Royal Netherlands
Meteorological Institute
Ministry of Infrastructure and Water Management

Automating air temperature Siting Classification of meteorological stations according to the World Meteorological Organization guidelines

J. Stuurman

KNMI Internal report IR-2019-01

Geographical Information Management and Applications

Automating air temperature Siting Classification of meteorological stations according to the World Meteorological Organization guidelines

Jelle Stuurman

Universiteit Utrecht
TU Delft
University of Twente.
Wageningen UR

March 1, 2019

Automating air temperature Siting Classification of meteorological stations according to the World Meteorological Organization guidelines

Jelle Stuurman

KNMI supervisors: Marieke Dirksen (MSc) & Marijn de Haij (MSc)
GIMA supervisor: Ph.D. Frank Ostermann
Professor: Prof. Ph.D. Menno-Jan Kraak

Colophon:

MSc GIMA Thesis
February 2019

This thesis research was performed and written at:
Royal Netherlands Meteorological Institute (KNMI).
De Bilt, The Netherlands

Author: Jelle Stuurman
UU solis ID: 3691233
UT student nr.: s6020488
jellejstuurman@gmail.com

KNMI supervisors: Marieke Dirksen (MSc) & Marijn de Haij (MSc)
GIMA supervisor: Ph.D. Frank Ostermann
Professor: Prof. Ph.D. Menno-Jan Kraak

Summary

This research thesis investigated the automatisation of the air temperature Siting Classification (SC) according to its WMO guidelines. SC values are qualitative numbers that shows how representative a meteorological observation is of its surrounding area. SC for air temperature is based on three characteristics at the location of the sensor: shading, presence of heat source and vegetation height. Instead of going into the field, the Automated SC (ASC) developed in this research thesis uses the Actueel Hoogtebestand Nederland (AHN) and Basisregistratie Grootschalige Topografie (BGT) open data datasets to determine the classifications for air temperature. The algorithm was first tested at the Automatic Weather Station (AWS) in De Bilt, after which it was applied to all the other 33 AWS sites owned and maintained by the Royal Netherlands Meteorological Institute (KNMI). There is need for an ASC to speed up the process but also to create a more reproducible approach than the current Manual SC (MSC) that is subject to human interpretations.

Each criterion within the air temperature SC guidelines is based on underlying scientific principles why they are included and each require different analyses to determine its SC. Shaded locations can have a different temperature than areas in the sun. Heat sources - including water bodies, buildings, roads or other artificial objects - can reflect or retain heat and therefore influence the air temperature. Vegetation causes cooling due to its evapotranspiration process and thus influence the observed air temperature.

The ASC model was developed using the programming language of R and made use of the AHN to detect the height of the surface, objects to determine shading. The BGT was used to determine the presence of heat sources through its topographic land use. These datasets were imported dynamically, making it possible to apply this model to any location on land in The Netherlands. Using a decision tree an air temperature SC Class 1 through 5 was determined for each AWS owned and maintained by the KNMI. The usability of the model was tested through visual validations based on photos or satellite imagery to check on its accuracy, and comparison validations based on the current MSC procedure done by the KNMI.

Results shows that an ASC model corresponds for 55.9% with the MSC values and showed that the model can be applied for a reproducible and consistent SC. In 41.2% of the cases, the ASC applies a stricter SC value compared to the MSC, resulting in a less representative SC. This is because the datasets interprets more shading and heat sources than through the MSC. This indicates that when the input data is correct, a SC is determined tat strictly follows the SC guidelines. Manual validations were done to check if the input data was not outdated or not interpreted due to a lack of information. The ability of performing an ASC can contribute for further research in improving the air temperature SC of existing, or determining new completely locations. The model can also be used to determine the air temperature SC of the vast of other sensors of third-party networks. Moreover, the model further substantiates the need to specify the different air temperature SC guideline criteria even more so that different surroundings are taken better into account.

Acknowledgements

This master thesis has been written to fulfil the graduation requirements of the Masters of Science degree in Geographic Information Management & Applications (GIMA). This marks an end of a long journey which I was looking forward to for or a long time, and now I can't realise just yet it has finally happened.

I would like to express my sincere gratitude to all my supervisors for giving me the tremendous amount of support, and feedback you have provided me. In particular my supervisors at the KNMI for investing their time to continuously help me to get the most out of this research. This includes Marieke Dirksen for helping me set up this research, guide me through the computer programming aspects and always providing feedback on the structure and direction of my thesis. Marijn de Haij for helping me to better understand the concepts of meteorological observations and provide extensive daily feedback on my thesis towards the end of my research. Frank Ostermann from University of Twente for guarding the overall research progress, making sure I always set the right priorities. Furthermore, I would like to thank the KNMI in general for giving me the opportunity to write my thesis at their institute. It has been a great pleasure to meet and interact with the many staff members and researchers who all work hard every day to help society in providing the best weather forecast and contribute to the science of seismology and climate change. I would like to also thank them for giving me the opportunity to prepare and present a poster at the TECO conference in Amsterdam organised by the World Meteorological Organization, meeting people from all over the world that are directly involved with the research I have done. It was certainly the biggest highlight during my whole period at the KNMI.

2018 was the first full year in which I took a new medicine that significantly could improve my health due to my chronic illness. In the first half of the year I had many good but also quite a few bad periods. It was not until the second half of 2018 when I really started to feel the benefits of this new medicine. Without this medicine, it might have been much harder to fulfil this research within a time period of a year. Therefore, I would like to dedicate this thesis to all the people who (in)directly have contributed to research in finding new medicines for my disease.

Last, but definitely not least, I would like to thank my family for always believing in me and giving me support during this research process, and more generally, during my time as a student. A special thanks goes to my sister, Anke, who has helped me to improve my thesis through proofreading.

February 28, 2019
Utrecht, The Netherlands

Contents

List of Figures

List of Tables

List of Abbreviations

AHN	Actueel Hoogtebestand Nederland
ASC	Automated Siting Classification
AWS	Automatic Weather Station
BGT	Basis Grootschalige Topografie
CBL	Convective Boundary Layer
CIMO	Commission for Instruments and Methods of Observation
DEM	Digital Elevation Model
DSM	Digital Surface Model
DTM	Digital Terrain Model
ECV	Essential Climate Variables
GIS	Geographic Information System
GML	Geography Markup Language
ISO	International Standard Organization
KNMI	Koninklijk Nederlands Meteorologisch Instituut
LGN	Landelijk Grondgebruik Nederland
LIDAR	Light Detection And Ranging
MSC	Manual Siting Classification
NWP	Numerical Weather Prediction
SC	Siting Classification
SEB	Surface Energy Balance
USCRN	United States Climate Reference Network
WFS	Web Feature Service
WMO	World Meteorological Organization
WIGOS	WMO Integrated Global Observing System
WSI	WIGOS Station Identifiers

1 Introduction

1.1 Motivation

To obtain meteorological observations of high quality it is important to know the uncertainties and biases on its measurements. Already two decades ago, research showed that observation errors associated with the surroundings of the site were often much larger than the measurement uncertainty of the sensors themselves. For air temperature, common surrounding characteristics that can influence on the measurement include the presence of (natural) objects that reflect and retain heat, or cause shading (Leroy, 2006). Official meteorological measurements are usually collected at Automatic Weather Stations (AWS), which are set up and maintained by national meteorological services and have to comply with a minimum set of recommendations by the Word Meteorological Organization (WMO). AWS sites are usually delineated areas where meteorological sensors are installed that automatically collect current weather measurements such as air temperature, pressure, wind and precipitation (figure 1.1). AWS sites are normally installed in a nationwide network targeted for different applications such as general weather forecasting, use in Numerical Weather Prediction (NWP) models and long-term climate monitoring. Observations are shared between countries under the umbrella of WMO (Estévez, Gavilán, & Giráldez, 2011; WMO, 2017). AWS are usually maintained by the different national meteorological services who are also responsible for its quality and operations of continuity. For air temperature it is important to collect measurements of high quality to inform society with correct information and substantiate the factors that may have an influence on climate change.

Figure 1.1: Automatic Weather Station (AWS) in Stavoren, the Netherlands. Meteorological observations are automatically collected and distributed to a central facility, including those of air temperature, pressure, wind, precipitation, relative humidity, global radiation, visibility, clouds, and present weather) *Source:* Geuko Boog

Before the 1990's, many national meteorological institutes often failed to follow their own guidelines on controlling uncertainty and exposures in their meteorological measurements. The instruments were often located in densely populated areas, while the observations were also used to represent the measurement of less urbanised surrounding area. This resulted in measurements that were not representative of the greater region. Furthermore, there was a lack of transparency and a standardised method for meteorological services to express their degree of compatibility with commonly accepted siting guidelines (Leroy, 1998). In the 1990's there was a serious reduction of observation instruments used by the Meteo France meteorological network. As a result Leroy (1998) suggested and developed a Siting Classification (SC) method that proposed an approach to note the representativeness of meteorological observations in the surrounding area ($<$ 500 metres) of third-party stations. A qualitative number scheme (usually 1-5) that classified the quality of specific variables like temperature, wind, pressure, and radiation. A few years later, it was found that this method could also be used to achieve improvements of meteorological observations within a network. Therefore, the SC have been developed in a recommendation classification scheme by WMO's Commission for Instruments and Methods of Observation (CIMO) (M. A. Wolff et al., 2014) in 2010 and eventually evolved into an International Standard Organization (ISO) standard in 2014 (nr. 9289:2014(E)) (WMO, 2017). These guidelines can be found in Appendix C as an annex and serves as the primary reference used in this research.

This SC scheme is especially useful for climate researchers, as it enabled them to obtain an indication an extent to which the measurements are influenced by the direct surroundings. The SC guidelines do not determine which meteorological measurements are good or bad. It rather indicates for which application a site or individual instrument is suitable for which observation. It shows how much a meteorological observation is representative of the surrounding area. Furthermore, the SC scheme also acts a parameter that enables comparisons between sites within a meteorological network or between observing networks (WMO, 2018). The SC guidelines allow better comparisons to be made as they are only meaningful if the sites have metadata that is collected in the same standardised way. The SC guidelines also provide important information for meteorological services to assess their network and allow them to set goals that will improve their observations. Currently, the SC scheme is implemented within a few national meteorological networks including Norway, Switzerland, The Netherlands and Japan. The Royal Netherlands Meteorological Institute (KNMI) has included the SC scheme in their most recent Observations strategy to further develop and improve meteorological observation networks (KNMI, 2015).

The SC standard is relatively new and only little research has been done to assess the value of the scheme and its practical use. So far, the focus of existing research is on the air temperature variable. Qualitative and quantitative methods to identify the measurement uncertainties are limited and often lack scientific basis (Kinoshita, 2014; M. Wolff et al., 2014). Furthermore, research done by different national meteorological institutes indicated that the SC guidelines do not always take site-specific conditions into account, making the classifications not representative of the greater surrounding area. Meteorological institutes of Nordic countries and Switzerland concluded that the combination of relatively low altitudes in mountainous regions and the varying azimuth angles of the sun lead to specific shading that results in a air temperature SC value that does not provide enough information for climatological research (Fisler, Kube, Stocker, Gr"uter, & Calpini, 2017; M. Wolff et al., 2014). Therefore, the Nordic and Swiss meteorological institutes have created an adjusted classification scheme that includes several adaptations to make the scheme more suitable for their own meteorological stations. In Japan, the air temperature SC is not even seen as suitable to use because of the many instruments that have installed in the dense urban areas. This would result in classification scores that do not represent meteorological measurements of the greater area if the SC guidelines would strictly be followed (Kumamoto et al., 2013). Moreover, the Belgian meteorological institute mentioned that the ranges of the classes are not specific enough to determine a certain Class, which may cause misleading representative values. (Sotelino, Coster, Beirinckx, & Peeters, 2016).

The issues found by the different meteorological institutes indicate that there is still room for substantial improvement in the current SC guidelines. The reproducibility of the underlying (manual) inspection process also poses an issue. Different data collection techniques are used to determine the SC because there is no standardised procedures or equipment to determine the classifications (Fisler et al., 2017; Kinoshita, 2014; Sotelino et al., 2016; M. Wolff et al., 2014; M. A. Wolff et al., 2014). All SC are currently done manually and carried out by inspection staff who visit each individual site for observations and measurements. The current SC criteria and procedures are subject to human interpretations and therefore may include more uncertainties on classifications than when this would be done automatically. Furthermore, more meteorological measurements are collected today through third-party sources and made publicly available. The SC provides a way to uniformly determine the quality of the measurements, but performing Manual SC (MSC) for the vast amount of data that is collected today costs too much time. Therefore, there is a high need to apply a SC procedure that allows a reproducible methodology through automation that additionally also speeds up the process.

A more reproducible and faster SC process would greatly improve overall reliability and interoperability of meteorological data. In this thesis research, an automated air temperature SC model is developed and tested on its usability in The Netherlands. Usability is seen as the extent to which the model can be used for any coordinates on land in The Netherlands. The model was developed using Dutch geospatial data and applied to the AWS sites operated and maintained by the KNMI in their national meteorological observing network.

1.2 Research objective

The primary objective of this thesis is to study and develop a model that automatically determines the air temperature (1.5 - 2.5 metres above the ground) classification based on the SC criteria set by the WMO. Furthermore, this research provides a scientific overview of the underlying principles and mechanisms to better understand the main aspects influencing the measured air temperature. This is important to understand better how the current SC are defined and thus contribute in the development process of the model.

To build an Automated Siting Classification (ASC) model, similar data and information input is required as the data that is currently collected for a manual classification from, which the information needed for the SC can be derived. Geographic Information System (GIS) tools will be used so that spatial datasets that coincides with the information that is needed to create a SC according to current WMO guidelines will be integrated into the model. In this research, this include Dutch geospatial datasets.

This will make the model independent from any field work on the site itself and thus make it possible to completely automate the process. The datasets that will be required for the air temperature criteria include height, solar angles and land use data. Using geospatial data as input, will make the model dependent on the resolution and accuracy of the datasets. Therefore, the impact of different data set (versions) on the reported output will be studied as part of this thesis. Because the automated calculations will be completely based on the existing SC criteria set by the WMO, the model usability will be assessed in a mutual comparison with the manual classification obtained in the field. Therefore, research will be done to understand the ins and outs of the manual procedure, including field work to observe this procedure. Manual Siting Classification (MSC) values obtained from field inspection reports will be compared to the automated results. During the development process, the model will continuously be adjusted to ensure more accurate and better comparable outcomes.

The model will follow the exact SC criteria for the air temperature variable set by the WMO guidelines. Calculations needed to classify the air temperature will be done in the same way as calculations based on manual observations. To automate the procedure, the model will include several (spatial) analyses to retrieve SC values as output. The model will first be tested for the AWS of De Bilt (WSI: 0-20000-0-06260) in The Netherlands after which the ASC was applied on all the other 33 sites. This will be done to analyse if the algorithm worked on other locations and evaluate if and what issues will be found. This will contribute in making conclusions if the model can be used for any other location on land in The Netherlands.

1.3 Research question

The main goal of this thesis project is to study and develop and ASC model for air temperature. This leads to the following research question:

Is it possible to automate the Siting Classification for the air temperature criteria according to the WMO guidelines using geospatial datasets and apply it to any location on land in The Netherlands?

This research question was split into 2 sub-questions that analyse the different parts of the model:

1. How can an ASC be developed for the air temperature variable for the AWS sites on land according to its criteria?
    - What are the scientific principles that underlie the Siting Classification for the air temperature variable and why are these important?
    - What geospatial datasets need to be imported and processed to generate air temperature SC values as output?
    - How does the ASC model cope with different the parts of the WMO SC criteria guidelines?
2. What is the accuracy of the ASC values and how do they compare to to the MSC values?
    - What aspects need to be taken into account when evaluating the accuracy of the ASC outputs?
    - How does the current MSC for air temperature procedure work and how does it differ to the automated procedure?
    - How do the automated outputs compare to the MSC values?

1.4 Scope

The air temperature variable was selected over the other variables (wind, precipitation, radiation and environment) because it is the most commonly used meteorological observation used in daily weather forecasting and climate change studies. Furthermore it is also considered as one of the Essential Climate Variables (ECV) and assumed to be affected by siting criteria (Bojinski et al., 2014). The air temperature variable was also selected because the data required for this SC primarily includes a variation of geospatial datasets and spatial analyses, making it more relevant and applicable for the context of this thesis research.

The model will be developed in such a way that it can be used for other purposes other than determining the SC of the air temperature at the current AWS sites. The location of an AWS sites is represented by a pair geographic coordinates. Therefore, the model will ensure SC will also be applied to any other location on land in The Netherlands. This requires a dynamic approach in collecting the data from the diverse sources and pre-processing all the information. Parts of the computations may also be used for determining the SC of other variables. The model could also be applied to classify the representativeness of air humidity sensors because they follow the same guidelines as that of air temperature. The model will exclude locations that are not on land, such as meteorological instruments in the middle of a lake or those that are placed on oil platforms at sea. No data is available of these areas or they have to follow different guidelines.

The main output is based on the exact same calculations and classification values that are currently defined in the WMO SC guidelines. One air temperature criteria within the guidelines was excluded in this research: the check if the AWS was "located on a flat, horizontal land" (WMO, 2017). This criterion is excluded because The Netherlands mainly consists of flat land and therefore the assumption was made all the sites are located on flat, horizontal land.

It is not the primary aim of this research to develop methods that improved the SC guidelines itself by defining the criteria better or make them more explicit. This research could provide information, however, which criteria applied in the SC for air temperature need to be defined better or made more explicitly to obtain results that are more reproducible and consistent to the guidelines. This research could also contribute in further standardisation the SC guidelines. Furthermore, no analysis will be done on how the results found could obtain better classification values if an AWS site would be located for a certain distance or installed completely somewhere else. The quality of the model will be evaluated by comparing results to all the 34 AWS sites in the Th Netherlands to its corresponding manual classifications values obtained from field inspection reports, (historic) photos of sites and satellite imagery.

1.5 Conceptual design

The different parts of this thesis leads to a conceptual design that determined the workflow of this research as shown in figure 1.2. At first the air temperature SC guidelines will be analysed to understand better what scientific principles the concepts are based on. Based on the criteria guidelines, the required datasets are selected and the analysis were written in code using R. The ASC model consists of both importing and (pre-)processing the data as well as performing all the analysis required to determine a classification. This will make the model dynamic to use for any location on land in The Netherlands. The model provides results which will be tested on its accuracy through visual and comparison validations. During the development process, the model will be continuously adjusted based on the interim outputs from the validation analyses done for all the AWS sites. This feedback loop is introduced to improve the meaning and quality of the results. The outcome of the validations analyses of all the sites leads to a conclusion on the extent the model is usable to create a SC for air temperature for any location on land in The Netherlands.

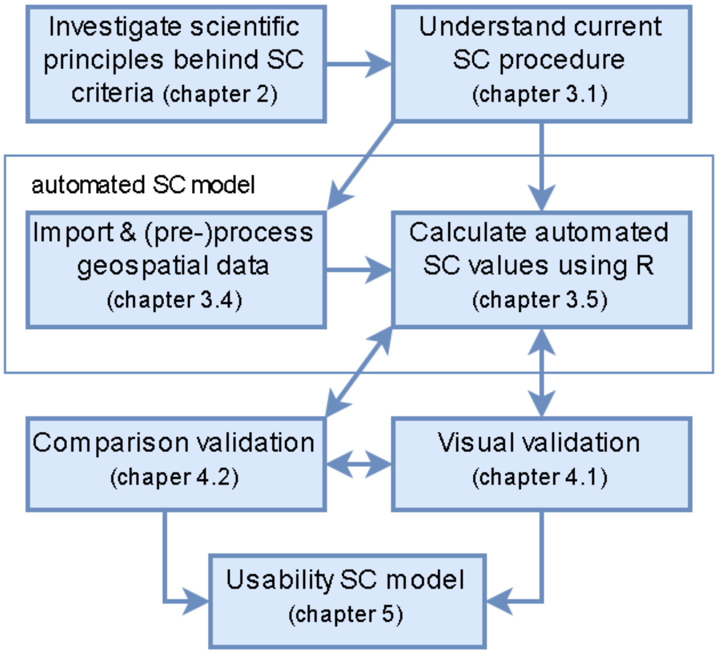


Figure 1.2: Conceptual Design of this thesis research to develop an ASC for the air temperature and test it on its usability.

1.6 Reading guide

This thesis used the current WMO air temperature SC criteria guidelines as starting point. Based on these guidelines, the underlying scientific principles relevant for the criteria are described in a scientific overview section in chapter 2. Chapter 3 provides an explanation of the methods applied in this research. A description is provided on how the current SC procedure works in chapter 3.1 and chapter 3.2.1 gives an overview of the selected AWS sites. An explanation is given on what data was imported and how they were processed in chapter 3.2. Chapter 3.3 describes how the different criteria are calculated leading to chapter 4 that includes the results obtained from the model including an visual model validation analysis (chapter 3.4.1) and a comparison validation analysis (chapter 3.4.2). Chapter 5 discusses the results obtained from the analyses to determine the usability of the developed ASC model. Finally, chapter 6 concludes this thesis expressing the usability of the ASC model that was developed in this research. This chapter also includes recommendations for further scientific research and for the KNMI to further develop and implement the ASC procedure by applying the model to other locations, using additional data or research an automatic method for other observation variables that are part of the SC guidelines.

2 Background information and relevant work

This chapter provides a scientific background behind the concepts that are found in the air temperature SC guidelines. Although the SC is clearly defined into a set of objectives and clear rules, there is some rather complex physics behind the world of air temperature monitoring. At first a description is given on the importance of measuring for the long run in terms of climate through air temperature monitoring. The chapter continues with a section that describe how heat interacts with the surface. In the third section of this chapter, the SC guidelines are subdivided into three different categories that each provide background information how the SC can be calculated according to the guidelines using different formulae and concepts or why they are included as criteria.

2.1 Air temperature monitoring

Besides real time meteorological measurements for daily use, AWSs are of great importance for monitoring of variations and changes in climate time series. This is known as climate monitoring and is the process of delivering as well as transforming data and information to describe the state or the changing state of climate (Karl et al., 1995). Proper climate monitoring is essential for a substantiated and objective debate on climate change. Today, climate change research focuses on its existence, ways to measure it, its possible causes and the modelling of future scenarios. Important element in this research includes climate monitoring to investigate any significant changes in the longer period of time. One of the most important variables in climate monitoring is the observations of air temperature; it is a key variable in climate monitoring. To collect data that is useful for long-term analysis, air temperature measurement of high and consistent quality are desired. It is therefore common for meteorological services to have and maintain climate reference networks that monitor and guarantee high quality of air temperature observations. In the Netherlands this includes stations that determine the Central Netherlands Temperature (van der Schrier, van Ulden, & van Oldenborgh, 2011).

Consistent air temperature monitoring can only be achieved if measurements are done the exact same way over the course of time, or if the changes in measurements of weather stations are completely documented. There are many guidelines to ensure a uniform way of air temperature and are all described in the Guide to Meteorological Instruments and Methods of observation composed by the CIMO of the WMO (WMO, 2017). This includes many guidelines about calibration, ventilation and many other technical aspects. For an air temperature sensor (figure 2.1) it is important the instrument is located at an height between 1.5 and 2 metres above the surface. Ensuring consistent data collection also includes the digitisation of all the available historical data (Mahmood et al., 2010; Parker, Jones, Peterson, & Kennedy, 2009). As an example, sudden jumps in air temperature series can indicate a change in the surroundings of a site or even move of the sensor itself. This can be corrected through a homogeneous adjustment if this has been documented. Although several homogeneity adjustments models have been developed to correct the changes and inconsistencies over time, there will always remain an error margin as long as there are different circumstances.

The AWS in De Bilt serves as an world wide important centennial station, and it is therefore important to keep assessing its representativeness and homogeneity of air temperature records (van der Schrier et al., 2011). This representativeness can be altered if the surroundings of the site (gradually) change or in case the site is relocated somewhere else. Therefore a study was done by the Royal Netherlands Meteorological Institute (KNMI) to investigate if the air temperature was still representative and if there were any inhomegeneities (Brandsma, 2011). In 2003 through 2005 parallel air temperature measurements were done at four different locations close to the operational air temperature sensor. Results showed that the operational site caused problems due to the nearby landscaped park with ponds which was developed over the years. This caused too much shading and sheltering to influence the wind and air temperature. Therefore the site was permanently moved in 2008 at its current location, that is also used for this thesis research. This is 200 metres east from its old position. The relocation of the operational site introduced an homogeneity which require a correction when the observed values are integrated in a time-series data. This research showed that (a change in) direct surroundings can have a direct influence on the observed measurements and therefore it is necessary to keep monitoring the surroundings on a regular basis to detect if there are any changes. The SC method can therefore serve as a method to keep these changes of surroundings better under control.

Figure 2.1: An air temperature and an air humidity sensor owned and maintained by the KNMI. This picture shows the shelters that are around the sensors.
Source: Raymond Sluiter

Many studies have been performed on how the quality of the sensors and the data they collect have an influence on air temperature monitoring and thus include any possible evidence that contributes to observed climate change. The influences on air temperature observations can be subdivided into three categories: the uncertainties of the sensor, the screen or shelter that protects the sensor and the surroundings of the meteorological instrument (WMO, 2017). Each category can introduce biases that has influence on the air temperature observation. When monitoring air temperature over longer period of time for climate research, additional biases can be introduced through different surrounding factors. Changes in observation instruments, objects surrounding a site, location of the sites or changes in way of observations including the time of day the measurement takes place, are factors that can introduce bias in the observations. This can lead to inhomogeneities on their turn and may result in erroneous characterisation of climate variability (Parker et al., 2009; Peterson et al., 1998; Pielke et al., 2007; Thorne et al., 2005). This makes it difficult to directly compare air temperature values over time. Using accurate data and meta-data collection protocols are therefore required to maintain high quality and observe any change in air temperature over climate time series.

The surroundings of an AWS partly determine the observational biases - the factors that make a meteorological observation less representative. It is therefore important to document with which surroundings the different measurements take place. Leroy's SC method can be used to classify AWS sites to give an indication with which surroundings the meteorological measurements have been taken. Using sites from the United States Climate Reference Network (USCRN), different published research have been done in the United States that make use of Leroy's SC method (Fall et al., 2011; Menne, Williams, & Palecki, 2010; Peterson, 2006; Pielke et al., 2007). By comparing the air temperature SC values, any significant air temperature changes were able to be observed. Results of these scientific studies were used to test existing or develop new homogeneity adjustment techniques to reduce or remove the biases introduced due to site exposure. While a few of these researchers claim a significant air temperature increase can be observed in both the sites that are classified as representative and unrepresentative of the region after homogeneity adjustments have been applied (Fall et al., 2011; Peterson, 2006), others claim this cannot be fully justified due to the methodology applied or lack of metadata on the sites (Menne et al., 2010; Pielke et al., 2007). For the purpose of this research, it is not relevant which conclusions are correct. Rather, it is important to mention that disagreement still exists about the applied methodology and associated homogeneity adjustments due to the (lack of) documentation of the (meta)data. While all researchers agree to use the SC method, discussions still exist how to deal with certain (meta)data that is missing or not consistently collected over time and required for a SC. All studies have manually classified sites by performing the classification on the sites themselves. A more consistent and reproducible approach could be found if a method was used that is able to classify all the sites in the same way by using the exact same data in an automated procedure. This could contribute to some discussions about the different applied methodologies of existing studies.

2.2 Surface energy balance

The land and the atmosphere interact with each other and are initiated by the solar radiation that hit the ground where an exchange of energy takes place. The presence of vegetation and heat sources therefore have an influence what happens with this radiation energy at the surface, and thus influencing the surroundings as a whole. This section provides a general description explaining the underlying physical and biological concepts behind the radiative heat exchange that takes place on the surface of the ground.

The exchange of heat at the surface is also known as the Surface Energy Balance (SEB). Part of the radiative energy that is present at the land surface enters the soil (soil flux, Q_G). The majority of the energy, however, is directly available to re-enter the atmosphere. This is done through two different types of turbulent fluxes: sensible heat flux (Q_S) and latent heat flux (Q_L). Sensible heat is the amount of energy that is absorbed or released in a phase change (eg. liquid to gas) and happens during surface evaporation. Moisture heat is the amount of energy needed to increase its air temperature without a change in phase. The amount of energy that is absorbed or released as sensible heat flux and moisture flux is dependent on the properties of the near-surface atmosphere. This land-atmosphere interaction explains the basics on how a heat air temperature measurement is determined (Heerwaarden, 2011; Kinoshita, 2014). This results in equation 2.1 that expresses the net radiation (Q_R) which is visualised in figure 2.2:

$$Q_R = Q_S + Q_L + Q_G \tag{2.1}$$

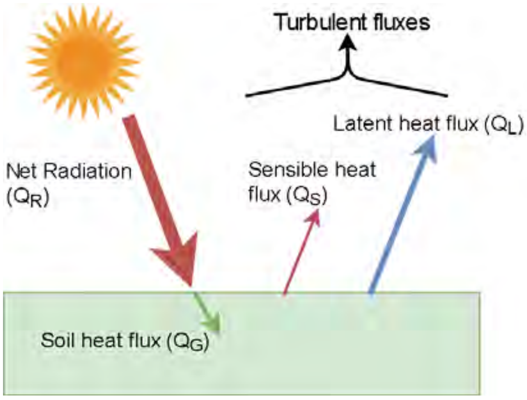


Figure 2.2: Visualisation of Net Radiation equation. The incoming energy at the surface comes from the sun. Some energy enters into the soil through soil heat flux abnd other is reflected through sensible and latent heat fluxes.
Adapted from: Budyko (1974)

.

The atmosphere near the surface is influenced by the properties of the land surface (the boundary) as well as the free atmosphere. Free atmosphere is the air that is higher than one kilometre above the ground. The air below is known as the Convective Boundary Layer (CBL). When heat is released from the land surface - known as convective turbulence - the heat rises within the CBL because these plumes of air are warmer and moister. This makes the air less dense than their surroundings. The air will stop rising when it reaches the free atmosphere where these turbulent plumes are no longer denser than the surrounding. The plumes will therefore drop again into the CBL, but will mix with the dry and stable air that is present in this part of the atmosphere through the process of entrainment. This process is shown in figure 2.3.

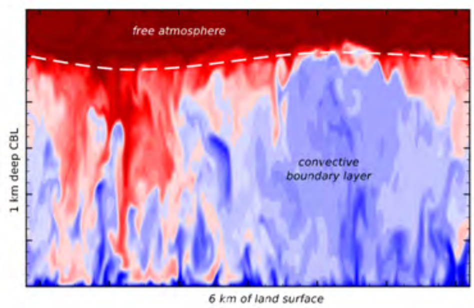


Figure 2.3: Visualisation of the Convective Boundary Layer showing the mix of of falling (red plumes) and rising (blue plumes) thermals in the lower 1 km of the atmosphere. *Source:* Heerwaarden (2011)

.

The properties of the CBL can vary over space because the land surface may vary and be heterogeneous. This results in various sensible heat and moisture fluxes over space that makes a horizontal dynamic CBL. Latent heat and moisture fluxes can therefore vary across space and result in a more complex model that defines the properties, and thus the air temperature, of the atmosphere. There are two types of heterogeneity that have an influence on the CBL: thermal and mechanical heterogeneity. The former defines the different heat and moisture fluxes in space while the latter defines the roughness lengths of the surfaces. Roughness of the surface has a strong impact on the wind and therefore on the flow of the air in the atmosphere. The flow of air determines the amount of heat there is present in the atmosphere at a certain location. It also impacts the amount of energy that can be released by plants through transpiration. The more air flow there is, the more energy the plants can release into the air. Therefore, the amount of wind has a crucial impact on the measured air temperature.

Air temperature is a measurement that determines how much heat is stored in the air. One of the major factors that influences this phenomenon is water. Water is present in the atmosphere in all three phases: solid, liquid and through vapour. Solid and liquid state water is seen in the atmosphere through clouds and precipitation. Water vapour is present through evaporation: the vaporisation process where water changes from its liquid state to a gas and determines the amount of moisture that is present in the air. This process can occur directly through water bodies or through transpiration from vegetation on land. All energy that is not used for evaporation is used to heat up the atmosphere and land surface. Consequently, evaporation has a large impact on the measured air temperature close to the land surface. (Heerwaarden, 2011)

The described equations and concepts of the SEB and CBL are not directly found in the temperate SC guidelines. These phenomenon do describe, however, the underlying reason why vegetation and the presence of heat sources have an influence on the air temperature in a certain area.

2.3 Air temperature Siting Classification guidelines

The air temperature SC guidelines (see appendix C) is part of a larger document that defines the classification criteria of different meteorological variables including air temperature, wind, precipitation, radiation and environment (WMO, 2017). These criteria mainly describe the influence of different types of direct surroundings on the measurements of the different meteorological instruments (WMO, 2017). The guidelines give an indication of the quality of the data collected by the instruments based on the presence of artificial and natural objects in the directs surroundings of the AWS site. The primary aim is to provide awareness to researchers and meteorological networks on how the direct surroundings of an AWS site influence the meteorological measurements in such a way that the end user of the data has sufficient metadata for correct interpretation of the observed values. The current guidelines do not provide absolute quantitative uncertainty of the measurements expressed in in number values although the aim is to achieve that in the future. Instead, it simply shows qualitatively how representative a site of a larger area through a number scheme by indicating the influence of natural and artificial objects in the nearby surroundings of a site. Representativeness in terms of the air temperature SC, as defined by the WMO, is the extent the direct surroundings of a site represents the characteristics of a small area (<2.5km ^{2}) (WMO, 2017). The SC guidelines only provides a broad indication of uncertainty around a measurement performed by the sensor. Each variable has its own classification score (usually between 1 and 5). The lower the value, the more representative the site is of that area. A variable with a Class of 5 is a site "where nearby obstacles create an inappropriate environment for a meteorological measurement that is intended to be representative of a wide area" (WMO, 2017).

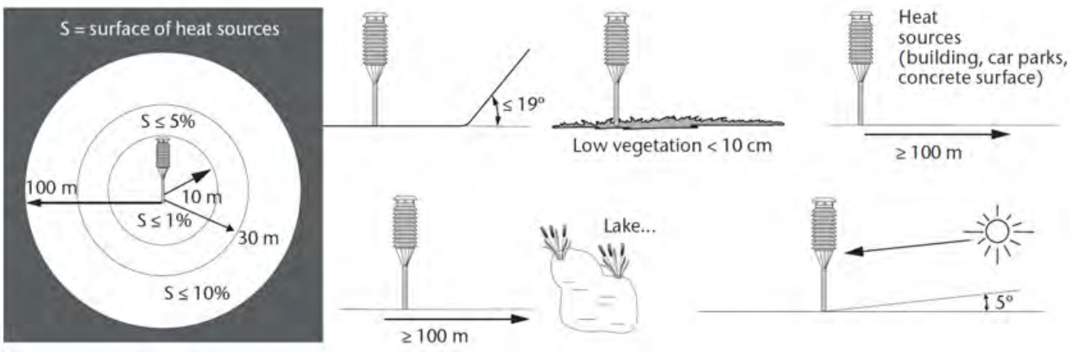


Figure 2.4: Visual overview of the different criteria set for the air temperature SC variable. These represent the criteria values for Class 1 and hence show the "ideal" situation for air temperature observations that are considered as most representative. *Source:* WMO (2017)

Various measurements and calculations are performed for the air temperature variable in order to determine the final SC Class number 1 through 5. The air temperature biases are influenced by three criteria:

1. The presence of shadings throughout the year due to tall objects in direct surroundings of the air temperature sensor
2. The presence of heat sources (natural/artificial objects and surfaces) in the direct surroundings of the air temperature sensor
3. The presence and height of the vegetation at the location of the air temperature sensor.

The criteria have their own rules and require their own set of calculations as shown in figure 2.4 for Class 1 and table 2.1 provides an overview of values for all classes. Each criterion has boundary values set for Class 1 through 5. This means that the highest Class found determines the overall SC value. This means that no weights of importance are applied to the individual criteria and that even a nearly perfect site will obtain a higher classification - and thus be a lower representation value - if it does not meet to just one criteria. This section describes these three criteria by providing some background information on why they are important and how these are measured.

Table 2.1: Class criteria values to determine the different Classes according to the WMO SC Guidelines. *Adapted from:* WMO (2017)

Criterion	Description	Class	Criteria value
shades	shadow angle	1-5	5, 7, 7, 20, NA
land use	inner buffer distance	1-5	100, 30, 10, 10, NA
land use	annulus distance	1-5	10-30, 5-10, NA, NA, NA
land use	outer buffer distance	1-5	0.1, 0.1, 0.1, 0.5, NA
vegetation height	vegetation height	1-5	0.1, 0.1, 0.25, NA, NA

2.3.1 Projected shadings

Air temperature values are different when measured in the shade than when they are measured in the sun, as the shortwave radiation from the sun directly influences the radiative balance of the air temperature sensor and its surrounding area. Therefore, it needs to be known if and when the air temperature sensor is shaded during what time of the day and year. Shadows are a result of objects that block the sunlight from shining on the surface of the earth. A shadow line can be characterised by its height (h_{shadow}) and its footprint ($dist$). The shadow height determines if a a point above the surface will be shaded and the shadow footprint determines the length of the shade on the ground. Both are dependent on the position of the sun, the height of the blocking object and the distance from a point to the blocking object.

Position of the sun

The position of the sun needs to be known first and is determined by the azimuth angle (α_{az}) as well as the elevation angle (α_{elev}) as shown in figure 2.5. The azimuth angle is the number of degrees the sun is away from north measured over east. The elevation angle is the amount of degrees the sun is elevated from the point of measurement. Both the azimuth and elevation angles can be calculated from every point on earth for every second during the year. This is done by using spherical trigonometry formulae. (Corripio, 2003)

Knowing the azimuth angle and solar elevation angles at constant intervals throughout one day for a certain location, provides all the elements required to plot a local sun chart. This line curve gives an overview what the minimum and maximum solar angles are throughout one day. A yearly overview is provided when several sun charts of different days throughout the year are plotted on one graph (see figure 2.6. Best practice for this year overview is to provide the sun chart of the shortest day (21$^{\mathrm{st}}$ of December) with a month interval until the longest day (21$^{\mathrm{st}}$ of June). This sun chart was used to determine if there is any shading based on the calculated shadow angles and thus determine the classification of the shadow criteria.

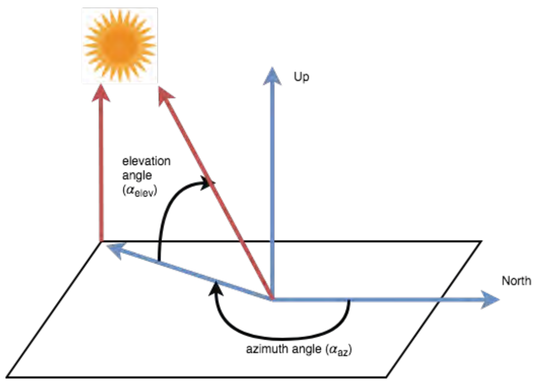


Figure 2.5: Solar angles determine the location of the sun from a certain point on earth. They include the azimuth angle and the elevation angle.
Adapted from: Blanco-Muriel et al. (2001)
.

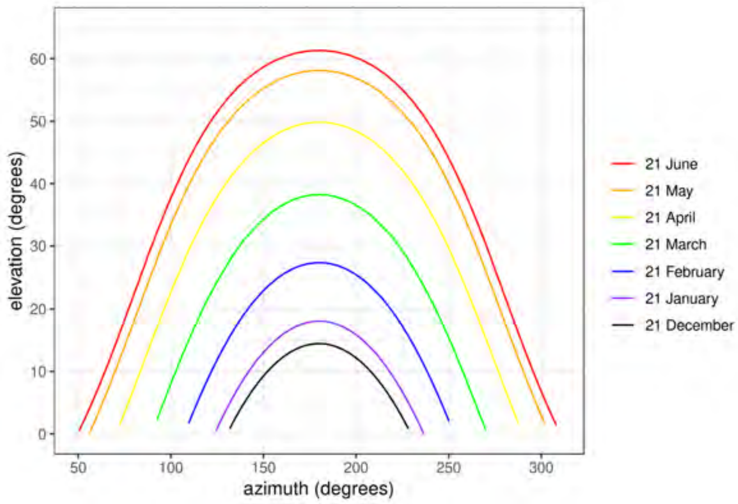


Figure 2.6: This sun chart shows the sun path of the shortest day of the year (21 December) through the longest day of the year (21 June) at the AWS in De Bilt.

Shadow height

With a known solar elevation angle (α_{elev}), height of the blocking object (h_{build}) and distance between the viewer's point and the object (dist), the height of a shadow (h_{shadow}) can be determined. The shadow height determines how high the viewer's point needs to be if it wants to stay out of the shade caused by the blocking object.

$$h_{\text{shadow}} = h_{\text{build}} - dist \times \tan(\alpha_{\text{elev}}) \tag{2.2}$$

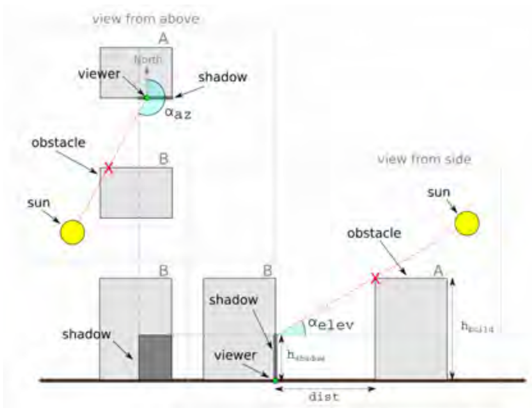


Figure 2.7: This diagram shows how shadow height can be determined. With a constant solar elevation angle, the distance from the viewer point to the blocking object and shadow height have an indirect relationship. *Source:* Dorman et al. (2017)

Shadow footprint

$$dist = \frac{h_{\text{build}}}{\tan(\alpha_{\text{elev}})} \tag{2.3}$$

Applying the distance to the outline of the obstacle that causes the shade results in a shadow footprint that is projected in the opposite direction of the solar azimuth. Shadow footprints are useful to calculate exactly what ground area is shaded at a certain moment in time. Plotting shadow footprints of different moments throughout the day gives a clear overview of the parts that are shaded at what time (figure 2.8). (Dorman et al., 2017)

When the elevation angle remains constant in both described formulas, the shadow height and distance will have an inverse relationship: the lower the shadow height, the greater the distance from the viewer's point to the blocking object and vice versa. This also means that the shorter the distance is to the blocking object from a viewer's point, the bigger the angle with the intersection point (red cross in figure 2.7), which is the top of the blocking object. This angle is known as the shadow angle. A viewer's point will be shaded as long as the shadow angle is greater than the solar elevation angle. It is this shadow shade angle that is used as a criterion in the SC guidelines.

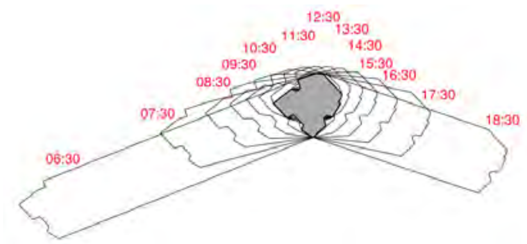


Figure 2.8: Example of shadow footprints during different moments of the day.
Source: Dorman et al. (2017)

Sading criteria guidelines

The SC shading criteria define a classification for when the air temperature sensor is shaded for any period of time during the day or year. The guidelines state that if a certain viewer's point is away from all projected shade with a solar elevation angle lower than 5 degrees, it receives a Class 1 for this individual criterion. On the other end of the scheme, a Class 5 is received for this individual criterion if any shadow angles is greater than 20 degrees. This means that if the all the shadow angles are smaller than 5 degrees, it receives Class 1 and if one or more shadow angles is greater than 20 degrees, it receives Class 5. The calculation of one solar angle and one shadow angle only allows for the determination of the projected shading at one moment in time. The year overview sun chart is used to determine the shadows angles throughout the whole year and thus know when a certain point is shaded. (WMO, 2017) Any shadow angle at any azimuth angle that is greater than its corresponding solar angle on the longest day (21 June) is not a seen as a shadow because the sun does not cross that path at any time during the year.

2.3.2 Heat sources

While shading may locally decrease the air temperature, artificial or natural objects (known as heat sources according to the SC guidelines) may warm up the measurement instrument as these objects may hold radiation much longer (WMO, 2017). On a warm sunny day in the summer, water bodies usually cause an cooling effect, and artificial bodies usually cause a heating effect (Heerwaarden, 2011). For a representative air temperature observation, it is desired that the area surrounding the sensor is away from heat sources that may attract or reflect radiation. These include water bodies, buildings, roads, antennas or other objects.

Important elements of influence on the air temperature measurement are size of the heat source and distance to the sensor. Also direction and amount of wind need to be taken into account (Jianxia et al., 2014; Kinoshita, 2014; Kumamoto et al., 2013; Shido, Yamamoto, Aoyagi, Seino, & Fujibe, 2016). It is therefore crucial that also wind is measured using high quality standards, by knowing what factors influence its speed and direction. (Brandsma, 2011). Small-scale experiments done on air temperature observations performed within one month in Japan showed that asphalt roads can have positive biases when air temperature is measured 0.5 or 1.5 metre above the ground. Bias up to 0.4 °in the summer with were greater than in the winter and only significant under conditions of little wind. Furthermore, this research showed that roads within 100 metres away from the thermometer can have an influence on the observed air temperature measurement (Kumamoto et al., 2013). Another experiment showed equivalent results implying that when the distance of the road to the thermometer is greater than 60 metres, the changes in air temperature become negligibly small (Kinoshita, 2014). Both experiments note, however, that it is very hard to mimic a situation that is applicable to all the observation sites. Each site is unique and has its own particular surroundings, thus making it hard to indicate what the exact causes of observed air temperature differences may be.

The locations at which meteorological observations are taken/performed, are usually selected mainly by the experience of trained staff with sufficient knowledge of the local conditions (Kinoshita, 2014; WMO, 1993). Placing sensors away from heat sources can be difficult to achieve in dense urban areas. Scientific research that tried to quantify the influence of artificial objects present in the direct surroundings of the air temperature sensor on the measurement has only been picked up ten years after the SC method has been developed by Leroy in 1998 (Kinoshita, 2014). This limited amount of research is an important reason why the SC method is only an indication of uncertainty that provides extra information and awareness about the conditions of the AWS and its surroundings and not a quantitative margin on a measurement.

The SC guidelines for the heat sources criterion consists of two elements: the distance to the nearest heat source and the relative amount of surface area the heat source has within certain radii. The guidelines first looks at if there are any heat sources objects presence within 100, 30 or 10 metres around the thermometer. If no heat sources are found within the radii, they respectively receive Class 1 through 3 for this criterion. If heat sources are found, however, the amount of surface area of the heat source within the associated radius is taken into account. If the surface area of all the heat sources together is smaller than the set criteria, then the associated heat objects are considered as negligible. If the surface area exceeds the set boundary value, the criteria are not met for that Class. Class 4 criteria only look at the amount of surface area within 5 and 10 metres radii. Class 5 is received when in case none of the criteria for Classes 1-4 are met (WMO, 2017).

2.3.3 Vegetation

Air temperature is highly influenced by the presence of vegetation because of photosynthetic processes plants undergo. Each type of vegetation differs in the amount of heat energy it absorbs from and reflect into the the air. Therefore the heat and moisture flux, and consequently the influence on the observed air temperature, is partially dependent on the type of vegetation. In general, vegetation has a cooling effect during the day and retains the heat during the night, especially in forestry areas. Due to evapotranspiration, vegetation causes cooling, thereby reducing the amount of moisture energy that is in the air and therefore also the air temperature. If there is moisture in the air, no evapotranspiration takes place and therefore also no cooling of the air temperature. This can causes droughts if no evapotranspiration takes place for a longer period of time (Dimoudi & Nikolopoulou, 2003; Heerwaarden, 2011; Petralli, Massetti, Brandani, & Orlandini, 2014). The amount of wind plays an important role in the amount of energy they absorbs and releases (Shido et al., 2016).

The vegetation criteria for the air temperature SC guidelines state that an air temperature sensor at an AWS should not be located between vegetation of significant height. This is because the presence of vegetation has an influence on the heat and moisture fluxes near the ground. Therefore, vegetation that is closer to an air temperature sensor can have a higher impact than plants or crops that are closer to the ground. Therefore the SC method sets a criteria on the height of the vegetation in the nearby area of the air temperature sensor. The criteria set different height boundaries to determine the Class. If the height of the vegetation is 10 cm or lower, it receives Class 1 and if it is higher than 25 cm, Class 4. Furthermore, the SC method notes that the surrounding ground of a site should have natural vegetation that is "representative of the region". The air temperature would not be representative if the sensor is located on a type of land that is not found in the greater area of the site (WMO, 2017).

3 Methodology

In this thesis research a model was developed that enabled an automatic SC of AWS sites based on the criteria guidelines by the WMO for the air temperature observations. This chapter explains what data is used and what steps are taken to develop the model and test it on its usability. The chapter starts off with a description of the current MSC procedure to understand the context of the model better. This also allowed a better analysis when the automated values were compared to the manual classifications. The chapter continues with an overview where all the AWS sites are located in The Netherlands. Also description of the different data that was included and (pre-)processed. Moreover, the chapter continues with the methods that were applied to develop and perform the different criteria analyses and ends with a section that describes how the model was tested on its usability.

The model was developed using R - a programming language that enables elaborate and powerful statistical computing (R, 2019). Within R, different GIS tools were used To perform spatial analysis and create geographic maps. The final model of this research was distributed online through GitHub as an R package called *temperatureSC* for future use, development and reference (Stuurman, 2019c). A R package is a set of reproducible R code that has functions that can be reused for different analyses. It also includes documentation on how the functions within the package work.

To automate the classification, a variation of spatial and mathematical analyses was needed. Different existing R packages were used that are capable of performing the necessary calculations for the model. Table A.1 in the appendix A provides an overview of the most important packages and the associated functions that were used to make a working SC model.

3.1 Current MSC procedure

The SC for the Dutch AWS is currently done manually by professional field inspectors from KNMI. They visit the AWS once a year and gather information. The results of this MSC process are processed and made available internally within the KNMI. They are considered in this research as the reference based on which the skills of the model developed in this study can be assessed. At first, it is assumed that the manual classification values are the correct ones, although the model may produce output that could lead to suggestions for improvement of the MSC procedure. It is therefore important to know how the current classification procedure works so that a better comparison can be made. This section describes how the current SC is performed manually and what procedure is carried out to obtain the final values for the SC classes. The description of this manual classification is based on several interviews held with two field inspectors of as well as being an observer as a participant during an AWS site field inspection. The described procedure below focuses mainly on the procedure carried out for the air temperature variable.

The current MSC method is performed by field inspectors who have a technical and meteorological background and is done according to standardised protocol. Besides the SC, they are also responsible to check if the instruments are clean, work properly and are calibrated. The SC is mainly based on three aspects: measurement of heights of terrain and instruments, visual interpretations and a conversation with the owner of the land where the AWS is located. Important part of the interpretation is the knowledge the inspector has of the (history of) the site and the surrounding area. Unlike in this research, field inspectors do a SC for all the variables that are described in the WMO guidelines and also perform actions to also include the wind, precipitation, environment, global radiation SC values. Collected information or a SC value of one variable may also be used as input to determine the classification of another. The data collected for the air temperature variable can also be used for the observation of humidity using the hygrometer. Determining the height of surrounding objects that cause shading on the air temperature sensor, for example, are also necessary to determine the influence on the wind.

The most important and time consuming part of the classification (other than the logistics of reaching the site) is measuring the height of all the objects in the surrounding area of the AWS site using a laser range finder binoculars (figure 3.1). This is required for the wind, precipitation and air temperature variables. Positioning the binoculars at a fixed position height of 1.5 metres above the ground and five metres south of the anemometer, the inspector makes a 360 degrees scan of the area. The inspector collects the exact coordinates from where the measurement location. Starting at north (0 degrees) going clockwise over east, the inspector points the binoculars to the top of every object found and presses a button to collect the measurement number, azimuth angle, elevation angle and the distance to the object. The azimuth angles and elevation angles are collected by a compass that is built within the binoculars. The distance to the object is determined by a laser that is sent to the object. The time it takes to return to the binoculars determines the distance. These objects can include trees, buildings, radio towers and meteorological measurement instruments. The laser scanner is attached to a on-site laptop computer where the information is directly stored. For a 360 degrees surrounding scan, the minimum number of measurements should be at least 360 (one for every degree) but this often includes more and reaches around 400. During this procedure, the inspector makes notes to describe the objects scanned. As such the field inspector can refer later which measurements (through its ID and azimuth angle) coincide with which objects.

Figure 3.1: Field inspector using laser binoculars to determine distance to, height and elevation angle of objects around an AWS site. *Source:* Jelle Stuurman

After scanning the heights of objects with the binoculars, a visual inspection is made of the land cover and use at the AWS and in its direct surroundings. Furthermore, the inspector checks if the vegetation is uniform in type and if it has the same height. The inspector also makes notes of any present heat sources. Heat sources include artificial objects such as buildings, roads, meteorological instruments, walls, but also any type of water body. Also (panorama) photos are taken of the site and parts of the sites that are important to document. Based on these visualisations, the inspector makes an estimate of what heat sources will influence the measurements. These estimates are mainly based on expertise knowledge about meteorology, previous experiences, and information of the site. As an example, if a small and shallow lake is found in the direct surroundings of the AWS site, it will not be taken into account as an influence on the measurement "because shallow and narrow water bodies have a insignificant influence on the air temperature observation", according to a field inspector. At last, the inspector asks the owner or manager of the land where the AWS is located, if there are any issues. The inspector also asks the routine of maintenance and gives any feedback about the status of the AWS if necessary.

Back in the office, the inspector writes a field inspection report, including the determination of the SC for each variable. The inspector reads in all the collected data from the binocular into a spreadsheet. Using a standard printed out sun path of 51, 52 or 53 degrees Latitude (that together cover the latitude range of The Netherlands), the most important shadow angles are drawn by hand with onto the sun path chart to see what values are higher than the criteria values provided by the WMO guidelines. Figure 3.2 shows the shadow angles determined at De Bilt drawn manually on a 52 degrees latitude sun chart.

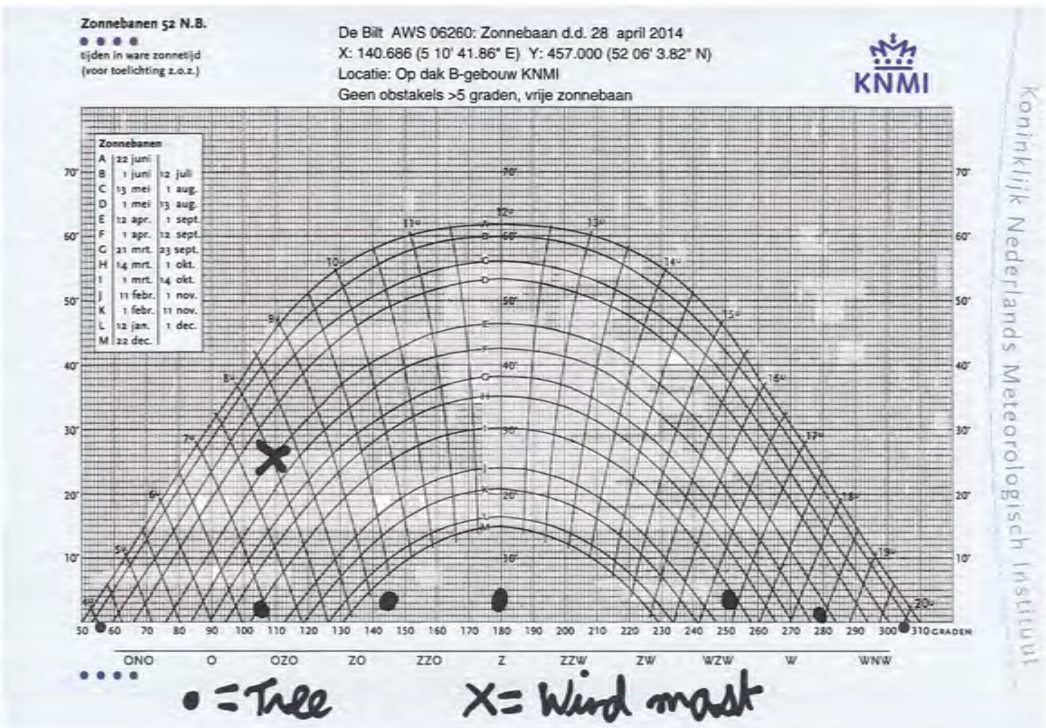


Figure 3.2: Sun chart including the drawn in shadow angles (black points) of the biggest or tallest objects in the surroundings. *Source: KNMI*

If any artificial objects or surfaces are found, the inspector checks the distances to these objects through a map viewer on the computer. Using a distance tool, an estimate is made how far these objects are located from the air temperature sensor and a guess is made if it covers a relative amount of surface area within a circle around the meteorological instrument. Based on the shadow angles compared to the sun paths, presence and amount of land use and state of the vegetation in the direct surroundings, a classification score is given for the air temperature variable. Often this includes a range of two adjacent numbers. This is especially the case when only one or a few shadow angles of very narrow objects are responsible for a higher Class value while the other criteria get a lower SC score. This often occurs due to the wind mast: a 10 m tall object to which the anemometer is attached and is is often nearby the air temperature sensor that causes shadow. The inspector writes down his findings into a report and uploads all the collected data into an online database. Performing field work and writing an inspection report to determine SC for all the variables may cost up to two full working days a year, especially if long travel time is required to reach the site. Field inspections are done every year for each site, as recommended by the SC guidelines from the WMO (WMO, 2017).

The results from the manual classification were compared after the developed ASC procedure through a comparison validation (see chapter 3.4.2). This was done because the model may have produced output that seems usable but if it does not completely coincide with the manual values, then the differences need to be analysed to see if the model required improvements, if the manual inspection was done differently, or if there are elements that need to be considered when comparing the two. While the automated procedure may include uncertainties due to uncertainties of the datasets, field inspections are subject to human interpretations that may lead to interpretation differences. This aspect was taken into account when comparing the sites to observe any trends of similarities and differences found.

3.2 Data description and pre-processing

Table 3.1: Overview of data(sets) used and their main characteristics. *Sources:* AHN; BZK; Kadaster; PDOK, 2018

dataset	Collection date	Renewal	Resolution	Accuracy
AWS coordinates	2019	yearly, if applicable	points	1 m
AHN2	2008-2012	5-6 years	0.5 metres	5 cm
AHN3	2012 - 2019		0.5 metres	5 cm
BGT	Continuous	Monthly but dependent by source holders	1:500 - 1:5000	20 cm
PDOK luchtfoto (Satellite image)	2018	yearly	25 cm	25 cm

Different datasets are required to create the ASC model for the air temperature variable. The SC is primarily based on and objects and surfaces that are in the direct surroundings of a site. This makes spatial data one of the most important data types to use as input in the model. Based on the current SC guidelines for the air temperature variable, the different datasets are divided into three main categories: height, solar angles and land use. Figure 3.3 visualises what specific information can be retrieved from which dataset. The starting point are the exact geographic coordinates of the air temperature sensor at the AWS sites. From there, different datasets are retrieved for the direct surroundings, including the Actueel Hoogtebestand Nederland (AHN) and Basisregistratie Grootschalige Topografie (BGT). Table 3.1 provides an overview of the used datasets with their general characteristics.

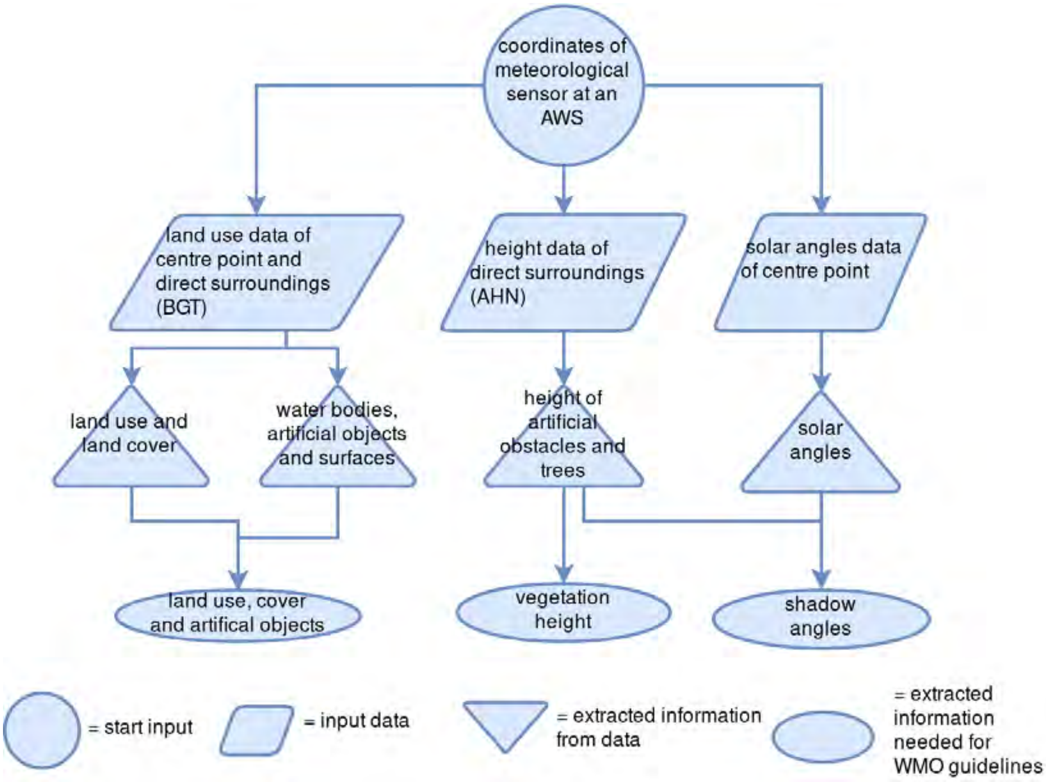


Figure 3.3: Data extraction from the different AHN, BGT and solar angles needed for the air temperature SC criteria.

The criteria for the air temperature variable are primarily based on the effect of the shading (height and solar angles) and of natural and artificial objects in the direct surroundings of a site (heat sources). Using an accurate altitude dataset that also has a high resolution, the altitude and distance from the AWS site of trees or other artificial object can be calculated. Together with the known solar angles, it can be estimated if and when the air temperature sensor will have any shading. Furthermore, height data was also used to calculate the altitude of the vegetation. The land use dataset enables to calculate how far the (artificial) objects, roads, buildings and water bodies are from the centre point and how much surface area they cover.

3.2.1 Studied AWS sites

Location studied AWS sites

The KNMI operates and maintains a nationwide meteorological observation network that contains 34 AWS providing real-time measurements for different purposes (KNMI, 2018). Besides the AWSs, extra or specific measurement instruments can be installed at harbours, airports or other special areas that require extra parameters or a (spatially) denser set of observational data. This research focuses only on the air temperature sensors that are located at the 34 sites and shown in figure 3.4. Among these AWSs, 14 of them are located at civil and defence airbases, thus representing 42.2% of all the sites.

Figure 3.4: Map of all the 34 AWS sites in the Netherlands. *Data:* KNMI

AWS in De Bilt

The site of De Bilt was used at first to test the model on before it was applied to the other AWSs. This site is also used in the rest of this chapter to describe the data that was used and then steps that were done. This site was selected as test site because - like many other researches done by the KNMI - this AWS was located next to the office of the meteorological institute so that own observations were easily made on the instruments and on the surrounding area to help and interpret the obtained results. This AWS is the most important station in The Netherlands as it is because of its central position close to KNMI headquarters and as it is used as the national indicator for reporting summer days or ice days etc.

Figure 3.5: AWS site in De Bilt. *Source:* Jelle Stuurman

The AWS itself is installed and surrounded by grassland in the direct surroundings. This follows the WMO criteria that an AWS needs to be located on a well vegetated and flat terrain (Steeneveld, Koopmans, Heusinkveld, van Hove, & Holtslag, 2011). The surrounding area (0.2 - 2 km) around the site mainly include grasslands but also include urbanised areas, a group of garden sheds and a landscaped park land with ponds. The urban areas can be found on the north, west and south side. The east side mainly includes rural areas going towards a a bigger rural and forestry area. When looking at the greater area around De Bilt, similar land use can be found. The north throughout the south area mainly have rural areas including the Utrechtse Heuvelrug, although urban areas are included as well (Zeist, part of Utrecht, Bunnik). At the other end of the Utrechtse Heuvelrug, the city of Amersfoort is located east. On the west and south west side of the AWS, there is mainly the city of Utrecht and rural areas beyond that.

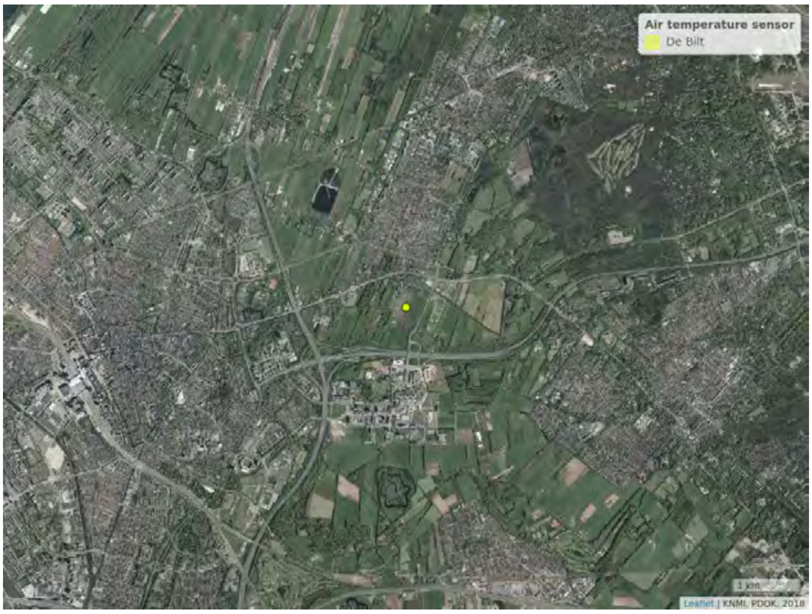


Figure 3.6: The greater area around the AWS in De Bilt mainly including rural areas, the city of Utrecht (west) and Utrechtse Heuvelrug (east). North is up.

Geographic coordinates

Accurate coordinates of an AWS site are required to determine the direct surroundings of an AWS site. To increase the accuracy of the results, the exact location of the air temperature sensor are used instead of the general coordinates of the site which is usually the pressure gauge. These instruments may be located at some distance from each other. The model required both the world standard coordinate system format - including that of the WMO - called WGS84 as well as the Dutch coordinate system, RD New. This is because some analysis calculations in the R packages are only developed based on the world format. To keep uniformity as much as possible, the Dutch RD New coordinates were used as much as possible, also because most input datasets are in this reference system. Appropriate spatial conversions were done throughout the analysis to obtain the right format of geographic coordinates, where needed.

3.2.2 Height data

AHN2/AHN3: altitude datasets

In the Netherlands the Actueel Hoogtebestand Nederland (AHN) datasets were used to determine the altitude of any location in the country (AHN, 2018). Its high resolution and high accuracy makes the AHN an excellent source to determine altitude of trees, buildings and any other large or tall artificial objects. The AHN is a high-resolution Digital Elevation Model (DEM) and is measured with Light Detection And Ranging (LIDAR) technology. It measured the altitude values relative to the Amsterdam Ordnance Datum (Normaal Amsterdams Peil, NAP) for one single moment in time. All the data is collected from early November to the end of March during which time plants and tree are not in their bloom, which could cause a decrease in altitude accuracy.

This research made use of two versions of the AHN: AHN2 and AHN3, both of which are made available by the Dutch government as open data. The AHN2 was collected between 2007 and 2012 and the AHN3 started in 2012 and will be finished by the end of March 2019. For this research the AHN2 and AHN3 were used. Both the AHN2 and AHN3 have a resolution of 0.5 metre with a altitude accuracy of five centimetres and an error margin of maximum of twenty centimetres. Objects that have a minimum width and length of two metres are guaranteed to be detected by the AHN2 and AHN3, although more narrow objects can be observed as well (van der Zon, 2013).

It is the combination of high resolution and accuracy that makes this a key dataset to use and make an ASC model possible. Without the high resolution and accuracy, the AHN dataset has, calculating shadow angles would be too inaccurate and questionable if an ASC would be feasible to achieve. For this research the AHN3 is primarily used as it has the most recent data to compare to MSC values. The AHN2 was used to analyse the shadow angles of the other sites to compare two moments in time and create a better understanding on how well altitude differences (such as trees) can be detected. For areas where the AHN3 is not available yet, only the AHN2 was used. Because both the AHN2 and AHN3 data are collected using similar LIDAR technique with the same resolution and accuracy, it is assumed that the two datasets only different is the moment of time they were collected. References made as AHN in this research report refer to either the AHN2 or AHN3.

Downloading and pre-processing the AHN

The AHN2 and AHN3 are available in raster and point cloud format. Each data format includes several different sub datasets. For this research the fifty centimetres resolution Digital Terrain Model (DTM) and Digital Surface Model (DSM) raster datasets are used. The DTM shows the altitude of the terrain and the DSM also includes the height of any obstacles such as buildings or other objects. The DSM is used to enable shadow angle calculations. The combination of the DTM and the DSM were used to calculate the height of the vegetation: it is difference in observed altitude between the two datasets. The term 'height' used in this research refers to a difference in detected altitudes within one or between the two types of sub datasets.

The DTM and DSM raster data were downloaded separately into R but first required a procedure as shown in the flow chart figure 3.7. This procedure was used for both the AHN2 and AHN3. Both the DTM and DSM are split up into separate raw raster

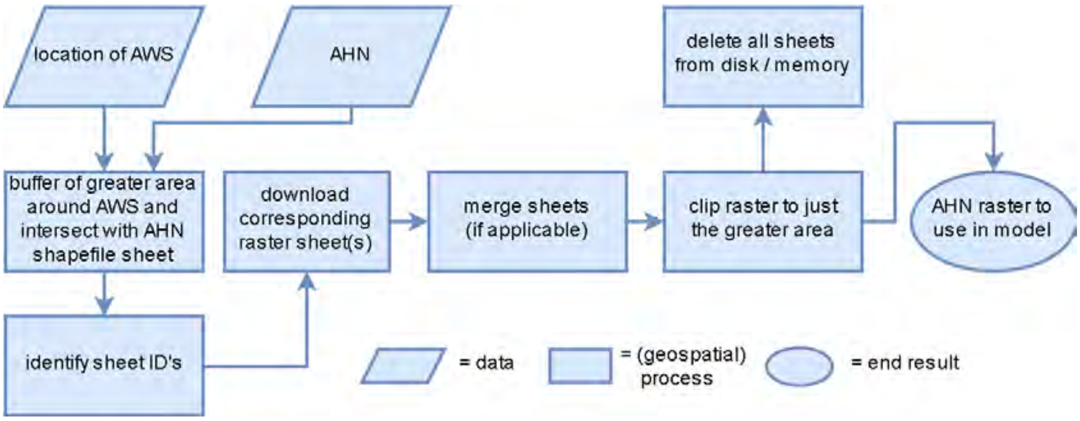


Figure 3.7: Methodology to extract the AHN datasets of the surrounding area around an AWS and import it into the model.

files (sheets) that each represents a rectangular area of around 30 km ^{2} (5 km x 6 km). These raw raster sheets are made available through the website of the PDOK (PDOK, 2018a, 2018b). The correct raw raster files needed to be identified first so that only the data was downloaded that fall within the region where the AWS and its greater area is situated in. The location of these raw raster files is made available as a separate shapefile through a Web Feature Service (WFS). A WFS is an interface standard to allow the reading of spatial data and its attributes across the internet (OGC, 2019). This shapefile includes an attribute value that has the unique ID's of all the sheets. With this shapefile, an intersection was made with the location of the AWS including a radius of one km around the site, the greater area of the weather station. As a result, the required unique ID's were identified. If more ID's were found during the intersection, all the sheets were downloaded onto the disk and merged at the end. Automatic downloading was possible because the download link always follows the same pattern and includes the raw raster unique ID.

The 500 metres radius was used to determine the final raster area that was used in the model. The rest of the downloaded data was deleted as it was not used for further processing and analysis. After the DTM and DSM were downloaded and pre-processed, the data was ready to be used for shadow angles and vegetation height calculations. Because this extract procedure can also be used for purposes other than this research, the script was published as a separate open source repository on GitHub (Stuurman, 2019a). This is done because no similar automatic extract tools existed that is capable to import the AHN data of a specific region using R.

Figure 3.8: DSM of direct surrounds of De Bilt AWS and its air temperature sensor based on the AHN3 dataset. Darker blue colour indicates a higher altitude.

3.2.3 land use data

BGT dataset

In the Netherlands the "Basisregistratie Grootschalige Topografie" (BGT) is the most detailed topographic map available in the country (Kadaster, 2018). It includes all the land use and artificial objects that can be found on land. The data is made available for free as open data by the Dutch government and maintained by the municipalities who are the source holders of the provided data. Each municipality is responsible and required by law to register all the objects and update the data stored in the BGT where needed or appropriate. A new release of this dataset is made available every month. Using this dataset gave the most detailed land use map as it includes all the smallest natural and artificial objects, as well as surfaces. Figure 3.8shows the BGT map of the direct surroundings of De Bilt that is used for the analysis.

Downloading and pre-processing the BGT

Similar as the AHN datasets, a WFS was used to retrieve all the relevant data for the BGT. Because this dataset is updated every month, downloading the data automatically through the model is an important step to always retrieve the most up to date topographic information. Figure 3.9 provides an overview how the BGT is selected and downloaded.

Only the relevant BGT data with a radius of 150 metres around the air temperature sensor of one site are downloaded into the model. This radius distance is enough to collect sufficient information within the direct surroundings of the site. This was also done because the current WFS only allows to download 100 spatial features per request, due to the high volume of information it receives. The BGT is a complex dataset that exists of different Geography Markup Language (GML) files that have different geometry types. Each GML file represents a land use type or type of object

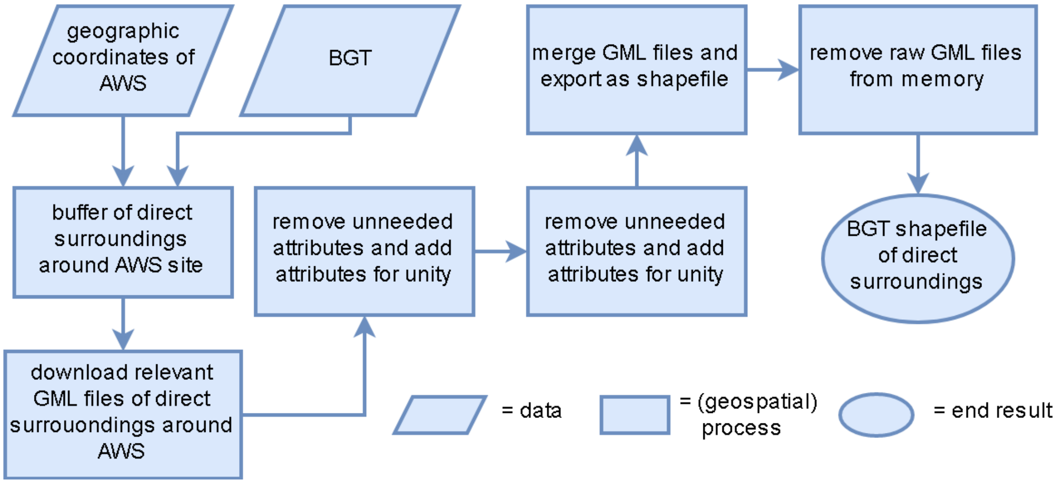


Figure 3.9: Methodology to extract the BGT datasets of the direct surroundings around an AWS and import it into the model.

(eg. water, roads, buildings, railway, radio tower etc.). Table 3.2 provides an overview of all the objects that were found in the direct surroundings. These are only a limited type of objects because AWS sites already are located in areas that have open spaces in the direct surroundings and therefore only contain certain objects.

Table 3.2: Overview of all the BGT objects (polygons) found in the direct surroundings of all the studied AWS sites.

Dutch BGT object names (polygons)	Description
begroeidterreindeel	vegetation
onbegroeidttereindeel	barren land
pand	building
wegdeel	road
waterdeel	water
scheiding	seperations (walls, fences)
overigbouwwerk	Other constructions

After all the GML files were downloaded through the WFS in the model, the files were merged together into one shapefile to allow efficient spatial analysis. Merging is only possible if all individual GML files have the exact same geometry type and attributes. Therefore, several steps were needed to allow the merge. At first, the 'separations' features were converted to polygons. Secondly relevant attributes were selected of each feature type. These attribute data may include specifications about the type of water body, road, building, barren or vegetation land. In this research, these specifications are not used as input within the model but only present to interpret the results better. These specifications about the features are excluded from the model because they are not always present and is dependent by the source holders to what extent they document everything. At this stage it is therefore not a consistent input variable to be used for the SC model. Lastly, empty attributes were added of the other feature types so that each object has the same attributes and allow a successful merge as a result. This makes the data ready to be used for spatial analysis according the criteria for the different type of heat sources. The whole download and merging procedure occur automatically every time the BGT is downloaded into the model. Similar to the AHN, this extract procedure has also been published as a separate open source repository on GitHub as it can be used for purposes other than this implementation within this model (Stuurman, 2019b).

Figure 3.10: BGT map of the direct surroundings of De Bilt.

3.3 Air temperature Classification Tree

The SC for air temperature consists of three criteria that each need to be calculated separately: shadow angles, land use and vegetation height and are found in the Appendix (C) as an annex. Table 2.1 already provided a summary overview of the used criteria values in this research. The results of all individual analysis leads to a final classification score. For each criterion, there are 4 different classes, each defining the requirements it needs to meet. Class 1 is the best scoring Class, indicating a representative air temperature observation. Class 5 is an observation that is not representative at all and is assigned when one or more of the Class 4 criteria are not met. In the model, the individual Class for each criterion type is derived separately. The end result, however, is a final air temperature SC score for the whole site. This led to a decision tree how and in what order the classes are determined as shown in figure 3.11. The final SC score is determined by its highest value SC criteria found. The decision tree only provides an overview how a SC is determined according to the guidelines using which datasets and does not directly follow the flow of analysis of the model. The model analyses each criterion separately (in the order of shading, heat sources , vegetation height) and determines its final SC value by the highest found Class number. The end result is, however, the same: one final SC score.

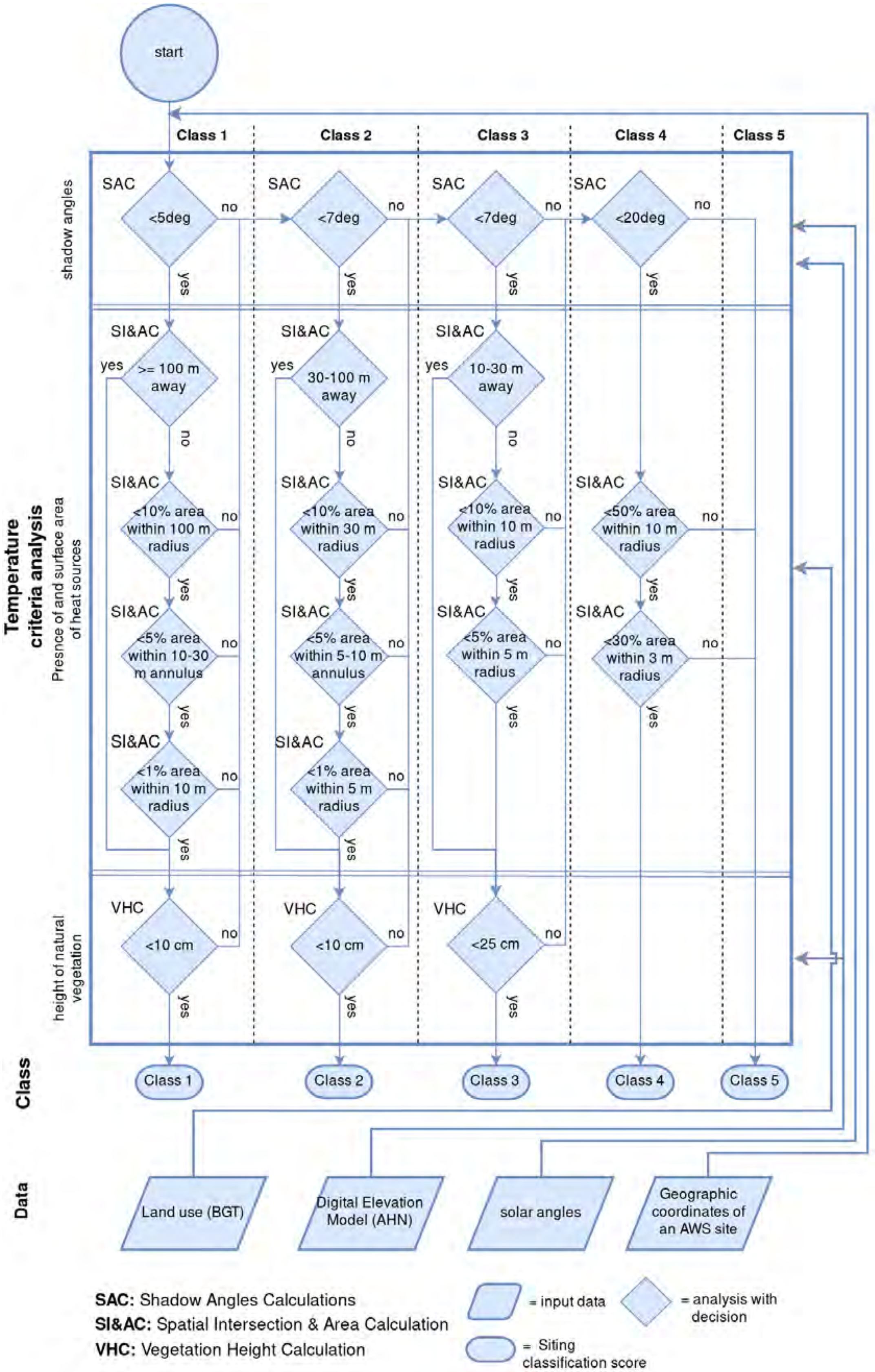


Figure 3.11: Decision tree to calculate the final SC value for the air temperature according to the WMO SC guidelines.

3.3.1 Projected shadings

Shadows have a significant influence on the measured air temperature as described in chapter 2.3.1 and are therefore included in the SC criteria of the air temperature variable. For the SC criteria, it is important to know if and under what solar elevation angle the air temperature sensor is shaded throughout the year. This is done in two steps: first, calculating the path of the sun to determine its elevation angles and second, the shadow angles that are caused by elevated (natural) objects. For each determined solar elevation angle, an associated shadow angle was derived as well.

Solar angles calculations

To calculate the solar and shadow angles of the different natural and artificial objects, the path of the sun around a certain geographic point needs to be known first. The solar angles include the azimuth and elevation. With a geographic location and a time stamp (date and time) the position of the sun (solar angles) can be calculated.

As the Earth makes a full spin around its own axis every 24 hours and around the sun every year, the solar angles are different at any point of time during the year. The air temperature SC guidelines do not make distinction between different seasons and therefore solar angles of different points of time during the whole year need to be known. To provide an overview of the solar angles throughout the whole year, it is best to create a sun chart. A sun chart plots different azimuth angles of one day on the x axis and their corresponding elevation angle on the y axis. This results in a curve that represents the path of the sun throughout one day. Combining the sun paths of several days throughout the year into one chart provides a year overview of the solar angles, as already shown in figure 2.6. The more solar angles used throughout the day, the more accurate the sun chart will be. It is common to plot the shortest and longest day of the year as well as one day for each month between these two days. In the model, the 21st day of every month was used from December throughout June. The months of July through November were skipped because these would give similar angles as January through May respectively.

For this research it is only relevant to consider the period between sunrise and sunset. Therefore, only positive elevation angles were used. A shadow angle can be derived from each determined solar angle. Due to the processing time of shadow angles, however, a limitation was made on the number of solar angles that are used throughout one day. The time resolution of the calculation has been set to 15 minutes. This means that the sun chart of one day is a plot of azimuth and elevation angles of 96 points in time (24 hours x 4/hour).

Shadow angles calculations

Shadow angles can be calculated based on the difference in detected altitude values (height) retrieved from the AHN DSM altitude dataset, and follow the formulas as described in 2.3.1. When the shadow angles are plotted together with the sun paths on a single chart, it can be identified when the location is shaded. It is estimated that objects that are located at a distance greater than 100 metres from an object, will not cause any shadow in 99% of the time (Dozier & Frew, 1990). The Class is determined if shadow angles are greater than their corresponding elevation angles of the sun at certain boundary angle values. When shadow angles are found between 5 and 7 degrees and they are higher than the solar elevation angle at the same azimuth angle, Class 2 or 3 is given according to the guidelines. Class 4 is given between 7 and 20 degrees. Above 20 degrees is associated to Class 5. Any shadow angle below 5 degrees is given Class 1 as its considered as landscape noise WMO (2017).

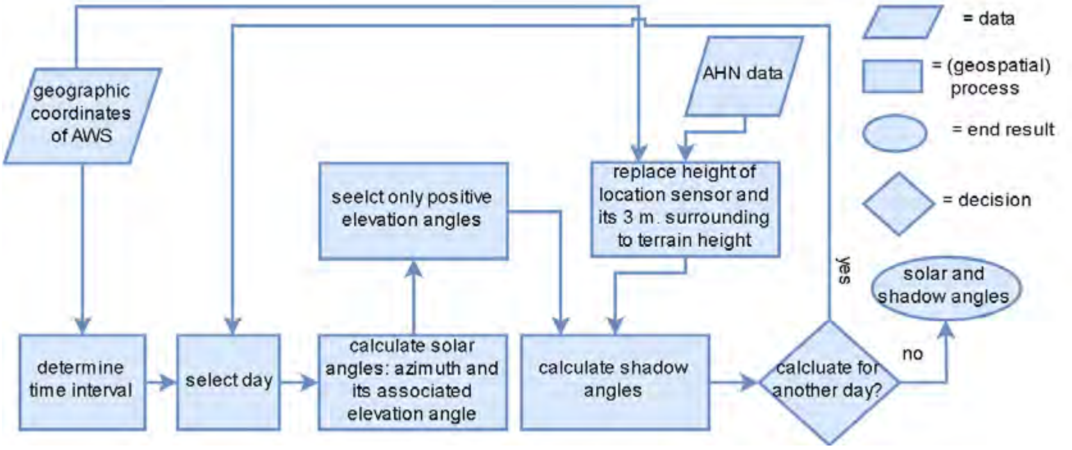


Figure 3.12: Methodology to calculate solar and shadow angles.

Due to possible inaccuracies of the location of the sensor, altitude data or with the error margins they bring along, the altitude at the location of the air temperature sensor has been adjusted so that all SC values at the different sites are calculated the same way. Not all locations detect the height of the sensor through the AHN altitude dataset. Therefore, the terrain altitude value is taken at the location of the sensor, to make sure the same type of value is used for each site. If the altitude of the sensor is detected by the AHN, however, inaccuracy of its location or the error margin of the AHN may also lead to wrong outcomes. This is because the altitude data of the sensor itself might be used in the shadow calculations. Therefore, also the surrounding 3 metres of the sensor is set to the minimum terrain altitude value found. A value of 3 metres was chosen as this is twice the width of the meteorological instrument and therefore ensures the whole sensor is brought along. The minimum terrain altitude value was selected to ensure it does not represent an altitude value that includes the height of an object, should the DTM have detected that by accident. The terrain altitude value was selected instead of an extra elevation of 1.5 metre (the height of the air temperature sensor) as the model produced unusable results when the latter approach was applied.

Calculating the shadow angles were done using the *horizonSearch* function from the Horizon package (Doninck, 2018). This function makes use of algorithms developed by Dozier and Frew (1990) to calculate slopes and horizons with a DEM as input. Determination of the shadow angles required pixel by pixel calculations and was therefore time consuming, especially if a distance up to 100 metres needed to be calculated with input pixels with a resolution of 0.5 metres. Therefore, two steps were implemented to allow a higher efficiency and less calculation time. First, only the shadow angles were calculated for each determined azimuth angle when the sun is above the horizon (elevation > 0 degrees). This is because shadows are only created on the surface when the sun is above the horizon. Second, only the pixels that are in line with the azimuth angles were used to calculate the shadow angles (see figure 3.13. This means that for each determined azimuth angle, the AHN dataset was masked to only the pixels that followed the line of the azimuth angle. The value calculated at the pixel where the location of the air temperature sensor was located, determined the shadow angle of that corresponding azimuth angle. The black pixels shown between the red dot (air temperature sensor) and green dot (anemometer) in figure 3.13 represents the shadow angle values created due to the latter. The yellow pixels indicate that there is no shadow angle and therefore no shade.

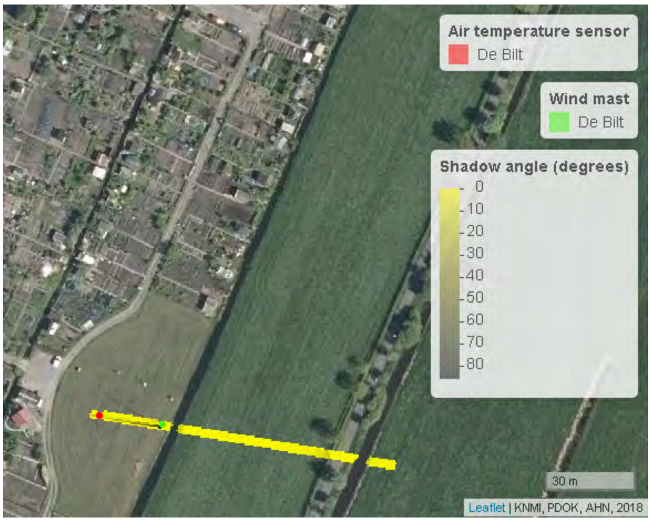


Figure 3.13: Shadow angle raster at an solar azimuth of 99 degrees. The shadow angle are represented by the black pixels between the temperature sensor (red dot) and the wind mast (green dot). The yellow represents that there is no shadow on the ground.

3.3.2 Presence and relative amount of land use

Different types of land use have an influence on the measured air temperature as described in chapter 2.2 and 2.3.2. These different types of land use are categorised as heat sources in the SC guidelines and include artificial or natural objects and surfaces. This includes buildings, roads, rivers, canals or lakes. The SC guidelines take heat sources within a certain distance from the air temperature sensor into account for the SC, as well as the relative amount of area covered by a specific heat source within a certain radius from the sensor. Consequently, the presence of a heat source alone does not mean it does not meet the criteria for a certain Class; this depends on a relative surface area the heat source covers that is defined in the guidelines.

To discover whether heat sources are within different distances from the centre point, different spatial intersections are done with the BGT dataset. At first an intersection is done with a circle around the centre point which has a distance radius defined for each Class by the SC guidelines. This gives an indication of the presence of heat sources within the set distance. In case heat sources are found through intersection, the relative amount of area that is occupied by the heat sources within the defined distance is calculated as a second step for that Class. Additional smaller circular buffers around the centre point are made after which new intersections are done. For each circular buffer, the relative amount of area is determined to see if it is smaller than the boundary values set by the SC criteria. For Class 1, if heat sources are detected, the relative amount of area cannot be larger than 10% within the 100-metre radius, 5% within the 30-100 metre annulus (area from 30 to 100 metre buffer area), or 1% within 10 metres from the centre point. These distances and the relative amount of surface area change across classes 2-4. Class 5 is met when the criteria set for Class 4 is not met. WMO (2017)

3.3.3 Vegetation height calculations

For air temperature, the height of the vegetation is also of importance, as also described in chapter 2.2 and 2.3.3. The height of the vegetation can be calculated by using two sub datasets of the AHN dataset: the DSM and the DTM. While the DSM determines the observed altitude, the DTM determines the ground level altitude. The documentation of the AHN2, mentioned that vegetation is not included in the the DTM and can be determined through the DSM (van der Zon, 2013). Subtracting the raster pixel values of the DTM from their corresponding DSM value determined the detected elevation at that moment of time when the AHN data was collected. This includes the elevation of buildings, or other objects but also vegetation. As a result, height of maize crops can be calculated. This relatively simple method is also applied by the National Institute for Public Health and the Environment in The Netherlands to determine many urban green indexes and is considered as a valid approach to calculate the height of the vegetation (Remme, Nijs, & Paulin, 2017). To avoid outliers - including the height of meteorological instruments - the median value of all the calculated height numbers that are in a radius of 10 metres around the air temperature sensor is gathered and used. The criteria guidelines do not specify within which area the height of the vegetation should be calculated. As a result, only the area which is 10 metres around the air temperature sensor was included in the analysis as this will have the most influence on the measurement. If the height of the vegetation is lower than 10 cm, the criteria is met for Class 1 and 2. Class 3 has a limit up to 25 cm. Class 4 does not have a criteria for the height of the vegetation WMO (2017).

3.4 Visual model and SC comparison validation methods

To test the model on its practical usability and performance, the results were analysed to see if the outcome produced coincide with reality and compare with the MSC values. This usability was divided into two different analyses: A visual model validation and a SC comparison validation analysis. Based on the outcomes from both of these analyses, the model was adjusted accordingly and conclusions on the performance f the model could be drafted. The aim of the visual validation analysis was to check if the model produced results that can be used for interpretations and corresponded to the real world, and the aim of the comparison validation analysis was to compare the results to manual classification methods, which were used as a reference.

The two different types of classifications may introduce uncertainties in their own form. The ASC may include input data is outdated or does not have sufficient information to be interpreted (correctly). The MSC is subject to human interpretations and may therefore introduce uncertainties to classifications due to different, but reasonable, judgements that are not necessarily specified in the SC guidelines.

3.4.1 Visual model validation method

After all the criteria were calculated through the model, the outputs were analysed to see if they produced results that can be interpreted and used. During the development of the model, the AWS site at De Bilt was always used as test case. Comparing the model outputs to (historic) satellite, photos and own observations at the AWS site of De Bilt gave a good indication of whether the data pre-processing went correctly. The shading charts were analysed by comparing the output generated from the AHN2 to the ones from the AHN3 to understand what caused any discrepancies in shadow angles and thus test its accuracy as a result. These shading angles were analysed by comparing historic satellite imagery or photos to see if there were any changes in artificial objects that caused height differences. For artificial objects, the BGT dataset was compared to satellite imagery and the reality to see if the dataset gave accurate information. After the whole model was tested for De Bilt, the model was run for the other sites to see if it still produced the same types of (expected) outputs when comparing to (historic) satellite imagery, photos and knowledge of the sites from the field inspectors. During the whole procedure the model was adjusted and optimised where needed. After analysing all the sites, conclusions were drawn whether and to extend to which the model produced accurate results.

3.4.2 Comparison validation with MSC as reference

After the visual model validation analysis was performed, the ASC values were compared to the manual classification that were derived through the MSC procedure as described in chapter 3.1 through a comparison validation. MSC were obtained from the inspection reports of the sites to understand why the MSC scores are given together with the input and knowledge from the field inspectors. For this it was important to select the inspection report that was closest to the date of collection of the AHN. Comparing a manual classification score from 2018 (the most recent inspection report) with an ASC score generated from an AHN dataset that was collected in 2012, could give a false comparison. This is because in between 2012 and 2018 trees might have grown significantly with a corresponding effect on the shadow angles impacting the AWS. This means, however, that for some the sites where no AHN3 is available yet a certain reservation needs to be made. The AHN2 data that was collected for these sites were between 2007 and 2012 and this was before a field inspection report included SC values. MSC were only performed as of 2013. As a result, no inspection report was available that includes a SC value and overlaps with the AHN collection year of that corresponding area. Therefore, the first field inspection report was selected in which a MSC took place (2013, 2014 or 2015) that was closest date to the AHN2 collection date. Manual checks were performed to see if these reports coincided with the report of the year of the AHN3 collection date to take into account of any major changes at or around the site. For all the other sites, the inspection report was used from the same year as the AHN3 dataset was created.

AHN altitude data was always collected between January and March while inspection reports usually take place between April and October. As a result, these inspection reports include the AHN altitude observations of that same year. For the BGT this return in time comparison was not possible as this dataset only provides the most up to date information. The BGT dataset is, however, less critical because the heat sources around a site will probably change less frequently. Based on the findings from the validations of the automated model, conclusions were made if and why the the automated and MSC scores showed the same values or why they were different. A match in SC classes was also considered when automated values fell within the range of 2 manual values. This is done because the MSC is subject to human interpretation and therefore can introduce an uncertainty on the SC.

4 Results

This chapter describes the results obtained from the computerised SC model and how they compare to the manual values. These results will help to determine the usability of the developed SC model. Each part that makes up the air temperature SC is analysed to see if the model produced any output that is meaningful to be interpreted and compared. At first, the model was tested on its accuracy through a visual model validation by analysing the different results for the AWS at De Bilt (WSI: 0-20000-0-06260) and described in chapter 4.1. Secondly, an overview is provided in chapter 4.2 of all the AWS sites, indicating how the outcomes are distributed over the 5 classes. Thirdly, in chapter 4.3, the visual validation was also applied to the other sites, describing trends how the SC Classes were obtained. Lastly, all the automated results are compared to the current manual classification values to analyse the similarities and differences found through a comparison validation and described in chapter 4.4.

4.1 Visual model validation of De Bilt

The model was initially developed based on De Bilt AWS as primary test site. The results obtained from this AWS site were analysed in detail to understand what the different parts of the model do. The outputs were compared together with KNMI inspection staff to satellite imagery, and to own observations done at the AWS site itself to discover if the results corresponded with the real world and if they were interpretable.

4.1.1 Presence of shading

Figure 4.1(a) and 4.1(b) show the calculated sun path with its corresponding shadow angles of the AWS at De Bilt derived from both the AHN2 (situation 2008) and AHN3 (situation 2014). Both charts show the shadow angle patterns that were expected based on satellite imagery and the real world. Furthermore, the high accuracy of the shadow angles is confirmed by the peak observed at an azimuth of 100 degrees and around 150 degrees. At 100 degrees is the location of the wind mast that measures wind speed and direction. At 150 degrees the differences in shadow angles was due to the growth of one tree. This tree grew more than two metres. This was derived from the differences in shadow angles the AHN datasets detected.

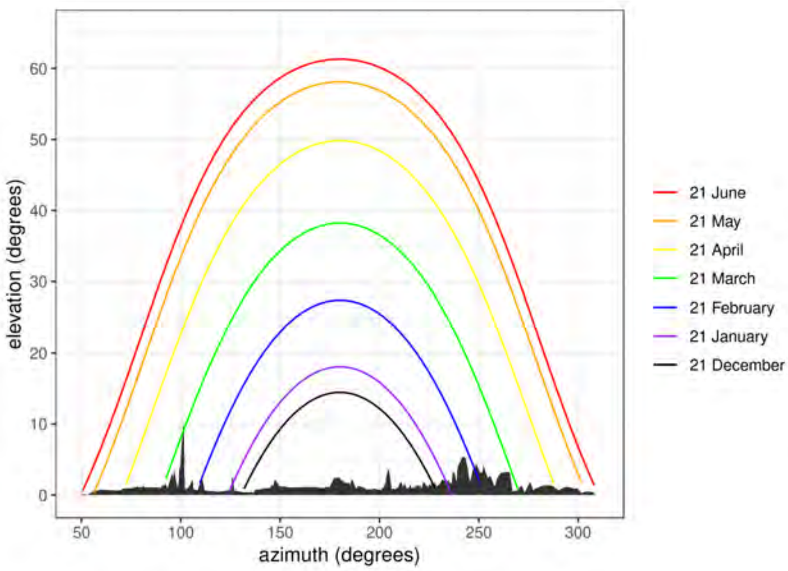
Although the mast carrying the wind sensor is a relative tall object (10 metres), with a width of only 0.5 metres it is also a very narrow object that may not necessarily be detected in the DSM. The detected height of 6.4 metres did not correspond, however, with the actual height of the wind mast. Furthermore, this instrument is 20 metres away from the air temperature sensor, which would result in a shadow angle of 26.5 degrees (as show in figure 3.2) if a wind mast height of 10 metres was taken into account. Instead, the AHN2 and AHN3 both showed a shadow angle of only 10 degrees, also resulting in an overall lower classification than with an expected angle of 26.6 degrees. The difference between what is expected from the real world and the results from the AHN2 or AHN3 can be caused due to a combination of any of the following errors: inaccuracy of the detected altitude, inaccuracy of the exact locations of the air temperature sensor and wind mast, due to resolution of the used datasets that is too low, or due to the used time increments of 15 minutes to determine the shadow angles.

These errors may have been introduced due to the error margins on the datasets or because the datasets are not accurate enough. The detected altitude by the AHN may not represent the actual height of the instruments as observed in the real world (after subtraction from the terrain altitude value). The error margin on the geographic coordinates of the meteorological instruments might determine shadow angles which are slightly off the location where needed. The resolution of the shadow footprint - which is the same as that of the AHN - may cause inaccuracies in the exact pixels that need to be considered as shaded. As a result, the pixel values (shadow angles) may represent numbers that are smaller than expected. The time increments of 15 minutes may cause the temperatures sensor not to be included at the selected points of time. This is also due to the resolution of the shadow footprint. This is especially caused by narrow objects, such as the wind mast. These errors responsible for the observed differences in shadow angles indicates a manual validation is needed to check if the outputs are correct.

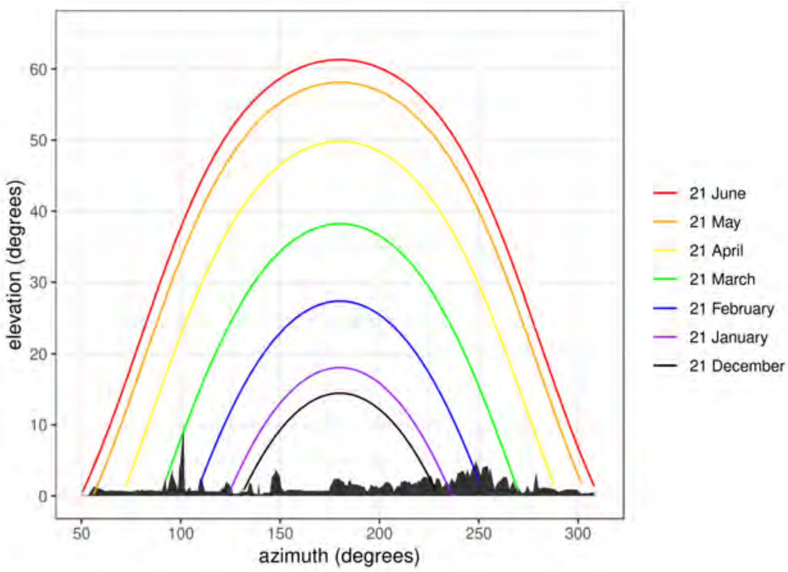
Table 4.1 provides a summary of how many values were found in which Class for the shading criteria for both the AHN2 and AHN3 when using a time interval of 15 minutes between sunrise and sunset for the selected seven days of the year. Overall the shading criteria for De Bilt got a Class of 5 for both charts because for only a few points in time the shadow angle (elevation in the chart) were above 20 degrees. For both the AHN2 and AHN3, the greater majority of the shadow angle are less than 5 degrees and would correspond to Class 1.

Table 4.1: Frequency of Class values observed with AHN2 and AHN3 dataset for De Bilt.

AHN	Shadow angle criteria Class count				
	Class 1	Class 2	Class 3	Class 4	Class 5
AHN2	336	0	0	1	0
AHN3	336	0	0	1	0



(a) AHN2 (2008)



(b) AHN3 (2014)

Figure 4.1: Automated sun paths (curved lines) and shading (black area) of 7 days throughout the year at the AWS of De Bilt based on the AHN2 (top) and AHN3 (bottom). Each day consist of increments of 15 minutes and only include the shadow angles between sunrise and sunset.

4.1.2 Presence of heat source

The BGT dataset was used to detect the land use type and any other artificial objects that are considered as heat sources according to the SC guidelines. The location of the air temperature sensor at De Bilt satisfied the criteria set for Class 1. Figure 4.2 shows the direct surroundings of the air temperature sensor with a radius of 100 metre. Although heat sources were found within a radius of 100 metres (water on the east and a building on the west side), their surface areas were sufficiently small to comply with the Class 1 criteria, as shown in table 4.2. When looking at the map carefully, however, it is observed that the provided information by the BGT does not coincide with the satellite imagery shown behind the data. Especially on the west side (left) of the sensor, the BGT data indicates there is only barren ground, while in the real world there are many garden sheds present that should be interpreted as heat sources. When looking at the provided attributes in the BGT, it specified that this barren ground was seen as a yard. The BGT does not specify more details about years. Although yards usually include a piece of land, it may also include buildings or other heat objects. This piece of 'barren' land use is therefore not categorised correctly or specified well enough to provide accurate information that is useful for the SC model. Furthermore, a paved road is missing in the BGT (near the shown building) that in the real world is present and goes around the west side of the AWS. The data has not been changed in this research in a way that it would better corresponds to the real world situation. Therefore, it remains questionable if Class 1 is a correct representation because in reality the area includes many more heat sources (garden sheds and road) within 100 metres around the sensor. The advantage of the BGT is that it is a very detailed land use and cover map, but this carries the risk of including errors and quickly becoming outdated.

It was important to choose the exact location of the air temperature sensor, and not the general coordinates of the site (the location of wind mast) because results may change rapidly with the high resolution the BGT dataset provides. When the location of the wind mast was selected, the nearby water bodies resulted in a classification score of 4 instead of 1.

Table 4.2: Relative surface areas within the different radii and annulus in the direct surroundings of De Bilt.

Objects within 100 m	**Relative surface area within**						Final Class
	100 m		10-30 m		10 m		
	measured	criteria	measured	criteria	measured	criteria	
7	5.77%	10.0%	3.7%	5.0%	0%	1.0%	1

Figure 4.2: Automatically determining the presence and surface area of heat sources in the direct surroundings of the air temperature sensor at De Bilt using the BGT dataset. According to the data, the area conforms to the criteria of Class 1. The satellite image indicates, however there are garden sheds present that should be seen as heat sources.

4.1.3 Detected height of vegetation

For De Bilt, no structural vegetation height differences were detected within a radius of 10 metres around the air temperature sensor (see table 4.3) when looking at its median value. This was done by subtracting the DTM from the DSM raster datasets. This was valid for both the AHN2 and AHN3 datasets. This means that it conforms to the vegetation criteria of Class 1 (< 0.1 metres). This is in accordance with SC guidelines to install the air temperature sensor in grassland that is maintained properly (WMO, 2017). The fact that a height greater than zero (see figure 4.3) was detected indicates, however, that altitude differences are detected between the DTM and the DSM. The location where this difference in height was detected coincided with the location of the air temperature sensor, confirming that small and narrow sensors (objects) can be detected by the DSM but not included in the DTM.

Table 4.3: Summary statistics of height difference values derived from the AHN2 and AHN3 for 10 metres around the air temperatures sensor located at De Bilt.

AHN	min height difference	1st quarter	median	3rd quarter	max height difference
AHN2	0	0	0	0	0.690
AHN3	0	0	0	0	0.652

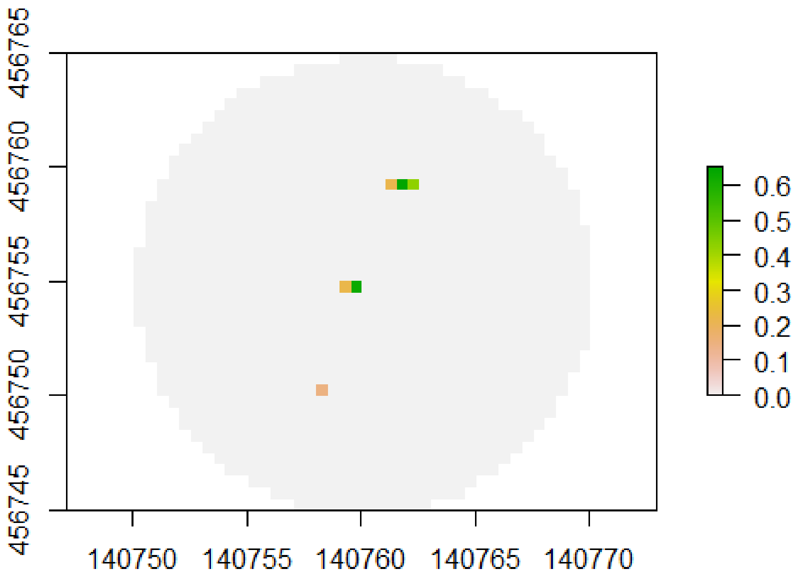


Figure 4.3: Detected height differences when the DTM is subtracted from the DSM raster within a radius 10 metres around the air temperature sensor. Detected height differences were the air temperature sensor (middle) and other meteorological instruments right above and below.

4.2 Determinants of SC of all AWS sites

After the model was applied at the AWS of De Bilt, the model was used to determine the SC of all the other AWS sites to see if it would give any outcomes that could cause the same type of accuracy results or deviations. Applying the ASC model to all the 34 sites showed that a majority of the sites were assigned to Class 4 (67.6%, n=23). There were no sites that retrieved a Class of 3: at none of the sites were there any shadow angles between 5 and 7, *and* no heat sources in the annulus of 10-30 metres around the AWS. The smallest represented was Class 2, with a relative amount of 5.9% (n=2). Class 5, a score that shows the meteorological observation is not representative, was given to 11.8% (n=4) of the sites. The SC values and its determinant criteria are summarised in figure 4.4.

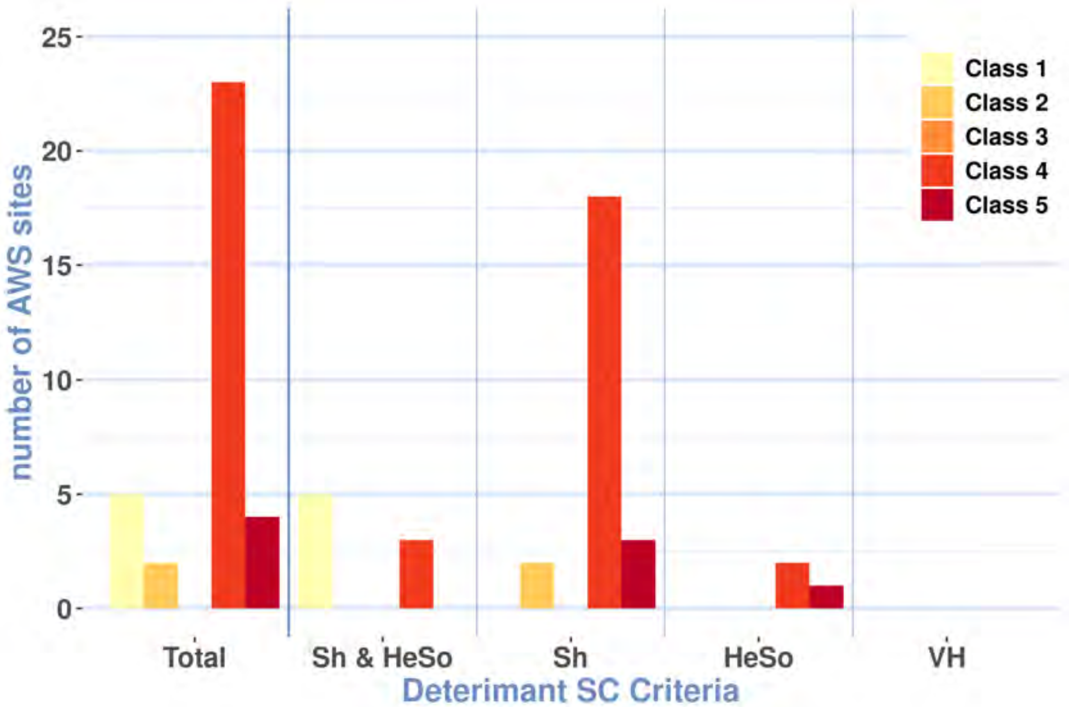


Figure 4.4: Class distribution of all the 34 AWS sites categorised by determinant SC criteria that caused the final SC value. The Total category (left) represents the total amount of sites that fall within each final SC Class. The Shadow (shadow) & HeSo (heat sources) represents the class where an equal class was found for the two criteria. No classes were found that were determined by the vegetation height (VH).

For 67.6% (n=23) of the sites, the shading angles was responsible for the highest Class value, while only in 8.8% (N=34) of the cases, the heat sources criteria was determining. For the remaining 23.5% (n=8) of the sites, the shadow angle and heat sources criteria showed the same score. Among the 23 sites that retrieved a Class 4, 17.4% of the sites (n=4), retrieved this Class due to the heat sources. In all the other Class 4 cases (82.6%, n=19), this was caused due to shading. All the sites obtained a Class 1 for the vegetation, meaning there was no vegetation detected that was higher than 10 centimetres.

The shadow angles were determined for seven days during the year with an time increment of 15 minutes as explained in chapter 3.3.1. This led to a total number of shadow angles that varied between 337 and 338 per site. One site has one more angle than the others due to the difference in the latitude, causing minimal changes between the north and south of The Netherlands at what time sun rises and sets. Each calculated shadow angle can be categorised into Class 1 through 5. Therefore a distribution can be made how many shadow angles for within each Class for each site, as also shown in table 4.1 for the AWS in De Bilt. The results showed that an increase in SC value for the shading criteria, was caused due to a small number of angles that did not meet the criteria for Class 1. The lowest percentage of shadow angles at one site that met Class 1, was 86.1% (N=337, n=290). This means that the air temperature sensor was only shaded due to shadow angles that were greater than 5 degrees during only 13.8% (n=47) time increments of 15 minutes. This is in total 705 minutes (11 hours and 45 minutes) throughout all the 7 selected days. On average this is only 1 hour and 40 minutes a day. This site was the AWS in Arcen and is known to be one of the sites that observes air temperature values that are among the least representative according to the SC guidelines due to surrounding of many heat sources and trees. The ASC therefore correspond with what would be expected in the real world.

Among all the AWS, an average of 97.3% of the shadow angles within one site fell in Class 1, ranging from 86.1% to 100.0%. With a low standard deviation (3.7%), this showed that the number of shadow angles that did not meet Class 1 criteria is relatively small within all the sites, if they were present at all. Only 7 sites (20.9%, N=34) had no shadow angles that were responsible for shading criteria that was higher than Class 1. The few angles per AWS that were responsible for a higher Class, were often due to meteorological instruments at the site that caused the shading. In the worse case among all the studied AWS, a shading criteria class of 4 was given only because the sensor was shaded for an average of 1 hour and 40 minutes through the 7 selected days. This showed that the shading criteria is sensitive for a change in Class and further substantiates the research done by M. Wolff et al. (2014) and Fisler et al. (2017) that this criteria needs to be revised so that a very few shadow angles a day or year are not responsible for a higher overall SC class.

The relative high number of sites whose final SC Class was determined by the shadow angle (67.6%, n=23) showed that the shading criteria was the factor that determined most final SC values. This can indicate that the sites take the influence of heat sources better into account than the influence of shadows. This assumes, however, that all the AHN and BGT data is accurate and correctly determined by the model. Therefore, the results need to be examined in detail to understand how the different classifications were determined. This was done through visual model validations for all the sites.

4.3 Visual model validation trends

An important aspect that determined an ASC value was the accuracy of the used datasets, as also shown for De Bilt. This section describes any observed trends how certain SC were determined throughout all the AWS sites. Examples of individual sites are used to explain the different Class determination.

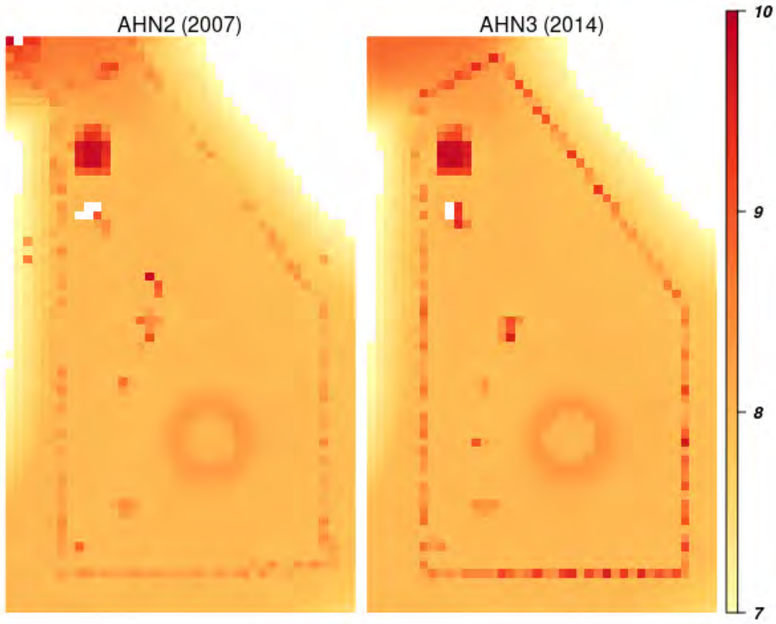
4.3.1 AHN2 and AHN3: small scale height objects

The AHN can detect very narrow objects; as seen at the site of De Bilt where the wind mast was detected. The successful detection of existing AWS infrastructure (sensors, masts, cabinets etc.) has been confirmed at many other sites. The SC value for the shading criteria was often increased because meteorological instruments - usually the wind mast with a height of 10 metres - were responsible for shading the air temperature sensor at one or more moments in time during the year. The detection of height differences did not always result in the correct shadow angle at the location of the air temperature sensor. This was, as also shown at the AWS in De Bilt - due to the inaccuracy the measured altitude of the air temperature sensor, the location of the weather instrument, or because of the resolution of the shadow footprint that determined the shadow angle.

Another object that often was detected were the fences that are around an AWS site to demarcate the site. These fences vary in height and type per site, but never completely block the sun. They either have holes in them, consists of vertical bars or only have one beam that is supported by many poles. The AHN datasets, however, detected these objects as an increase in altitude (compared to the terrain altitude) and therefore as something that would cause shading. Many air temperature sensors are located close to an edge of an AWS and therefore would be 'shaded' by the fences. A good example is the site of Vlissingen where a metal fence was replaced by one with vertical bars (see figure 4.5(a) and 4.5(b)). This was done in 2008, one year after the data for AHN2 was collected. As a result the AHN3 dataset of 2014 showed an increase in height of fences, and thereby also the shadow angles (figure 4.5(c)). The detection of fences was for 5 sites (partially) responsible for an increase in the shading criteria Class, especially in areas where no other major other shadow angles were detected.

(a) 2007

(b) 2014



(c) Raster maps of the AWS site in Vlissingen

Figure 4.5: AWS site of Vlissingen in 2007 (left) compared to 2014 (right) shows a change in fences both on the photos (a, b) and the raster maps (c). The change in pixel colour in the raster diagrams show an increase in detected altitude. From this the fences can clearly be determined and seen they have raised in height. *Source:* KNMI

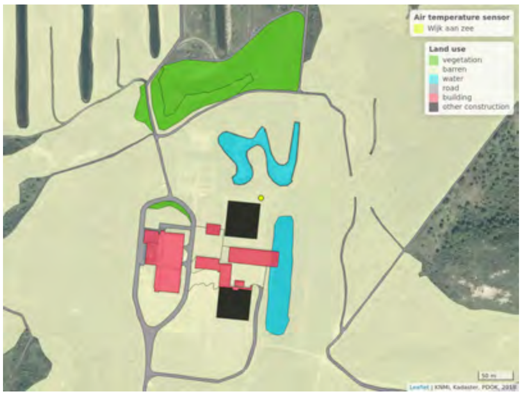
4.3.2 Differences between the AHN2 and AHN3

At many sites a small increase in shadow angles was detected at specific azimuth angles when the AHN2 was compared to the AHN3. Through photos and satellite imagery, it was identified that these increases were caused by to trees that had grown between the time data was collected for the two datasets. This confirms that having the most up to date information is crucial for an accurate current SC, especially for trees which can grow significantly within one year. For the site of Rotterdam (WSI: 0-20000-0-06344) the growth of trees was responsible for the ASC to increase from Class one to Class 2 when comparing the AHN2 to the AHN3.

4.3.3 BGT: variation in level of detail

The presence and Class/type of objects based on the BGT, often corresponded with the real world. Some sites, however, included wrong/outdated or no information at all, as previously shown at the site of De Bilt.

For the site at Wijk aan Zee (WSI: 0-20000-0-06257; figure 4.6), water was shown at the north side of the site in the BGT data set, while from satellite imagery and photos this was not observed. The field inspectors indicated that this area was only filled with water after periods with dense rainfall. It is therefore unclear why the BGT data set marked this area as a water.



(a) BGT map (2018)

(b) Satellite image (2018)

Figure 4.6: BGT map and satellite image of the AWS site in Wijk aan Zee in 2018. A water body was detected north of the temperature sensor (yellow pont) but not shown in the real world.

It is tricky to determine the influence of heat sources on sites located near water. The site of Vlissingen (WSI: 0-20000-0-06310) is located near the sea and, as expected, water in a radius of 100 metres of the site was identified. As a result, the model saw these water bodies as a heat source because no distinction was made in the model between the different types of heat sources. In the real world, however, this water can be considered as a 'representative element' of the area and therefore not determined as a heat source (WMO, 2017). The model did not take this 'representative element' into consideration which would filter out these water bodies as heat sources and produce results that would follow the SC guidelines even better.

The level of detail shown in the BGT varied per site. Most sites were located on pieces of land that were considered as vegetation (grasslands), as expected. Sometimes, however, the data also included the location of the meteorological instruments to be seen as a separate spatial feature. This was sometimes at the location of the rain gauge (circular area consisting of bricks) or location where the electricity and hardware was installed that controls and sends the collected data from the different sensors. This feature was often barren ground but that is not seen as a heat source by the model. Because the specifications of the data did not always mention it was something artificial (such as pavement), these objects were not seen as heat sources. In reality, however, one could see these artificial objects as potential heat sources. The field inspectors indicated they do not consider the meteorological instruments as heat sources, unless if they are very large and cause a lot of shading. This varies per AWS and is also dependent on the site specific surroundings and circumstances.

It varied per site if these 'barren' locations had specification values in the attribute data. Therefore, they were not included at all in the model to maintain consistency throughout the SC. A better specification of what to include in the SC process as obstacles and heat sources (including the AWS infrastructure itself such as fences, buildings, cabinets, masts), is needed.

In other two cases, such as Gilze-Rijen (WSI: 0-20000-0-06350) and Leeuwarden (WSI: 0-20000-0-06270), the path going from the main road to the AWS in the middle of a piece of land was detected as a 'road' (Figure 4.7. The specification shown in the attributes of the BGT indicated this was a footpath. The footpath at these sites is often made of sand, and paths made of sand should not be considered as heat sources. However, the BGT does not specify if the path is paved or is made of sand. As a result, these sites got a high classification because too much area within the inner radius around the temperature sensor was detected as heat source(s). Because these paths are made of sand in the reality, these do not need to be determined as heat sources in the real world.

(a) BGT map of the AWS in Leeuwarden (b) Satellite image of AWS in Leeuwarden

Figure 4.7: The path towards the AWS was detected as a road and therefore a higher air temperature SC score was given. In the real world these paths are not present (anymore) or are made of sand and therefore do not need to be considered as heat sources.

4.3.4 Detection of vegetation height

No vegetation height was detected at any AWS site where more than half of the observed height were higher than 0 metres. This was also the case when looking at the 25th percentile height value. Only a few raster pixels at each site showed altitude differences between the DTM and DSM. Sometimes this was because of nearby meteorological instruments. As a result, no vegetation height was detected in the AHN2 and AHN3 datasets and therefore all sites gained Class 1 for the vegetation criteria.

4.4 Comparison validation: ASC vs. MSC

This section describes how the ASC values are compared to the MSC values. The MSC results are taken from the inspection reports of the year that matches best with the collection of the AHN dataset used. An overview of the ASC vs. MSC values is provided in table 4.4. This section describes why different SC values were found and whether identical/matching results found for automated and MSC were properly determined for the same reason. When a MSC had a range of two values, the value closest to the ASC was used to determine the difference in Classes.

Table 4.4: Distribution f ASC and MSC values. In total there were 19 AWs sites that had matching ASC and MSC values.

		MSC					**Total count**
		1	**2**	**3**	**4**	**5**	
ASC	**1**	4	0	1	0	0	5
	2	0	2	0	0	0	2
	3	0	0	0	0	0	0
	4	5	5	1	12	0	23
	5	0	0	2	1	1	4
Total count		9	7	4	13	1	**34**

4.4.1 Different ASC and MSC values

There was a mismatch between the ASC Class and the MSC Class for 44.1% (n=15) of all the AWS sites (N=34). For only one site (Twenthe; WSI: 0-20000-0-06290) the ASC Class was lower than the MSC scores (Class 1 vs. Class 3-4). The field inspection report indicated that the wind mast was responsible for a SC of Class 3-4, whilst this pole to which the wind sensor is attached was not detected in the AHN. This may be due to some fidelity issues when the AHN2 data was collected for this site. A new analysis should of this station should be performed when the AHN3 data becomes available for this AWS in March 2020. For all the other mismatches, the ASC Class was higher than the MSC Class. This shows that the ASC applied a stricter SC than the MSC. The reason for these stricter SC can be divided into three categories: shading of meteorological instruments, shading due to fences, too much detailed information provided by the BGT.

In 20.6% (n=7) of the AWS, a mismatch was found due to the detection of objects that caused shading by the automated procedure but not written down as an influence in field report for the MSC. In one case (2.9%) this led to a Class difference of 3 (Schiphol; WSI: 0-20000-0-06240). In 14.7% (n=5) of the situations, a mismatch was found because of the detection of fences in the automated classifications. In three cases (8.9%) this was even the cause for a difference of 3 classes. The last case where a Class difference of 3 was found, was because of too many spatial features were detected by the BGT (as shown in figure 4.7).

4.4.2 Same ASC and MSC values

For 55.9% (n=19) of all the sites, the ASC Class was the same as the MSC Class or fell within the range of MSC values. A Class 1 match was found for 21.5% (n=7) of the sites because in neither SC methods, any significant shading or heat sources were detected or observed. In the remaining 44.1% (n=15) matching cases, both methods indicated shading and/or heat sources as the cause for higher Class values. In 8.9% (n=3) of the situations, however, an additional note was made about vegetation in the MSC: either this was too high or not present at all (no grassland). Overall this means that the cases that did match, corresponded for the same reason.

5 Discussion

In this research it was shown that it is possible to perform an ASC for air temperature at AWS sites using geospatial datasets. The automated model was developed based on the AWS at De Bilt, and later applied to all the other 33 AWS in the Netherlands. The model used the AHN dataset for the detection of altitude needed for shading as well as vegetation, and the BGT dataset for the determination heat sources. The results obtained from the model were compared to those obtained through MSC and they matched for 55.9% (n=19) of the sites. This chapter discusses what the use and accuracy of the dataset meant for the usability of the model to be applied at any location on land in The Netherlands. The chapter starts off with a reflection of the computational performance of the model in terms of time it took for the model to run.

5.1 Computational performance of the model

An important goal of the ASC model was to evaluate its speed compared to a MSC procedure. Calculating an ASC for one site is much dependent by the computing power available. During this research calculating the SC for one AWS site took around 50 minutes. Therefore, calculating ASC values for all the sites took just over 2 full days (50 hours). The model started with determining the location of the site. This was done by using geographic coordinates as input. This means applying the model to any other list of other locations on land in The Netherlands is easy to achieve. In the next step the required AHN and BGT datasets of the surroundings were retrieved automatically. This step took took a few minutes for each dataset. The disadvantage of automatically retrieving the required datasets is that the model is dependent on the access and availability of these online resources. The deduction of shadow angles took the longest during the whole procedure. In this research no surroundings were taken into account that were further than 100 metres way from the air temperature sensor. As a result, calculating of the shadow angles took around 45 minutes per site. This computing time would be increased if the radius distance from the sensor was increased. This is because the shadow angles of more pixels would need to be calculated. Determining the heat sources and height of the vegetation were calculated instantly. The model automatically produced different output tables, charts, and maps of all the information required to perform a visual and comparison validation analyses. These analyses required expertise and human interpretations to understand the results of the model. This was time consuming because each site has its own characteristics for which the automated and manual results needed to be interpreted. Manual checks were needed to see if the AHN and BGT data corresponded with the real world and the surroundings described in the field inspection reports.

5.2 Accuracy of the AHN

The high resolution and accuracy of the AHN was the primary reason why this dataset was used in the model but this had both positive and negative consequences for the outputs as a result. The high accuracy was confirmed when the AHN2 was compared to the AHN3, showing that changes in heights of trees were found and correctly observed. The high resolution of the AHN dataset also showed it has the possibility to detect the height of meteorological instruments present at the site. This was important because the shadow angles that are caused by these sensors are often considered during the manual classification and often even caused an increase in classification score when the WMO guidelines are strictly followed. The measured height of the sensor did not always correspond, however, to the actual height. This caused errors in the size of the shadow angles. This was either due to the fact the actual height was not properly detected by the AHN or because of the error margin from the derived shadow angles due to the resolution of the shadow footprint. Furthermore, detection of the height of the surrounding sensors did not always occur and varied per site. This creates an inconsistency in the AHN dataset and thus an influence when the model is applied to the different sites.

The high resolution and accuracy also caused some shadow angles that are not (completely) considered in the real world. This occurred at sites where the fence around an AWS was detected as a increase in altitude, and thus an object causing complete shadow. In reality, these fences have holes in them and do therefore not cause complete shadows. Field inspectors indicated that fences can have a little influence on the observed measurements, but they are not always written down as heat sources in the field inspection reports because they do not cause complete shadows. The high resolution and accuracy of the AHN showed that it is feasible to calculate shadow angles, but for each site it needs to be checked if the outcomes make sense due to the input data. A possible solution would be the use of additional datasets that has the demarcated areas of the AWS so that they can filtered out as height data. The BGT dataset would be able to do this at the locations where the area of the AWS is seen as a separate spatial feature. Therefore combining the AHN and BGT could possibly improve the outputs of the model to better coincide with the real world.

Using the DSM and DTM datasets to detect the height of the vegetation did not lead any results that were considered inaccurate. This was because most surrounding areas of sites showed a median height value of 0, resulting in a classification score of 1 for this criterion. Although this may be because all the researched sites are located at maintained grasslands (like most others), it is uncertain whether the used approach to detect the height of the vegetation can be considered as a valid. Calculating this criterion was included in the model as it is mentioned in the WMO SC guidelines but because vegetation height may change on a weekly basis, using outdated geospatial datasets is not the best approach to perform this particular calculation.

5.3 Accuracy of the BGT

The high resolution of the BGT was also the reason why this data was used for this research. The high resolution was confirmed with the detection of the smallest roads and objects. As a result, this led to very accurate results when checking for the presence and surface area of heat sources. In some cases, this was too accurate resulting in higher classification values due to heat sources that are not even considered by the field inspectors. This was especially true when the paths or roads towards and inside the AWS were shown in the BGT dataset.

The level of accuracy of the BGT did, however, vary per site and thus created an inconsistency. At some AWS sites, data was missing where expected or it was because of outdated information. Although this dataset has many standardised data collection methods, it may sometimes lack the detail desired for this SC model. This is especially the case when the specifications of the land use type need to be known in order to have information that is valuable for the model. This level of detail is difficult to standardise as this dataset is maintained by all the individual municipalities in The Netherlands. The BGT has the ability to receive feedback from any user to indicate incorrect or missing data (Kadaster, 2019). As a result, the model could be improved if feedback was send specifying what information needs to be changed and thus increase the accuracy of the ASC outputs

Applying the model to the other sites has shown that the datasets can be used to perform the calculations that are required for a SC. The results showed, however, that the AHN and BGT datasets may have wrong, outdated, misleading or no information at all about the altitude or land use. This can have a significant influence on determining an accurate classification that is valid today. Therefore, a manual validation of the outcome will always be required in order to determine if the results are realistic to be used.

5.4 Model performance compared with MSC

The accuracy analysis showed that usable classification values can be obtained from the ASC model that also compare for 55.9% of the cases with the MSC. An important aspect of the MSC is the interpretation done by a field inspector. The reason that led to a MSC value may therefore be different than the automated classification, even though they show the same score. The ASC produced results that can be deducted from the model, while the MSC was dependent on the field inspection reports and the information gained from different field inspectors. Simply comparing the results does not provide an explanation as to why the output are the same or different. Expertise of both the model and performing a field inspection is therefore always required to compare the two results.

Using the field inspection reports that were of the year when the AHN of that area was collected were considered as accurate comparisons. Otherwise height values of one year are compared to those of another year, leading to misleading and possibly even wrong comparisons. This is especially the case for the growth of trees or change in building landscape that cause more/less shading. A possible solution to make AHN data more up to date so that it could correspond better with the height of the surroundings is by introducing an algorithm that automatically detects the trees and adjust their height according to a growth rate.

Ideally, comparing BGT data from the same years should also be done, but this dataset did not allow to do that and therefore only included the most recent available information. Comparing land use from the same year was considered less critical because the 100 metre direct surroundings of a site is not expected to change very quickly. The use of data from different years in the ASC vs. the MSC procedures may have led to a skewed comparison that would not have occurred if all the information came from around the same time. It is therefore important that it is exactly known from what time the data in the ASC comes from to produce outputs that represent or represented the the real world.

Both types of classifications required a form of interpretation to be made and are therefore time consuming in their own way. Efforts have been made so that the model produced outputs that can immediately be interpreted and compared with the MSC. This was done by producing charts, tables and maps that summarise and explain the results. The comparisons showed, however, that if the datasets are accurate, usable and valuable SC can be calculated automatically using a consistent approach. The model has also shown that the ASC will usually generate a stricter classification than the MSC, indicating that - if the datasets are correct - the automated method provides results that have a higher chance of complying with the guidelines to the criteria of that Class.

6 Conclusions and Recommendations

6.1 Conclusions

The primary aim of this research was to develop a ASC model that determined the SC of air temperature at AWS according to the guidelines set by the WMO. This resulted in the following research question:

Is it possible to automate the Siting Classification for the air temperature criteria according to the WMO guidelines using geospatial datasets and apply it to any locationb on land in The Netherlands?

The SC guidelines for air temperature exist of three different criteria based on shading, the presence of heat sources and the height of vegetation. Each criterion has its own justification for being included in the guidelines. Locations that are shaded have a different temperature than areas that are in the sun. Heat sources - including water bodies, buildings, roads or other artificial objects - can reflect or retain heat and therefore influence the air temperature. Vegetation is able to cause cooling due to its evapotranspiration process and therefore also have an influence on the observed air temperature.

An ASC model was developed that included the shading, heat sources and vegetation criteria. This was done by using the AHN and BGT spatial datasets, with resolutions high enough to meet the different air temperature SC criteria. A decision tree was developed to determine the SC of each criteria as well as an overall SC. The geospatial datasets used are available for the whole of the Netherlands, which made it possible to develop a dynamic model that automatically imported the data of the requested areas. As a result, an automated air temperature SC was determined for all 34 AWS sites owned and maintained by the KNMI.

The results showed that the ASC method is capable to determine a reproducible SC that followed the criteria for the shading and heat sources consistently and strictly. At 55.9% of the 34 analysed AWS sites, the ASC values corresponded with the MSC indicating that it was possible to automate the air temperature SC process using Dutch geospatial datasets. In 41.2% of the cases, a stricter ASC was applied compared to the MSC. This was primarily because the datasets detected more height differences that created higher shadow angles or because of heat sources that were outdated or not interpreted due to a lack of information. Another important aspect of uncertainty to take into account was the date when the digital datasets were collected. This determines on what altitude and land use data the information was collected and hence the ASC results were derived. Therefore, manual validations were required to see if the input for each site was up to date and correct. Knowledge about the data was always required and a form of human interpretation is unavoidable to apply the

ASC method in an operational setting. Because a majority of ASC output coincided with the MSC indicates, however, the developed model gives a good indication that the SC guidelines can be used for an automatic classification.

6.2 Recommendations

This research showed that an ASC for air temperature SC is possible to achieve. Preliminary results of this research were already presented and discussed at the biennial WMO-CIMO Technical Conference in Amsterdam through a poster. A copy of this poster can be found in appendix B as an annex (Stuurman, Dirksen, de Haij, Boog, & Rozeboom, 2018). The final results of this research further confirmed that the model can be used as a method for ASC. The process can be used as a tool to improve the observing network by finding more suitable sites for meteorological sensors in more representative locations for the larger area. The model can also be used for any location on land in The Netherlands and therefore it could also be applied in further research to assess air temperature measurements from third-party networks, which generally provide larger data volumes. Moreover, this model includes aspects that can be used for other SC variables such as the AHN and BGT data as well as the algorithm to determine the shadow angles. This research may therefore contribute or substantiate the need for development of similar automated procedures of the other SC variables, starting of with the wind variable. For this SC, more specification is needed about terrain roughness. Therefore land use geospatial datasets like the Landelijk Grondgebruik Nederland (LGN) could be used to help and obtain the necessary information.

Further research is needed, however, to improve the model in detecting the height of the vegetation and finding ways for automated quality control of the results, i.e. to determine automatically whether the automated SC output can be considered as valid. Furthermore, studies should be performed in the possibilities to integrate the results of the MSC within the ASC model, so that cross validation could also be done automatically and the two types of classification processes can be combined. The ASC model showed that differences found between the ASC and the MSC are partly because the WMO guidelines do not specify the criteria well enough. The results also indicated that a SC value increases when improving the situation for just one or a very few shadow angles. Therefore, it is recommended that WMO specifies the (very strict) SC criteria guidelines so that the determination SC Classes of the shading criteria is determined based on the distribution of shadow angles. The MSC model developed in this research could help to determine how these adjusted guidelines should be formulated.

References

AHN. (2018). *Actueel Hoogtebestand Nederland.* Retrieved 2018-07-30, from http://www.ahn.nl/index.html

Blanco-Muriel, M., Alarcón-Padilla, D. C., López-Moratalla, T., & Lara-Coira, M. (2001, jan). Computing the solar vector. *Solar Energy*, *70*(5), 431–441. doi: 10.1016/S0038-092X(00)00156-0

Bojinski, S., Verstraete, M., Peterson, T. C., Richter, C., Simmons, A., Zemp, M., ... Zemp, M. (2014, sep). The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy. *Bulletin of the American Meteorological Society*, *95*(9), 1431–1443. doi: 10.1175/BAMS-D-13-00047.1

Brandsma, T. (2011). *Parallel air temperature measurements at the KNMI-terrain in De Bilt (the Netherlands) May 2003-April 2005* (Tech. Rep.). KNMI. Retrieved from http://wap.knmi.nl/onderzk/klimscen/papers/Hisklim7.pdf

Budyko, M. I. (1974). *Climate and life.* New York: Academic Press.

BZK. (2018). *Basisregistratie Grootschalige Topografie.* Retrieved 2018-07-30, from https://www.geobasisregistraties.nl/basisregistraties/grootschalige-topografie/basisregistratie-grootschalige-topografie

Corripio, J. G. (2003). Vectorial algebra algorithms for calculating terrain parameters from dems and solar radiation modelling in mountainous terrain. *International Journal of Geographical Information Science*, *17*(1), 1–23. doi: 10.1080/713811744

Dimoudi, A., & Nikolopoulou, M. (2003, jan). Vegetation in the urban environment: microclimatic analysis and benefits. *Energy and Buildings*, *35*(1), 69–76. doi: 10.1016/S0378-7788(02)00081-6

Doninck, v. J. (2018). *Horizon R package.* Retrieved 2018-07-07, from https://cran.r-project.org/web/packages/horizon/index.html

Dorman, M., Vulkan, A., Erell, E., & Kloog, I. (2017). *shadow : R Package for Geometric Shadow Calculations in an Urban Environment.*

Dozier, J., & Frew, J. (1990). Rapid Calculation of Terrain Parameters for Radiation Modeling From Digital Elevation Data. *Transactions on Geoscience and Remote Sensing*, *28*(5), 963–969. doi: 10.1109/36.58986

Estévez, J., Gavilán, P., & Giráldez, J. (2011, may). Guidelines on validation procedures for meteorological data from automatic weather stations. *Journal of Hydrology*, *402*(1-2), 144–154. doi: 10.1016/J.JHYDROL.2011.02.031

Fall, S., Watts, A., Nielsen-Gammon, J., Jones, E., Niyogi, D., Christy, J. R., & Pielke, R. A. (2011). Analysis of the impacts of station exposure on the U.S. Historical Climatology Network temperatures and temperature trends. *Journal of Geophysical Research Atmospheres*, *116*(14), 1–15. doi: 10.1029/2010JD015146

Fisler, J., Kube, M., Stocker, C., Gr"uter, E., & Calpini, B. (2017). *The MeteoSwiss Classification Procedure for Automatic Weather Stations* (Tech. Rep.). WMO. Retrieved from https://library.wmo.int/index.php?lvl=notice_display&id=19954{#}.WYqss4R96Uk

Heerwaarden, C. C. V. (2011). *Surface evaporation and water vapor transport in the convective boundary layer.* (Unpublished doctoral dissertation). Wageningen University.

Jianxia, G., Yanhua, G., Xuefeng, S., Jiangfeng, G., Zhiya, T., Dongxia, T., ... Yuan, Y. (2014). Experiments and simulations of sitting classification for wind and temperature observation. In *Wmo/cimo teco 2014.* St. Petersburg.

Kadaster. (2018). *BGT.* Retrieved 2018-05-10, from `https://www.kadaster.nl/bgt`

Kadaster. (2019). *Verbeter de kaart.* Retrieved 2019-02-02, from `https://verbeterdekaart.kadaster.nl/`

Karl, T. R., Derr, V. E., Easterling, D. R., Folland, C. K., Hofmann, D. J., Levitus, S., ... Withee, G. W. (1995, dec). Critical issues for long-term climate monitoring. *Climatic Change*, *31*(2-4), 185–221. doi: 10.1007/BF01095146

Kinoshita, N. (2014). An Evaluation Method of the Effect of Observation Environment on Air Temperature Measurement. *Boundary-Layer Meteorol*, *152*, 91–105. doi: 10.1007/s10546-014-9918-2

KNMI. (2015). *Observations Strategy KNMI 2015-2024.* Retrieved from `https://bibliotheek.knmi.nl/knmipubmetnummer/knmipub233.pdf`

KNMI. (2018). *Automatische weerstations.* Retrieved 2018-05-03, from `https://www.knmi.nl/kennis-en-datacentrum/uitleg/automatische-weerstations`

Kumamoto, M., Otsuka, M., Sakai, T., Hamagami, T., Kawamura, H., Aoshima, T., & Fujibe, F. (2013). Field Experiment on the Effects of a Nearby Asphalt Road on Temperature Measurement. *Sola*, *9*.

Leroy, M. (1998). Meteorological measurement representativeness nearby obstacles influence. In *Wmo/cimo teco 1998* (pp. 51–54). Cassablanca.

Leroy, M. (2006). Documentation of Surface Observation. Classification for Siting and Performance Characteristics. *Meteo-France*, 3–7.

Mahmood, R., Pielke, R. A., Hubbard, K. G., Niyogi, D., Bonan, G., Lawrence, P., ... Syktus, J. (2010, jan). Impacts of Land Use/Land Cover Change on Climate and Future Research Priorities. *Bulletin of the American Meteorological Society*, *91*(1), 37–46. doi: 10.1175/2009BAMS2769.1

Menne, M. J., Williams, C. N., & Palecki, M. A. (2010, jun). On the reliability of the U.S. surface temperature record. *Journal of Geophysical Research*, *115*(D11), D11108. doi: 10.1029/2009JD013094

OGC. (2019). *Web Feature Service.* Retrieved 2019-02-07, from `https://www.opengeospatial.org/standards/wfs`

Parker, D. E., Jones, P., Peterson, T. C., & Kennedy, J. (2009, mar). Comment on "Unresolved issues with the assessment of multidecadal global land surface temperature trends" by Roger A. Pielke Sr. et al. *Journal of Geophysical Research*, *114*(D5), D05104. doi: 10.1029/2008JD010450

PDOK. (2018a). *AHN2 download atom feed links.* Retrieved 2018-07-07, from `http://geodata.nationaalgeoregister.nl/ahn2/atom/ahn2_05m_ruw.xml`

PDOK. (2018b). *AHN3 downloads.* Retrieved 2018-07-07, from `https://www.pdok.nl/nl/ahn3-downloads`

PDOK. (2018c). *hogere resolutie luchtfoto als open data bij PDOK.* Retrieved 25-01-2019, from `https://www.pdok.nl/-/nieuw-hogere-resolutie-luchtfoto-als-open-data-bij-pdok`

Peterson, T. C. (2006). Examination of potential biases in air temperature caused by poor station locations. *Bulletin of the American Meteorological Society*, *87*(8), 1073–1080. doi: 10.1175/BAMS-87-8-1073

Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., ... Parker, D. (1998, nov). Homogeneity adjustments of in situ atmospheric climate data: a review. *International Journal of Climatology*, *18*(13), 1493–1517. doi: 10.1002/(SICI)1097-0088(19981115)18:13<1493::AID-JOC329>3.0.CO;2-T

Petralli, M., Massetti, L., Brandani, G., & Orlandini, S. (2014, mar). Urban planning indicators: useful tools to measure the effect of urbanization and vegetation on summer air temperatures. *International Journal of Climatology*, *34*(4), 1236–1244. doi: 10.1002/joc.3760

Pielke, R., Nielson-Gammon, J., Davey, C., Bliss, O., Doesken, N. C., Ming, F., ... Kenneth, G. (2007). Documentation of Uncertainties and Biases Associated with Surface Temperature Measurement Sitesfor Climate Change Assessment. *American Meteorological Society*, *June*(February), 913–928. doi: 10.1175/BAMS-88-6-913

R. (2019). *The R Project for Statistical Computing.* Retrieved from `https://www.r-project.org/`

Remme, R., Nijs, T. D., & Paulin, M. (2017). *Natural Capital Model: Technical documentation of the quantification, mapping and monetary valuation of urban ecosystem services* (Tech. Rep.). RIVM. Retrieved from `https://www.rivm.nl/publicaties/natural-capital-model-technical-documentation-of-quantification-mapping-and-monetary`

Shido, F., Yamamoto, A., Aoyagi, T., Seino, N., & Fujibe, F. (2016). An observational study of the influence of nearby plants and artificial structures on the surface air temperature . In *Wmo/cimo teco 2016.*

Sotelino, L. G., Coster, N. D., Beirinckx, P., & Peeters, P. (2016). Environmental Classification of RMI Automatic Weather Station Network..

Steeneveld, G. J., Koopmans, S., Heusinkveld, B. G., van Hove, L. W. A., & Holtslag, A. A. M. (2011, oct). Quantifying urban heat island effects and human comfort for cities of variable size and urban morphology in the Netherlands. *Journal of Geophysical Research*, *116*(D20), D20129. doi: 10.1029/2011JD015988

Stuurman, J. (2019a). *AHN extract.* Retrieved from `https://github.com/JelleSt/rAHNextract`

Stuurman, J. (2019b). *BGT extract.* Retrieved 2019-02-22, from `https://github.com/JelleSt/rBGTextract`

Stuurman, J. (2019c). *temperatureSC R Package.* Retrieved from `https://github.com/JelleSt/temperatureSC`

Stuurman, J., Dirksen, M., de Haij, M., Boog, G., & Rozeboom, R. (2018). Towards automated CIMO siting classification using geospatial datasets. In *Wmo/cimo teco 2018.* Amsterdam. Retrieved from `www.wmocimo.net/wp-content/uploads/P1_1-Abstracts-for-website-65.pdf`

Thorne, P. W., Parker, D. E., Christy, J. R., Mears, C. A., Thorne, P. W., Parker, D. E., ... Mears, C. A. (2005, oct). Uncertainties in Climate Trends: Lessons from Upper-Air Temperature Records. *Bulletin of the American Meteorological Society*, *86*(10), 1437–1442. doi: 10.1175/BAMS-86-10-1437

van der Schrier, G., van Ulden, A., & van Oldenborgh, G. J. (2011, may). The construction of a Central Netherlands temperature. *Climate of the Past*, *7*(2), 527–542. doi: 10.5194/cp-7-527-2011

van der Zon, N. (2013). *Kwaliteitsdocument AHN2* (Tech. Rep.). Retrieved from `http://www.ahn.nl/binaries/content/assets/hwh---ahn/common/wat+is+het+ahn/kwaliteitsdocument_ahn_versie_1_3.pdf`

WMO. (1993). *Siting Exposure of Meteorological Instruments* (Tech. Rep. No. 55). Author. Retrieved from `https://library.wmo.int/index.php?lvl=notice_display&id=9887`

WMO. (2017). *Guide to Meteorological Instruments and Methods of observation.* Author. Retrieved from `https://library.wmo.int/index.php?lvl=notice_display&id=12407#.WDVYPdXyuM8`

WMO. (2018). *Siting Classification for Surface Observing stations on Land.* Retrieved 2018-04-12, from `http://www.wmo.int/pages/prog/www/IMOP/SitingClassif/SitingClassif.html`

Wolff, M., Haapa, I., Huuskonen, M., Johansson, P., Kielland, G., Krabbi, M., ... Wittskog, C. (2014). Air Temperature Sensor Siting Classification in Nordic Countries. In *Wmo/cimo teco 2014.* WMO.

Wolff, M. A., Torfoss, T., Grinde, L., Hofstra, M., Solvang, R., Kielland, G., ... Nygård, H. D. (2014). Site Classification of the Observation Network of the Norwegian Meteorological Institute : Progress and Challenges. In *Wmo/cimo teco 2014.* WMO.

A Overview of R packages used

Table A.1: Overview of R Packages used including their most important functions.

R Package	Important functions	Descriptions
Insol	sunpos();	azimuth and elevation angle of the sun
Horizon	horizonSearch()	calculate shadow angles
rgdal	various	general library required for read and writing spatial objects
GdalUtils	ogr2ogr()	Read WFS
raster	raster(), writeraster()	read and write raster files
sp	coordinates()	create spatial point
sf	st_ead(), st_rite(), st_ntersect() st_rea()	read and write shapefiles intersection of polygons calcluate area of polygons
leaflet	leaflet()	create and show interactive map within R
ggplot2	ggplot()	create sun and shadow chart

B Annex: Copy of poster presented at WMO-TECO 2018

Towards automated CIMO siting classification using geospatial datasets

Jelle Stuurman, Marieke Dirksen, Marijn de Haij*, Jitze van der Meulen, Andrea Pagani, Geuko Boog, Rene Rozeboom

The surroundings of meteorological observing stations have to be monitored carefully to assess their compliance to (inter)national requirements regarding measurement quality. This helps to ensure the representativeness of the measurements and the homogeneity of data sets, which is essential for applications like climate monitoring. Nowadays, Siting Classifications (SC) are usually performed manually. Here an automated SC model is presented that determines the classification of an Automatic Weather Station (AWS) site based on the CIMO guidelines (WMO No.8). An automated SC makes the interpretation more objective and saves the inspectors time. Moreover, there is an increasing demand for classification of 3rd party observing sites, which requires a more efficient inspection process as their volume is large. Our SC uses a detailed Digital Elevation Model, land use map and solar angles. First results show similar output as manual classifications.

Introduction

The CIMO Siting Classification (SC) for surface observing stations on land has been implemented as part of the yearly station inspection in the KNMI surface network by the end of 2014. This network consists of approximately 50 Automated Weather Stations (AWSs), serving different end users (synoptic meteorology, climatology, aviation). The AWSs on land are manually inspected using several tools, such as rangefinder binoculars to measure the distance, height and direction of nearby obstacles. Inspection reports are stored centrally and include a description of the technical status of an observing site as well as (changes in) the direct surroundings, and proposed measures for siting improvement. However, this involves a lot of manual work and the output does not fully meet user needs in terms of fit-for-purpose applicability.

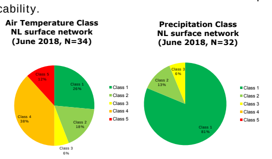


Figure 1: *a summary of the Air Temperature/Humidity and Precipitation classes for the Dutch observing network reported in June 2018.*

The current classification scheme is used by KNMI inspection staff for several years now. From their perspective, however, there is still a need within CIMO to review the existing guidance and criteria used in the classification scheme. There are still several ambiguities per variable, that should be clarified and reconsidered to optimize the added value of the process. For example, multiple small obstacles around an observing site, the impact of obstacles in different directions or short periods per day should be treated as well.

Case Study: Automated Temperature SC

The Automated SC for the temperature variable (Fig 2.) is presented for the AWS in De Bilt, The Netherlands (Fig. 3 & Fig. 4). For temperature the following criteria are of importance:

- Shading
- Land use of surroundings
- Vegetation height

The guidelines are translated into automatic calculations using the required spatial data. For this case study we made use of open datasets. Code is made available online through GitHub.

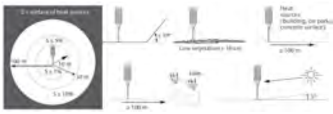


Figure 2: *Siting Classification criteria for temperature. [1]*

Figure 3: *Satellite imagery of the AWS De Bilt and surroundings.*

Figure 4: *AWS of De Bilt.*

https://github.com/Jellest/wmoSC

Data

- **Digital Elevation Model**: *"Actueel Hoogtebestand Nederland 2"* (AHN2) [2] with a resolution of a 50 cm and an accuracy of 5 cm (Fig. 5).
- **Land use**: *"Basisregistratie Grootschalige Topografie"* (BGT), [3] being the most detailed topographic map of The Netherlands, is used as land use data (Fig. 6).
- **Solar angles and azimuth**: R's *insol* library [4].

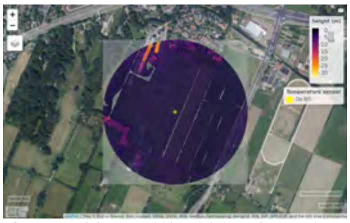


Figure 5: *Digital Elevation Model (AHN2) dataset in the surroundings.*

Figure 6: *BGT land use data 100 m around the temperature sensor.*

Data Processing

- The data is processed according to the decision tree in Figure 7.
- Variable only gains a certain class if all associated criteria are met.
- AHN2 is used in combination with solar angles to determine shadings (Fig. 5 & Fig. 9)
- AHN2 is used to determine vegetation height.
- BGT land use data is used to determine the size of surrounding heat sources (including buildings, road surfaces and water bodies) and its distance from the temperature sensor.

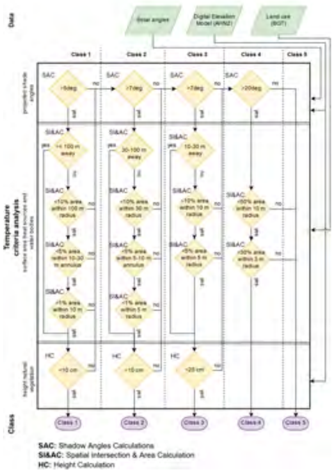


Figure 7: *Decision tree of all the WMO criteria guidelines set for the temperature variable.*

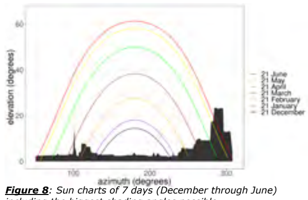


Figure 8: *Sun charts of 7 days (December through June) including the biggest shading angles possible.*

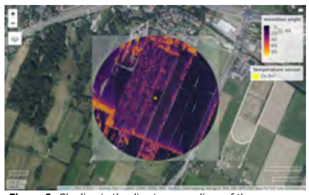


Figure 9: *Shading in the direct surroundings of the temperature sensor at an azimuth of 180 degrees (South).*

Results

- Shading angles are found above 20 degrees but for most of the time below 10 degrees.
- Several buildings, roads and water bodies are detected in the near surroundings but are not significant in size.
- From the raw and filtered elevation model no significant vegetation height difference was found.

Type of Criteria	Key Findings	Meet criteria
Projected shade	• Found >20 degrees shading angles	Class 4
Heat sources in direct surroundings	• Heat sources found within 100 meters • No significant area coverage of heat sources within 10 m and 100 m radii, as well as 10-30 m annulus	Class 1
Height of vegetation	• No significant height difference	Class 1

Discussion

- The automated SC is dependent on the quality of the different datasets.
- The AHN2 has a high resolution but may not be accurate enough to measure the height of the vegetation.
- The AHN2 is a measurement of only one moment in time, possibly influencing the height of trees due to seasonal changes.
- Although the land use map has a high resolution, it may be outdated or miss out information that can only be interpreted by the inspectors in the field.
- Visual interpretations in the field still need to be made.

Conclusion

- **An automated Siting Classification for temperature is presented**
- **Methods can be applied to other locations and are less subjective than manual inspections**
- **Results are similar to manual Siting Classification**

Future Research

- Including SC for: radiation, wind, precipitation, environment
- Applying the SC model to all the AWS on land in The Netherlands.

References

[1] WMO (2010). Annex 1.B. Siting Classification for surface observing stations on land.
[2] AHN2
https://www.ahn.nl
[3] BGT
https://www.kadaster.nl/bgt
[4] insol
https://cran.r-project.org/web/packages/insol

*Royal Netherlands Meteorological Institute (KNMI), R&D Information and Observation Technology, PO Box 201, 3730 AE De Bilt, The Netherlands, E-mail: haijde@knmi.nl

C Annex: Copy of WMO air temperature SC guidelines

Source: WMO (2018)

ANNEX 1.B. SITING CLASSIFICATIONS FOR SURFACE OBSERVING STATIONS ON LAND

(The text of the common ISO/WMO standard 19289:2014(E))

INTRODUCTION

The environmental conditions of a site[1] may influence measurement results. These conditions must be carefully analysed, in addition to assessing characteristics of the instrument itself, so as to avoid distorting the measurement results and affecting their representativeness, particularly when a site is supposed to be representative of a large area (i.e. 100 to 1 000 km^2).

1. SCOPE

This annex[2] indicates exposure rules for various sensors. But what should be done when these conditions are not fulfilled?

There are sites that do not respect the recommended exposure rules. Consequently, a classification has been established to help determine the given site's representativeness on a small scale (impact of the surrounding environment). Hence, a class 1 site can be considered as a reference site. A class 5 site is a site where nearby obstacles create an inappropriate environment for a meteorological measurement that is intended to be representative of a wide area (at least tens of km^2). The smaller the siting class, the higher the representativeness of the measurement for a wide area. In a perfect world, all sites would be in class 1, but the real world is not perfect and some compromises are necessary. A site with a poor class number (large number) can still be valuable for a specific application needing a measurement in this particular site, including its local obstacles.

The classification process helps the actors and managers of a network to better take into consideration the exposure rules, and thus it often improves the siting. At least, the siting environment is known and documented in the metadata. It is obviously possible and recommended to fully document the site, but the risk is that a fully documented site may increase the complexity of the metadata, which would often restrict their operational use. That is why this siting classification is defined to condense the information and facilitate the operational use of this metadata information.

A site as a whole has no single classification number. Each parameter being measured at a site has its own class, and is sometimes different from the others. If a global classification of a site is required, the maximum value of the parameters' classes can be used.

The rating of each site should be reviewed periodically as environmental circumstances can change over a period of time. A systematic yearly visual check is recommended: if some aspects of the environment have changed, a new classification process is necessary.

A complete update of the site classes should be done at least every five years.

In the following text, the classification is (occasionally) completed with an estimated uncertainty due to siting, which has to be added in the uncertainty budget of the measurement. This estimation is coming from bibliographic studies and/or some comparative tests.

The primary objective of this classification is to document the presence of obstacles close to the measurement site. Therefore, natural relief of the landscape may not be taken into account, if far

[1] A "site" is defined as the place where the instrument is installed.
[2] Whereas this is referred to as an annex in the WMO *Guide to Meteorological Instruments and Methods of Observation* (WMO-No. 8), it is referred to as a standard in the ISO document.

away (i.e. > 1 km). A method to judge if the relief is representative of the surrounding area is the following: does a move of the station by 500 m change the class obtained? If the answer is no, the relief is a natural characteristic of the area and is not taken into account.

Complex terrain or urban areas generally lead to high class numbers. In such cases, an additional flag "S" can be added to class numbers 4 or 5 to indicate specific environment or application (i.e. 4S).

2. AIR TEMPERATURE AND HUMIDITY

2.1 General

Sensors situated inside a screen should be mounted at a height determined by the meteorological service (within 1.25 to 2 m as indicated in the WMO *Guide to Meteorological Instruments and Methods of Observation* (WMO-No. 8)). The height should never be less than 1.25 m. The respect of the higher limit is less stringent, as the temperature gradient versus height is decreasing with height. For example, the difference in temperature for sensors located between 1.5 and 2 m is less than 0.2 °C.

The main discrepancies are caused by unnatural surfaces and shading:

(a) Obstacles around the screen influence the irradiative balance of the screen. A screen close to a vertical obstacle may be shaded from the solar radiation or "protected" against the night radiative cooling of the air, by receiving the warmer infrared radiation from this obstacle or influenced by reflected radiation;

(b) Neighbouring artificial surfaces may heat the air and should be avoided. The extent of their influence depends on the wind conditions, as wind affects the extent of air exchange. Unnatural or artificial surfaces to take into account are heat sources, reflective surfaces (for example buildings, concrete surfaces, car parks) and water or moisture sources (for example, ponds, lakes, irrigated areas).

Shading by nearby obstacles should be avoided. Shading due to natural relief is not taken into account for the classification (see above).

The indicated vegetation growth height represents the height of the vegetation maintained in a "routine" manner. A distinction is made between structural vegetation height (per type of vegetation present on the site) and height resulting from poor maintenance. Classification of the given site is therefore made on the assumption of regular maintenance (unless such maintenance is not practicable).

2.2 Class 1

(a) Flat, horizontal land, surrounded by an open space, slope less than ⅓ (19°);

(b) Ground covered with natural and low vegetation (< 10 cm) representative of the region;

(c) Measurement point situated:

    (i) At more than 100 m from heat sources or reflective surfaces (buildings, concrete surfaces, car parks, etc.);

    (ii) At more than 100 m from an expanse of water (unless significant of the region);

    (iii) Away from all projected shade when the sun is higher than 5°.

A source of heat (or expanse of water) is considered to have an impact if it occupies more than 10% of the surface within a circular radius of 100 m surrounding the screen, makes up 5% of an annulus of 10–30 m, or covers 1% of a 10 m radius area.

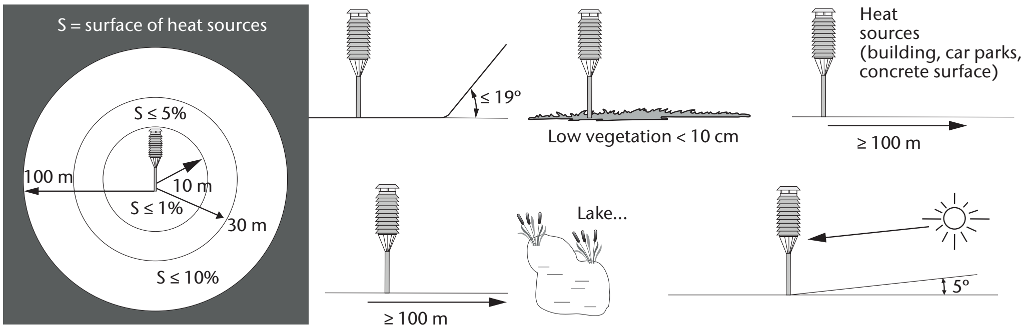


Figure 1.B.1. Criteria for air temperature and humidity for class 1 sites

2.3 **Class 2**

(a) Flat, horizontal land, surrounded by an open space, slope inclination less than ⅓ (19°);

(b) Ground covered with natural and low vegetation (< 10 cm) representative of the region;

(c) Measurement point situated:

(i) At more than 30 m from artificial heat sources or reflective surfaces (buildings, concrete surfaces, car parks, etc.);

(ii) At more than 30 m from an expanse of water (unless significant of the region);

(iii) Away from all projected shade when the sun is higher than 7°.

A source of heat (or expanse of water) is considered to have an impact if it occupies more than 10% of the surface within a radius of 30 m surrounding the screen, makes up 5% of an annulus of 5–10 m, or covers 1% of a 5 m radius area.

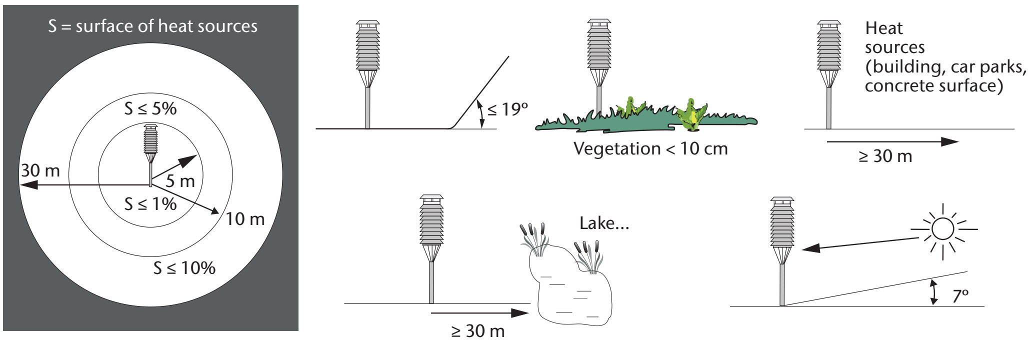


Figure 1.B.2. Criteria for air temperature and humidity for class 2 sites

2.4 **Class 3 (additional estimated uncertainty added by siting up to 1 °C)**

(a) Ground covered with natural and low vegetation (< 25 cm) representative of the region;

(b) Measurement point situated:

(i) At more than 10 m from artificial heat sources and reflective surfaces (buildings, concrete surfaces, car parks, etc.);

(ii) At more than 10 m from an expanse of water (unless significant of the region);

(iii) Away from all projected shade when the sun is higher than 7°.

A source of heat (or expanse of water) is considered to have an impact if it occupies more than 10% of the surface within a radius of 10 m surrounding the screen or makes up 5% of a 5 m radius area.

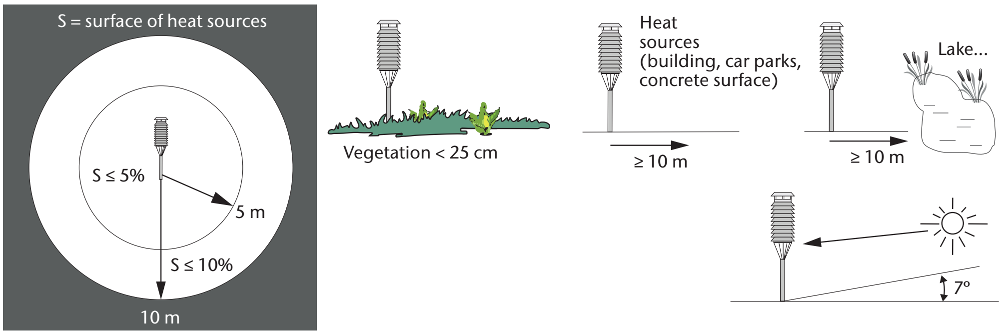


Figure 1.B.3. Criteria for air temperature and humidity for class 3 sites

2.5 **Class 4 (additional estimated uncertainty added by siting up to 2 °C)**

(a) Close, artificial heat sources and reflective surfaces (buildings, concrete surfaces, car parks, etc.) or expanse of water (unless significant of the region), occupying:

(i) Less than 50% of the surface within a 10 m radius around the screen;

(ii) Less than 30% of the surface within a 3 m radius around the screen;

(b) Away from all projected shade when the sun is higher than 20°.

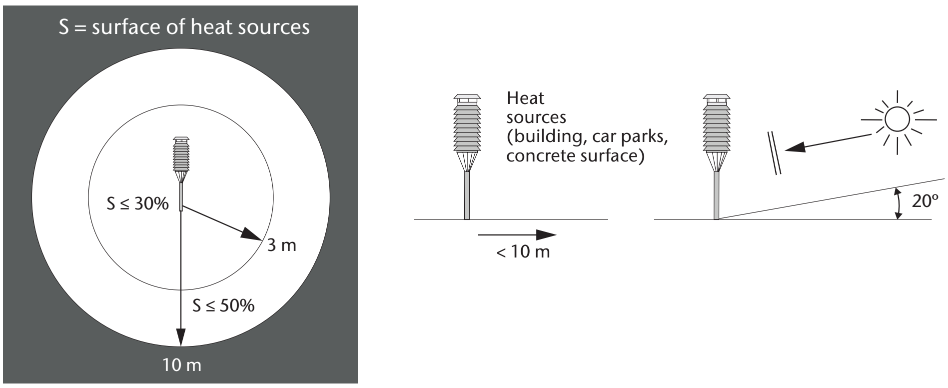


Figure 1.B.4. Criteria for air temperature and humidity for class 4 sites

2.6 **Class 5 (additional estimated uncertainty added by siting up to 5 °C)**

Site not meeting the requirements of class 4.

Royal Netherlands Meteorological Institute

PO Box 201 | NL-3730 AE De Bilt
Netherlands | www.knmi.nl