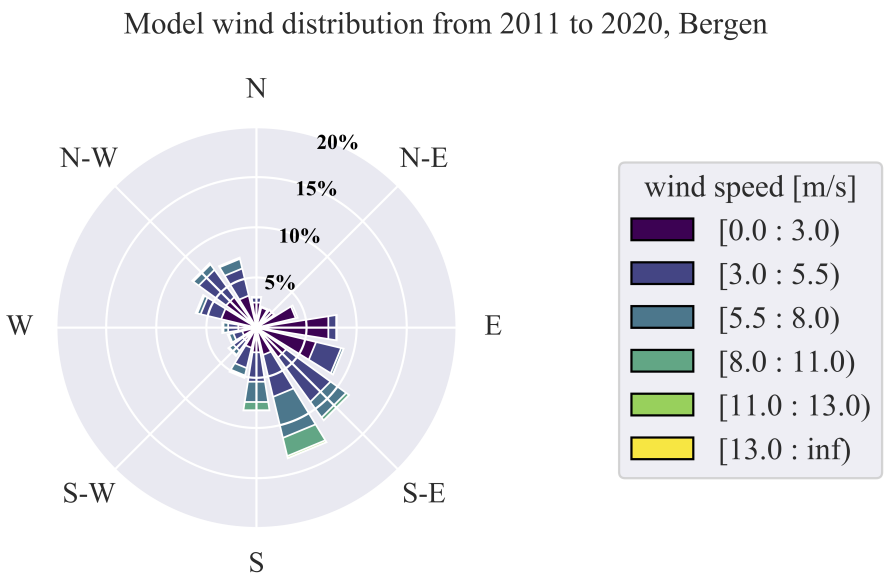


Figure 4.36: The model wind distribution for the overall period from 2011 to 2020

The modelled seasonal variation is shown in Figure 4.37. Many of the same features found in the observations can be seen, but just like for the wind distribution in Figure 4.36, there is an abundance of winds from the southwest to the south region and from the northeast to the east region that is not shown in the observations, for all seasons. In the winter season the model show wind speeds at 5.5 m/s or less arriving from the southeast/east and strong winds reaching a speed between 8 and 11 m/s arriving from the south. Like for the observations, the highest portion arrives from southeast/east, but the model overestimates the wind speed. The fastest wind speed in the model reaches a maximum of 18 m/s.

In the spring, more winds arrive from the northwest region, just like we saw in the observations, with 10% arriving from the northwest, 9% arriving from the northwest/north, and 8% arriving from northwest/west. The overestimation of winds coming from south and east are still present, and where the observations show that 25% of the winds arrive from southeast/south, the model only shows that around 12% arrives from this direction.

More wind arrives from the northwest in the summer. Though the observations showed a higher percentage arriving from northwest, the model performs well in replicating the wind speed of these winds. The same can be said for the winds arriving from the southeast region. They are similar to the observations in terms of speed, but the wind direction is more distributed. In the autumn, the distribution shifts so that most of the wind is arriving from the south-southeast-

east region. Like for all the seasons, the winds arriving from east and south in the model are overestimated.

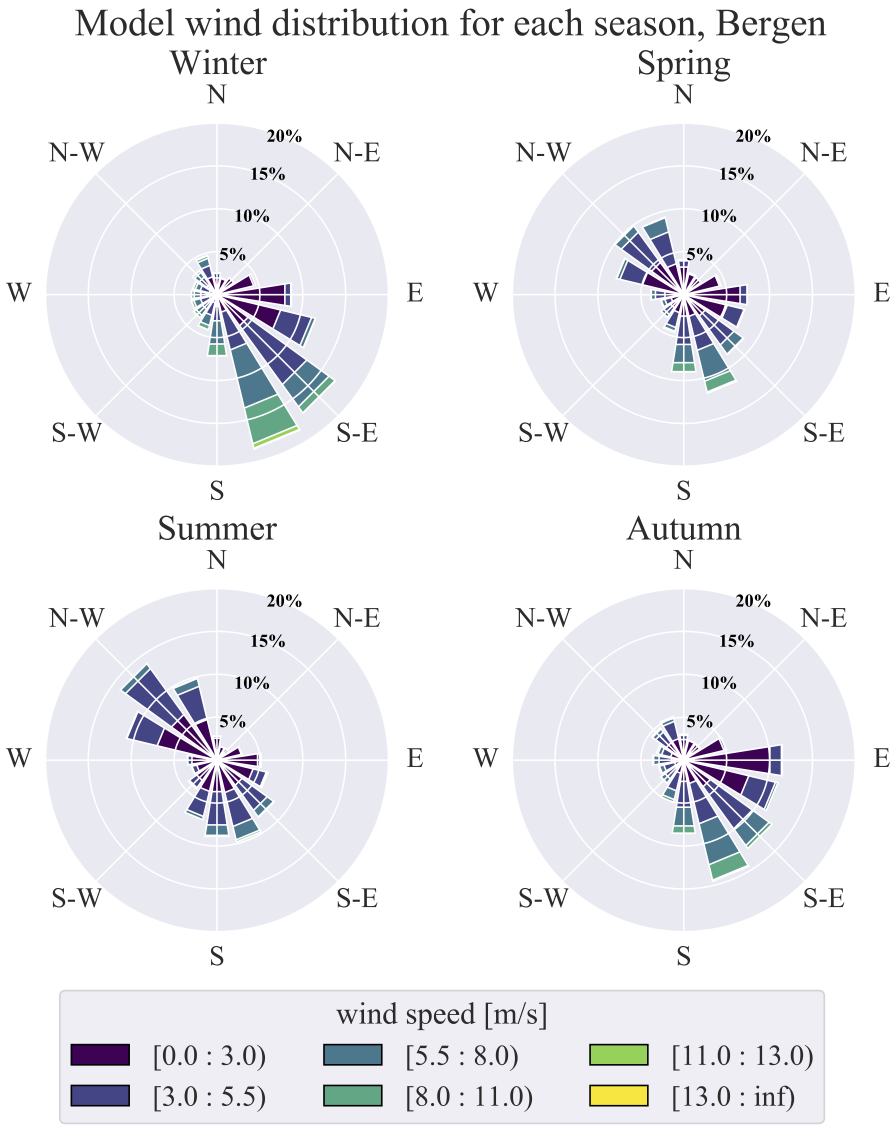


Figure 4.37: The model wind distribution for each season in Bergen.

To see how well the model performs in replicating the seasonal variations of dT/dz and the inversions for Bergen, we study Figure 4.20. The upper plot showing the multiyear average dT/dz for each month reveals that the model shows a bias towards positive dT/dz. The observed multiyear average for January to March and from October to December all show a negative value, while the model shows a positive value. From April to August, the model show weaker negative value than the observations, and in September, the multiyear average of dT/dz is zero. The model performs best in May and June.

From the lower plot of Figure 4.38, we gather that the model overestimates the multiyear monthly average intensity from January to March, November and December, and underestimates the intensity from May to September. In April and October, the modelled and observed multiyear monthly average are similar. It is interesting to note that both the observations and modelled multiyear monthly average is strongest around April and October. The underestimation during the summer months may be due to either not managing to reproduce the abundance of those short lived, shallow summer inversions, or that the summer inversions in the model are weaker.

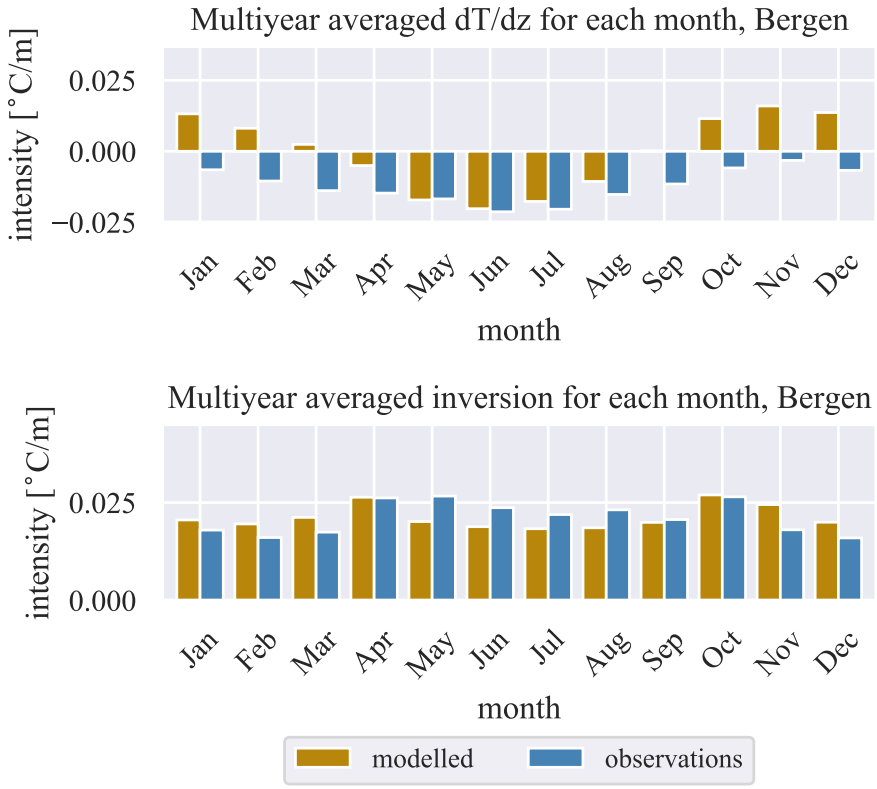


Figure 4.38: The multiyear monthly average dT/dz for each month (top), and the multiyear monthly average inversion intensity for each month (bottom) from the model output (yellow) and the observations (blue)

In Figure 4.39, we find the scatter plot of modelled and observed dT/dz. The correlation between the temperature gradients is 0.64, while the correlation between the inversions is 0.29. In the scatter plot we observe more clearly that the model overestimates the weaker inversions, but is better at replicating the intensity of the temperature gradients.

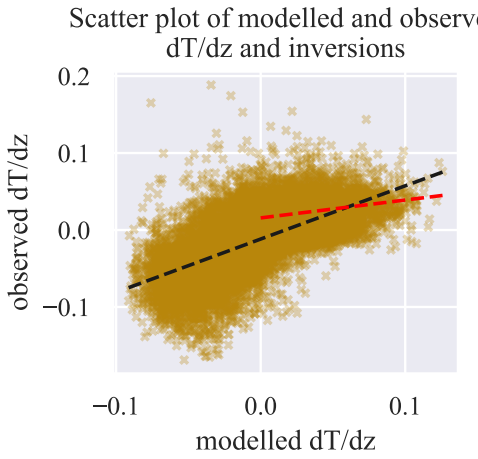


Figure 4.39: Scatter plot of the observed dT/dz and modelled dT/dz, Bergen. The black line shows the linear relationship between the observed and modelled dT/dz, while the red line shows the linear relationship between observed and modelled inversions.

Figure 4.40 shows the yearly average inversion. The observed average intensity for 2011 and 2020 has been discarded, but the model shows the same variation from 2012 to 2019 as the observations. The largest difference can be found in 2016, where the modelled yearly average overestimates by approximately 0.025 $°C/m$. The model shows a yearly average inversion intensity of 0.0265 $°C/m$, while the observations shows a yearly average inversion intensity of 0.023 $°C/m$. From 2013 to 2015, the model decreases less than the observations. However, the overall conclusion is that the modelled yearly average inversions are similar to the observed yearly averages, both in terms of variation and magnitude.

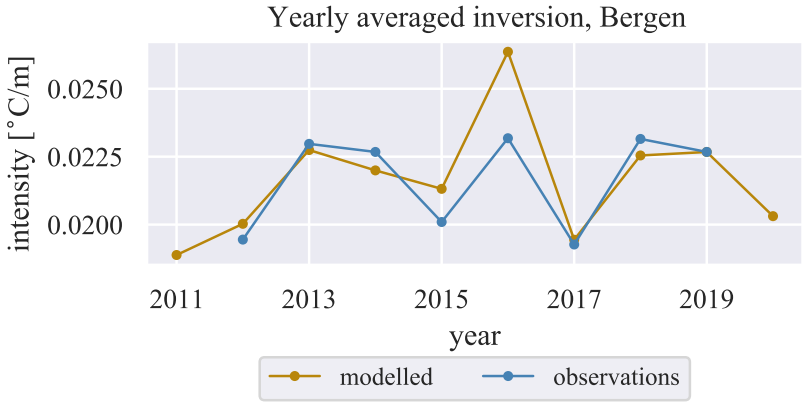


Figure 4.40: The modelled (yellow) and observed (blue) yearly average inversion intensity for Bergen.

In Figure 4.41 we can compare the average of the observed and modelled

inversions for each season. The upper plot, displaying the average for the winter season, shows that the model is good at replicating the variation we see in the observations, with the strength of the modelled average inversion intensity being slightly stronger. The largest difference is found in the winter of 2015/2016, where the difference is approximately 0.01 $°C/m$, but both the observed and modelled inversions show the same increase from the winter of 2014/2015 to 2015/2016, and decrease from the winter of 2015/2016 to 2016/2017.

The second plot from the top shows the yearly averaged inversions for spring. From 2012 to 2014, the model output shows similar variations as the observations, but in 2015 however, the observed average drops, while the modelled average increases. The model average in 2016 is close to the observed average. From 2016 to 2018, both the model output average and the observed average decreases in 2017 and increases in 2018, but the model shows stronger variations. In 2019 and 2020, the model underestimates compared to the observed averages. All in all, the model is able to reproduce averages similar to the observations, but is not equally good at replicating the small variations.

The plot displaying the averages for the summer months, shows that the model underestimates the summer averages obtained from the observations. We find a similarity in the variations, with both the model and observations having increasing averages from 2011 to 2014. In 2015, the model average decreased while the observed average increased. From 2016 and to 2020, the variation is again similar. Like for spring, the variations throughout the observational period is small. Finally, for autumn in the lower plot, we observe that the model performs well from 2011 to 2018, both in terms of variation and magnitude of the autumn averages. In 2019, the model shows suddenly a much higher average of 0.03 $°C/m$.

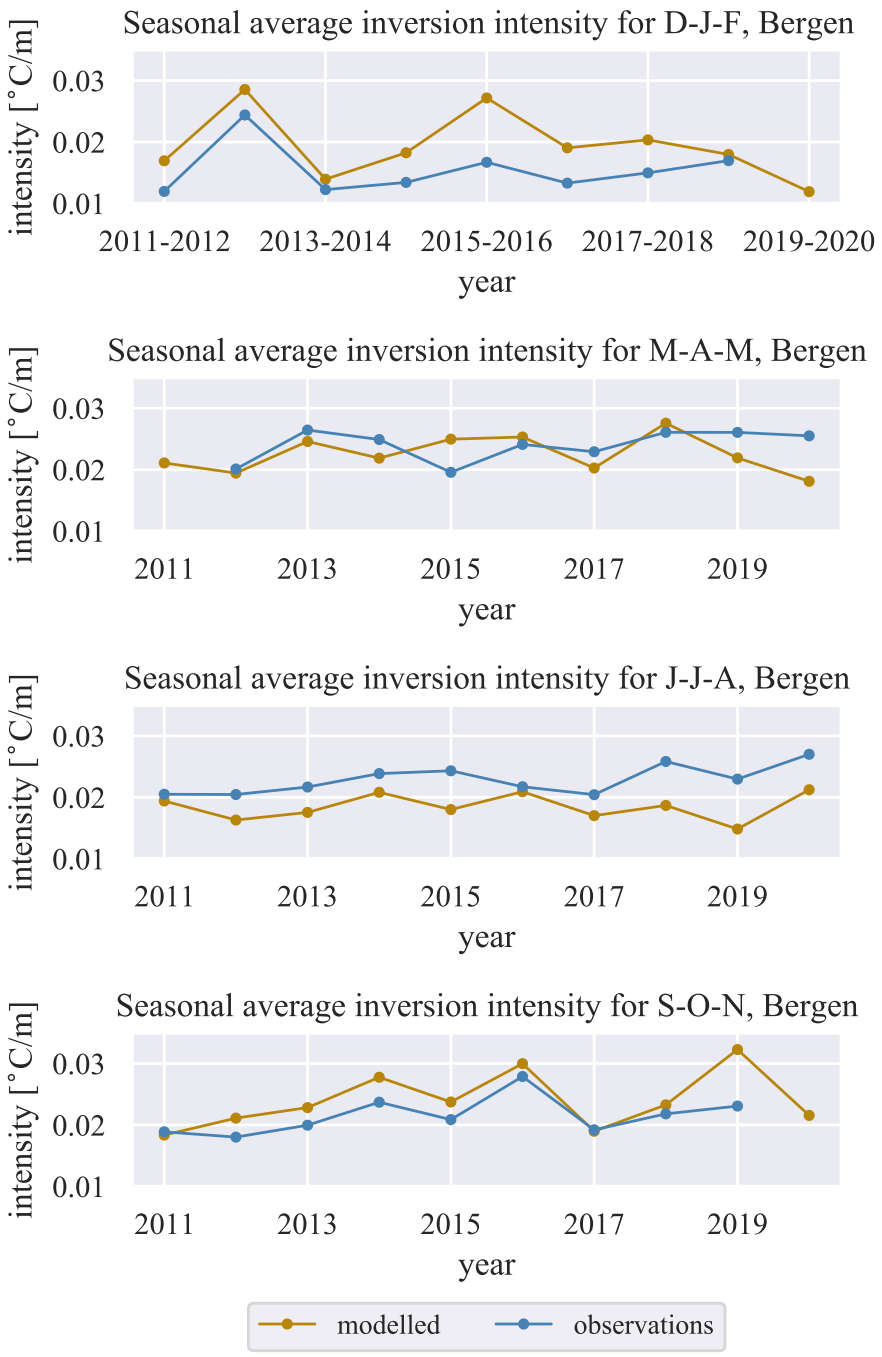


Figure 4.41: The model (yellow) and observed (blue) seasonal average inversion intensity for Bergen. From the top: winter (December, January, February), spring (March, April, May), summer (June, July, August) and autumn (September, October, November).

4.3 Analyzing the surface energy balance for selected inversion episodes

For this section, we will investigate the SEB in the model. We will look into two periods, one where the intensity of the modelled inversions are similar to the observed, and one where the modelled inversion intensity differ from the observations. Because there is no available observations of the SEB fluxes, we can only compare the model fluxes for the two periods with each other.

Oslo

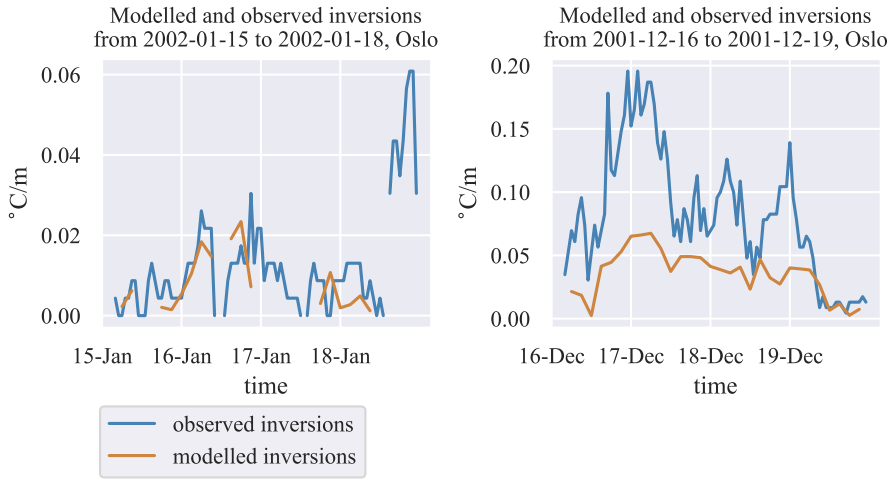


Figure 4.42: The observed inversion intensity with a time resolution of 1 hour (blue), and the modelled inversion intensity with a time resolution of 3 hours (orange). To the left we is a period where the model and the observations are similar, and to the right is a period where they differ.

The plot to the left of Figure 4.42 shows the modelled and observed inversions for a period where the intensities are similar, while the plot to the right shows the same, but for a period where the modelled inversions are weaker in intensity than the observations. It is important to point out that the modelled inversion intensity have a time resolution of 3 hours, while the observed inversion intensity have a time resolution of 1 hour.

As we recall from previous results, the model was not able to reproduce some of the stronger inversion intensities in Oslo. In the period shown to the left of Figure 4.42, both the observed and modelled inversions are weak in intensity. Ignoring the observed inversion during the 18th of January, the observed inversions have a maximum intensity of 0.31 $°C/m$, while the modelled have a maximum of 0.24 $°C/m$. In the period shown to the right, the observed inversions are stronger than the modelled inversions. The observed inversions reach an intensity of 0.2 $°C/m$, while the model only reaches to approximately 0.07 $°C/m$. The inversions are coherent throughout the period in 2001, but not for the episode period in 2002.

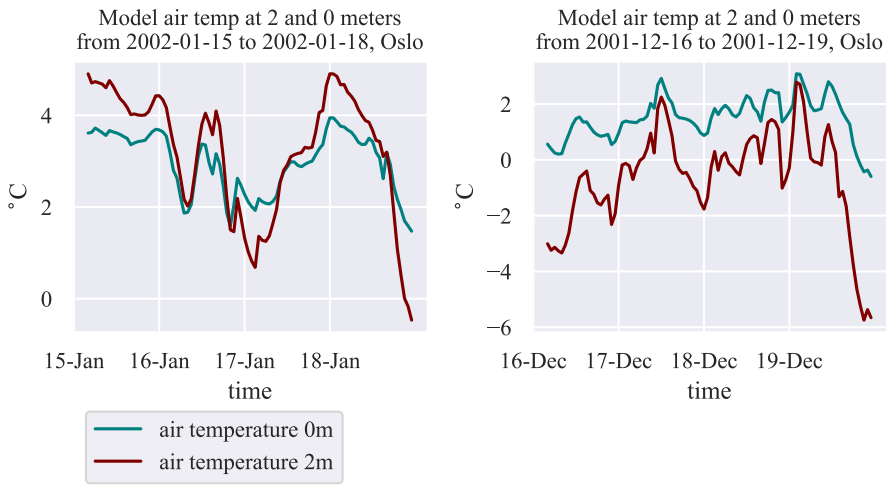


Figure 4.43: The air temperature at 2 meters (red) and the skin temperature (green). The plot to the left is for the period where the modelled and observed inversions are similar, and the plot to the right is for the period where the modelled and observed inversions differ.

Figure 4.43 displays the temperature at 0 meters, also called the *skin temperature*, and the air temperature a 2 meters. Firstly studying the plot to the left and comparing it to the inversions for the same period reveals that the inversions are present when the 2 meter air temperature is higher than the skin temperature. Figure 4.42 shows the inversion intensity increasing from 18:00 on the 15th of January to 12:00 on the 16th. For the same period, Figure 4.43 shows that the temperature difference between 2 and 0 meters decrease, while both temperatures drop. The 2 meter temperature decreases from 4.5 $°C$ to 2 $°C$, and the skin temperature decreasing from 3.5 $°C$ to just below 2 $°C$. At approximately 10:00 on the 16th, the temperature profile would show little temperature change between 0 and 2 meters, and an increase in temperature with altitude from 0 to 25. We observe that the inversion occur during night, and breaks up at sunrise when both the skin temperature and the 2 meter air temperature increases. The same mechanism happening from 15:00 the same day, to around 18:00; the temperature drops and an inversion occurs. For the weak inversion episode observed from 15:00 on the 17th to around 12:00 on the 18th, the temperature does not drop on the night of the 18th, but the 2 meter air temperature is higher than the skin temperature. Notice that the temperature gradient at 00:00 on the 18th between 0 and 2 meters have an intensity of approximately 0.5 $°C/m$, while the temperature gradient between 2 and 25 meters have an intensity of oonlu 0.005 $°C/m$. The inversion disappears when the temperature at 2 meters approaches the same temperature as the ground.

To the right in Figure 4.43, we find that the 2 meter air temperature is below the skin temperature for the whole period. Because Figure 4.42 shows that there is an inversion episode for almost the whole period, one would expect the air temperature to be warmer than the skin temperature, but this suggests

that the inversion lies between 2 meter and 25 meters caused by advection of warmer air at higher altitudes. To check the consistency of the model data, we calculated the change in outgoing longwave flux, ΔF, by using the Stefan Boltzmanns law.

$$\begin{aligned}\Delta F &= \sigma(T_{S_2}^4 - T_{S_1}^4) \\ &\approx \sigma(272.5^4 - 276^4) \\ &= -16.37\ W/m^2\end{aligned}$$

Where T_{s_1} is the skin temperature at 12:00 the 19th of December, and T_{s_2} is the skin temperature at 00:00 the 20th. By comparing it to the surface outgoing longwave flux (shown in Figure B.2), were the the ΔF is approximately -15 W/m^2, and we can conclude that the plotted temperatures matches with the radiative fluxes.

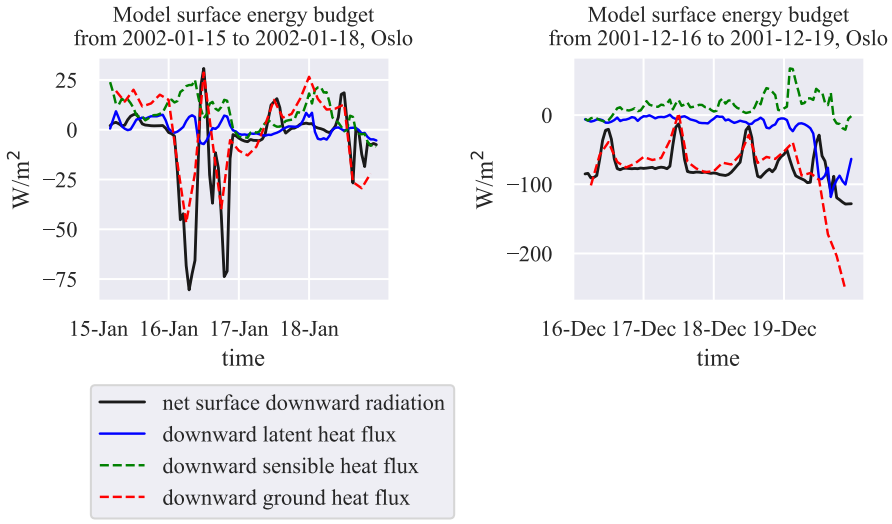


Figure 4.44: Surface energy budget showing with the net radiation (black), the latent heat flux (blue), the sensible heat flux (green, dashed), and the ground heat (red, dashed). The plot to the left is for the period where the modelled and observed inversion intensity are similar, and the plot to the right is for the period where the modelled and observed inversion intensity differ.

From the plot to the left showing the SEB in Figure 4.44, we find that for the first inversion episode (from 18:00 on the 15th to approximately 10:00 on the 16th), both the net surface downward radiation and the downward ground heat flux drops. The net radiation drops to approximately -75 W/m^2 and the ground heat flux drops to almost -50 W/m^2. The sensible heat flux increase, indicating that the flux is pointing downwards to the ground, and the latent heat flux is weak but pointing downwards most of the inversion episode as well. We observe similar variations for the second inversion the same day. The fluxes for the third inversion period, from 12:00 on the 17th to 12:00 on the 18th, show little variations. The ground heat flux is approximately 25 W/m^2 and the sensible heat flux is just below 25 W/m^2, due to the air temperature being

higher than the skin temperature, as we saw in Figure 4.43. The latent heat flux throughout the period from the 15th to the 18th of January fluctuates very little around 0 W/m^2.

In the plot to the right in Figure 4.44, the net surface downward radiation is below zero throughout the whole period. We find the diurnal variation to be more apparent here, but all the downward fluxes, except the latent heat flux, decreases from 00:00 the 19th of January to 00:00 the 20th. The fluxes correspond with the drop in the temperatures from Figure 4.43. The ground heat flux drops the most, increasing outwards from 100 W/m^2 to 250 W/m^2. This indicates that the isolation of the ground, most likely snow in this case, has disappeared. The latent heat also drops, suggesting that the air becomes more dry.

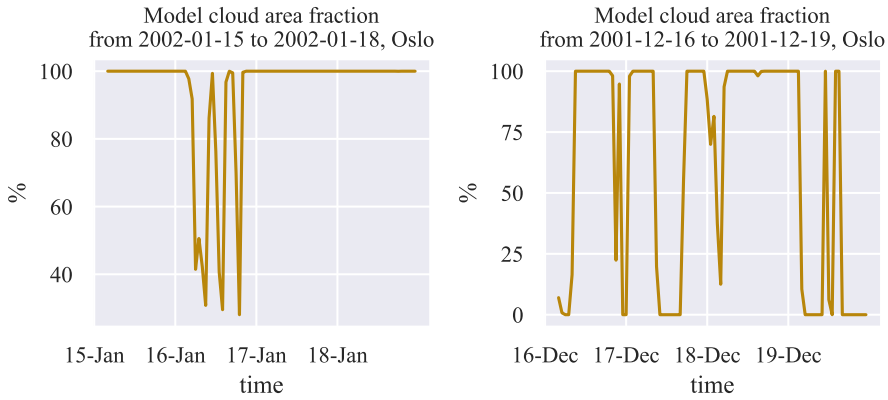


Figure 4.45: The cloud area fraction from the model output. The plot to the left is for the period where the modelled and observed inversions are similar, and the plot to the right is for the period where the modelled and observed inversions differ.

The cloud area fraction for the period in 2002 shows a 100% cloud coverage except during the 16th, explaining why we observe so little diurnal variation except in the period with little cloud coverage, allowing for more incoming solar radiation. The plot to the right, displaying the cloud area fraction from the 16th to the 19th of December in 2001, shows that the cloud coverage varies, and as a result, the surface experiences a stronger diurnal variation in the incoming solar radiation. During the period where the fluxes dropped, the clouds disappear, and less longwave radiation is emitted back towards the surface (seen in Figure B.2).

Lastly, we will look at how the relative humidity varies for these two periods. The plot to the left for the period in 2002 reveals that the relative humidity is stable, mostly only fluctuating between 100% and 95%. The relative humidity drops to 90% on the night of the 16th of January, which is consistent with the drop in the temperatures and the net downward radiation fluxes. We also see the relative humidity dropping for the two other observed inversion episodes as well. For the right side plot of Figure 4.46, the relative humidity is decreasing

from the start of the period, and decreasing more rapidly from the late hours of the 18th down to approximately 50%.

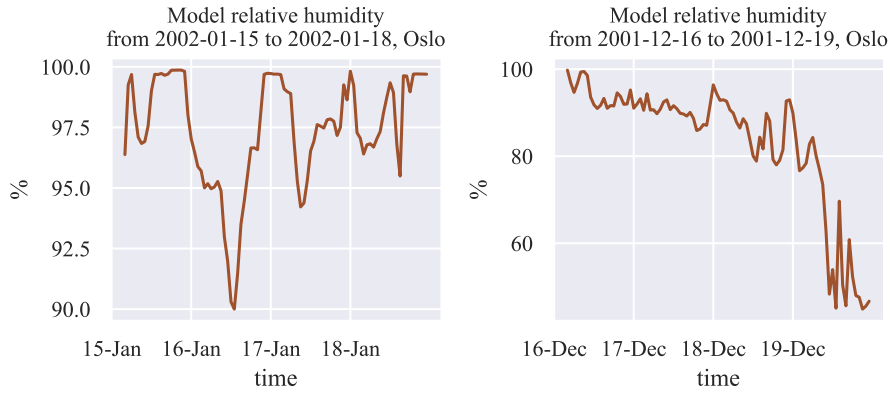


Figure 4.46: The relative humidity at 2 meters from the model output. The plot to the left is for the period where the modelled and observed inversions are similar, and the plot to the right is for the period where the modelled and observed inversions differ.

Bergen

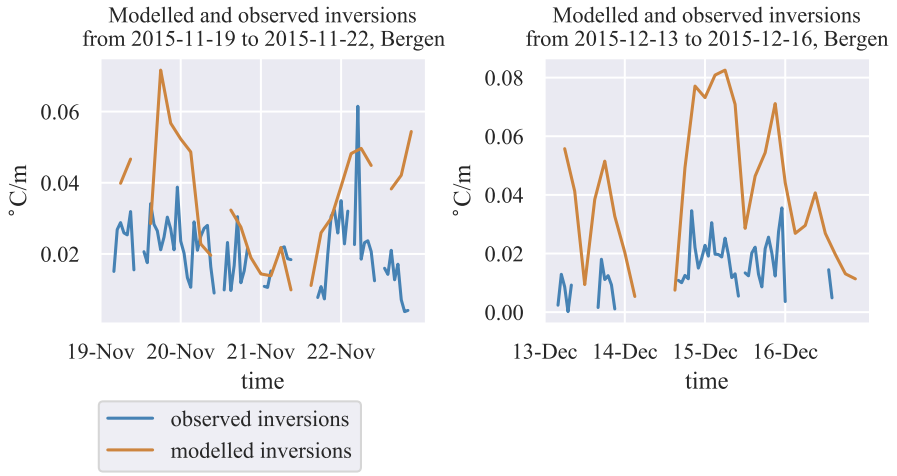


Figure 4.47: The observed inversions with a time resolution of 1 hour (blue),and the modelled inversions with a time resolution of 3 hours (orange). To the left we is a period where the model and the observations are similar, and to the right is a period where they differ.

A similar analysis of the model output was done for Bergen. We chose to look at periods from 19th to the 22nd of November, 2015, when the model inversions are similar to the observed, and from the 13th to the 16th of December, 2015, for when the model inversions have a higher intensity than the observations.

The plot to the right showing the period in November, reveals that the model overestimates the intensity of the inversion episode occurring in the afternoon on the 19th to a a few hours before noon on the 20th, but manages to capture the magnitude of the inversion episodes between 12:00 on the 20th and 12:00 on the 22nd. These inversions are night inversions, just like the ones we observed for the period in January, 2002 in Oslo. For the period in December, where there is higher difference between the model output and the observations, we note that the difference is approximately 0.05 $°C/m$, which is small compared difference in the period we looked at for Oslo. In Bergen, the model overestimates the inversion intensity compared to the observations, while for Oslo, the model was underestimating the inversion intensity compared to the observations. The inversions are more coherent in the December episode, than in the November episode.

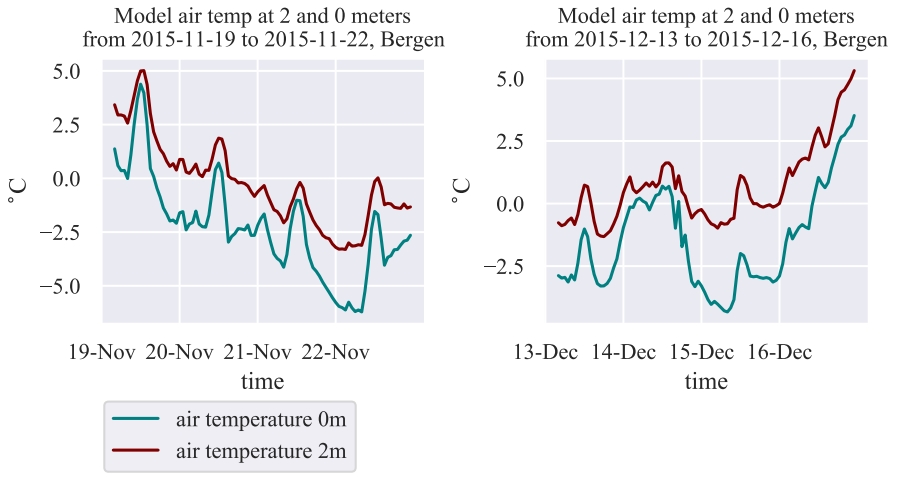


Figure 4.48: The air temperature at 2 meters (red) and the skin temperature(green). The plot to the left is for the period where the modelled and observed inversions are similar, and the plot to the right is for the period where the modelled and observed inversions differ.

The air temperature for the period in November shows that the skin temperature is always colder than the 2 meter air temperature. We note the diurnal variation, showing warmer temperatures around 12:00 for all days. The inversions are night time/evening inversions, and starts when the sun disappears. The temperature is decreasing throughout the period, but start to increase from noon on the 22nd. The period in December show diurnal variations for both temperatures, peaking around noon, with the exception of the 16th. The skin temperature and the temperature at 2 meters increases from noon on the 15th, and throughout the rest of the period. Here too, we find that the skin temperature is colder than the 2 meter temperature. In the period where the model is not showing any inversions, from a few hours after 00:00 to the afternoon on the 14th of December, we observe that the temperature difference between the skin temperature and the 2 meter temperature decrease. When the temperatures decreases and the temperature difference increases from noon on the 14th to

noon on the 15th, the inversion appears again, and becomes weaker as both the skin temperature and the 2 meter temperature increases from 12:00 on the 15th and throughout the remainder of the study period.

Studying the surface energy budget to the left of Figure 4.49, the first thing we notice is the diurnal variation in the net radiation, the sensible heat and the ground heat flux. During daytime, the net radiation reaches a value of approximately 25 W/m^2, and is pointed downwards to the ground, while during nighttime, the net radiation lies at approximately -80 W/m^2 and points out from the ground. For the sensible heat, during evening and night, the value lies around 40 to 50 W/m^2, as the skin temperature is colder than the 2 meter temperature, while during the presence of sunlight, it drops to just below 0 W/m^2, when the skin temperature and the 2 meter temperature becomes approximately the same. The latent heat flux is stable throughout the period, lying just below 0 W/m^2. We note that the latent heat flux is small, and might indicate that the air is dry. The magnitude increases slightly during the daytime. The ground heat is positive during the day, pointing downwards into the ground, and negative during night. In order to balance out the values for the net short- and longwave radiation, the magnitude of the ground heat flux is large. In Figure 4.48, we saw that the temperature decreased throughout the period, but there is no visible change in the variations in the radiation. This hints at winds arriving and thereby causing the colder temperatures.

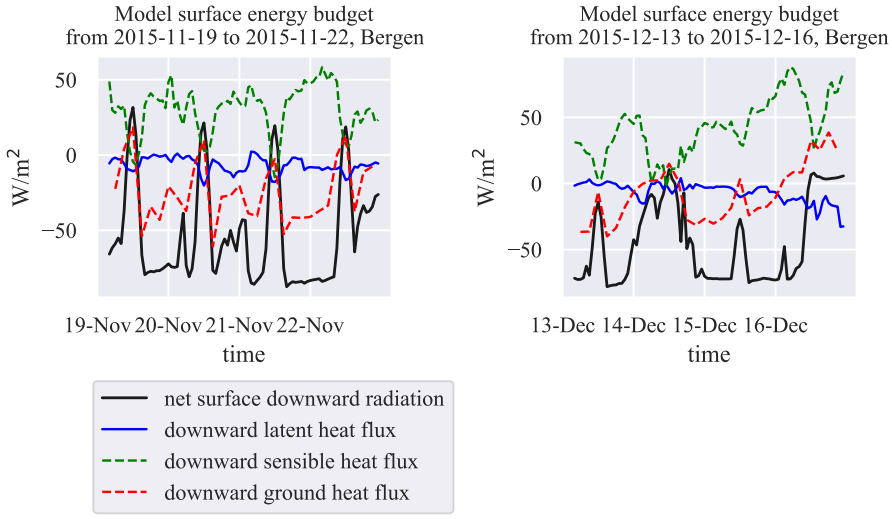


Figure 4.49: Surface energy budget showing the net radiation (shortwave + longwave) in black, the latent heat flux in blue, the sensible heat flux in green,dashed, and the ground heat flux in red, dashed. The plot to the left is for the period where the modelled and observed inversions are similar, and the plot to the right is for the period where the modelled and observed inversions differ

For the December period, shown to the right in Figure 4.49, we find that the magnitude of the net short- and longwave radiation decrease during the 14th is

causing the jump in temperature we observed for the same period in Figure 4.48. Because the skin temperature and the 2 meter temperature were approximately the same during that period, the sensible heat flux decreases to almost 0 W/m^2. When the temperatures increases, so does the downward sensible heat flux, but it dips around 12:00 on December the 16th. The magnitude of the latent heat flux for this period is also small, pointing upwards. The ground heat flux increases from the 15th, becoming small when the net short- and longwave radiation goes towards zero, except during the afternoon of the 16th. There, the net radiation is approximately zero, while the ground heat flux and the sensible heat flux increase, and the latent heat flux decrease. The decrease in downward latent heat flux implies that the relative humidity increase.

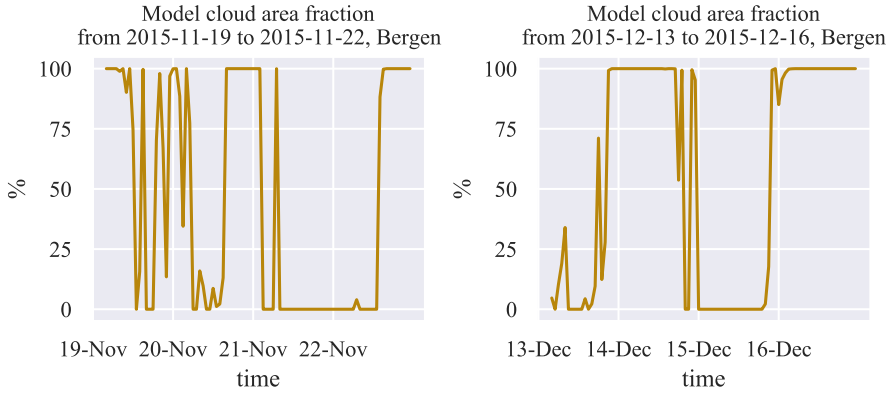


Figure 4.50: The cloud area fraction from the model output. The plot to the left is for the period where the modelled and observed inversions are similar,and the plot to the right is for the period where the modelled and observed inversions differ.

Plotting the cloud area fraction for these two periods (Figure 4.50), we find that from the 19th to the 22nd November, the cloud coverage varies. During the day, the cloud area fraction is small, allowing for radiation to penetrate down to the ground, allowing for the diurnal variation in the modelled surface energy budget in Figure 4.49. At the end of the period, the cloud area fraction increases to 100%, which causing the decrease in the outgoing net radiation, as the clouds absorb and re-emit long wave radiation, they do not reflect the radiation back to the surface. For the period from 13th to 16th of December, we find that from 00:00 on the 14th to 00:00 on the 15th, the plot shows full cloud coverage, explaining why we see the decrease in the outgoing net radiation in the surface energy budget. We see the same tendencies from 00:00 on the 16th and throughout the rest of the study period.

Lastly, we look at the relative humidity, as plotted in Figure 4.51. For the first study period, we find that the relative humidity decreases. There are large variations in the relative humidity. Figure 4.49 showed that the latent heat flux was negative, but small in magnitude. The variation of the latent heat flux was also small due to the cold temperatures. For the second study period, taking place during December, we observe that the relative humidity drops from a few

hours after 00:00 on the 14th to 12:00 on the 15th, but there is not much change in the latent heat flux, also, again, due to the cold temperatures. Because there is not much change in the temperature during this period, an explanation as to why we see this drop in relative humidity might be the appearance of winds.

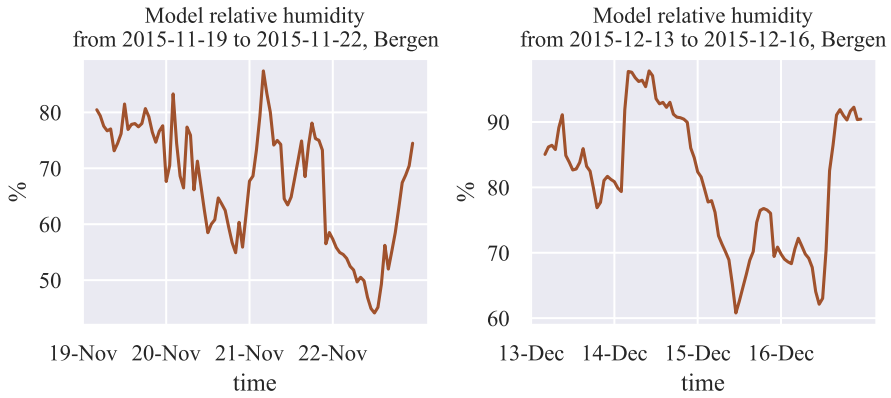


Figure 4.51: The relative humidity at 2 meters from the model output. The plot to the left is for the period where the modelled and observed inversions are similar, and the plot to the right is for the period where the modelled and observed inversions differ

4.4 Analyzing trends of a 20 year period

Lastly, we will look into the trends of the NO_2 concentrations and the modelled inversions from NORA3. This time, we will look into a study period of 20 years for both study areas, from 2001 to 2020. We will also study the trends in the the seasonal averages, and test there are any significant trends using a t-test.

Oslo

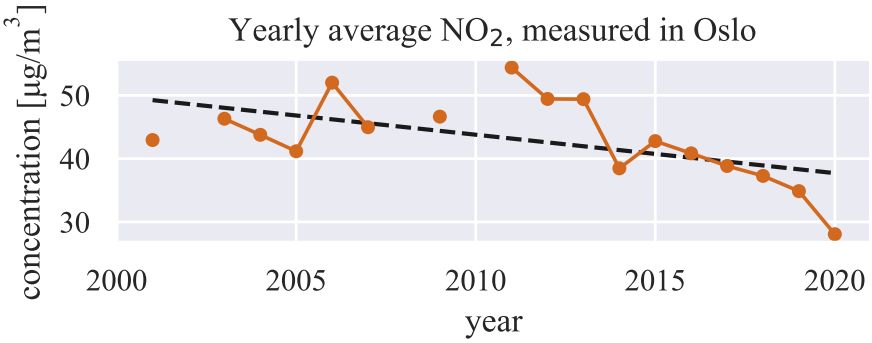


Figure 4.52: Yearly average NO_2 measured at Alnabru in Oslo, from 2001 to 2020, with trend line.

The yearly average of the NO_2 concentrations show an increase from 2001 to 2011, where it reaches a maximum of approximately 55 $\mu g/m^3$, and declines from 2011 to 2020. In 2020, the yearly average was just below 30 $\mu g/m^3$. The trend for the overall period is declining, as shown in Figure 4.52, where the decline is approximately at a rate of 0.6 $\mu g/m^3$ per year. A t-test was performed numerically. Our null hypothesis states that the trend is not significant (the slope of the linear regression is equal to zero), but the result showed that the trend indeed is significant.

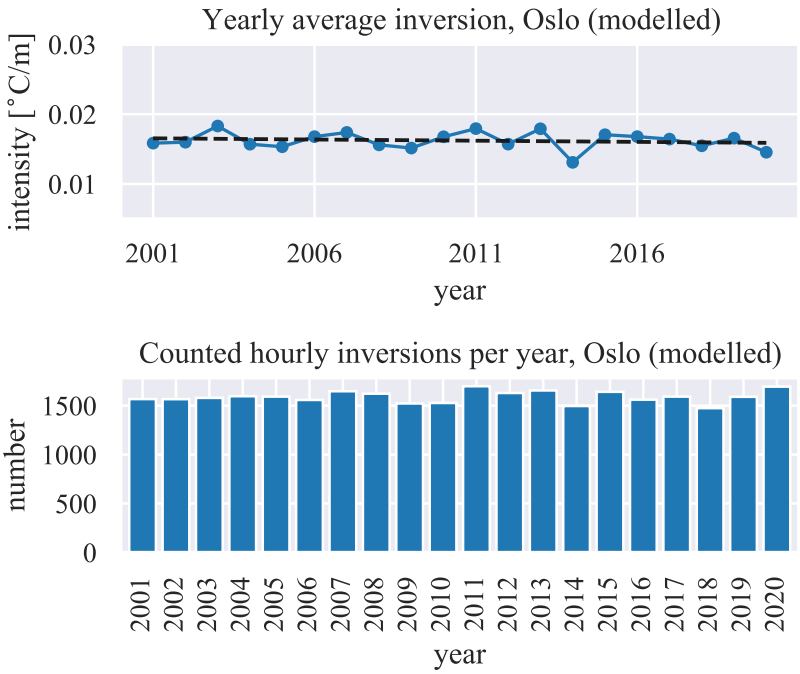


Figure 4.53: The modelled yearly average inversion intensity (top), and the number of hourly counted inversions per year (bottom), for Oslo.

When studying the yearly average, we find no evidence of a significant trend, and the values fluctuate between 0.01 and 0.02 $°C/m$. When performing a t-test, we calculate the absolute value of the statistical t-value to be $|t_s| = 0.69$. Comparing it to the critical t-value, $t_{\text{crit}} = 2.10$, we find that since $t_s < t_{\text{crit}}$, the null hypothesis can not be rejected. Thus, the t-test agrees with our assessment. The lower plot of Figure 4.53, presenting the number of yearly counted inversions, shows no evidence of trend either. The yearly count lies between 1500 and 2000.

The yearly average inversion intensity for each season, shows no significant trend for any of the seasons either. In the upper plot of Figure 4.54 showing the winter average inversion intensity, the yearly average fluctuates between 0.023 and 0.017 $°C/m$ and is quite stable from the winter of 2001/2002 to the winter of 2012/2013 and from the winter of 2014/2015 throughout the rest of the period. In the winter of 2013/2014, the yearly average dropped to almost

0.01 $°C/m$. The absolute of the statistical t-value is $|t_s| = 0.44$, which is smaller than $t_{\text{crit}} = 1.72$ (from the t-distribution table), and so we cannot reject the notion that the trend is insignificant.

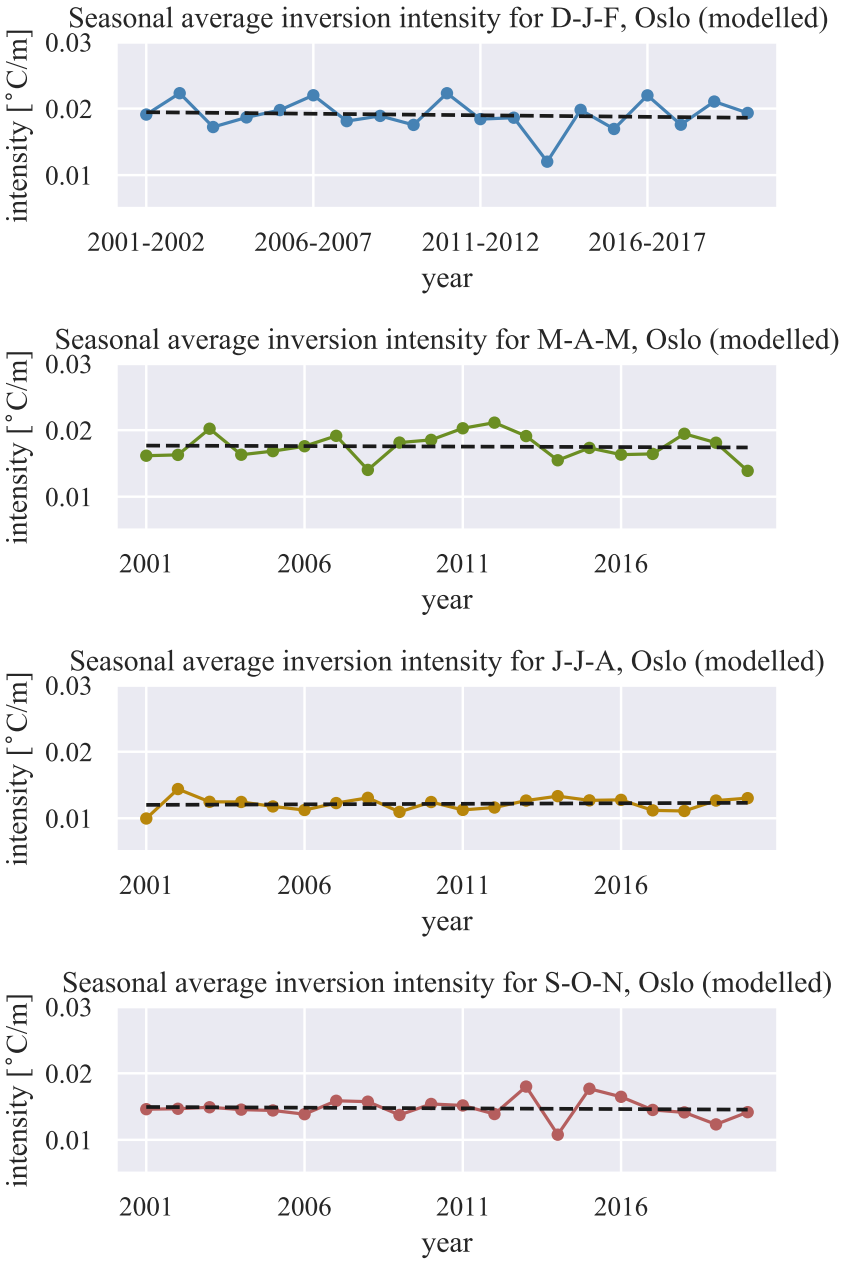


Figure 4.54: The model seasonal average inversion intensity for Oslo in a 20 year period. From the top: winter (December, January, February), spring (March, April, May), summer (June, July, August) and autumn (September, October, November).

The same applies for the other seasons as well, with the statistical t-value for spring and summer being 0.19, and for autumn being 0.40. This suggests that the negative trend we observe for the NO_2 is mainly due to reduction in emissions, and not because of changes in the meteorology. An interesting thing to point out is that the yearly average NO_2 concentrations increased from 2001 to 2013 in Figure 4.52, and we find that there is an increase in the spring and autumn inversions from 2001 to 2013, which can be affecting this. The yearly average NO_2 concentrations drops in 2014 to just below $\mu g/m^3$, and the model output also show a drop in the yearly average inversion intensity for the same year. The yearly average NO_2 concentrations also show a drop in 2005, but there is no evidence from the model output suggesting that the inversions are a cause of this.

Bergen

The data for the concentrations of NO_2 at Danmarks plass was only available from the beginning of 2003, and the yearly average is presented in Figure 4.55. The negative trend is significant ($|t_s| = 4.35$), and decreases with approximately 1 $\mu g/m^3$ per year in the study period. There was an abrupt increase from 40 $\mu g/m^3$ in 2009 to above 55 $\mu g/m^3$ in 2010, before it again decreased in 2011. In 2020, the yearly average reached just above 20 $\mu g/m^3$.

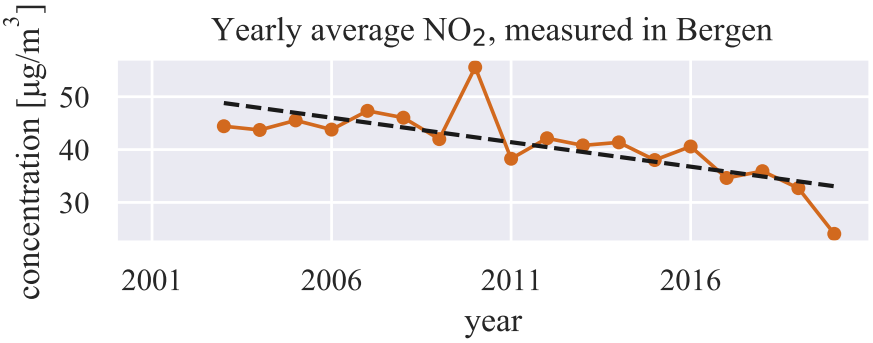


Figure 4.55: Yearly average NO_2 measured at Danmarks plass in Bergen, from 2001 to 2020, with trend line.

For the yearly average inversion in Bergen, there is a significantly negative trend throughout the study period. This is shown in Figure 4.56. Here, the plot also reveals an increase in 2010, suggesting that the increase in the NO_2 concentrations the same year is a response to the increase in the yearly average inversion intensity. The model output also shows a small increase in 2016, where the yearly average NO_2 concentrations also increased. For the t-test, the absolute of the statistical t-value was calculated to be 1.55. The number of counted inversions show little variations, and there is no evidence of a trend. The plot below, showing the number of hourly counted inversions per year, suggests that there is no trends in abundance of the inversions.

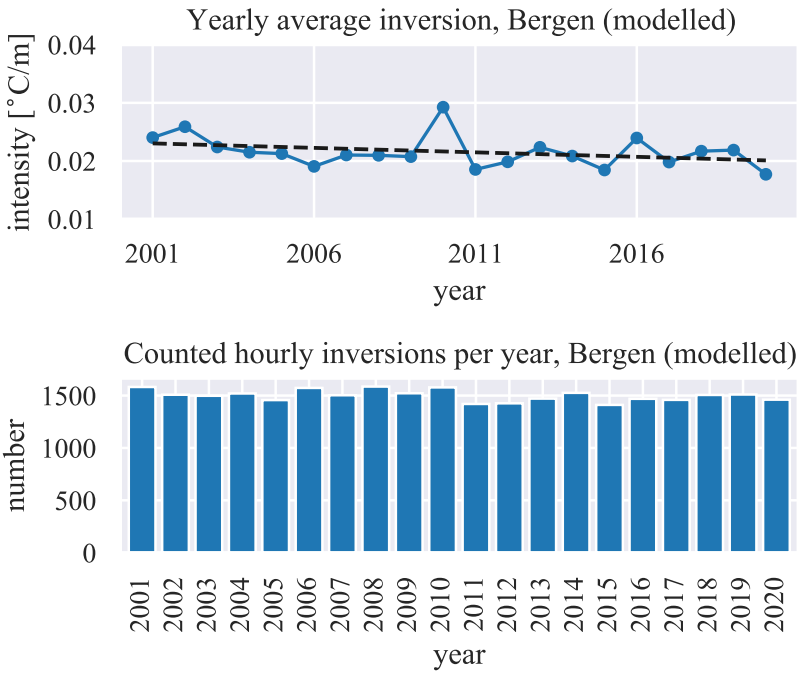


Figure 4.56: The modelled yearly average inversion intensity (top), and the number of hourly counted inversions per year (bottom), for Bergen.

Studying the trends for the seasonal average in Figure 4.57, we find that there is a negative trend for the winter season, the spring season and the autumn season. Applying a t-test we find that the absolute value of the statistical t-value is 1.60, 0.96, 0.60 and 3.98 for winter, spring, summer and autumn, respectively.

The negative trend for the winter season is the most prominent one. The linear regression reveals that the trend decreases with an intensity of 0.008 $°C/m$ for the entire period. We observe that the variations in the seasonal average in is also most prominent during winter. The trend for the spring averages decreases very little, and the variation is smaller compared to winter. The model output shows a rapid decrease from 2018 to 2020. The summer average inversion intensity decreases more rapidly than the spring average. The linear regression reveals that the trend line decreases with 0.004 $°C/m$ for the entire period. The variations in the summer averages are small. The t-test revealed that the trend in the autumn averages is not significant, but we observe that the variations are of the same magnitude as the winter variations.

Observe that the peak of inversion intensity in 2010 is caused by increase in the winter and autumn inversion intensities. The same can be observed for the increase in yearly average inversion intensity in 2016. The negative trend seen in the yearly average is mostly due to negative trend in the winter averages, but also due to negative trend in the spring and summer average.

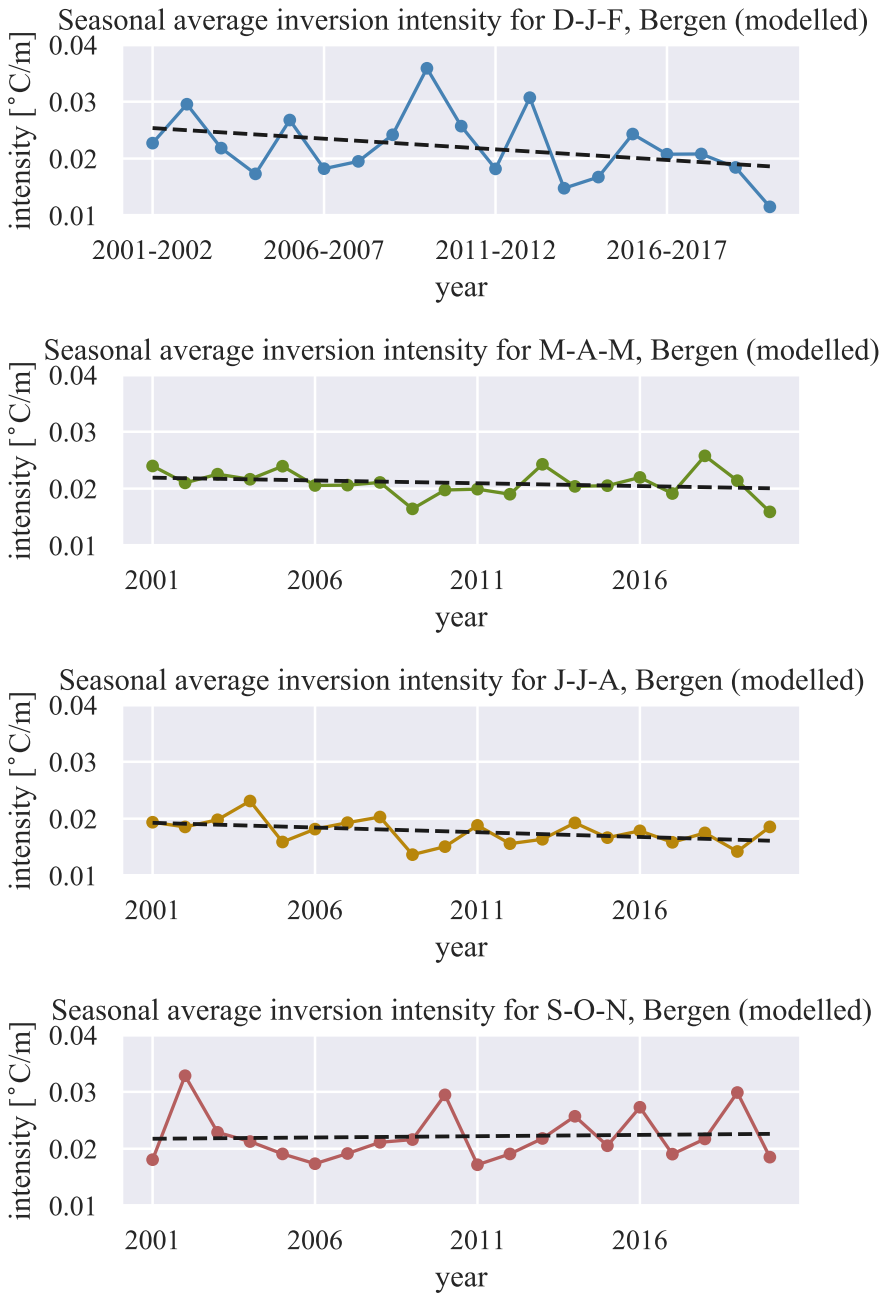


Figure 4.57: The model seasonal average inversion intensity for Bergen in a 20 year period. From the top: winter (December, January, February), spring (March, April, May), summer (June, July, August) and autumn (September, October, November).

CHAPTER 5

Summary, conclusion and outlook

An analysis of the output from the observational stations in Oslo and Bergen was made. In Oslo, the meteorological data were obtained from Valle Hovin, while the NO_2 concentrations were measured at Alnabru. The data was filtered to only contain measurements from 07:00 and 08:00 on Mondays and Thursdays in order to make the NO_2 concentrations independent from the variability in emissions. The NO_2 concentrations and the measured inversions between 2 and 25 meters had a correlation of 0.7, while the NO_2 concentrations and the measured inversion between 8 and 25 meters had a correlation of 0.5. A regression model was made, using the inversion intensity between 2 and 25 meters, the wind speed and wind direction as explanatory variable, to understand the impact the meteorological variables have on the NO_2 concentrations. The R^2 of the model was 0.6, showing that the model was able to capture a large proportion of the variance of the NO_2 concentrations, but plotting the model and calculating its statistical properties showed that it did not manage to reproduce the episodes with the highest NO_2 concentrations.

In Bergen, the meteorological data were measured at Florida and Geofysen, and the NO_2 concentrations were measured at Danmarks plass. The temperature gradient and inversion intensity were measured between 2 and 40 meters. The data was filtered to only contain measurements from 07:00 and 08:00 from Monday to Friday. The correlation between the NO_2 and the wind speed was the strongest, with a value of -0.5, while the NO_2 concentrations and the inversion intensity had a correlation of 0.3. This suggests that the station observing the NO_2 concentrations and the stations observing the inversion are too far apart, as the result does not correspond with theory. Like for Oslo, we made a regression model. Analysis revealed that the inversion intensity and wind speed were the only explanatory variables with statistical significance. The regression model gave an R^2 value of 0.36, implying that the model is not suited to predict the NO_2 concentrations as it does not explain enough of the variations observed.

However, it is important to point out the roughness of assuming the NO_2 concentrations in the filtered data for Oslo and Bergen to be independent of emissions. This is especially relevant to take into account when evaluating the regression model, as the episodes showing high concentrations of NO_2 can be caused by episodes of increased pollution.

We compared the winds, temperature gradient and inversions with data

from NORA3, a high-resolution hindcast produced by downscaling the ERA5 reanalysis in the numerical weather prediction model HARMONIE-AROME. The results for Oslo showed that the NORA3 winds were shifted more towards north-south direction, suggesting that the vertical resolution of the model was not high enough to capture the topography of the Grorud Valley. The modelled winds were also weaker than the observed winds. The seasonal variation of the temperature gradient intensity showed good compliance with the observed, but with a negative bias. The seasonal variation of the modelled inversion intensity underestimated the observed inversion intensity for all months, but the difference was largest in early summer. This is also shown in the yearly average inversion intensity, where we found that the model managed to capture the variation, but underestimated the intensity by almost 0.02 $°C/m$. The seasonal average inversion intensity showed that the model varied similarly to the observations for spring, summer and autumn, but deviated from the observed winter variation, but the observed data coverage was sparse.

The data in Bergen showed that the wind direction and wind speed were closer to the observations, compared to the model performance in Oslo. However, a small portion of weak winds were arriving from northeast and southwest that were not present in the observations. The seasonal variation of the modelled temperature gradient intensity showed a bias towards positive values, with the exception of in May and June where the NORA3 output and observations were similar. The NORA3 output for Bergen's seasonal variation of the temperature inversion intensity showed a small overestimation from January to March and in November and December, and a small underestimation from May to August, compared to the observations. The average intensity was similar for April, September and October. The modelled yearly average inversion intensity was similar to the observations, both in magnitude and variation. The model overestimated the winter average with an average of approximately 0.005 $°C/m$ throughout the period, while the modelled spring average was similar to the observations in magnitude, but the variations slightly differed. For summer average inversion intensity, the model underestimated slightly (approximately 0.005 $°C/m$ or less), and for the autumn average inversion intensity, the model overestimated slightly (again, approximately 0.005 $°C/m$ or less). The variety for those two seasons showed resemblance to the observations,

We only assessed the model output for one grid point against one measuring station for both Oslo and Bergen. This is not enough to get a well rounded assessment of the model performance for the study areas, and an idea for continuing the work on this thesis would be to interpolate the model output from multiple surrounding grid points, weighted by the topography, to the location of the measuring stations. It would also improve the result if we were able to gather observations of temperature inversions from multiple sources distributed evenly in the study areas.

Further, we analyzed the surface energy balance (SEB) for two periods of four days at each of the study areas. For Oslo, a period in January, 2002, was chosen to study where the modelled and observed inversion intensities were similar. The inversions in the study period where night inversions, disappearing during the day. The modelled short- and longwave flux were between ± 25 W/m^2, except for during the 16th, were the cloud coverage changed to allow for more

solar radiation to the ground, and outgoing radiation reach up to 75 W/m^2. In the period December, 2001, the model underestimated the observed inversion intensity, though the variation showed similar characteristics. The inversion episode was continuous for the study period, and the 2 meter air temperature was colder than the skin temperature, suggesting that the inversions were caused by warmer air arriving in higher altitudes. A drop in temperature was observed during the 19th, and the SEB revealed that the loss of heat was mostly transferred in the outgoing ground heat flux.

The modelled and observed inversion intensities were similar for a period in November, 2015 in Bergen. Here too, the inversions were night inversions disappearing at sunrise. The 2 meter and skin temperature decreased, and the cloud coverage was sparse. This allowed for a prominent diurnal variation in the SEB. In the period in December, 2015, the inversions were more coherent, and NORA3 overestimated the intensity compared to observations. The temperatures, as well as the downward sensible heat flux, increased from the 15th and throughout the period. The ground heat flux switched from upward to downward direction at midnight on the 16th. We also observe the net upward short- and longwave radiation becoming weak as cloud coverage increases.

To assess the models ability to parameterize the SEB, it would be ideal to gather observations of the radiative fluxes in the same area as the observed inversions. It is also not enough to compare two inversion episodes over four days to draw any conclusion for the true SEB performance in the model. But the period underestimating the inversion intensity in Oslo, and the period overestimating the inversion intensity in Bergen, are periods with continuous inversions, while the episodes the managed to capture the intensity for Oslo and Bergen were sporadic night inversions that disappeared during sunrise. It would be interesting to further investigate more episodes of both cases, to see if we there are any connections to the overestimation/underestimation in the model.

Lastly, we studied the averages over a 20 year period from 2001 to 2020. The plot of the yearly average NO_2 concentrations showed that the trend was declining for both Oslo and Bergen. The yearly average inversion intensity in Oslo showed no significant trend, and neither did any of the seasonal averages. In Bergen, the trend of the yearly average inversion intensity was significantly negative. Plots of the seasonal averages revealed that the trend was strongest negative for the winter season, but that there was a significant negative trend for the spring and summer averages too.

The aim of this thesis was to investigate if there are any trend in inversion intensity affecting the trend in NO_2 concentrations. The result suggests that the trend in the NO_2 concentrations is not affected by trend in the inversion intensity in Oslo, while in Bergen, the NO_2 concentrations trend is affected by the negative inversion intensity trend. However, can we fully conclude that this is the case? The yearly average showed that NORA3 was able to yield some of the variations of the observed inversions, but underestimated the intensity. We also saw that NORA3 was equally good at reproducing the variation for the winter averages, which is when the strongest inversion episodes occur. In order to truly argue that there is no trend in the inversion intensities affecting the negative NO_2 trend, we would have to make a bias correction in NORA3.

We also saw that the correlation between the modelled and observed inversions were weak, and this suggests that we cannot use NORA3 to predict specific episodes of strong inversions for Oslo.

For Bergen, we cannot conclude that the negative trend we observe in the modelled yearly averages impacts the negative trend in the NO_2 concentrations, because even though the model was able to reproduce the yearly average inversion in both variation and intensity, the correlation between the observed inversion intensity and NO_2 concentrations are were weak. A station observing the inversions is needed closer to station at Danmarks plass. Due to lack of time, we were not able to make a regression model for the twenty year period from 2001 to 2020, using the modelled variables as explanatory variables and NO_2 concentrations as the dependent variable. This would have been interesting to investigate, as it would give us a better idea of how well the model might be able to predict episodes with high NO_2 concentrations. The results in this study show promise, but the conclusion is confined by the restricted number of observed variables and observational stations.

APPENDIX A

Summary tables

A.1 Oslo

OLS Regression Results

Dep. Variable:	NO2	R-squared:	0.396
Model:	OLS	Adj. R-squared:	0.396
Method:	Least Squares	F-statistic:	903.6
Date:	Mon, 20 Dec 2021	Prob (F-statistic):	4.43e-153
Time:	09:34:23	Log-Likelihood:	-6748.4
No. Observations:	1380	AIC:	1.350e+04
Df Residuals:	1378	BIC:	1.351e+04
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	59.2284	0.954	62.076	0.000	57.357	61.100
inversion	1066.8082	35.489	30.060	0.000	997.190	1136.427

Omnibus:	324.799	Durbin-Watson:	1.157
Prob(Omnibus):	0.000	Jarque-Bera (JB):	2464.578
Skew:	0.881	Prob(JB):	0.00
Kurtosis:	9.306	Cond. No.	40.9

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Figure A.1: Summary table for simple linear regression **model 1**, Oslo

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                    NO2   R-squared:                       0.503
Model:                            OLS   Adj. R-squared:                  0.502
Method:                 Least Squares   F-statistic:                     336.5
Date:                Mon, 14 Feb 2022   Prob (F-statistic):          1.21e-101
Time:                        14:51:00   Log-Likelihood:                -3312.0
No. Observations:                 667   AIC:                             6630.
Df Residuals:                     664   BIC:                             6644.
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     83.1045      3.077     27.010      0.000      77.063      89.146
inversion    913.4255     46.321     19.720      0.000     822.472    1004.379
wind_speed    -6.7581      0.742     -9.111      0.000      -8.214      -5.302
==============================================================================
Omnibus:                      137.354   Durbin-Watson:                   1.249
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1132.433
Skew:                           0.663   Prob(JB):                    1.25e-246
Kurtosis:                       9.244   Cond. No.                         125.
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure A.2: Summary table for multiple linear regression **model 2**, Oslo

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                    NO2   R-squared:                       0.509
Model:                            OLS   Adj. R-squared:                  0.507
Method:                 Least Squares   F-statistic:                     229.2
Date:                Mon, 14 Feb 2022   Prob (F-statistic):          5.21e-102
Time:                        14:51:01   Log-Likelihood:                -3308.1
No. Observations:                 667   AIC:                             6624.
Df Residuals:                     663   BIC:                             6642.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      89.0570      3.729     23.882      0.000      81.735      96.379
inversion     926.7507     46.331     20.003      0.000     835.778    1017.724
wind_speed     -6.8166      0.738     -9.234      0.000      -8.266      -5.367
wind_direction -0.0450      0.016     -2.795      0.005      -0.077      -0.013
==============================================================================
Omnibus:                      138.885   Durbin-Watson:                   1.240
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1150.375
Skew:                           0.673   Prob(JB):                    1.58e-250
Kurtosis:                       9.291   Cond. No.                     5.51e+03
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 5.51e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Figure A.3: Summary table for simple linear regression **model 3**, Oslo

A.2 Bergen

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                    NO2   R-squared:                       0.298
Model:                            OLS   Adj. R-squared:                  0.297
Method:                 Least Squares   F-statistic:                     248.7
Date:                Mon, 14 Feb 2022   Prob (F-statistic):           5.75e-47
Time:                        22:52:06   Log-Likelihood:                -2910.4
No. Observations:                 588   AIC:                             5825.
Df Residuals:                     586   BIC:                             5834.
Df Model:                           1
Covariance Type:            nonrobust
================================================================================
                   coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
Intercept      113.1850      1.543     73.360      0.000     110.155     116.215
log_wind_speed -27.7972      1.763    -15.771      0.000     -31.259     -24.336
==============================================================================
Omnibus:                       38.529   Durbin-Watson:                   0.961
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               55.490
Skew:                           0.514   Prob(JB):                     8.92e-13
Kurtosis:                       4.099   Cond. No.                         1.57
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure A.4: Summary table for simple linear regression **model 1**, Bergen

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                    NO2   R-squared:                       0.356
Model:                            OLS   Adj. R-squared:                  0.353
Method:                 Least Squares   F-statistic:                     161.3
Date:                Mon, 14 Feb 2022   Prob (F-statistic):           1.57e-56
Time:                        14:56:49   Log-Likelihood:                -2885.3
No. Observations:                 588   AIC:                             5777.
Df Residuals:                     585   BIC:                             5790.
Df Model:                           2
Covariance Type:            nonrobust
================================================================================
                   coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
Intercept        99.7241      2.379     41.922      0.000      95.052     104.396
log_wind_speed  -25.6369      1.716    -14.936      0.000     -29.008     -22.266
inversion       724.9046    100.311      7.227      0.000     527.891     921.919
==============================================================================
Omnibus:                       13.789   Durbin-Watson:                   0.961
Prob(Omnibus):                  0.001   Jarque-Bera (JB):               15.678
Skew:                           0.295   Prob(JB):                     0.000394
Kurtosis:                       3.540   Cond. No.                         83.1
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure A.5: Summary table for multiple linear regression **model 2**, Bergen

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                    NO2   R-squared:                       0.356
Model:                            OLS   Adj. R-squared:                  0.353
Method:                 Least Squares   F-statistic:                     107.8
Date:                Mon, 14 Feb 2022   Prob (F-statistic):           1.56e-55
Time:                        14:56:49   Log-Likelihood:                -2884.9
No. Observations:                 588   AIC:                             5778.
Df Residuals:                     584   BIC:                             5795.
Df Model:                           3
Covariance Type:            nonrobust
================================================================================
                   coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
Intercept         97.0218      3.936     24.649      0.000      89.291     104.752
inversion        732.1942    100.689      7.272      0.000     534.437     929.951
log_wind_speed   -25.5938      1.718    -14.901      0.000     -28.967     -22.221
wind_direction     0.0133      0.015      0.862      0.389      -0.017       0.044
==============================================================================
Omnibus:                       14.729   Durbin-Watson:                   0.954
Prob(Omnibus):                  0.001   Jarque-Bera (JB):               17.027
Skew:                           0.303   Prob(JB):                     0.000201
Kurtosis:                       3.572   Cond. No.                     1.57e+04
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.57e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Figure A.6: Summary table for simple linear regression **model 3**, Bergen (the addition of wind direction is not significant)

APPENDIX B

Radiative fluxes

B.1 Oslo

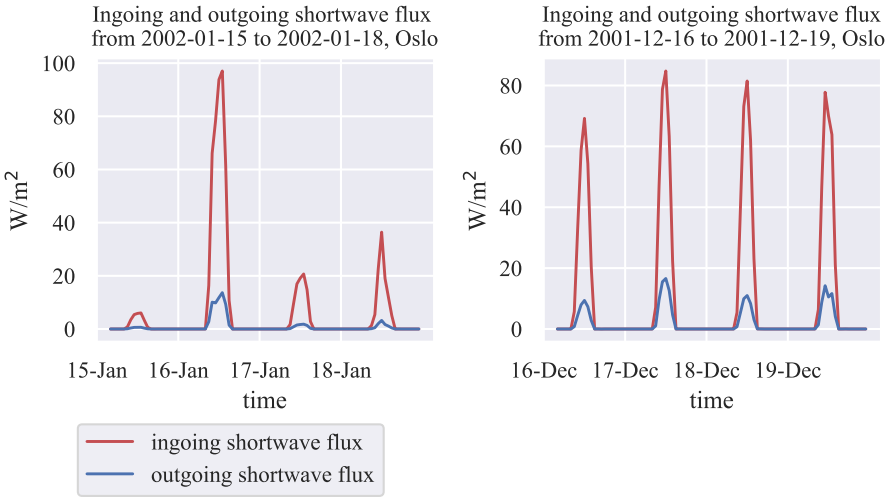


Figure B.1: Shortwave fluxes for the period where modelled and observed inversions were similar (left) and for the period where modelled and observed inversions differed (right), in Oslo. Positive downwards.

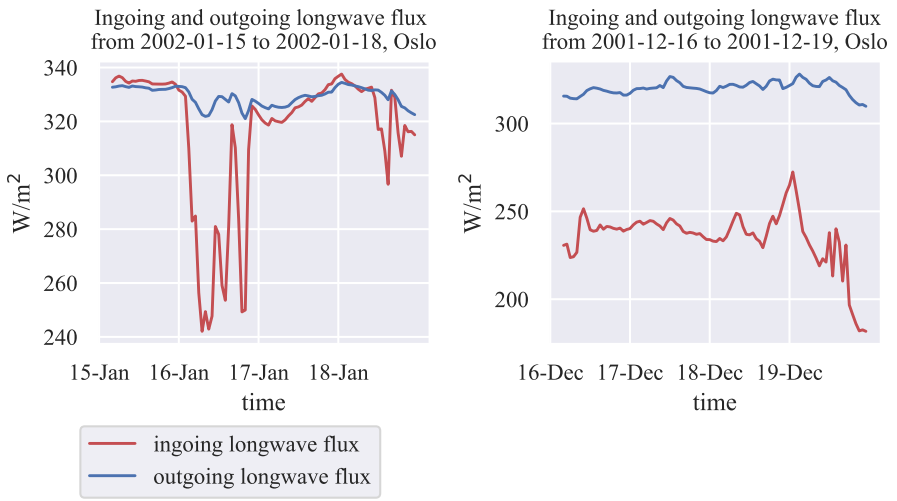


Figure B.2: Longwave fluxes for the period where modelled and observed inversions were similar (left) and for the period where modelled and observed inversions differed (right), in Oslo. Positive downwards.

B.2 Bergen

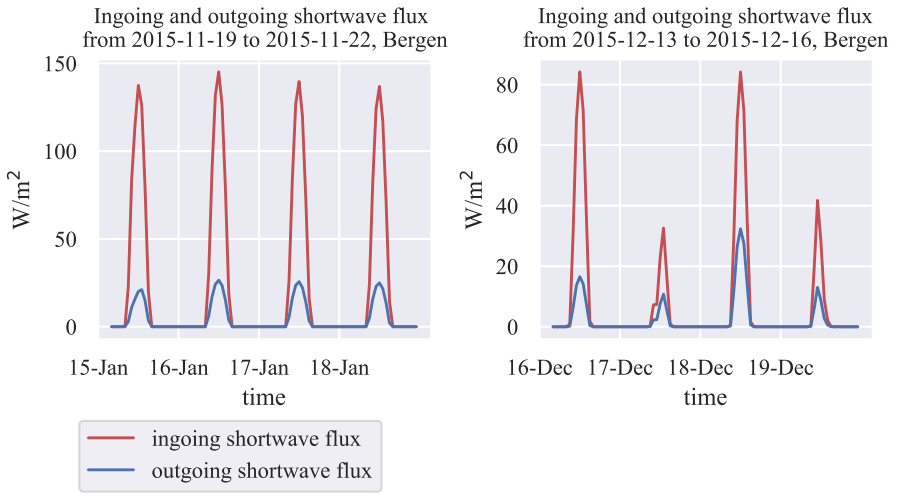


Figure B.3: Shortwave fluxes for the period where modelled and observed inversions were similar (left) and for the period where modelled and observed inversions differed (right), in Bergen. Positive downwards..

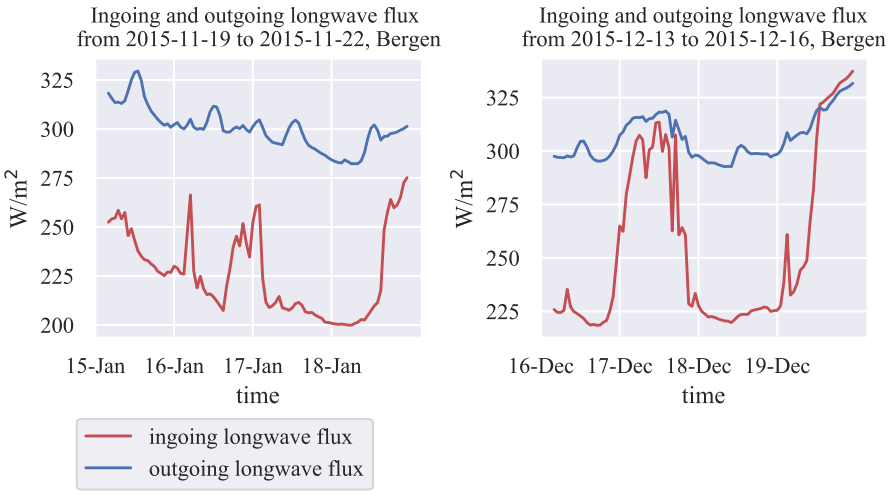


Figure B.4: Longwave fluxes for the period where modelled and observed inversions were similar (left) and for the period where modelled and observed inversions differed (right), in Bergen. Positive downwards.

Bibliography

(MET), N. M. I. (2021). *Seklima*. URL: seklima.met.no (visited on 04/02/2022).

Bailey, A., Chase, T. N., Cassano, J. J. and Noone, D. (2011). 'Changing Temperature Inversion Characteristics in the U.S. Southwest and Relationships to Large-Scale Atmospheric Circulation'. In: *Journal of Applied Meteorology and Climatology* vol. 50, no. 6, pp. 1307–1323.

Caserini, S., Giani, P., Cacciamani, C., Ozgen, S. and Lonati, G. (2017). 'Influence of climate change on the frequency of daytime temperature inversions and stagnation events in the Po Valley: historical trend and future projections'. In: *Atmospheric Research* vol. 184, pp. 15–23.

Ha, J., Choi, Y., Lee, S. and Oh, K. (2020). 'Diurnal and Seasonal Variations in the Effect of Urban Environmental Factors on Air Temperature: A Consecutive Regression Analysis Approach'. In: *International Journal of Environmental Research and Public Health* vol. 17, no. 2.

Haakenstad, H., Breivik, Ø., Furevik, B. R., Reistad, M., Boglinger, P. and Aarnes, O. J. (Oct. 2021). 'NORA3: A Nonhydrostatic High-Resolution Hindcast of the North Sea, the Norwegian Sea, and the Barents Sea'. In: *Journal of Applied Meteorology and Climatology* vol. 60, no. 10, pp. 1443–1464.

Jacob, D. J. (1999a). 'Oxidizing Power of the Troposphere'. In: *Introduction to Atmospheric Chemistry*. Princeton, NJ: Princeton University Press, pp. 200–220.

— (1999b). 'Simple Models'. In: *Introduction to Atmospheric Chemistry*. Princeton, NJ: Princeton University Press, pp. 24–41.

— (1999c). 'Simple Models'. In: *Introduction to Atmospheric Chemistry*. Princeton, NJ: Princeton University Press, p. 2.

Jay L. Devore, K. N. B. (2012a). 'Regression and Correlation'. In: *Modern Mathematical Statistics with Applications*. Ed. by Mendelsohn, R. and Neumann, J. E. New York, NY: Springer, New York, NY, pp. 662–664.

— (2012b). 'Regression and Correlation'. In: *Modern Mathematical Statistics with Applications*. Ed. by Mendelsohn, R. and Neumann, J. E. New York, NY: Springer, New York, NY, p. 614.

— (2012c). 'Regression and Correlation'. In: *Modern Mathematical Statistics with Applications*. Ed. by Mendelsohn, R. and Neumann, J. E. New York, NY: Springer, New York, NY, p. 682.

Jay L. Devore, K. N. B. (2012d). 'Tests About a Population Mean'. In: *Modern Mathematical Statistics with Applications*. Ed. by Mendelsohn, R. and Neumann, J. E. New York, NY: Springer, New York, NY, pp. 443–444.

Ji, F., Evans, J. P., Luca, A. D., Jiang, N., Olson, R., Fita, L., Argüeso, D., Chang, L. T.-C., Scorgie, Y. and Riley, M. (2019). 'Projected change in characteristics of near surface temperature inversions for southeast Australia'. In: *Climate Dynamics* vol. 52, pp. 1487–1503.

Jones, G. R., Proudfoot, A. T. and Hall, J. I. (1973). 'Pulmonary effects of acute exposure to nitrous fumes'. In: *Thorax* vol. 28, no. 1, pp. 61–65.

Kukkonen, J. et al. (2005). 'Analysis and evaluation of selected local-scale PM10 air pollution episodes in four European cities: Helsinki, London, Milan and Oslo'. In: *Atmospheric Environment* vol. 39, no. 15. Fourth International Conference on Urban Air Quality: Measurement, Modelling and Management, 25-28 March 2003, pp. 2759–2773.

Lenderick, G. and Holtslag, A. A. (Oct. 2004). 'An updated length-scale formulation for turbulent mixing inclear and cloudy boundary layers'. In: *Quarterly Journal of the Royal Meteorological Society* vol. 130, no. 604, pp. 3405–3427.

Lowry, T. and Schuman, L. M. (1956). '"SILO-FILLER'S DISEASE"—A SYNDROME CAUSED BY NITROGEN DIOXIDE'. In: *Journal of the American Medical Association* vol. 162, no. 2, pp. 153–160.

Marshall, J. and Plumb, R. (2008). 'Convection'. In: *Atmosphere, Ocean and Climate Dynamics*. Ed. by Dmowska, R., Hartmann, D. and Rossby, H. Burlington, MA, USA: Elsevier Inc, pp. 31–57.

Masson, V. et al. (July 2013). 'The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes'. In: *Geosci. Model Dev.* vol. 6, pp. 929–960.

Meijgaard, E. van, Ulft, L. van, Lenderink, G., Roode, S. de, Wipfler, L., Boers, R. and Timmermans, R. (2012). *Refinement and application of a regional atmospheric model for climate scenario calculations of Western Europe*. KvR report 054/12. Dutch National Research Programme Climate changes Spatial Planning.

Milionis, A. E. and Davies, T. D. (2008). 'A comparison of temperature inversion statistics at a coastal and a non-coastal location influenced by the same synoptic regime'. In: *Theoretical and Applied Climatology* vol. 94, pp. 225–239.

Müller, M. et al. (2017). 'AROME-MetCoOp: A Nordic Convective-Scale Operational Weather Prediction Model'. In: *Weather and Forecasting* vol. 32, no. 2, pp. 609–627.

Norwegian Environment Agency (2021). *Miljøstatus*. URL: https://miljostatus.miljodirektoratet.no/tema/forurensning/sur-nedbor/nitrogenoksid-nox/ (visited on 30/01/2022).

Norwegian Environment Agency, the (2022). *Tiltak for bedre lokal luftkvalitet*. URL: https://luftkvalitet.miljodirektoratet.no/artikkel/artikler/hva-kan-jeg-gjore/ (visited on 06/02/2022).

Norwegian Government (2007). *4 Internasjonale avtaler og forpliktelser*. URL: https://www.regjeringen.no/no/dokumenter/nou-2007-8/id473567/?ch=5 (visited on 06/02/2022).

Norwegian Institute of Public Health (FHI) (2020). *Nitrogendioksid*. URL: https://www.fhi.no/nettpub/luftkvalitet/temakapitler/nitrogendioksid2/ (visited on 03/01/2022).

Norwegian Petroleum Directorate (2020). *Norges miljø- og klimaforpliktelser*. URL: https://www.npd.no/fakta/publikasjoner/rapporter/rapportarkiv/kraft-fra-land-til-norsk-sokkel/2---utslipp-til-luft-fra-petroleumssektoren/norges-miljo--og-klimaforpliktelser/ (visited on 06/02/2022).

Scaife, A. A., Knight, J. R., Vallis, G. K. and Folland, C. K. (2005). 'A stratospheric influence on the winter NAO and North Atlantic surface climate'. In: *Geophysical Research Letters* vol. 32, no. 18. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2005GL023226.

Seilen, K. (2018). 'Temperature Inversions in the Bergen Valley and Testing of Low-Cost Temperature and Air Quality Measuring Instruments'. In.

Solbrekke, I. M., Sorteberg, A. and Haakenstad, H. (Nov. 2021). 'The 3 km Norwegian reanalysis (NORA3) – a validation of offshore wind resources in the North Sea and the Norwegian Sea'. In: *Wind Energy Science* vol. 6, pp. 1501–1519.

Statistics Norway (SSB) (2021a). *Bilpark*. URL: https://www.ssb.no/statbank/list/bilreg (visited on 04/02/2022).

— (2021b). *Oslo*. URL: https://www.ssb.no/kommunefakta/oslo (visited on 04/02/2022).

— (2021c). *Statistikkbanken*. URL: https://www.ssb.no/statbank (visited on 30/01/2022).

Stull, R. (1988). 'Practical Meteorology - An Algebra-based Survey of Atmospheric Science'. In: University of British Columbia, p. 2.

Sundqvist, H., Berge, E. and Kristjánson, J. E. (Aug. 1989). 'Condensation and Cloud Parameterization Studies with a Mesoscale Numerical Weather Prediction Model'. In: *Monthly Weather Review* vol. 117, no. 8, pp. 1641–1657.

United Nations Economy Commission for Europe (UNECE) (1999). *Gothenburg Protocol*. URL: https://unece.org/gothenburg-protocol (visited on 06/02/2022).

United States Environment Protection Agency (EPA) (2021). *Nitrogen Dioxide Pollution*. URL: https://www.epa.gov/no2-pollution/basic-information-about-no2 (visited on 06/02/2022).

University Corporation for Atmospheric Research (UCAR) (2014). *Ozone in the Troposphere*. URL: https://scied.ucar.edu/learning-zone/air-quality/ozone-troposphere (visited on 09/12/2021).

Wallace, J. and Hobbs, P. (2006a). 'Atmospheric Boundary Layer'. In: *Atmospheric Science*. Ed. by Dmowska, R., Hartmann, D. and Rossby, H. Burlington, MA, USA: Elsevier Inc, pp. 375–417.

— (2006b). 'Atmospheric Thermodynamics'. In: *Atmospheric Science*. Ed. by Dmowska, R., Hartmann, D. and Rossby, H. Burlington, MA, USA: Elsevier Inc, pp. 63–111.

Wolf, T., Esau, I. and Reuder, J. (Sept. 2014). 'Analysis of the vertical temperature structure in the Bergen valley, Norway, and its connection to pollution episodes'. In: *Journal of Geophysical Research: Atmospheres* vol. 119, no. 10, pp. 645–662.

World Health Organization (WHO) (2021). *Air Pollution*. URL: https://www.who.int/health-topics/air-pollution#tab=tab_1 (visited on 09/12/2021).

Xu, C. (2019a). 'Correlation and simple regression'. In: *Statistical and Stochastic Methods in Hydrology (Lecture notes)*. Blindern, Oslo, Norway, p. 94.
— (2019b). 'Multiple regression analysis'. In: *Statistical and Stochastic Methods in Hydrology (Lecture notes)*. Blindern, Oslo, Norway, p. 109.
Zhan, Q. and Lan, Y. (Nov. 2017). 'How do urban buildings impact summer air temperature? The effects of building configurations in space and time'. In: *Building and Environment* vol. 125, no. 3, pp. 88–98.