On high resolution wind, shear and cloud vertical structure : preparation of the aeolus space mission

ON HIGH RESOLUTION WIND, SHEAR AND CLOUD VERTICAL STRUCTURE

-PREPARATION OF THE AEOLUS SPACE MISSION-

Ph.D. DISSERTATION

Karim Houchi

The research presented in this thesis was carried out at Weather Department of the Royal Netherlands Meteorological Institute, in the frame of a collaboration between Technische Universiteit Eindhoven and the École Polytechnique de Paris.

This research was supported by a grant from the Dutch Ministry of Transport, Public Works and Water Management (Ministerie van Verkeer en Waterstaat) in support of the European Space Agency, Aeolus mission.

Cover by Ramdane Alkama and Karim Houchi
Printed by PrintSupport U4, Meppel, The Netherlands
A catalogue record is available from the Eindhoven University of Technology Library
ISBN: 978-90-386-4047-1

ON HIGH RESOLUTION WIND, SHEAR AND CLOUD VERTICAL STRUCTURE

-PREPARATION OF THE AEOLUS SPACE MISSION-

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr.ir. F.P.T. Baaijens, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op dinsdag 5 april 2016 om 16:00 uur

door

Karim Houchi

geboren te Tizi-Ouzou, Kabylia, Algerije

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

Voorzitter:	prof.dr. H.J.H. Clercx
1e promotor:	prof.dr. H.M. Kelder
2e promotor:	prof.dr.ir. G.J.F. van Heijst
co-promotor:	dr.ir. A. Stoffelen (KNMI)
leden:	prof.dr. P.F. Levelt (TUD)
	prof.dr. E. Kallen (University Stockholm)
	prof.dr.ir. G.M.W. Kroesen
	dr. P.H. Flamant (CNRS)

Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

La véritable inspiration suppose du souffle. Ce n'est souvent que du vent.
Gilbert Keith Chesterton

Genius is one percent inspiration, ninety-nine percent perspiration.
Thomas Edison

il n'y a que le travail qui peut vous récompenser demain.
Lounes Matoub

Summary

The context of this PhD thesis is the Atmospheric Dynamics Mission, now called Aeolus, of the European Space Agency (ESA), planned for launch in 2017. ADM-Aeolus will provide atmospheric wind profiles up to about 30 km height, using a Doppler Wind Lidar (DWL). Wind profiles observations are expected to improve the initialization of Numerical Weather Prediction (NWP) models to benefit weather forecast skill, as well as provide an improved characterization of the atmospheric dynamics, mixing and transport on both short and long time scales of months or years. The thesis focuses on the use of available world-wide high-resolution radiosonde data to characterize the vertical structure of the atmosphere. Radiosondes provide high resolution vertical profiles of wind, temperature and humidity and are a unique resource to depict the motion in the atmosphere on the smaller scales and prepare for the Aeolus mission. Moreover, radiosondes depict the clouds in association with the motion in the atmosphere. Radiosonde wind and cloud properties are studied, which are both relevant for the simulation of the Aeolus observations.

First, an effective statistical quality control method is developed to clean the high-resolution radiosonde data from unrepresentative observations (outliers). Also, differences in the accuracy of various types of radiosonde wind-finding systems (radio theodolite, LOng Range Navigation (LORAN) and Global Positioning System (GPS) are highlighted. Indeed, the difference in systems accuracy measuring wind and the ascent height of the radiosonde balloon may affect the derivation of wind shear, as well as the rate of ascent of the balloon and the vertical wind component. It is mainly found that though radio-theodolite-based winds are less accurate than winds based on the other wind-finding systems, they provide after quality control [1] a good statistical representation of horizontal wind and of the vertical shear of the horizontal wind, here simply called vertical wind-shear.

[1] Houchi, K., A. Stoffelen, G.J. Marseille, J. De Kloe (2015). Statistical Quality Control of High-Resolution Winds of Different Radiosonde Types for Climatology Analysis. Journal of Atmospheric and Oceanic Technology, 32, 1796-1812.

Secondly, collocation of radiosonde observations with the short-range forecasts of the European Centre for Medium-range Weather Forecasts (ECMWF) NWP model allowed a detailed description of the atmospheric dynamics and its representation in global NWP models. Collocated radiosonde and ECMWF horizontal wind and vertical wind-shear climatologies were compared, including spatial and temporal variabilities [2]. The wind shear (mean and variability) is found to be a factor of 2.5 (zonal) and 3 (meridional) larger in the observed wind profiles than in the NWP model profiles. It is demonstrated consequently that the effective ECMWF vertical model resolution is about 1.7 km, failing thus to well resolve the atmospheric phenomena that occur below this scale, such as turbulence, convection, gravity waves, etc.. With these results a global (from ECMWF model) and realistic (from radiosondes observations) wind database has been created for Aeolus Doppler Wind Lidar (DWL) simulations. The dynamical characteristics inferred from radiosondes and the ECMWF NWP model are complemented by the atmospheric optical properties from CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation) data [3,4].

Radiosonde database is further used to study combined effects of clouds and wind shear. Such combined effects are important for the interpretation of the Aeolus DWL data, but also of relevance in physical parameterizations used in atmospheric modeling. A preliminary assessment of the vertical cloud structures has been made on the basis of high-resolution radiosondes (chapter.5). The results are compared to collocated ECMWF model and collocated CALIPSO data sets, in terms of vertical distribution and frequency of occurrence of clouds. This work aims to identify optically realistic atmospheric scenes with different vertical cloud structure configurations and associated wind dynamics to simulate such scenes with Aeolus DWL. The recent use of this database [5] confirmed some of our results, but it is further used to optimize the vertical sampling of Aeolus and its wind retrieval performance.

Keywords: Atmospheric Dynamics Aeolus, Wind, Wind Shear, Climatology, Statistics, Quality Control, Wind-Finding System, Cloud Vertical Structure

[2]Houchi, K., Stoffelen, A., Marseille, G.J., de Kloe, J. (2010). Comparison of wind and wind shear climatologies derived from high-resolution radiosonde and ecmwf model. Journal of Geophysical Research, 115, D22123.

[3]Marseille, G.J., Houchi, K., Stoffelen, A. & de Kloe, J. (2011). The definition of an atmospheric database for ADM-Aeolus. Atmospheric. Measurement Techniques, 4, 67-88.

[4]Stoffelen, A., H. Körnich, G.J. Marseille, K. Houchi, and J. De Kloe (2009), Assessment of Optical and Dynamical Atmospheric Heterogeneity, available through KNMI, ESA's Report, Ref: AE-TNKNMI-VAMP-002_v7_KNMI_150609, version: 0.8, 15 June 2009.

[5] Sun, X. J., R. W. Zhang, G. J. Marseille, A. Stoffelen, D. Donovan, L. Liu, J. Zhao, The performance of Aeolus in heterogeneous atmospheric conditions using high-resolution radiosonde data, Atmos. Meas. Tech., 7, 26952717, 2014 www.atmos-meas-tech.net/7/2695/2014/ doi:10.5194/amt-7-2695-2014

Samenvatting

De context van dit proefschrift is de atmosfeerdynamica missie, nu genaamd Aeolus, van de European Space Agency (ESA), gepland voor lancering in 2017. Aeolus zal atmosferische windprofielen tot ongeveer 30 km hoogte meten, met behulp van een Doppler Wind Lidar (DWL). Windprofielwaarnemingen zullen naar verwachting de initialisatie van numerieke (NWP) weermodellen verbeteren, hetgeen de kwaliteit van weersverwachtingen ten goede komt, evenals resulteren in een verbeterde karakterisering van de dynamica van de atmosfeer en haar transport- en mengingsprocessen zowel op de korte als de lange tijdschalen van maanden of jaren. Het proefschrift richt zich op het gebruik van wereldwijd beschikbare hoge-resolutie radiosondegegevens en exploreert met name de kleinschalige verticale structuur van de atmosfeer. Radiosondemetingen bieden verticale profielen van wind, temperatuur en vochtigheid op hoge resolutie en een unieke bron voor de karakterisatie van kleinschalige beweging in de atmosfeer en is daarmee een goede voorbereiding voor de Aeolus missie. Radiosonde wind- en wolkeneigenschappen zijn bestudeerd, welke beide relevant zijn voor de simulatie van Aeolus waarnemingen.

In de eerste plaats is een efficiënte statistische methode voor kwaliteitscontrole ontwikkeld om de hoge-resolutie radiosonde gegevens te ontdoen van niet representatieve waarnemingen (outliers). Verder zijn verschillen in de nauwkeurigheid van de verschillende soorten radiosonde meetsystemen onderzocht (Radio theodoliet, Long Range Navigation (LORAN) en Global Positioning System (GPS)), welke verschillen met name zeer belangrijk zijn voor de afleiding van windschering, maar ook voor het afleiden van de stijgsnelheid van de radiosondeballon en de verticale windcomponent. Het is vooral te vinden dat, hoewel radiotheodoliet-windmetingen minder nauwkeurig zijn dan winden op basis van de andere meetsystemen, ze na kwaliteitscontrole [1] een goede afspiegeling zijn van de statistische verdeling van horizontale wind en van de verticale schering van de horizontale wind, hier verder verticale windschering genoemd.

Ten tweede, collocatie van radiosondewaarnemingen met de korte-termijn prognoses van het European Centre for Medium-range Weather Forecasts (ECMWF)

NWP weermodel geven een gedetailleerde vergelijking van de atmosferische dynamica in radiosondes en haar representatie in wereldwijde NWP modellen. De radiosonde en ECMWF horizontale wind en verticale windschering klimatologie werden vergeleken, met inbegrip van hun ruimtelijke en temporele variabiliteit [2,5]. De windschering (gemiddelde en variabiliteit) blijkt een factor 2.5 (zonaal) en 3 (meridionaal) groter in de waargenomen windprofielen dan in de NWP-model profielen. Het is verder aangetoond dat de effectieve verticale ECMWF modelresolutie ongeveer 1,7 km is en het model atmosferische verschijnselen die zich voordoen op kleinere schaal, zoals turbulentie, convectie, zwaartekrachtsgolven, etc. niet goed oplost. Met deze resultaten is een wereldwijd (middels het ECMWF model) en een realistische (op basis van radiosonde karakteristieken) winddatabase gemaakt ten behoeve van Aeolus DWL simulaties. De dynamische eigenschappen afgeleid uit radiosondes en het ECMWF NWP model zijn aangevuld met de atmosferische optische eigenschappen van Calipso (Cloud-Aeolsol Lidar en Infrarood Pathfinder Satellite Observation) [3,4].

Radiosondegegevens worden vervolgens gebruikt om gecombineerde effecten van wolken en verticale beweging te bestuderen. Dergelijke gecombineerde effecten zijn van belang voor de interpretatie van de Aeolus DWL gegevens, maar ook van belang in de fysische parameterisaties die gebruikt worden in atmosferische modellering (NWP). Een eerste analyse van verticale wolkenstructuren is gemaakt op basis van hoge-resolutie radiosondes [5]. Resultaten zijn vergeleken met gecolloceerde ECMWF modelvelden en Calipso gegevens in termen van optische eigenschappen van wolken en hun verticale verdeling, dit in een aantal verschillende klimaatregios. Het identificeren van optisch-realistische atmosferische scnes met verscheidene verticale wolkstructuren en bijbehorende complexe winddynamica is van belang voor het simuleren van dergelijke scnes met de Aeolus DWL. De resulterende atmosferische database wordt gebruikt voor het optimaliseren van de verticale en horizontale bemonstering met het Aeolus meetsysteem en de verfijning van de Aeolus windberekeningen.

.

[1]Voor de referenties 1, 2, 3, 4 en 5, zie de voetnoot onder de Engelse samenvatting

Résumé

Cette thèse s'inscrit dans le cadre de la mission de l'Agence Spatiale Européenne (ESA), nommée Aeolus Mission, dont le lancement est prévu pour 2017. La mission Aeolus consiste principalement à mesurer à partir de l'espace les profils du vent atmosphérique jusqu'à environ 30 km d'altitude, en utilisant un lidar à vent à effet Doppler (DWL). Grâce à ces mesures des profils des vents, des avancées sont attendues dans l'initialisation des modèles numériques de prévision météorologique (NWP) pour améliorer les prévisions météorologique, ainsi que la description de la dynamique, le mélange et le transport atmosphèriques sur des échelles de temps courtes et longues, des mois ou des années. La thèse se focalise sur l'usage de données de radiosondages à haute résolution disponibles sur l'ensemble du globe pour caractériser la structure verticale de l'atmosphère. Les radiosondes fournissent des profils verticaux à haute résolution du vent, de pression, de température et d'humidité, et qui sont une ressource unique pour caractériser le mouvement dans l'atmosphère à petites échelles et préparer ainsi la mission Aeolus. Les propriétés du vent et des nuages, indispensable pour la simulation des observations Aeolus, sont étudiées.

Premièrement, une méthode statistique efficace de contrôle de qualité est développée pour éliminer les observations peu répresentatives (points abérrants) des données de radiosondages. De plus, les différences dans la précision de mesure du vent de divers types de systèmes radiosonde; communément appelé "Windfinding"; (Radiothéodolite, LORAN (LOng Range Navigation) et le GPS (Global Positioning System), sont mises en évidence. En effet, la difference dans la précision des systemes de mesure du vent et de la hauteur d'ascension du ballon de radiosonde peuvent fausser la dérivation des cisaillements du vent, ainsi que celles du taux d'ascension du ballon et de la composante du vent vertical. Il a été principalement conclu, que même si les données basées sur le radio-theodolite sont moins précises que les données basées sur d'autres systèmes "Windfinding", elles fournissent une bonne représentation statistique du vent horizontal et des cisaillements du vent vertical après le contrôle de qualité [1].

Deuxièmement, la co-localisation des observations de radiosonde avec les prévisions à court-terme du modèle numérique d'ECMWF a permis une déscription détaillée de la dynamique de l'atmosphére et sa représentation dans les modèles numériques globaux. Les climatologies du vent horizontal et cisaillement du vent vertical à partir des radiosondages et du modèle d'ECWMF colocalisés ont été comparés, y compris les variabilités spatiale et temporelle [2,5]. Un cisaillement amplifié du vent (moyenne et variabilité), s'avère être d'un facteur de 2,5 (zonal) et 3 (Méridien) plus grand dans les profils observés que dans les profils du vent du modèle numérique d'ECMWF. Il est par conséquent démontré que la résolution verticale effective du modèle d'ECMWF est d'environ 1,7 km, ne permettant pas ainsi de résoudre les phénomènes atmosphériques pouvant se produire en dessous de cette échelle, comme la turbulence, la convection, les ondes de gravité, etc. Avec ces résultats, une base de données du vent, globale (à partir du modèle d'ECMWF) et réaliste (à partir des observations de radiosondes) a été créée pour les simulations du Lidar à vent d'Aeolus. Les caractéristiques dynamiques déduites à partir des radiosondes et du modèle d'ECMWF sont complétées par les propriétés optiques atmosphériques mesurées par CALIPSO (Cloud-Aeolsol Lidar and Infrared Pathfinder Satellite Observation) [3,4].

Les données de radiosondages sont en plus utilisées pour étudier les effets combinés de nuage et de cisaillement vertical du vent. De tels effets combinés sont importants pour l'interprétation des données du DWL d'Aeolus, mais aussi importants dans les paramétrisations physiques utilisées dans la modélisation atmosphèrique. Une évaluation préliminaire des structure verticales des nuages (CVS) a été elaborée à partir des données de radiosondages (chapter.5). Les résultats sont comparés avec celles des jeux de données collocalisés du modèle d'ECMWF et de CALIPSO, en termes de fréquence d'occurence et de distribution verticale des nuages. Ce travail a pour objectif d'identifier des scènes atmosphériques optiquement réalistes avec des configurations de structures de nuage verticales différentes et de la dynamique du vent associée. Les caractéristiques des nuages combinées à la dynamique du vent seront utilisés pour simuler de telles scènes avec Aeolus-DWL. L'utilisation récente de cette base de données [5] a confirmé certains de nos résultats, mais il est en outre utilisée pour optimiser l'échantillonnage vertical d'Aeolus et sa performance dans l'extraction du vent.

Mots-clés : ADM-Aeolus, Vent, Cisaillement du Vent, Climatologie, Statistique, Contrôle de Qualité, Wind-Finding System, Structure Verticale des Nuages.

[1]Pour les references 1, 2, 3, 4 et 5, voir en bas du résumé en Anglais.

CONTENTS

CHAPTER 1

Introduction

1.1 Motivation of the Thesis

Weather is one of the fundamental earth processes that affected humankind throughout its past and probably will affect its future. In fact, weather shapes the earth and modifies or affects populations, their behavior, their environment, etc. Weather can affect the human being in complex ways leading to catastrophic situations in case of extreme weather events (floods, hurricanes, etc.). But also in simpler way, i.e., in the every daily life, causing discomfort and possible negative effects on the human body in case of extreme high temperatures, humidity and wind. Furthermore, weather influences many aspects of our lives, in commercial activities such as farming, fishing, energy, construction and transport, etc. Wind in particular, has played a major role in human civilization. Indeed, wind has not only inspired mythology, but has also influenced the events of history, expanded world trade and warfare, provided energy for mechanical work, electricity, recreation, etc. In addition, wind is responsible for a variety of aeolian processes such as formation of fertile soils, erosion, transport of dust from large deserts and of pollutants, the spread of wildfires, dispersion of seeds from various plants, as well as bird migration and flying insect populations, etc. Consequently most these phenomena related to wind lead to the shaping of land-forms and the transformation of the environment. For all these reasons, human have always sought to control the weather and to predict it as accurately as possible, to aid anyone in the planning of daily affairs and anticipate extreme events than can put in danger lives and properties. Despite the important advances recorded in weather forecasts in recent years, the missing observations of wind profile data over a large part of the globe hampers a significant improvement in the accuracy of the forecasts.

Besides weather forecasting, there is furthermore a great interest in climatological records. In fact, the Intergovernmental Panel On Climate Change (IPCC) stated that the earth climate is likely being changed by humankind. This means that weather

events that are extreme today may become less extreme and more frequent in the next decades. Moreover, the earth processes and phenomena linked to the weather may change systematically, such as for example erosion and vegetation, but it also affects human activities in energy, water protection, transport and agriculture for example. Indeed, wind is one of the most important meteorological variables to determine the atmospheric dynamics and thus the absence of such information in weather forecasting and climate research is detrimental. The Atmospheric Dynamics Mission (ADM-Aeolus mission) of the European Space Agency (ESA), or simply known as ***Aeolus*** [1] (see chapter 2) will provide wind profile measurements from space. As such, the Aeolus satellite (Reitebuch *et al.*, 2010), due for launch in 2017, is expected to remedy this lack of data, by providing reliable, continuous and global wind-profile measurements from space using a Doppler Wind Lidar [2] (DWL). Section 1.2 focuses on wind and dynamics in weather and climate. Section 1.3 explains the importance of wind information in the Global Observing System (GOS), in addition the other meteorological variables such as temperature, pressure and humidity denoted PTU. But, we focus on the necessity of fulfilling the wind measurements lacking in many areas on the globe, and that will be beneficial for Data Assimilation (DA) process and NWP and climate models. Section 1.4 recall how wind measurement are done throughout the history, from Beaufort scales until the latest modern measurements from satellites. Section 1.5 gives an overview of radiosonde instrument used to collect the data and ECMWF model. The last section give an overview of the content of each chapter of the thesis and the addressed research questions.

1.2 Wind

Wind is the movement of the atmosphere, a mass of gas that can be found in many planets of the solar system, such as in Neptune and Saturn where the winds are the most violent. On earth, the unequal distribution of solar heating over its surface generates differences in air temperature, density and pressure, causing thus the air to move, i.e., creating wind. The greater these differences in air pressure between two regions are, the stronger the wind will be. Wind will continue to blow until the pressure differences is reduced, i.e., pressures of two zones, previously very different, become more equal when the wind blows. However, due to earth rotation the Coriolis force is in effect, which deflects the wind away from the depression, relaxing the filling of it. Areas of low surface pressure, called cyclones, are generally associated with rising warm air, whilst areas of high surface pressure, known as anticyclones, are associated with subsiding air. The Coriolis force (which increases with latitude from zero at the Equator to a maximum at the poles) explains why winds generally blow perpendicular to the direction of the pressure difference; it acts at right angles to the direction of motion, so as to cause deflection to the right in the Northern Hemisphere and to the left in the Southern Hemisphere (Holton, 1992). In meteorology, contrary to oceanography, wind directions are conventionally specified as being the compass point from which the wind originates, counted clockwise. So, for instance wind blowing from

[1] Aeolus in Greek mythology means the "keeper of the winds"
[2] LIght Detection And Ranging

the west to the east, are called westerly winds and are associated with a wind direction of 270 degrees.

The average movement of air constitutes the general circulation of the atmosphere, as the prevailing marine currents define the oceanic circulation. Both circulations are described by laws of fluid dynamics which are implemented in NWP, ocean and climate models in discrete form, which are mutually intimately coupled. Concerning the atmospheric circulation, heat is transported away from equatorial regions towards the poles and returning cooler air to the tropics. This leads to various spatial and temporal large-scale circulations on the globe, known as Global Wind Patterns (GWP), as schematized in figure 1.1. It shows the Hadley, Ferrel and Polar cells at different latitudes ; the common wind patterns along the longitude: NE and SE trade winds on each side of the Inter-Tropical Convergence Zone (ITCZ), Westerlies in the mid-latitudes and Easterlies in the polar regions, including the strong subtropical and polar jets (westerly winds) that form around the tropopause where air masses of the different cells meet (Ahrens, 2011). These meandering jet streams are narrow ribbons of fast (reaching speeds of up to 500 km/h) and continuously moving air, which are associated with large temperature gradients. The jet currents blowing from west to east in both hemispheres are typically thousands of kilometers long and several hundred kilometers wide, but only a few kilometers thick, and their positions vary from one day to another.

Local and regional circulations of long-duration are also formed due the inhomogeneous and various types of the earth surface (e.g., sea, land, mountains, grass, ice, etc.). Some of these regional and local wind patterns and their periodicities are know and their occurrences may be predicted (though not always accurately), such as monsoons within tropical regions, i.e., thermal low circulations over terrain and high plateaus (Keshavamurthy & Rao, 1992), diurnal (daily cycle) breezes near coastal and mountainous areas (sea/valley breezes during the day, and land/mountain breezes during the night), etc.. Detailed description with illustrations of all these phenomena may be found in the book of Ahrens (2007). The map in figure 1.2 gives an overview of the most common and known local winds and their assigned names over the globe, such as the Siricco, Mistral, etc. Contrary to long-duration winds, local high speed winds over short periods, such as gusts, squalls, tornadoes, etc, are highly unpredictable and can occur in different locations depending on weather conditions, though their occurrence is known to be different from one climate region to another. So, globally wind occurs on a range of scales, extending from gusts that die out quickly after blowing over a few meters, to thunderstorm flows lasting tens of minutes, to local breeze circulations lasting a few hours, and finally to global winds traveling generally thousands of kilometers in steady patterns lasting for weeks.

Atmospheric winds are not always uniform flows, so generally they present a large variability in the horizontal and vertical, and can also be observed in different forms or phenomena, such as turbulence, gravity waves, shear, etc. It is also well known that vertical wind gradients (vertical wind shear) are generally large near the surface and at jet level, where areas of wind shear caused by various weather phenomena can

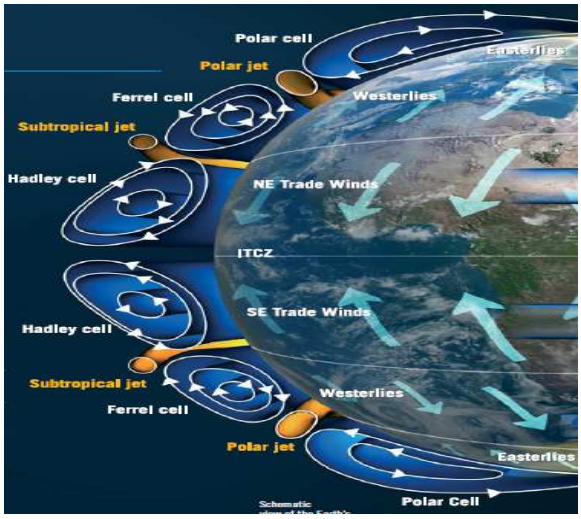


Figure 1.1: The major global earth wind patterns. Amospheric circulation in each hemisphere consists of three cells - the Hadley, Ferrel and polar cells. High-speed wind fields known as "jets" are formed between these cells and which are associated with large temperature gradients. (Courtesy of ESA).

for instance lead to dangerous situations for aircraft, buildings, etc. In meteorology, and NWP modeling in particular, an accurate representation of wind shear (therefore wind) is very important, since its inaccurate representation is a source of errors in predicting the locations, development, and paths of weather events. So, high resolution measurements are necessary to accurately capture wind and shear variability through the whole atmosphere. Moreover, regions with occurrence of strong vertical wind shear are of particular interest.

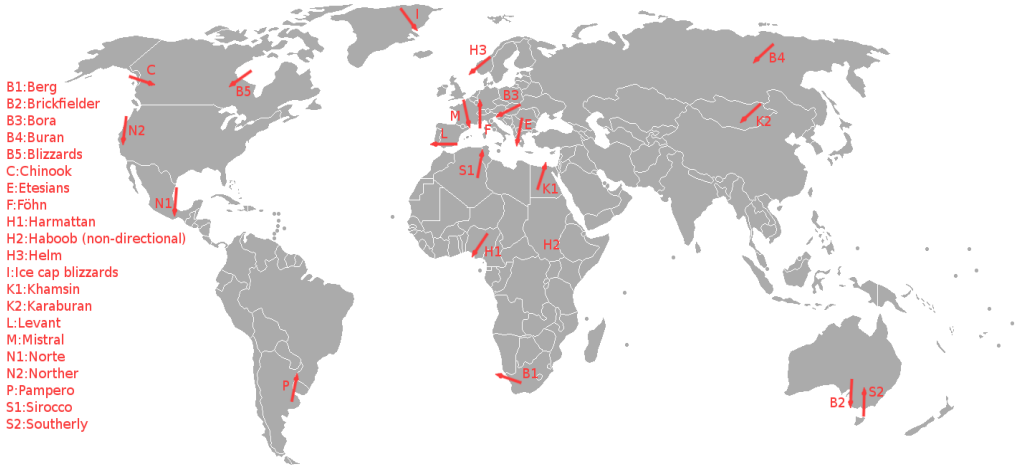


Figure 1.2: Map and list of some well-known local winds. The map is based on an image from the book "Wereldwijzer" by Marshall & Cavendish.

1.3 The Need for Wind Profiles

Human beings have always tried to tame the wind, by avoiding it or taking advantages from it, e.g., sailing, aviation, windmills, etc. They have also always tried to develop accurate instruments for measuring it. Winds generally are classified according to their spatial extent, their speed and their geographic location, but also according to the forces that it exerts and its effects (section 1.1). The first estimation of the strength of the wind has been introduced in 1806 by an English naval officer, Francis Beaufort, known since as the Beaufort scale. This scale, originally intended for sailors and which consists in using visual features of the sea to estimate wind strength, has been extended to land. Wind strength on land has been estimated via the physical effects on vegetation and structures or using wind-socks, etc., before the invention of modern instruments. Since the invention of the barometer and hygrometer in the 18th century, which constitutes the foundation of meteorology, numerous clever mechanical devices were invented to measure and record any and every meteorological (weather) parameter conceivable: ombroscope, rain gauge, thermometer, mechanical anemometer, wind speed indicator, remote readout wind vane, pole star recorder, radiosonde, etc (DeFelice (1998) and Brock & Richardson (2001)).

Meteorologically speaking, developments in NWP and climate models and related climate research are highly dependent on the quantity and the quality of the measurements that define the actual state of the atmosphere and earth surface. Wind speed and direction observations and mass distribution observations including the thermodynamic quantities (PTU) over the whole globe and at different atmospheric levels, constitute the basic and essential variables. The ensemble of these meteorological measurements obtained from ground-based techniques, airplanes and from satellites, are coordinated by the World Meteorological Organization (WMO), both with respect to methods and facilities, known as the Global Observing System (GOS). The GOS is of prime importance to the WMO for observing, monitoring and reporting on the weather, climate and the related natural environment conditions throughout the world. The GOS is used for the preparation of weather analyses, operational forecasts, hazard warnings, traffic advice, or provide advice on other services, etc. It is also a key asset for the earth's science community for climate monitoring and environmental activities, which are in general carried out under WMO programs and by other relevant international organizations. In NWP, GOS observations play a key role in the DA process, where an 'a priori' atmospheric state (forecast) is combined with the current observations in order to determine the most optimal or probable estimate of the atmosphere, known as the *analysis* (Daley (1993) and Ide (1997)). The analysis will then serve as an initial condition for a new forecast. The DA process is cycled over observation time windows, where an analysis is made in each window, followed by a short-range forecast to be used in the next observation window. This is explained and illustrated in detail in chapter 2.

For the effective development of the GOS, it is of interest to determine the expected relative benefit of mass and wind observations. The Rossby radius of deformation, R, defines the horizontal length scale above which the wind and mass field are in

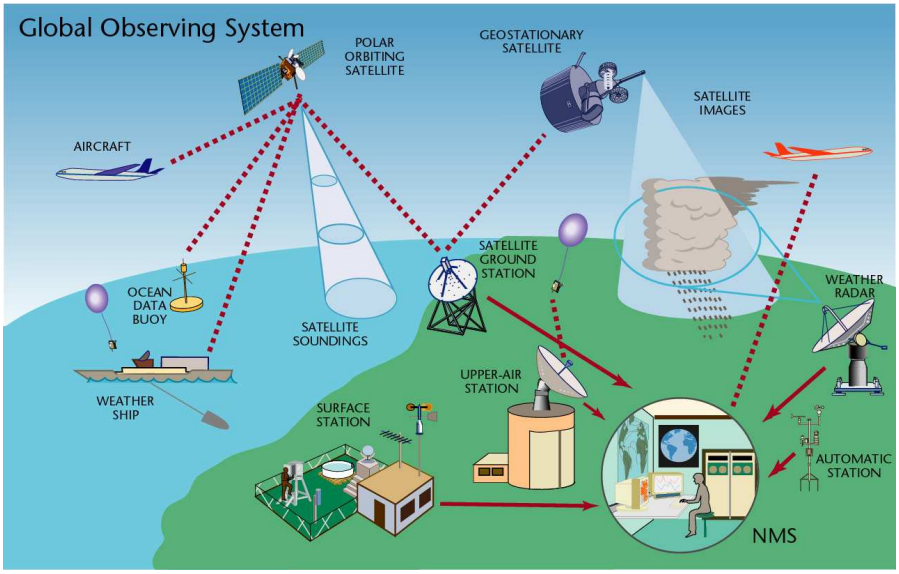


Figure 1.3: Overview of the different measurement instruments comprised in the Global Observing System (GOS)(credit to WMO).

approximate geostrophic balance. The limit is given as function of the vertical scales, as shown for instance in figure 1.4 for the mid-latitude case, where $R \sim 1000$ km. For the dynamic evolution of systems of length scales narrower than the R-limit, the wind and mass field are not directly coupled and independent wind observations are crucial. On these small scales the wind is the essential information and the atmospheric mass field adjusts to it. For motion systems of length scale larger than the R-limit, wind and mass field are coupled, so the wind adjusts to the mass field (Holton, 1992). In some regions, such as a tropics, wind is independent of the mass field, therefore wind cannot be derived from mass and direct measurements of wind are essential. The results of observation impact studies and the potential impact of new wind measurements can be interpreted physically using the Rossby radius of deformation (R), as defined in equation 1.1. The importance and deficiency of wind profile measurements in certain regions that necessitates direct measurements can be understood better through the physical analysis of the mathematical concept of the Rossby radius, as detailed below.

$$R = \frac{\sqrt{gh}}{2\Omega sin(\Phi)} \tag{1.1}$$

where g is the acceleration due to gravity, h the vertical scale of an atmospheric motion system, Ω the angular velocity of the Earth and Φ the latitude.

We can conclude with the following general statements: *i)* R goes to infinity at the equator, so information on the wind vector field is essential in the tropics as it governs tropical dynamics on all length scales. *ii)* In the extra-tropics and for small horizontal

scale features (length scales $L << R$; ~ 1000 km at the mid-latitudes) and deep vertical structures (large h), wind data are also the primary source of information. For large horizontal scale features ($L >> R$) and shallow vertical structures (small h), mass field information is most important. But it is important to measure the three-dimensional wind field directly also on very large scales in the extra-tropics even when mass field adjusts to mass field. Indeed, for very large planetary scales the geostrophic relation between mass and wind is not exactly valid, so small errors in mass field can generate large errors in the derived wind. *iii)* However, large scales evolve slowly, whereas the smaller scales evolve relatively fast. Therefore, to achieve more detailed meteorological analyses, besides the fact that we need wind profile observations, we would also need them more frequently in both time and space.

Wind dynamics and transport properties are essential in understanding weather and climate evolution. They are also essential in understanding small synoptic scale, mesoscale, and ultimately microscale exchange processes, such as convection, cloud formation, aerosols composition, transport of pollutants, etc. So, the provision of more complete global wind observations for the improvement of atmospheric flow analysis are recognized by the WMO as a prime need (WMO, 2002). The WMO has defined different wind profile measurement requirements (WMO, 2003) depending on whether they should be used for NWP or for the modeling of climate change and other applications. In addition, the continuous progress in detail of NWP models associated with the exponential progress in computing power leads to an increasing demand for high-density wind observations. Currently, most of three-dimensional wind measurements comprised in the GOS, are based on various instruments, as illustrated in figure 1.3. They can be classified in three data types (ESA, 2008): **i)** Single level surface data (e.g., synoptic reports from land stations and ships, drifting or moored buoys, scatterometer winds from satellites, etc.); **ii)** Multi-level upper-air observations such as radiosondes, profilers and operational polar orbiting satellite sounder data (e.g. ATOVS satellite). Notice that satellite sounders provide a global coverage of radiance data, which can only be used indirectly to define the mass field (temperature and humidity); **iii)** Finally, the single-level upper-air data which comprise mainly aircraft reports and Atmospheric Motion Vector (AMV) winds derived from geostationary satellite imagery, though aircraft data (wind and temperature) tend to be multilevel at airports, since more and more measurement are being made during takeoff and landing.

Among all the components of the current GOS, radiosondes and wind profilers are the only instruments that can provide vertical profiles of the wind field. Wind profilers constitutes still a rudimentary network, because of their cost and the difficulty in their deployment over the whole globe. The radiosonde launch sites are however providing regular measurements twice a day, at 00 and 12 UTC, and in addition at 6 and 18 UTC for some stations. Radiosondes can reach the higher atmosphere (up to ~ 35 km) while profiler measurements are generally below ~ 12 km. Due to cost, the radiosonde network has been reduced by about 50% in the nineties of the last century and spatial converge is scarce and insufficient to get a global and accurate picture of the dynamic state of the atmosphere. Indeed, the available data are mainly limited

to the continents and to the northern hemisphere, see coverage map in figure 3.1. Although the radiosonde network has been gradually reduced in recent years, the way of using the instrument has shifted from a launch at fixed location to a targeted launch at some locations, i.e., by targeting weather events using mobile stations. Galvin (2003). There is a tendency to automate the radiosonde launch, in particular in the developed countries.

Winds from satellite clouds (or water vapor) in images are derived from the motion of identified targets (e.g., cloud structures) where it is assumed that the target is advected by the atmospheric flow (wind). This intrinsic assumption is not always accurate, while in addition there is significant uncertainty in the determination of the height level of the observations. From basic physics (Rossby radius), it can thus be understood how Aeolus wind information will be important, in particular in some climate regions, as the tropics and extra-tropics. Indeed, this is because wind observations are in these cases more important than mass observations and wind profile data are very sparse in most of these regions. The major impacts are mainly expected on forecasting in the tropics, in particular of severe events such as tropical cyclones, and the prediction of mesoscale structures in the extra-tropics. For more details on the potential impacts of having additional wind information, in terms of weather and climate, see chapter 2.

1.4 Wind Profiles from Space: Aeolus

For a high measurement frequency, such as needed in weather applications, other techniques may be used to measure wind: propagation speed of ultrasound signals, by the effect of ventilation on the resistance of a heated wire, an anemometer based Pitot tubes, etc. The most common instrument for wind velocity or speed measurement is a cup anemometer, an instrument with rotating cups or propellers. An electrical device records the revolutions of the cups and calculates the wind velocity. Wind direction is reported by the direction from which it originates. For example, a southerly wind blows from the south to the north. However, remote sensing techniques constitutes the most considerable and major advancement in measuring the atmospheric, the oceanic and environmental parameters, since these techniques allow not only more precise measurements, but also continuous and dense spatial and temporal coverage. Remote sensing methods for wind include SODAR [1], Doppler LIDARs and RADARs[2]. They can measure the Doppler shift of electromagnetic radiation scattered or reflected off suspended aerosols or molecules. Radiometers and radars for instance, can be used to measure the surface roughness of the ocean from space or from airplanes where ocean roughness can be well associated with the wind vector close to the sea surface.; Geostationary satellite imagery can be used to estimate the winds throughout the atmosphere via feature tracking of, e.g., clouds or water vapour structures.

[1] SOund Detection And Ranging
[2] RAdio Detection And Ranging

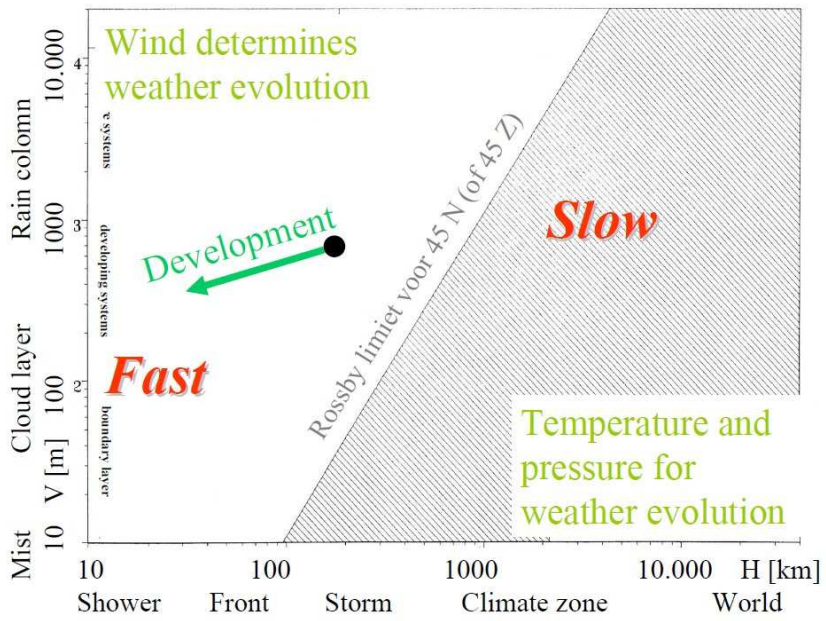


Figure 1.4: Vertical and horizontal scales of extra-tropic atmospheric Rossby waves. The Rossby radius of deformation (R) for a latitude of 45 ° (defined by the straight line) as a function of the vertical (depth) scale (h), defines the frontier between two ranges of atmospheric motions: Left side of R-line, the range within which the wind field dominates the atmospheric dynamics and thus three-dimensional wind measurements are important; Right side of R-line, the range where mass and wind are in geostrophic balance (Credit to ESA (2008)).

In terms of global ground-based wind profiles (mainly balloon radiosondes and very few radars), only a poor network exists. Due to the limited coverage (mostly Northern Hemisphere extra-tropics over land), measurements by satellite are essential to get a more uniform global coverage. The possibility to measure global wind profiles from space was first studied by NASA in the eighties of the last century, where it was finally concluded that an active optical system (lidar) could provide global data of the required accuracy. This led to the Laser Atmospheric Wind Sounder (LAWS) (WMO (1987) and Baker *et al.* (1995)) concept, with the main objective to provide wind Vector profiles for the benefit of NWP. However, it turned out soon that this LAWS concept would be extremely difficult to build and launch into space. In Europe, scientists followed similar ideas and, e.g., BEST, Bilan Energétique du Système Tropical, CNES (1988) has been proposed. At the same time, scientific studies were performed assessing the impact of DWL observations in DA systems widely used for NWP (Courtier *et al.* (1992) and Lorenc *et al.* (1992). The results of the latter study supported the basic idea that single-component wind measurements provided by a

space wind lidar would positively impact data assimilation and complement other existing observations. All the efforts supported by ESA during the past 20 years, to gather and combine knowledge from various studies concerning the evaluation of the prospects of using space-borne DWLs to measure global wind fields and impact NWP, gave birth to a space mission, commonly known as the Aeolus mission.

Three- and four-dimensional variational DA techniques (Courtier *et al.* (1998), Rabier *et al.* (1998) and Andersson *et al.* (1998)), in particular, are used in modern NWP. The main idea is that all observational information can be analyzed in a dynamically-consistent manner complementing the a priori information on the dynamics provided by the NWP system. The DA system is important, since it provides the horizontal wind component normal to the measured Horizontal Line-Of-Sight (HLOS) wind component probed by Aeolus, which is mainly the East-West component. This make it possible to derive a complete dynamic field. The prior NWP information also allows to correct for the relatively small temperature and pressure dependence (due to Rayleigh-Brillouin collision broadening of molecular light scattering, see Chapter 4 in ESA (2008) at the initial step of the assimilation, by using the temperature information provided by the short-range NWP model forecast. A dynamically-based assimilation system is thus a necessary requirement to produce meteorologically useful information from the observations provided by the Aeolus DWL system. A number of issues remain in the Aeolus mission, among them the optimal vertical sampling of the Aeolus DWl, which inspired most of the work in this thesis. See for a detailed description of the Aeolus mission chapter 2. The main issues of the Aeolus mission that are related to this thesis are summarized below.

Aeolus will sample the atmosphere from the surface to about 30 km height. This will be done with two receiver channels, one dedicated to measuring the Doppler shift of cloud and aerosol particles and the other dedicated to measure the molecular Doppler shift (chapter 2). Both channels are limited to 24 vertical measurement bins. This constraint in the vertical resolution of the Aeolus DWL may affect the beneficial impact of Aeolus products on NWP, particularly with the progressively increasing NWP model sampling in the vertical and the horizontal due to computing power growth. In order to optimize the wind retrieval and therefore gain the maximum benefit from Aeolus wind products, a detailed knowledge (i.e., including the small scales structures) of the atmospheric dynamics (wind and shear) and optics (clouds and aerosols) is needed. It will facilitate the selection of the most useful distribution of the 24 vertical range-bins after conducting simulations of Aeolus DWL measurements. Another aspect, which is related to the expected beneficial NWP impact, is to investigate the vertical resolution of the NWP model. GCMs are discussed later on in this chapter. One should generally strive to obtain an observation sampling that fits the GCM vertical resolution, which generally differs from the GCM grid size.

Given the complex heterogeneous optical and dynamical properties of the atmospheric in which the Aeolus DWL will operate, spatially unrepresentative samples may be obtained and errors in the wind measurements may be induced in some atmospheric conditions. The spatial and temporal distribution of aerosol and cloud

particles is essentially heterogeneous in the atmosphere and will moderate the Aeolus Ultra-Violet (UV) Lidar signal. Since UV can penetrate only thin clouds, it is important to investigate the effects of the Cloud Vertical Structures (CVSs) on the Aeolus DWL measurements. To study aerosol and CVS in association with dynamical atmospheric structures, various data sources have been used, mainly balloon radiosonde, the European Centre for Medium-Range Weather Forecasts (ECMWF) model and the Cloud-Aerosol Lidar and Infrared Pathfinder satellite Observation (CALIPSO) data sets.

1.5 Balloon-borne Radiosonde Systems and GCM

1.5.1 Balloon Radiosondes

A balloon sonde (or radiosonde) is an aerostat that is used in the fields of meteorology and of astronautics to characterize the thermodynamic and dynamic state of the atmospheric air. The sonde device with mainly three sensors placed on board, is carried aloft by a balloon filled with hydrogen, to measure the thermodynamic variables: pressure, temperature and humidity (or simply denoted as a PTU). Wind is derived from the successive positions of the balloon radiosonde, via the Global Positioning System (GPS), radio navigation, radar tracking when a radar reflector is integrated in the sonde or any other or combined radio-location system, known as Wind-Finding System (WFS), see section 1.5.2. The images shown in figure 1.5 give an overview of the different radiosonde instrument components. The collected weather observations at different atmospheric levels are transmitted by radio to ground stations, and reported in a format of alphanumerical codes by the World Meteorological Organization (WMO), known as TEMP (WMO designates here the numerical codes FM-35 and FM-36, successively, for surface and ship-based TEMPs). More details and formulas are given in the appendix A.3.

The first true radiosonde that sent precise encoded telemetry from weather sensors was invented in France by Robert Bureau, who gave the name "radiosonde" and flew the first instrument on January 7, 1929 (Javelle, 2000). A year later, on January 30, 1930, Pavel Molchanov flew a new design radiosonde which he developed independently. Molchanov's radiosonde became a popular standard because of its simplicity. For example, it converts sensor readings into Morse code Pope (1888), which makes it easy to use without special equipment or training. Notice that radiosondes designed only for wind speed and direction measurement (without PTU sensors) are called *rawinsonde* and radiosondes that can be dropped from aircraft are known as *dropsondes*. They are usually equipped with a parachute rather than a balloon. One of the most important characteristics of a radiosonde as compared to the other conventional observations is that it can reach altitudes of ~ 35 km or higher, covering the full troposphere and a large part of the stratosphere. With recent developments, radiosondes are more and more miniaturized and also become less expensive with respect to Rocket sondes, profilers or satellites, due market competition and the diversification of new manufacturers. Most of the modern radiosondes are mainly based on GPS WFS, and many issues that have been part of the earlier radiosondes

are resolved or improved, such as the localization accuracy, bias in measured temperature due to sun radiation heating the metallic armature that warms up the air before entering in the thermometer, humidity bias due to freezing, etc. The functioning of the balloon-borne radiosonde system from the release of the balloon until its burst in the stratosphere (generally up to $\sim$ 30 km) and then falling to the ground, is mainly based on three important components (figure 1.6): **i)** the radiosonde box with its PTU sensors; **ii)** the ground system, including a PC computer which is connected to a sounding processor subsystem (SPS311 or SPS220 in case of VAISALA systems) via a network adapter. The passive sounding controller contains the processor cards for PTU, GPS and/or VLF wind finding, and appropriate connections to the required antennas. **iii)** Finally the sounding software which is a radio-sounding data analysis, archival and relaying software for the sounding system (e.g., VAISALA MW31). The software consists of a main program and several optional extensions.

1.5.2 Wind-Finding Systems (WFS)

Besides the different radiosonde brands, radiosonde instruments may be categorized according to the tracking system used to measure wind components, i.e., WFS. For the data collected and used in this thesis work, we identified mainly 3 three types of WFS as shown in the right side of figure 1.6 and which are described in more detail in chapter 3. First, LORAN-C (LOng RANge Navigation -C type) which is a terrestrial navigation system using lower frequency radio transmitters. The Loran-C receiver is based on a cross-chain approach, by using transmitters from two Loran-C chains simultaneously. It is usually used to determine the position of a ship or aircraft; A second system is using the continuous worldwide navigation system, known as GPS; which is maintained by the US government. GPS WFS comprises a subsystem which is the GPS processor unit that calculates wind information from the GPS data processed and transmitted by the radiosonde. The GPS processor unit uses the GPS remote (from sonde) and GPS local (received at the station) signals to derive winds; The third tracking system consists in following the radiosonde balloon by a RADAR (RAdio Detection And Ranging) system. This system, known as radiotheodolite (e.g., Radiotheodolite RT20 of VAISALA in figure 1.6) and following the optical theodolite, operates in the frequency band 1600-1700 MHz. It is mostly used in the USA. For this WFS, wind is derived from the successive positions of the balloon at known time intervals. The position is obtained by measuring the azimuth angle and any two of the following parameters: elevation, slant range and height A.3. So, the accuracy of the wind product depends on the measurement accuracy of these parameters.

Table 1.1 shows the different characteristics of the measurements made by each system. Details on the measurement principle, including the results of data analysis from these three WFS types are presented in chapter 3. The most abundant high-resolution radiosonde data are from the SPARC (Stratospheric Processes And their Role in Climate) project and are based on radiotheodolite, and mainly VAISALA RS-80-57H (Finland) and Sippican (USA) brands. The other datasets, from BADC (British Atmospheric Data Centre UK), from AMMA (African Monsoon Multidisciplinary Analysis) project and De Bilt station (Netherlands) are VAISALA RS80 and

RS92 radiosonde types, based successively on LORAN-C and GPS WFSs. A preliminary analysis leading to the selection of the most useful data sets for the different studies reported in this thesis are shown in the appendix A.2. The overall coverage of the selected stations is shown in the map 3.1.

Radiosonde type **Features**	GPS receiver (RS92-SGP)	LORAN-C receiver (RS92-KL)	Radiotheodolite (RT20)
Telemetry			
Frequency band	403 Mhz	403 Mhz	300 kHz
Navigation accuracy			
Positioning uncertainty horizontal	10 m	/	/
Velocity measurement uncertainty	0.15 ms^{-1}	0.7 ms^{-1}	Better than 1 m/s with EL $angle >$ 17°
Directional measurement uncertainty	2 degrees		
Meteorological Sensors			
Temperature sensor			
Total uncertainty in sounding	0.5 °C	0.5 °C	
Resolution	0.1 °C	0.1 °C	
Humidity sensor			
Total uncertainty in sounding	5 % RH	5 % RH	
Resolution	1 % RH	1 % RH	
Pressure sensor			
Total uncertainty in sounding	1 hPa	1 hPa	
Resolution	0.1 hPa	0.1 hPa	

Table 1.1: Characteristics of different radiosonde types distinguished by the WFS used, GPS, LORAN-C or radiotheodolite. The characteristic numbers shown here are mainly for the VAISALA (Finland) brand, but are quite similar to the other radiosonde brands (e.g., Sippican, USA) of the same WFS type.

1.5.3 General Circulation Model (GCM)

A General Circulation Model (GCM) is a numerical model of mathematical equations describing the general circulation of a planetary atmosphere or ocean. It is mainly

based on the Navier-Stokes equations on a rotating sphere with thermodynamic parameters for various energy sources and sinks: radiation, latent heat, etc. All these equations constitute the basis of complex computer programs that are commonly used to simulate the atmosphere or ocean of the Earth. The atmosphere in GCMs is divided into layers. We briefly mention here that the model used in this thesis work is the ECMWF NWP model with its two versions, L60 and L91 (see figure 1.7) as defined in the ECMWF Integrated Forecasting System (IFS). Model levels (also often called σ-levels) contrary to standard pressure levels, are continuous atmospheric levels since they follow the contours of the Earth's surface in the lower and mid-troposphere. However, the effect of the surface orography decreases with altitude until the vertical layers in the upper atmosphere (i.e., in the upper stratosphere and mesosphere) become identical to pressure levels.

In February 2, 2006, the ECWMF model gained additional levels, passing from the L60 to the L91 version, and horizontal resolution, with their respective horizontal grids of 40 km (T511) and 25 km (T799). T511 and T799 denote a spectral truncation of spherical harmonics that span the earths atmosphere, where the numbers are inversely proportional to the horizontal grid (mesh) size. It is important to recall that the grid size (vertical or horizontal) is different from the "effective" resolution that defines the scales that can be resolved by the model and described realistically. For instance, it is shown in Stoffelen *et al.* (2008) that the "effective" horizontal resolution of the ECMWF L91 model is more than 5 times the horizontal grid resolution of 25 km. This difference in the effective resolution and grid size can reach a factor of 7 for some other models and depending on definition. In a NWP model the vertical resolution will generally be degraded with respect to the vertical sampling due to artificial high-order horizontal and vertical diffusion schemes. These are needed to avoid numerical instability in the discrete GCM representation. More details on the ECMWF model characteristics may be found in section 4.2.1 of chapter 4.

1.6 Overview of the Thesis

The topics addressed in this thesis concern wind and wind shear profiles, with a special focus on small-scale (i.e., high resolution) dynamics, including the cloud vertical structures. These are relevant in the framework of ESA's Aeolus mission, in line with the thesis title "*On high-resolution wind, shear and cloud vertical structures -Preparation of the Aeolus space mission-*". As mentioned in the previous sections, continuous and regular wind profile observations over a large part of the globe are still missing. Following the importance of the observed wind information for several domains and applications, e.g., aviation, sailing, etc., it is particularly relevant to improve NWP analyses, forecasts and climate models in order to better service these domains and applications. Moreover, providing better understanding of atmospheric dynamics will improve weather and climate forecasts. ESA's *Aeolus* space mission (described in details in chapter 2) aiming to provide global and continuous wind profiles using a DWL, is expected to fill this gap in wind profile data. However, a number of issues related to the principle of measurement of the DWL and the operational environment are studied within several ESA projects (Stoffelen *et al.*, 2009), such as: the optical

and dynamical heterogeneities of the atmosphere in the horizontal and vertical, the measurement Signal-to-Noise Ratio (SNR), the calibration stability, processing and DA capabilities, expected beneficial impact. All these issues are primordial and at interplay in the determination of the optimal strategy for the spatial and temporal sampling of the Aeolus DWL.

Among these issues that inspired most of this PhD thesis work, is the vertical sampling of the Aeolus DWL. Indeed, this issue is very important, because significant errors in the Aeolus wind measurements may occur, mainly due to the limited number of vertical atmospheric bins that can be sampled by the DWL and given the complex optics and dynamics of the atmosphere where the DWL will operate (see chapter 2 for more details). The spatial and temporal distribution of the aerosol and cloud particles is essentially heterogeneous in the atmosphere and thus moderates the Aeolus UV Lidar signal. So, this constraint in the vertical sampling of both Mie (scattering from clouds and aerosol) and Rayleigh (scattering from molecules) channels, may affect the beneficial impact of Aeolus products on NWP, particularly with high resolution models. However, in order to optimize the wind retrieval and therefore gain maximum benefit from Aeolus wind products, a detailed knowledge of the atmospheric dynamics (wind and shear) and optics (clouds and aerosols) has been gathered at KNMI. It will facilitate the selection of the most useful distribution of the 24 Mie and Rayleigh vertical range-bins after conducting simulations of Aeolus DWL measurements, wind processing and assessing impact. Technically, the range-bin distribution may be changed up to 8 times per orbit, thus allowing different settings in some different climate regions.

Aeolus DWL simulation software packages are developed within ESA projects. The outputs of the DWL simulations may be processed and used as inputs in the Aeolus L2A algorithm (developed at LMD/MF/DLR) dedicated to cloud and aerosols retrievals, and the Aeolus L2B algorithm being developed for wind retrieval (at ECMWF/KNMI/MF). In this perspective, the study of high-resolution wind profiles and Cloud Vertical Structures (CVS) as presented in this thesis is very relevant for the simulation of the Aeolus observations. This will in fact help to identify dynamically realistic (i.e., at small scales) atmospheric scenes with different vertical cloud structure configurations to be simulated with the Aeolus DWL simulator. This Aeolus data base is being used to optimize the vertical sampling of Aeolus and its wind retrieval performance. Thus the main research questions addressed in this PhD thesis are:

1. What are the best and most useful observations needed to investigate the atmospheric dynamics and optics at high resolution, which are relevant for the preparation of ESA's Aeolus space mission ?

2. What is the accuracy (bias) and precision (variance) of the various selected data sets? Is quality control needed ?

3. What is the difference in the representation of the atmosphere dynamics in terms of spatial and temporal variabilities between the observations (high resolution radiosondes) and GCM (ECMWF model) ? Similarly, the differences in the observed and modeled CVS by including CALIPSO satellite data ?

4. Which data sets can be used to build a global atmospheric data base at high resolution needed to simulate the Aeolus DWL realistically; GCMs, observations, or both ?

5. Can we determine reliable CVS profiles from radiosonde data which may be used to investigate CVS effects on the estimation of wind measured with Aeolus DWL at the different vertical height ranges ?

To address these main questions and many others, this thesis is presented in seven chapters including the general introduction and conclusion. The introduction describe the motivation of this PhD work that includes: the basic physics of wind dynamics, its properties, its representation and its importance in the GOS and GCMs and many other domains, and its measurement instruments. In the second chapter, we present the context of the PhD study, which is ESA's Aeolus space mission. A special focus is given to the vertical sampling issue of the Doppler Wind Lidar (DWL), that inspired most of the thesis work. In the third chapter, an effective statistical quality control (QC) method is described that has been applied on the high-resolution radiosonde data, extensively exploited in this thesis. This QC method has been developed to clean the data from unrepresentative observations (outliers) and is presented and illustrated graphically. Also, differences in the accuracy of various types of radiosonde WFSs (radio theodolite, LORAN and GPS) are highlighted, which is in particular very important for the derivation of wind shear but also for deriving the vertical wind component.

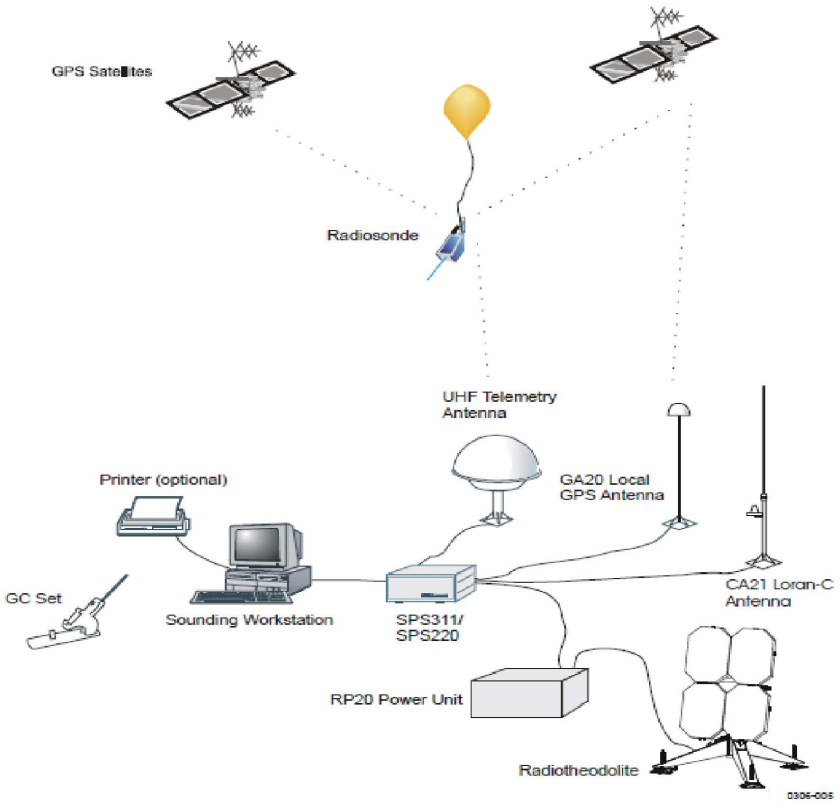
In chapter four, we present the collocation procedure of radiosonde observations with short-range forecasts of the ECMWF model that allowed a detailed description of the atmospheric dynamics and its representation in global NWP models. Spatial variability of wind, wind shear and radiosonde-balloon drift over different climate regions (tropics, subtropics, mid latitudes and the poles), in terms of averages (mean and median) and variances (σ, percentiles, etc.), are studied. Also, systematic vertical sampling effects on wind and wind shear are Investigated, both for single radiosonde profiles and statistically. An original method is developed that quantifies the substantial difference between observed and modeled wind and shear and determines the "effective" vertical resolution of the ECMWF model. The method is based on the wind shear representation of radiosonde profiles on various scales as compared to the ECMWF model. Chapter five is an extension of chapter four and covers the temporal (monthly, seasonably and yearly) variability of wind and wind shear of both radiosonde observations and ECMWF model in different climate zones, including their respective quantifications. This includes the quantification of the differences in wind and shear variabilities between land and sea and between dawn and dusk. We present also how to combine the high-resolution radiosonde observations, which have a limited temporal and spatial coverage, with the global but relatively smooth ECMWF model. The purpose of this datasets combination is to create a global atmospheric dynamics data base at high resolution (i.e., including small scale structures) needed for Aeolus simulations. This data base includes atmospheric optical properties from CALIPSO data collocated with ECMWF model dynamics, which is enhanced to provide the observed variability.

In Chapter six we present an assessment of vertical cloud structures as a function of altitude, in terms of vertical distribution and occurrence frequencies. This is important for Aeolus, since the DWL cannot penetrate thick cloud layers. A methodology is proposed to estimate the effect of CVS on wind measurements over different vertical atmospheric range-bins by combining the information of wind and CVS profiles from radiosonde. However based on CVS profiles, this methodology suggest the derivation of the Cloud Liquid Water Content and Cloud (CLWC) Cloud Ice Water Content (CIWC), which is crucial. The overall results are summarized in the conclusions, with a focus on their applications and their implications in Aeolus-DWL simulations to optimize the vertical sampling and the wind retrieval. This is followed by an outlook.

Figure 1.5: Images illustrating the following: the radiosonde package and its various components (mainly pressure, temperature and humidity (PTU) sensors), and the different stages of preparation of the radiosonde and balloon until the launch

.

a) Balloon-Radiosonde system components



b) Overview of balloon radiosonde in operation

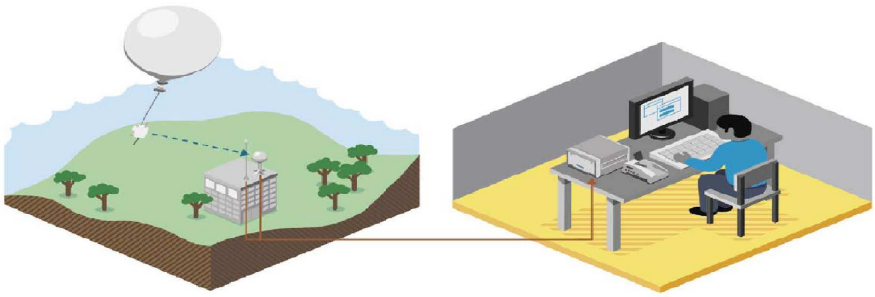


Figure 1.6: a) An overview of the various components that compose the balloon-borne radiosonde system: the balloon and radiosonde package (see figure 1.5), the ground processing system (PC computer connected to a sounding processor subsystem via a network adapter) and radiosonde sounding software; notice at the right of the picture the two other types of systems that can be used for wind retrieval (LORAN-C and radiotheodolite), in addition to GPS. These are known as Wind-Finding Systems (WFS). b) Balloon radiosonde in operation (Courtesy of ESA).

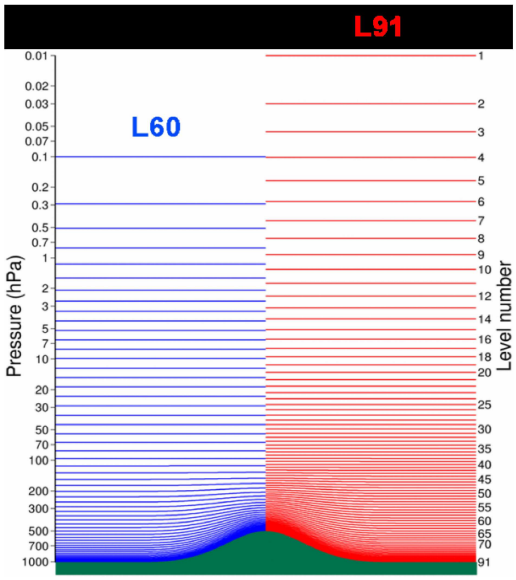


Figure 1.7: Correspondence between the two versions of the ECMWF GCM vertical model levels, L60 and L91. The new L91 GCM which is operational since February 2, 2006, shows an enhanced sampling compared to L60, in particular in the troposphere

Aeolus Atmospheric Dynamics Mission

CHAPTER 2

Aeolus Atmospheric Dynamics Mission

Abstract
A space-borne atmospheric dynamics mission , called Aeolus, is under preparation at the European Space Agency (ESA) since 1989. It was selected in October 1999 as the second core Earth Explorer mission within ESA's Living Planet Programme. The mission is designed to make direct measurements from space of wind profiles from the earth's surface up to about 30 km height using a Doppler Wind Lidar (DWL). The launch is scheduled for early 2017. One of the main objectives of the mission is to remedy the lack of wind data over a large part of the earth's atmosphere, data which are necessary to improve Numerical Weather Prediction (NWP) and better understand climate dynamics (Stoffelen *et al.*, 2005). In this chapter we give a detailed overview of the mission, principle of and theoretical background on the measurement technique of a DWL, and the potential benefit of Aeolus data products on NWP and climate modeling and related researches. We focus in particular on the vertical sampling of the Aeolus DWL, which is the issue that inspired most of this thesis work. In fact, retrieving accurate atmospheric wind may be challenging in cases of complex optical and dynamical heterogeneity of the atmosphere, mainly because of a limitation in the number of vertical range bins to 24. On the other hand the range bins measured with the Aeolus DWL can be set flexibly and may be changed about 8 times per orbit. This issue is addressed by DWL simulations and discussed in more detail at the end of this chapter.

2.1 Introduction

Remote Sensing (RS) systems may be used on ground, but also on airplanes or spacecraft, where the latter category is generally referred to as Earth Observation (EO). EO uses electromagnetic (e.m.) waves with wavelengths from ultraviolet (UV) to microwave (MW). E.m. waves may be sensed from natural sources by so-called passive instruments, but also from synthetic sources by active instruments. The Aeolus DWL

is an *active sensor* and emits UV radiation to the earth's atmosphere and analyzes the signal that is returned. Active instruments employ a reflection or *scattering technique* since the return signal is either reflected from a mirror surface or scattered back to the satellite by a rough surface or atmospheric particles. A main interest in active techniques is in the possibility of sensing the signal strength as a function of the time delay between transmission and reception of the signal, a so-called *ranging technique*. Radar, Radio Detection And Ranging, and Lidar, Light Detection And Ranging, are both used in EO RS as active scattering techniques to profile resp. rain or cloud with MW and cloud, aerosol and/or molecules with UV or optical wavelengths. Both Radar and Lidar techniques may in addition sense the wavelength change, i.e., so-called Doppler shift, between emitted and received signals, which is a measure of the movement of the scattering particles in the direction away from (or towards) the instrument. The Aeolus DWL senses both the strength of the backscattered signal and the Doppler shift or movement of air particles.

The potential of DWL in measuring wind fields has been first assessed in the frame of the atmospheric laser Doppler instrument report (ESA, 1989). In a joint collaboration between the US and Europe, some ideas on a space-borne DWL were carried forward and presented in a workshop in 1995 (ESA, 1995a). These two reports lead towards the foundations of the Aeolus atmospheric dynamics mission, first expressed in the Aeolus "Report for Assessment" (ESA, 1996), in which a first feasibility assessment on the successful use of a space-borne DWL for meteorological application was provided. The ESA Aeolus mission consists in a satellite carrying a DWL called ALADIN (Atmospheric LAser Doppler INstrument Lidar) and is planned for launch in 2017. The High Spectral-Resolution Lidar (HSRL) will measure global wind-component profiles from the earth's surface up to about 30 km altitude. The instrument is based on a direct-detection Lidar incorporating a fringe-imaging receiver (analyzing aerosol and cloud backscatter) and a double-edge receiver (analyzing molecular backscatter), see table 2.2 for an overview and the components and features of the Aeolus DWL. By processing the backscatter signals, line-of-sight (LOS) wind-component profiles along the satellite track may be produced. Aeolus will as such provide consistent observations of wind profiles over the whole globe, including remote areas lacking ground-based weather stations, such as over the oceans, tropics and southern hemisphere. This will provide meteorologists with much desired information for improving weather prediction. This will be possible above thick clouds or down to the surface in clear sky. In thin clouds or at the top of thick clouds wind information will also be available. In addition to wind information, information on other elements like clouds and aerosols can also be extracted. The Aeolus data will be disseminated to the main NWP-centres in near real-time (NRT). This wind data will improve significantly both coverage and accuracy of wind profile observations.

In section 2.2, we first recall the relevant theoretical background of some physical phenomena involved in the measurement technique of a DWL . In the same section we then give more general information about the Aeolus space mission, and more specifically about the operation of the Aeolus DWL. In section 2.3, we focus on the issues related to the principle measurement of the DWL, mainly the vertical sampling

		PBL	Troposphere	Stratosphere
Vertical Domain	[km]	0-2	2-16	16-30
Vertical resolution	[km]	0.5	1.0	2.0
Horizontal Domain		global		
Number of profiles	$[hour^{-1}]$		> 100	
Horizontal Integration length	[km]		50	
Accuracy (HLOS Component)	[m/s]	1	2	3
Timeliness	[hour]		3	
Length of Observational Data Set	[yr]		3	

Table 2.1: Main Aeolus observation requirements as derived from World Meteorological Organization (WMO) requirements (ESA, 2008).

of the DWL, that inspired most of the work done in this thesis and in the framework of the VAMP (Vertical Aeolus Measurements Positioning) project. Finally, in section 2.4 we note expected benefits from Aeolus in NWP and climate models and related research.

2.2 Doppler Wind Lidar (DWL)

In this section we recall the relevant theoretical background related to the measurement techniques of a DWL in general, i.e. Lidar (or Laser) profiling, scattering theory and Doppler effect phenomena. Then, more specifically, we summarize the measurement principle and features, etc., for the case of the Aeolus DWL.

2.2.1 Laser Profiling

One of the simplest and a recurrent application of the Lidar is laser profiling (or laser Altimetry), which is conceptually straightforward. The instrument emits a short pulse of light (UV, visible or near-infrared radiation) towards the earth's surface, and its reflected "echo" is detected some time later. By recording the time delay and knowing the speed of propagation of the pulse, the range (distance) from the instrument to the surface can be determined. The transmission of a continuous stream of pulses allows to build up an along-track profile of the range, and if the position of the platform as a function of time (track) is accurately known the surface elevation profile may then be been deduced. To construct a typical laser profiler as shown in figure 2.1, the measurement system should be composed of three components: the *transmitter*, the *receiver* and the *interval timer*. The first component is typically a semiconductor laser, usually Nd:YAG (Neodymium: Yttrium-Ammonium-Garnet) or GaAs (Gallium- Arsenide), that is capable of producing a short (of the order of 1 ns) and intense pulse with a small angular width. While the second component (the receiver) is a photodiode or Charge-Coupled Device (CCD) that detects the return signal (pulse). Finally, an interval timer with a resolution of 1 ns, which is activated by the generated signal and deactivated by the return signal, giving thus the measure

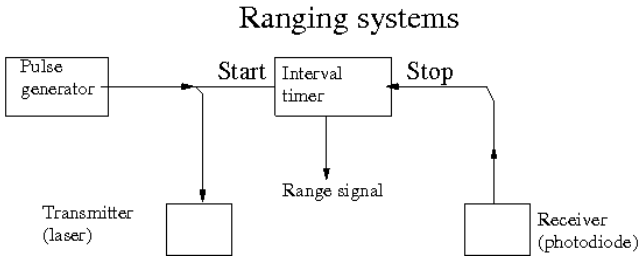


Figure 2.1: Schematic construction of a laser profiler

of the time delay (eq. 2.1). All these steps are summarized in a simple schema in figure 2.1

$$T = \frac{2H}{v_g} \tag{2.1}$$

where H is the range and v_g is the group velocity of the pulse.

The error caused by setting $v_g \approx C$ in eq. 2.1 is small. In fact the propagation in dry air of the group velocity for optical and near-infrared radiation differs from the speed of light (C) in vacuum by at most 0.03% (near the surface). The travel time accuracy ΔT_t, is normally governed by the rise time t_t of the received pulse and its signal-to-noise ratio SNR. if V_s is the voltage amplitude of the received pulse and V_n the amplitude of its variation due to noise, the signal-to-noise ratio (voltage) is defined as $SNR = V_s/Vn$. So, the greatest accuracy with which the timing of the received pulse can be determined is given by eq. 2.2:

$$\Delta T_t = \frac{t_t}{SNR} = \frac{t_t}{V_n} V_s \tag{2.2}$$

Generally, the expected features of such system is to achieve a high spatial resolution at the surface, i.e, the sampled points should be close together, high range resolution, and a sensitivity that is great enough to detect signals returned from surfaces of weak reflectivity. A more extensive description of laser profiling may be found in the text book of Rees (2001).

2.2.2 Scattering Theory: Light and Atmosphere Interactions

In scattering theory, light scattering can be divided into three types depending on the electromagnetic wavelength (λ) and the size of the particles interacting with the light. This may be characterized with a dimensionless size parameter, α which is defined in eq. 2.3 as follows:

$$\alpha = \frac{\pi . d_p}{\lambda} \tag{2.3}$$

where $\pi.d_p$ is the particle circumference and λ is the wavelength of incident radiation. Based on the value of α, three domains are defined as follows:

$\alpha << 1$: Rayleigh scattering (small particle size compared to the wavelength of the incident light)

$\alpha \approx 1$: Mie scattering (particle size about the same as wavelength of light)

$\alpha >> 1$: Geometric scattering (particle size much larger than wavelength)

Both Mie and Rayleigh scattering are considered elastic scattering processes [A1]. In elastic scattering the wavelength or frequency of the light is not substantially changed. However, electromagnetic radiation scattered by moving scattering particles does undergo a frequency Doppler shift, which can be detected and used to measure the velocity of the scattering particles by lidar and radar techniques . The Aeolus DWL design is based on these two types of scattering, Mie and Rayleigh, measured by two dedicated receivers within the ALADIN instrument 2.2. The complimentary characteristics of the Mie and Rayleigh return signals will enable an accurate wind profile measurement over the entire altitude range. The Doppler frequency shift can be measured by two types of measurement techniques, namely, coherent heterodyne systems and the direct detection, interferometric, systems. In the case of the Aeolus mission, the latter technique appears to be more suitable (accurate). This is due to the fact that for the direct-detection interferometric systems, the accuracy depends only on the total scattered signal energy, while heterodyne detection depends on the energy of individual pulses, requiring high-energy laser technology (Rees, 2001). There are in fact very profound differences of principle in their respective operation. Coherent heterodyne systems operate by beating the scattered and Doppler shifted radiation with an optical laser oscillator at the surface of a detector. The resultant electrical beat-frequency signal is analyzed post-detection to produce the Doppler frequency. On the other hand, in the direct detection methods the optical signal field is analyzed and dispersed in an interferometric filter prior to detection. Both systems require interferometric precision in the optical manipulation of the signal beam. However, due to the very different physics of these two schemes, the performances are very different in principle. For heterodyne systems the key parameter is shown to be the photon degeneracy - that is the number of photo-detections per optical mode (i.e., in a single coherence area and coherence time). In low-backscatter conditions, this requires that the available laser power to be distributed into pulses of the largest possible energy. However, for direct detection, the accuracy depends only on the total scattered signal, possibly integrated over many pulses.

2.2.3 Doppler Effect

The Doppler effect (or Doppler shift) is a phenomenon that was proposed for the first time by the Austrian physicist Christian Doppler in 1842 (Doppler, 1842). It is the change in frequency of a wave for an observer moving relative to the source of the wave. it can be observed whenever there is relative motion between a source of waves (e.g. sound, water or light waves) and the observer . In 1845, the hypothesis was tested

[1] In the case of light waves, the Doppler effect is equal if either the source or the observer is moving. But for sound, because sound waves require a medium for propagation, there is a difference

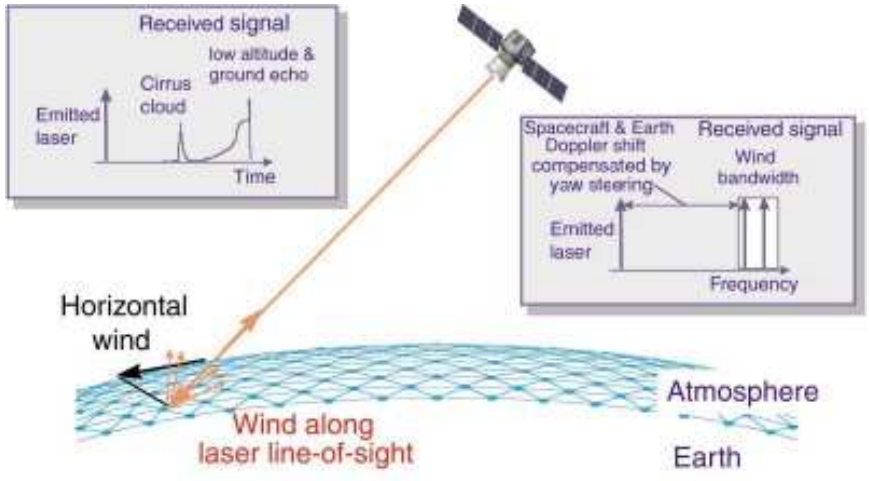


Figure 2.2: Doppler Wind Lidar principle (courtesy ESA)

and confirmed for sound waves by Buys Ballot, and in 1948 for the electromagnetic waves by Hippolyte Fizeau. The shift of a wave's frequency is caused by the relative motion of an observer and the wave source. This motion causes the frequency of the wave to increase (e.g., higher sound's pitch in case of sound wave) as the source and observer approach each other and to decrease (e.g., lower sound's pitch) as they recede. In the limit where the wave's speed is much greater than the relative speed of the source and observer, and which is often the case with electromagnetic waves (e.g., light, $V << C$), the relationship between observed (the resultant) frequency f (eq. 2.4) and emitted frequency f_0 is given by Werner (2005):

$$f = f_0(1 + \frac{V}{C}) \tag{2.4}$$

where V is the velocity at which the observer is approaching to or receding from the source, and C is the speed of light. So, under a Doppler shift, the optical frequency of light is shifted by a factor of V/C.

The Doppler shift actually detected by a DWL system is the result of two Doppler shifts. The first shift in frequency is induced by the scattering air particles being investigated, which constitute a moving observer w.r.t. the DWL. The second shift arises because the particles moving with the air subsequently act as moving sources, scattering the Doppler-shifted light. Since the sources are moving with respect to the lidar system (now a stationary observer), a second Doppler shift, f' (Eq. 2.5) is seen on the already Doppler-shifted light with frequency f.

$$f' = f(1 + \frac{V}{C}) \tag{2.5}$$

whether the source or the observer is moving.

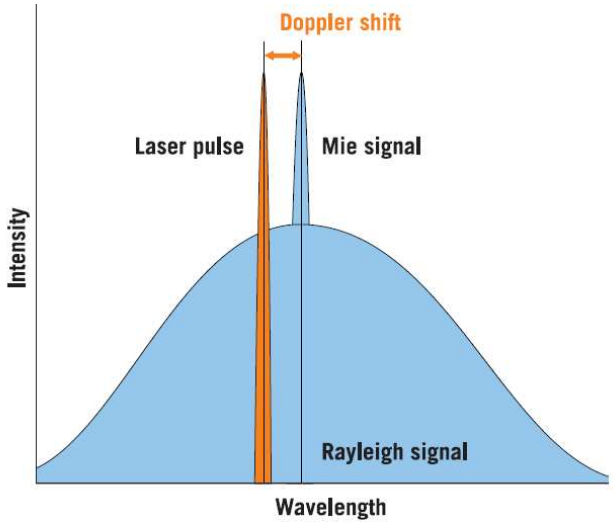


Figure 2.3: Mie and Rayleigh Doppler shift (courtesy ESA)

The Doppler frequency shift detected back at the source (Eq. 2.6) is given by $\Delta f_D = f' - f_0$. Since $V << C$, the shift Δf_D may be written:

$$\Delta f_D = 2 f_0 \frac{V}{C} = \frac{2V}{\lambda_0} \tag{2.6}$$

where λ_0 is the laser wavelength.

Figure 2.3 illustrate an example of wavelength shifted Rayleigh and Mie spectrum.

2.2.4 Aeolus-DWL Measurement Principle

The Aeolus satellite will carry a high Resolution Spectral Lidar (HRSL) flying in a low sun-synchronous orbit at 400 km height and with a 6:00 LST equator crossing time will always be close to a dusk or dawn position. The lidar measures the wind component along the laser line-of-sight (LOS), using a slant look angle of 35° with nadir, as illustrated in figure 2.4a. The wind will be measured at 90° to the satellite ground track towards the night side of the Earth. It has been demonstrated that with this fixed viewing geometry most of the scientific requirements could be met (ESA, 1995b). Notice that in autumn of 2010 it has been decided to move from operation in "burst" mode (i.e., measurement of 50 km ($\sim 7s$) track at the start of every BRC of 200 km ($\sim 28s$) as shown in 2.4a,b), to "continuous" mode; knowing that the satellite velocity is about 7 km/s. Figure 2.4b shows one 50 km-long wind observation in more detail, which is divided into individual measurements of 3.5 km (0.5 s), leading to 14 measurements. One measurement corresponds to 50 laser pulses, while one observation corresponds to 700 laser pulses. However, this discontinuous measurement mode, known as "burst" mode is abandoned by ESA in late 2010 and "continuous" measurement mode (i.e. the laser remains always switched on) is now favored due to more stable laser performance.

As described previously in this section, the measurement principle of the DWL is based on three fundamental physical notions, laser profiling (or ranging technique), scattering and Doppler effect. In the case of the Aeolus mission two types of scatterings are involved, Mie and Rayleigh. The first refers to light scattering by particles of about the same size as the laser wavelength (λ), such as cloud, aerosols, hash, etc.; the Aeolus DWL operates in the Ultra-Violet (UV) range of the electromagnetic spectrum at 355 nm. The second refers to light scattering from particles much smaller than λ. Mie and Rayleigh scattering are targeted in two distinct receiver channels. The vertical resolution in both receiver channels is limited to a maximum of 24 vertical range bins, but flexible (i.e., bin sizes may be modified up to eight times per orbit on a weekly basis). An example of a vertical distribution of Mie and Rayleigh range-bins (vertical layers) is shown at the left side of figure 2.4a, where we can see that the molecular (Rayleigh) channel reaches higher than the Aerosol/cloud (Mie) channel because of the absence or rarefaction of aerosol and clouds above a certain altitude.

The general principle of the functioning of Aeolus DWL (ALADIN) is illustrated in 2.5. The DWL emits an ultraviolet laser light pulse (355 nm) towards the atmosphere with a known Pulse Repetition Frequency (PRF = 100 Hz) and which interact with air particles and molecules that meet in their paths. A small fraction of the laser light is scattered back at a shifted frequency (Doppler effect, see section 2.2.3) towards the instrument and detected by a receiving telescope. The comparison of the frequencies of the emitted and received signals, allows the determination of the Doppler shift. In addition to the Doppler shift due to wind velocity (scattering particles driven by winds) along the LOS, the received laser signal is Doppler-shifted also relative to the emitted laser due to spacecraft movement and earth rotation. Both frequency shifts are however compensated by utilizing a proper satellite attitude-control scheme (and a satellite yaw steering mode). After subtraction of the known satellite and earth motion components the residual shift in the frequency is proportional to the speed of the backscattering particles/molecules moving with the wind, which may be determined as follows:

$$V = \frac{\lambda_0}{2}(f_0 - f') \tag{2.7}$$

with λ_0 and f_0 resp. the wavelength and the frequency of the emitted laser light and f' the received frequency of the backscattered light, i.e., following Eq. 2.6.

The frequency shift is related directly to the wind velocity along the laser beam, see Werner (2005).This is a measure of the wind component along the laser Line-Of-Sight (LOS), i.e., $V = V_{LOS}$. The Horizontal-LOS wind, V_{HLOS}, is then derived as follows:

$$V_{HLOS} = \frac{V_{LOS}}{sin(\varphi)} \tag{2.8}$$

with φ the local incidence angle. Using the LOS azimuth angle, V_{HLOS} may be expressed in the easterly and northerly wind components, resp. u and v, as follows:

$$V_{HLOS} = -u.sin(\psi) - v.cos(\psi) \tag{2.9}$$

As shown in the next section, V_{HLOS}, may be estimated at each vertical range bin.

Parameters	*Value*
Transmitter	
Wavelength	355 nm
Pulse energy	130 mJ (120 mJ goal)
Repetition rate	50.5 Hz
Line width	50 MHz
Duty cycle	100%
Receiver	
Fizeau line width (Mie)	184 MHz
Double Fabry-Perot (Rayleigh)	1.666 GHz
Line width	2 GHz
Spacing 2.3 pm	5.48 GHz
Detector quantum efficiency (Mie/Rayleigh)	82%
Signal detection (Mie receiver)	Silicon CCD detector in accumulation mode
Signal detection (Rayleigh receiver)	Silicon CCD detector and two read-outs
Signal Processing	
Altitude range (Mie + Rayleigh)	-1 to +26.5 km (extendable)
Vertical resolution	1 km (adjustable)
Horizontal accumulation length	3.5 km (adjustable)
Opto-mechanical Subsystem	
Telescope diameter	1.5 m

Table 2.2: Technical specifications of the Atmospheric LAser Doppler INstrument (ALADIN) of ADM-Aeolus mission.

2.2.5 From Laser Light to LOS Wind Profiles

As illustrated in figure 2.2 and stated in the previous section, Aeolus-DWL wind measurements start by emitting laser pulses with a known frequency towards the atmosphere. A small part of the light is scattered back and frequency-shifted (Doppler effect) by air molecules and/or particles moving with the wind velocity at different vertical layers. The backscattered light is detected by a large telescope (see 2.2) and directed to two dedicated channel receivers, the particle (Mie) and molecular (Rayleigh) channels. The Mie channel consists of a Fizeau spectrometer and an Accumulation Charge-Coupled Device (ACCD) detector, and is mainly sensitive to the backscatter from aerosol and cloud particles. The Rayleigh channel employs a dual-filter (known also as double-edge) Fabry-Perot (FP) interferometer and an ACCD

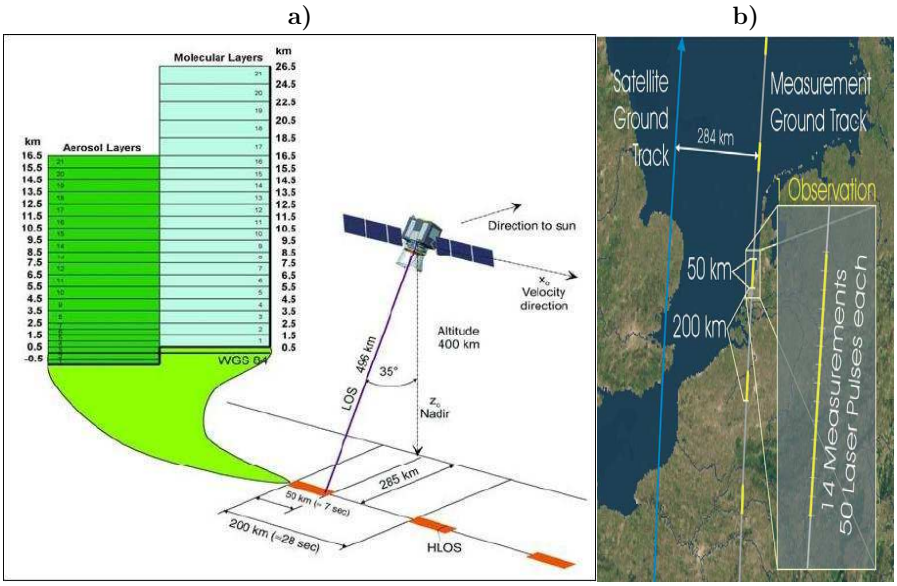


Figure 2.4: Principle measurement profile and sampling concept of the ALADIN instruments of ADM-Aeolus mission: a) Baseline Aeolus measurement geometry; b) Ground measurement sample (courtesy ESA). This measurement sampling mode is currently outdated, because ESA decided late 2010 to abandon this discontinuous mode (known as "burst" mode) and to switch to "continuous" mode.

detector, and is mainly sensitive to the molecular backscatter. Both the interferometers act like spectral filters and the spectra are imaged at the ACCDs detectors to determine the shift in the frequency (Doppler shift, see 2.7) and the LOS wind (2.9).

Using light-scattering particles in the air (cloud/aerosols and molecules) as tracer, LOS wind speed is determined as a function of height ranges (bins) by integrating over vertical bins (see eq. 2.10) of varying depth.

$$V_m(i) = \frac{\int_{z_i}^{z_{i+1}} S(z)V_t(z)dz}{\int_{z_i}^{z_{i+1}} S(z)dz} \tag{2.10}$$

with V_m the mean measured wind inside the bin, V_t is the "true" wind, and $S(z)$ is the total energy of the return signal as a function of height z, which can be seen as the number of photons detected at the ACCD for each channel, either Mie or Rayleigh. Recall Table 2.1 with the WMO requirements for the Aeolus DWL measurement vertical resolution and the maximum of 24 vertical range bins (ESA, 1995b). Given this limitation in vertical sampling for both channels and given the often complex

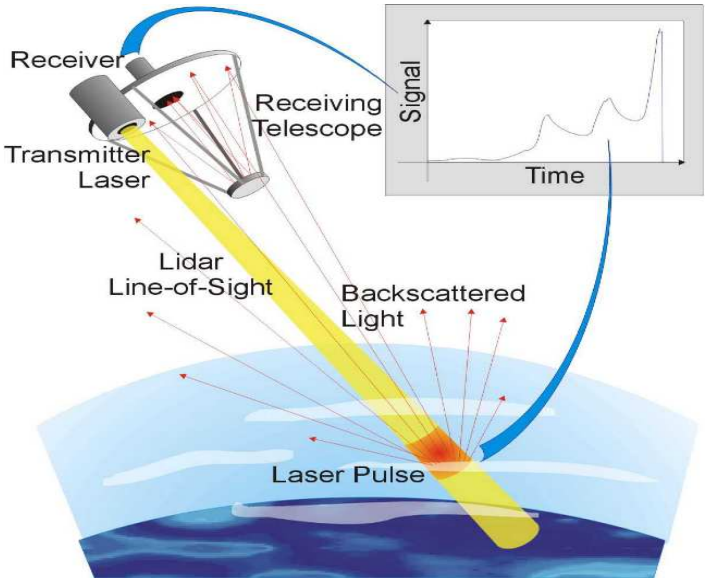


Figure 2.5: Measurement principle of Doppler wind lidar (DWL); DWL emits a laser pulse towards the atmosphere, then collects samples and retrieves the frequency of the backscattered signal by clouds, aerosols, air molecules and the Earth's surface. The lidar measures the wind component along the laser line-of-sight (LOS), using a slant angle off nadir (courtesy ESA).

optical and dynamical heterogeneity of the atmosphere, errors in wind estimation inside the bin may occur. This is discussed in section 2.3.

2.3 Optimal Vertical Sampling by Aeolus

2.3.1 Optical and Dynamical Atmospheric Heterogeneity

The atmosphere as shown in figure 2.6 is globally heterogeneous in terms of dynamics (wind, shear, turbulence, etc.) and in terms of optical properties (molecules, cloud, aerosol, smoke, etc.). These air molecules and particles will scatter and Doppler-shift the light signal emitted by Aeolus-DWL from space. The small fraction of light scattered back towards the instrument detected by a telescope is exploited to determine the wind over different range bins sizes with a resolution varying with altitude (2.1). An example of a distribution of these range bins (for Mie and Rayleigh channels), as established in the frame of VAMP project among other possible distributions, is shown in figure 2.7. Regarding the principle measurement of Aeolus and under specific dynamically and optically heterogeneous atmospheric conditions, Aeolus wind measurements will be challenging. In fact, very sharp vertical gradients of the horizontal wind (i.e., wind shear) will be a major phenomenon that Aeolus may fail to

represent well, mainly because of the limited vertical sampling of the DWL (i.e., can not measure more than 24 vertical range gates in both Mie and Rayleigh channels). The jets close to the tropopause for instance, are typically narrow in the vertical and associated with a large variability of the horizontal wind in the vertical. So, higher resolution (shallower) range bins than the proposed 1 km depth (2.1) is necessary to well depict realistic jet patterns. Aeolus wind measurements will be possible above thick clouds or down to the surface in case of cloud-free conditions. Also, in thin clouds or at the top of thick clouds wind information will be available, as well as information on clouds and aerosol.

It is shown in Stoffelen & Marseille (1998) that cloud and vertical wind shear occurrences are not statistically correlated, but this may need further elaboration at the cloud scale. In the presence of combined wind shear and cloud and/or aerosol, large spatial representation errors may occur due to vertically non-uniform $S(z)$ in Eq. 2.10. Depending on the position of the cloud and/or aerosol inside the bin, the return signal from Mie and Rayleigh may originate from different heights (Height assignment errors), and both channels may observe different wind in case of significant wind shear. It is already shown in the ESA MERCI project (Stoffelen *et al.*, 2002) that within a 1-km bin, mean Mie signal height may vary significantly with an average RMS variation of 100 m against 25 m for Rayleigh. For instance, largest height assessment errors will occur when optically thick cloud/aerosol exist within the range gate. For instance, considering a common value of the tropospheric wind shear of about 50 ms^{-1} per km combined with the 100 m RMS Mie height assessment uncertainty, this would result in wind shear errors of 5 ms^{-1}, which can be quite detrimental in NWP analyses and forecasts. In this context, characterizing the climatology of wind and wind shear (dynamics) paves the way to define strategies to optimize the distribution of the vertical bins for an optimal use of Aeolus wind information for NWP and climate. This starts with the identification of the most complex and challenging scenarios from global climatologies that describe most accurately the dynamical and optical properties of the atmosphere, as provided by different data sources. In the frame of the ESA VAMP project (see next section various data sources are used to provide dynamical and optical climatologies: high-resolution radiosonde observations, ECMWF model, CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation) and Cloud Resolving Models (CRM). Identified and selected scenarios are used for Aeolus DWL simulation, e.g., the End-to-End simulator (E2S) (ASTRIUM, 2011a), LIPAS (Veldman *et al.*, 2000), in performance simulation studies in order to prepare the processing algorithms for the Aeolus mission.

2.3.2 Atmospheric Database and Aeolus DWL Simulations

As stated previously, the limited vertical sampling capability of the Aeolus DWL to 24 range gates make wind measurements very challenging in some specific complex atmospheric conditions. These challenging situations, that may cause errors in the retrieved wind measurements, concern mainly regions of strong vertical wind shear combined with high optically heterogeneous atmosphere, as detailed in the previous

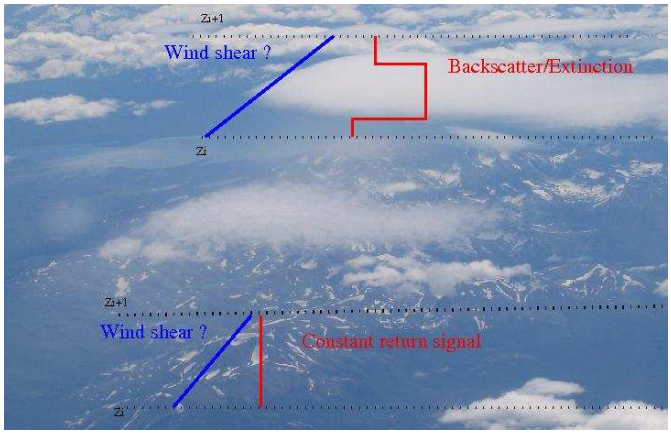


Figure 2.6: Illustration of the potential effect of clouds on Aeolus wind measurements, in particular in the presence of strong wind shear. In this last case, high vertical resolution is suitable to minimize the errors in wind measurements. Image from the plane over Halifax mountains, Canada

section 2.3.1 However, one of the useful features of the Aeolus instrument is the possibility of re-sizing of the vertical bins, i.e. increasing or decreasing the resolution, up to 8 times per orbit. The sampling scheme must be predetermined and is uploaded to the satellite on a weekly basis. This allows to optimize the sampling of different climate zones by varying the distribution and size of the range bins taking into account the typical climatological geophysical conditions. To do so, the Aeolus DWL lidar End-to-end Simulator (E2S), has been developed within ESA project. Other simulator software from previous projects is also available including the DE2S developed at DLR (Germany) and the lidar performance analysis simulator (LIPAS) developed at KNMI (Netherlands) (Marseille & Stoffelen, 2003). However LIPAS simulator is proven to be more efficient and faster and therefore it is extensively used in VAMP studies for a statistical evaluation of retrieved winds from many orbits. The output of the E2S is used as input to develop and test the Aeolus L1B/L2A/L2B processing algorithms. The objective of the L1B/L2B processing software is to retrieves the Aeolus mission primary wind product and is developed by DLR, ECMWF and KNMI. And the objective of the L2A processor, developed at LMD, France, is to retrieve mission secondary products such as cloud and aerosol detection and their properties. The L2A and L1B/L2B algorithms, which use the outputs of the E2S, as inputs, are assessed and adjusted in order to improve their performances. A description of the steps that form the Aeolus simulation and processing software chain (E2S/L0/1A/1BP/L2A/L2B/L2C) can be found in the respective reference documents (ASTRIUM (2011a), ESA (2011), ASTRIUM (2011b), Tan *et al.* (2007b) and Tan *et al.* (2008)). The L2C actually is a part of the ECMWF IFS model. The test system constructed at KNMI to run the complete simulation and processing chain is described in de Kloe (2011) and the latest results are documented in de Kloe (2012). A preliminary selection of the possible vertical sampling scenarios is given in de Kloe *et al.* (2007).

Besides the simulation tools for testing the Aeolus wind quality, an accurate representation of the dynamics and optical properties of the atmosphere is required to conduct the Aeolus simulations in order to optimize the spatial and temporal sam-

pling of Aeolus DWL. Notice that a special focus is on the vertical sampling because of the flexibly that the DWL instrument offers for the vertical binning, as stated above. Various datasets have been considered to design a database that well represents the dynamical and optical properties of the atmosphere, including ECMWF model fields, radiosondes, CALIPSO backscatter data and Cloud Resolving Models within the ESA VAMP project (Stoffelen *et al.*, 2009). This resulted in the construction of a consistent and detailed atmospheric database with a global coverage of atmospheric dynamics(from the ECMWF model) combined with detailed atmosphere optical properties (from CALIPSO). The relatively smooth model winds were enhanced by adding the variabilty of small scales structures as derived from a statistical intercomparion of wind-shear obtained from radiosondes and the ECMWF model, see (Marseille *et al.*, 2011). From this database, challenging regions with complex heterogeneous atmospheric scenes that may induce errors in the retrieved winds, and their frequency of occurrence have been identified. In particular, the combined occurrence (correlation) of large wind-shear and backscatter variability within the observation scene (typically 50 km along the satellite track for the Aeolus burst mode concept and flexible accumulation range for the selected continuous mode concept and 1 km vertical bins) gives large wind errors. In order to minimize the errors in Aeolus winds in such scenes and thus maximize its performance for the retrieval of good quality wind, different vertical sampling scenarios has been determined accordingly. For instance, cirrus clouds in the tropics are found at altitudes up to 18 km. To minimize wind errors in these regions, the Mie channel (measuring aerosol and cloud particles) bins should be available here at the highest possible resolution (minimal bin size). But this should not be at the detriment of the vertical coverage of the other parts of the atmosphere, thus a compromise should be always made in order to get an accurate wind measurement over the whole atmospheric levels. Different vertical sampling scenarios have been simulated and assessed with the simulation tools and the most promising selection is discussed in Marseille *et al.* (2010). Furthermore, the impact of the various vertical sampling scenarios for Aeolus on NWP, with a special focus on stratospheric dynamics, has been investigated in the operational ECMWF environment (Körnich, 2010). This study showed mainly how different vertical sampling scenarios may improve the forecast quality, e.g., a better description of the stratospheric dynamics might improve intra-seasonal forecasts of European winters. Indeed adapting the vertical sampling of Aeolus to typical geophysical conditions improves the overall wind quality, which in turns will maximize the benefit of Aeolus products. It is noted that conclusions from the VAMP study are currently revisited due to the change of Aeolus burst mode operation to continuous operation (Stoffelen *et al.*, 2008).

Further investigations that relates high-resolution wind profiles and Cloud Vertical Structures (CVS) are discussed in this thesis, and which is partly implemented in 6. This is very relevant for the simulation of the Aeolus observations and testing of the L2B processor, since it will in fact help to identify dynamically realistic (i.e., at small scales) atmospheric scenes with different vertical cloud structures to be observed by Aeolus. CVS are needed to quantify how often correlated wind variability and cloud backscatter variability occur within Aeolus bins. Because correlated wind and backscatter variability within an Aeolus bin results in errors in retrieved winds due to

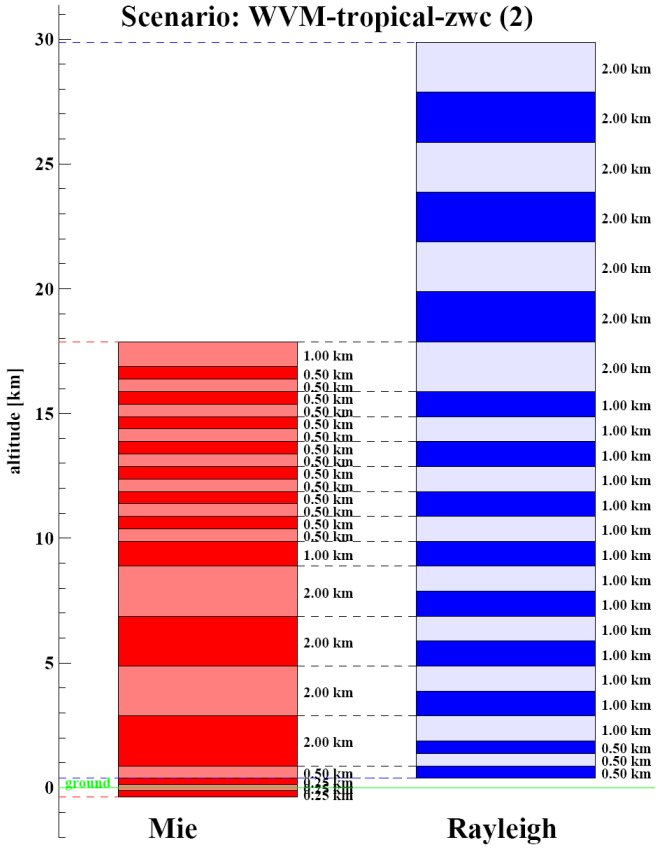


Figure 2.7: An example of a distribution of Mie and Rayleigh channels, where the focus (i.e., increase of resolution) is in the PBL and at the tropopause for Mie and in the troposphere for Rayleigh

the fact the level L2B processor assumes the cloud to be homogeneously distributed within the bin. The database of combined high-resolution wind and cloud backscatter can be used to quantify the error in Aeolus wind processing, and can be thus used to improve the level L2B processor designed for Aeolus wind retrieval.

2.4 Expected Aeolus Benefits

Atmospheric wind vector data are one of the most important meteorological assets in our daily life, but they are the most important information lacking in the current Global Observing System (GOS), as described in section 1.3. Measuring winds from space is fundamental for the improvement of weather and climate prediction and related research, as demonstrated in Baker *et al.* (1995), Stoffelen *et al.* (2005) and in many other publications. Some of the important benefits expected from ADM-Aeolus mission are summarized below.

2.4.1 Numerical Weather Prediction (NWP)

In Numerical Weather Prediction (NWP), discritized mathematical models describe an initial state of the atmosphere (weather) and its evolution to future states, usually within the short-term of few days [1]. The estimation of the initial state is provided through a data assimilation process [2] consisting of *Analysis cycles*, where observations within the analysis time window are combined with the *a priori* forecast (Daley, 1993). The result of this procedure, called *Analysis* and which is considered as 'the best' estimate of the initial state of the system, is used as initial conditions to run a new forecast. The new forecast is then used in the next *Analysis cycle*. An example of the most used and advanced data assimilation technique, i.e., the 4DVAR, is illustrated in figure 2.8. Any credible procedure of preparing the initial state of the atmosphere for NWP models is highly dependent on the number and the quantity of the observations available. From this perspective, global observations of wind profiles of Aeolus, will have a major impact by improving the NWP models in many aspects, as already demonstrated in many studies: Stoffelen *et al.* (2006), Žagar *et al.* (2008), Tan *et al.* (2007a), Marseille *et al.* (2008c), Marseille *et al.* (2008b) and Marseille *et al.* (2008a). In general, the main expected benefits from ADM-Aeolus on NWP are summarized here:

- better accuracy in the determination of the initial state and improved skill in weather prediction.
- improved tropical circulation will lead to better description of the energy and water cycle and will improve our understanding of the interactions between wind, temperature and humidity, therefore improve, e.g., the description of the Hadley circulation in the tropics.
- by improving the accuracy of the initial state in NWP and with improved modeling of atmospheric dynamics, it will be for instance possible to estimate the position and intensity of (tropical) cyclones with a better precision, thus improving extreme weather forecasts.
- the percentage of failures in the short-range forecasts will be dramatically reduced with the introduction of ADM-Aeolus observations. Some remaining failure cases in predicting synoptic events in the Northern hemisphere, e.g., such as strong mid-latitude storms, are likely to be resolved by ADM-Aeolus observations. Also, small-scale details of intense wind events will improve for short-range forecasts because of the earlier detection of their development.
- Medium-range forecasts will be improved in both hemispheres with a better definition of the planetary-scale waves, and more uniform coverage of the Earth by wind profiles, when Aeolus provides such observations in otherwise data sparse areas, particularly over the oceans, in the tropics and southern hemisphere.

[1] The limited number and the uneven spread of observations available around the globe leave always some uncertainty in the estimate of the initial state. In practice weather prediction skill is generally limited to about a week, due to the error growth in NWP models due to chaotic behavior in the atmosphere.

[2] 4D-VAR is the most common data assimilation technique used in NWP nowadays

– ADM-Aeolus may serve the quality control for other sources of remote sensing data, e.g., against the old or new generation of vertical temperature and humidity sounders. This will lead, indirectly, to improvements in forecasting skill, in the Southern Hemisphere in particular, where remote-sensing data are the primary source of information.

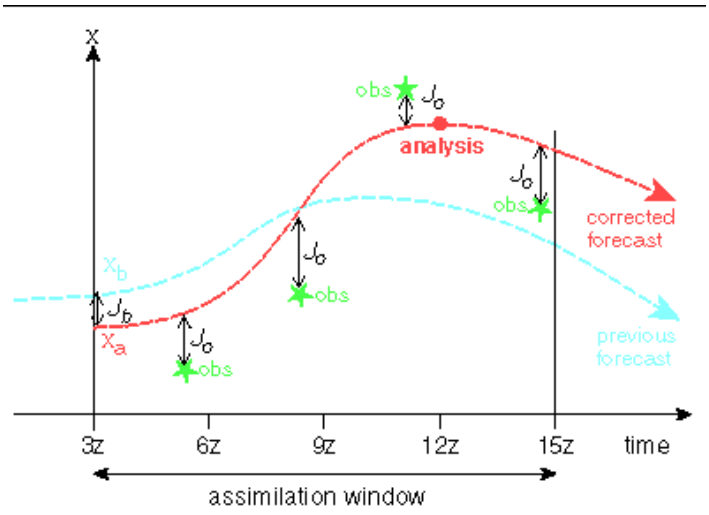


Figure 2.8: A schematic illustrating the 4DVAR analysis in the assimilation windows of 12 hours (courtesy ECMWF).

2.4.2 Climate Modeling and Research

A General Circulation Model (GCM), also known as a global climate model, uses the same equations of motion as a NWP model, but their purpose is to numerically simulate changes in climate as a result of slow changes in some boundary conditions (such as the solar constant) or physical parameters (such as the greenhouse gas concentration). While NWP models are used to predict the weather in the short (1-3 days) and medium (4-10 days) range, GCM's are run for longer periods, for years or long enough to learn about the climate in a statistical sense (i.e., the evolving mean and variability). A GCM should accurately depict the movement and evolution of disturbances; such as frontal systems and tropical cyclones, as good as a NWP model may do. But due to model error growth, after a few days the model output becomes useless as a weather forecast. The quality of a GCM is judged, mainly, by the quality of its tropical or extratropical climatology. Aeolus will provide due to its wind information, data concerning the transport of atmospheric constituents like ozone or aerosol, e.g., their stratosphere/troposphere exchange, which will provide for example a more accurate picture of the Earth's global energy budget. Improved transport modeling will be valuable for climate modeling and research , e.g., better

understanding of the global atmospheric circulation and related features such as precipitation systems, the El Niño and the Southern Oscillation phenomena. Some other expected impacts of Aeolus winds on climate studies are summarized below:

- improving the recording and monitoring of climate in different parts of the world
- Build better circulation models to predict climate based on improved modeling of dynamical processes; leading to improved projections of global warming and the effects of pollution on climate, etc.
- Aeolus provides globally horizontally-uniform height-resolved information on wind speed and direction and at different layers of the atmosphere. Aeolus provides also information on cloud and aerosol layers in the atmosphere. This information offers scope for climate studies.
- with improved representation of transports in atmospheric models the spatial distribution of atmospheric constituents becomes more accurate. This will aid in the validation and calibration of sources and sinks in atmospheric chemistry.
- as addressed in CLIVAR (Climate Variability and Predictability) research programme within the World Climate Research Programme (WCRP), more wind observations such as from Aeolus will help the understanding of physical processes in the climate system which are responsible for climate variability on time scales ranging from seasons to centuries. For example, Aeolus vertical wind profiles will be very useful for the study of: the El Niño and Southern Oscillation system (ENSO) and monsoon systems in the tropics, and
- ADM-Aeolus products will help in scientific diagnostics of large-scale processes, in particular in the tropics where the wind field is a critical dynamical variable (independent of the mass field), and because of the otherwise almost complete lack of direct observations in this region.

Statistical Quality Control of High-resolution Winds of different Radiosonde types for Climatology Analysis

CHAPTER 3

Radiosonde Quality Control and Windfinding Systems

Abstract
Quality control (QC) is among the most important steps in any data processing. These steps are here elaborated for high vertical resolution radiosonde datasets, which were gathered and analyzed to study atmospheric winds. The data base is composed of different radiosonde wind-finding systems (WFSs), including Radiotheodolite, LORAN-C and GPS. Inspection of this data base, particularly for wind, wind shear and ascent height increments (dz), showed a non-negligible amount of outliers in Radiotheodolite data as compared to the two other WFSs, thus denoting quality differences between the various systems. An effective Statistical QC (SQC) is then developed to isolate and eliminate outliers from the more realistic observations. Improving the accuracy of the Radiotheodolite WFS is critical to the derivation of the vertical motion and the vertical gradients of the horizontal wind, i.e., wind shear, mainly because of the direct dependence of these quantities on dz. Based on the climatological distribution of the QC-ed dz, a new approach is suggested to estimate these wind quantities for Radiotheodolite data. The approach is validated with the high-quality modern WFSs (LORAN-C and GPS). Although initially of reduced quality, applying SQC and using the climatological mean dz of $12s$ smoothed Radiotheodolite profiles, shows very good improvement in the climatological wind analyses of Radiotheodolite WFSs. Notably, the climatologies of ascent rate, vertical motion, horizontal wind and vertical shear now look comparable for the various WFSs. Thus, the SQC processing steps prove essential and may be extended to other variables and measurement systems.

3.1 Introduction

Radiosonde wind vector (speed and direction) are derived from the successive horizontal positions of the radiosonde balloon. Thus, the accuracy of the observed wind is intimately related to the accurate measurement of the radiosonde position. Position calculations are performed using different Wind-Finding Systems (WFS).

Here we focus mainly on three WFS types: Radiotheodolite, LORAN (LOng RAnge Navigation) and GPS (Global Positioning System), which are used to collect the high-resolution radiosonde database used in the context of the preparation of the European Space Agency's (ESA) mission -Aeolus Atmospheric Dynamic Mission- (Stoffelen *et al.*, 2005). This database presented here and illustrated in figure 3.1, followed a quality control (Houchi *et al.*, 2015) and is utilized in Houchi *et al.* (2010) and (Marseille *et al.*, 2011). An overview of the measurement principle of the three WFSs is given in subsection 3.23.2.2.

Although SPARC (Stratosphere-troposphere Processes And their Role in Climate) Radiotheodolite data are subject to a strict Quality Control (NCDC, 1998) at the University Corporation for Atmospheric Research and Joint Office of Science Support (UCAR/JOSS), a non-negligible amount of unrealistic observations (outliers) are observed in these data, as compared to LORAN-C and GPS radiosonde WFSs. This indicates mainly the difference in the accuracy between these three radiosonde WFSs. SPARC data are widely used in many studies, (Wang & Geller, 2003), (Gong *et al.*, 2008), etc., but with limited references to Quality Control (QC) aspects. The goal in this chapter is to present a Statistical QC (SQC) method for winds, wind shear and ascent height increments (dz) of high-resolution radiosondes, and demonstrate its effectiveness in isolating and removing outliers from a realistic probability-density distribution (PDF) of the observations. Because of the moderate accuracy of the Radiotheodolite data as compared to LORAN-C and GPS WFS data, the derivation of the meteorological quantities depending on the ascent height increments (dz) is very challenging. Thus, to estimate such quantities for Radiotheodolite-based SPARC data, two new approaches are proposed in this chapter. The first approach is to estimate the vertical motion based on the ascent rate statistics, and the second is to establish wind shear climatology. Both approaches which use the statistics (mean and percentiles) of the QC-ed dz, are verified and validated against more accurate measurements of LORAN-C and GPS WFSs; from AMMA (Africa Monsoon Multidisciplinary Analysis) projects, BADC (British Atmospheric Data Center) and De Bilt station (figure 3.1).

The different high-resolution datasets used and their related WFSs are presented in section 3.2, this includes an overview of the initial QC of SPARC data applied at UCAR/JOSS. In section 3.3, subsequent analysis of the different radiosonde WFS types shows the pronounced differences in the distribution of vertical differences (du, dv) of the horizontal wind components (u,v) and ascent height increment, dz (i.e., the sampling interval). Then, sets of typical observed raw profiles of horizontal wind and their corresponding shear profiles for various WFSs are shown. In section 3.4, the developed SQC method is described in detail, while it is graphically illustrated in section 3.5 for the different analyzed quantities: u and v components, their vertical shears and dz. This includes the comparison of z and dz statistics for the different WFSs. In section 3.6, approaches are proposed to statistically estimate i) the ascent rate and vertical motion and ii) wind shear. Overall results and benefits of the SQC processing steps are summarized in section 3.7.

Figure 3.1: The geographical locations of analyzed high-resolution radiosonde datasets: SPARC (circles), BADC (hexagrams), AMMA (diamonds) and De Bilt (Square) as function of climate regions, successively for the tropics (red), subtropics (blue), midlatitudes (black) and Polar (magenta). Note the orography (brown) in the map which may explain the appearance of lee waves for some stations, in particular over the Rocky Mountains. The right legend bar from zero meter and up indicates the altitude of the earth surface from sea-level; below sea level is masked white.

3.2 High-resolution Radiosonde Data Base

In this section, we present the high-resolution radiosonde database that we utilized and the different WFSs used to track and collect data from balloon radiosondes. Special attention is paid to the different WFSs, thus, derivation methods of vector winds for each WFS are described. An overview of the quality of the ensemble datasets is presented.

3.2.1 High-resolution Radiosonde Data and QC Status

The KNMI (Royal Netherlands Meteorological Institute) high-resolution database is being developed in the framework of the ESA Aeolus mission with radiosondes, initially from the SPARC project dataset (Chanin, 1995), and subsequently extended to more datasets from: BADC, AMMA and De Bilt. But it is still being extended with other data sources, e.g., from Méteo-France and Finland. In this study, the first four datasets have been exploited, i.e., SPARC, BADC, AMMA and De Bilt station. The SPARC dataset is based on Radiotheodolite, while there is mixing of LORAN-C and GPS WFSs for the other datasets. Radiosonde data are available at high temporal resolution: 6 s (~ 30 m) for SPARC data, 2 s (~ 10 m) for BADC and AMMA, and 10 s (~ 50 m) for De Bilt in 2006. Values between brackets denote the corresponding spatial vertical resolution, by considering the approximate radiosonde-balloon ascent rate of 5 m s^{-1}. The balloon is designed to rise at a more or less constant rate, because as the air density and thus friction per area decreases, its cross section gets larger (Brock & Richardson (2001), WMO (2008b)).

Moreover, SPARC data available for the period 1998-2006 are quasi-regularly given at 00 and 12 UTC in ASCII files. The data records contain 22 fields as follows: time from release, pressure, temperature, dew point, relative humidity, u and v wind components, wind speed and direction, ascent rate, balloon position data and altitude. BADC and AMMA radiosoundings of $2s$ raw resolution constitute a collection of about 21 stations each. The former are mainly dedicated to operational weather prediction, while the latter are dedicated to a campaign study of the African monsoon. Contrary to the analyzed SPARC period (1998-2006), these two datasets appear very irregular and, e.g., unavailable for several years. Thus, after assessment of the distribution in time of all available BADC and AMMA radio-soundings, only few stations appear to be useful for climatology studies, i.e., those sounding stations that are regular and continuous in time. Over 2006, for instance, only 9 BADC and 6 AMMA

stations (available at both 12 and 00 UTC) were selected as valuable. At 6 and 18 UTC, the number of useful stations is even much lower in 2006, 2 to 6 stations for BADC and only 1 station for AMMA, depending on the range of the targeted period. Also the data from station De Bilt are analyzed, which are regular. This station uses both LORAN-C (during the night) and GPS (during the day) WFSs.

For SPARC Radiotheodolite data sets of both radiosonde brands (Sippican and Vaisala), used to derive winds at $6s$ sampling, it is observed that elevation and azimuth angles are biased by elevation angle oscillations, which occur occasionally. Consequently, this leads to large oscillations in wind velocity, at low elevation angles in particular. Therefore, these measured parameters, together with the PTU (pressure, temperature and relative humidity) quantities, are QC-ed at the UCAR/JOSS, to remove outliers before deriving the wind components. A full description of the QC method is available in NCDC (1998), but we succinctly summarize the most important information on the joint UCAR-JOSS QC. The data underwent a two-stage QC process as follows:

1. QC-1. An internal consistency check that includes two inspections: "reasonable" limit checks on all parameters as defined by WMO (DiMego *et al.*, 1985); and rate-of-change checks on temperature, pressure and ascent rate (Schwartz & Govett, 1992) and (DiMego *et al.*, 1985).

2. QC-2. Each sounding endured a visual examination verifying those parameters that are too variable for automatic checks: wind speed, wind direction and relative humidity. In addition, the verification of the former automatic QC flag (QC.1) is allowed at this stage.

Regarding the QC of the other WFS data sets, based on LORAN-C and GPS, no quality control has been applied on these raw data. Nevertheless, at first view their quality appears to be much better than the Radiotheodolite SPARC data, as can be seen in the next section 3.3. So, the limited accuracy in most of data based on Radiotheodolite, called for quality enhancement measures and the development of a new Statistical Quality Control (SQC) to get rid of outliers. Although initially developed for SPARC data, this statistical method will also be applied to the more sparse non-SPARC datasets.

3.2.2 Radiosonde Wind-Finding Systems

As mentioned previously, radiosonde WFSs use different ways to determine the horizontal position of the balloon and, consequently, the derivation of the wind components (u, v, w) is also different. For the vertical position, apart from GPS WFS, the balloon vertical ascent height is generally determined by using the hypsometric relation (Eq. 3.1):

$$z_{i+1} - z_i = \frac{R.T_v}{g}.ln(\frac{P_i}{P_{i+1}}) \tag{3.1}$$

where z_i is the geopotential height at level i [m], R the specific gas constant for air (287.06 [J kg^{-1} K^{-1}] for dry air), $T_v = [T_v(i+1) + T_v(i)]/2$ is the mean virtual

temperature [K] for the layer between two successive isobaric surfaces at pressure levels P_i and P_{i+1} and g is the gravitational acceleration [m s^{-2}], with a standard value of 9.80665 according to WMO technical regulations. Note that Eq. 3.1 allows the determination of the ascent rate and vertical wind after removing the estimated net balloon lifting (see end of section 3.5).

The Theodolite WFS is one of the first and oldest WFS. Initially the instrument was optical, which makes the height to which the balloon can be tracked dependent on many factors, such as the speed of winds aloft, obstruction in the light path between the balloon and the observer, mainly cloud, and the bursting point of the balloon. During a night observation, radiosonde tracking is accomplished by attaching a small self-illuminating light stick or a small battery-powered lighting unit to the balloon (OFCM, 2006). Fortunately, since the introduction of the Radar, the balloon tracking has become much more simple and accurate, hence the name Radiotheodolite. The Radiotheodolite system requires the determination of the azimuth angle and two of the following three quantities, height, elevation angle, and range, during the balloon ascent through the air until it bursts. With this information, the balloon position at different atmospheric levels can be determined and, consequently, the wind components. However, the measurement of the range is only possible when the radiosonde is equipped with a reflector, to be tracked by the radar.

LORAN is a globally-distributed terrestrial radio-navigation system using low-frequency radio transmitters, introduced in 1958. The detection system uses multiple transmitters, known as hyperbolic positioning or multi-lateration. In case of the radiosonde, 3 transmitters are generally used to determine the position and the speed of the balloon receiver (Fang, 1986). LORAN-C operates in the low frequency range of 90 to 110 kHz, and it is widely used in many countries; to determine for instance the location of an aircraft or a ship. LORAN, which started to lose popularity with the invention and growing use of GPS, currently tends to become more popular again; mainly to serve as backup and land-based alternative to GPS or any other Global Navigation Satellite Systems (GNSS). LORAN-C is the latest version before the recent introduction of the modernized version, named E-LORAN (Enhanced-LORAN). A prototype of E-LORAN was introduced in 2008 (Basker & Williams, 2010)). The interest for this type of navigation system grows with the introduction of the new modernized version, E-LORAN, which is independent, dissimilar and complementary to GNSS (Int. Loran Ass., 2007).

GPS is currently the most widely used radiosonde WFS. The principle of GPS is quite similar to LORAN-C, but with the multi-lateration method applied spatially in three dimensions. GPS receivers locate at least 3 satellites and estimate the range to each. This information is then used to deduce the receiver location in the earth's atmosphere, with an accuracy of $0.5m$ in the horizontal and $30m$ in the vertical (Brock & Richardson, 2001). In case of radiosonde balloons, the procedure is repeated at regular time intervals to determine the successive locations of a balloon and thus derive the wind vector. Many available networks, using radiosonde WFSs different from GPS, are switching progressively to this modern GPS WFS. In the United States for instance, the entire Radiotheodolite radiosonde network is progressively replaced

with modern GPS radiosondes in recent years, as planned in the Replacement Network System Project (Fitzgibbon *et al.*, 2005). More detailed information on all existing WFSs can be found in WMO (2008a). The vertical wind shear vector, **s**, is defined as the gradient of the horizontal wind vector $\mathbf{v} = (u, v)$ with respect to vertical height z (Eq. 3.2):

$$\mathbf{s}_i = \frac{\mathbf{v_{i+1}} - \mathbf{v_i}}{z_{i+1} - z_i} \tag{3.2}$$

with i indicating the level number of wind or wind shear (each shear level is in between 2 wind levels), $i = 1, ..., m-1$ for shear levels, and m the total number of vertical wind levels, which depends on the vertical sampling used, u and v are the zonal and meridional wind components. Wind shear levels are plotted as a function of shear heights, $(z_i + z_{i+1})/2$.

3.3 Wind-Finding Systems Comparison

3.3.1 PDF Profile of Wind and Height Differences

To get a first idea about the measurement accuracy of each WFS used in our data base, we derive the difference between successive values in the ascent profiles of zonal (du) and meridional (dv) winds and ascent height (dz), denoted here 'difference profiles'. Then, the probability distribution functions (PDFs) of individual profiles of these differences are determined, using bin sizes of 1 m s $^{-1}$ for du and dv, and $10m$ for dz. For a fair comparison, both BADC and AMMA datasets of $2s$ are re-sampled at the vertical resolution of SPARC data of $6s$. However, De Bilt data are kept at their raw vertical resolution of $10s$. The right plots of figures 3.2a,b illustrate these PDFs, where the relative frequencies of occurrence of u, v and z differences within the mentioned range bin sizes are depicted. From these figures, we can clearly see a large dispersion in SPARC Radiotheodolite data as compared to the other WFSs (e.g., shown for GPS in De Bilt in figure 3.2b) for the three analyzed quantities. We can see in particular that only in SPARC radiotheodolite data outliers exist, i.e., many values that fall far away from the main mode of the PDF. Though some dispersion may be attributed to meteorological events, such as extreme weather conditions in some climate regions, gravity waves at high altitude, etc., many values appear unrealistic, i.e., are outliers. It is important to mention that these findings for each WFS are very typical for most individual profiles. This indicates that Radiotheodolite data suffer from a relatively large number of outliers; and therefore an additional QC is necessary. It appears that above about 22 km the accuracy in dz measurements is severely decaying due to a digitization issue in the SPARC Radiotheodolite data, as show in figure 3.2a. It is verified that this digitization issue originates from the pressure profiles, which are used together with temperature profiles to derive geopotential heights (ascent heights), z (see Eq. 3.1). The Radiotheodolite WFS is further known to have a limited accuracy for elevation angles below 17 degrees (Vaisala, 2002). This is due to the interference between signals received directly from the radiosondes and those received by reflection from adjacent surfaces, generally termed multipath interference (WMO, 2008b). The

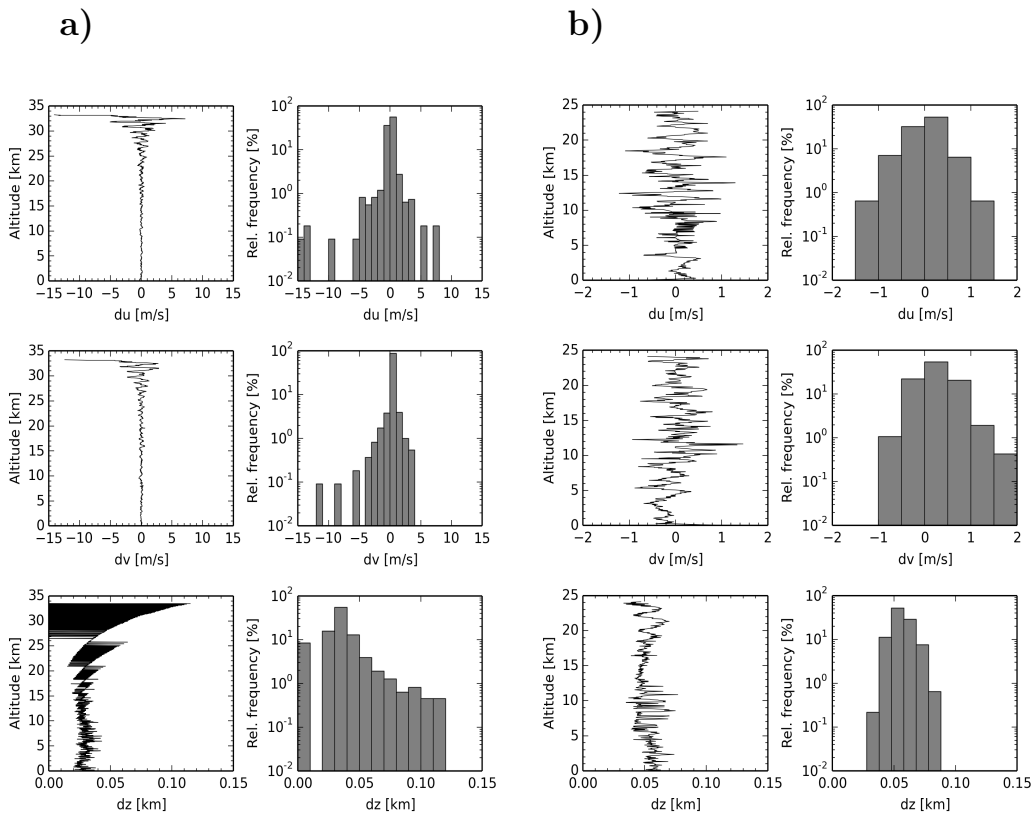


Figure 3.2: a) Typical radiosonde 'difference-profile', successively for zonal (du) and meridional (dv) winds and ascent height increment (dz) (left three column-plots) and their corresponding relative distributions over range bins of 1 m s^{-1} for du and dv and 10 m for dz (right three column-plots); b) Same as figure 3.2a, but now for a GPS WFS from De Bilt data, with a vertical resolution of 10 s (~ 50 m). This result is typical for a single profile for modern WFSs (GPS and LORAN-C) sampled at 6s (~ 30 m) or more.

amount of multipath interference depends very critically on the positioning of the antenna relative to adjacent reflecting surfaces, e.g., whether the Radiotheodolite is positioned on a roof or on the ground.

3.3.2 Distributions of Wind and Wind Shear Profiles

We show some examples of the common raw observation profiles of zonal wind and wind shear as function of ascent height, at z and $(z_i + z_{i+1})/2$, respectively, first for SPARC Radiotheodolite and subsequently for the other types, LORAN-C and GPS at station De Bilt and for the BADC and AMMA campaigns. Figure 3.3a

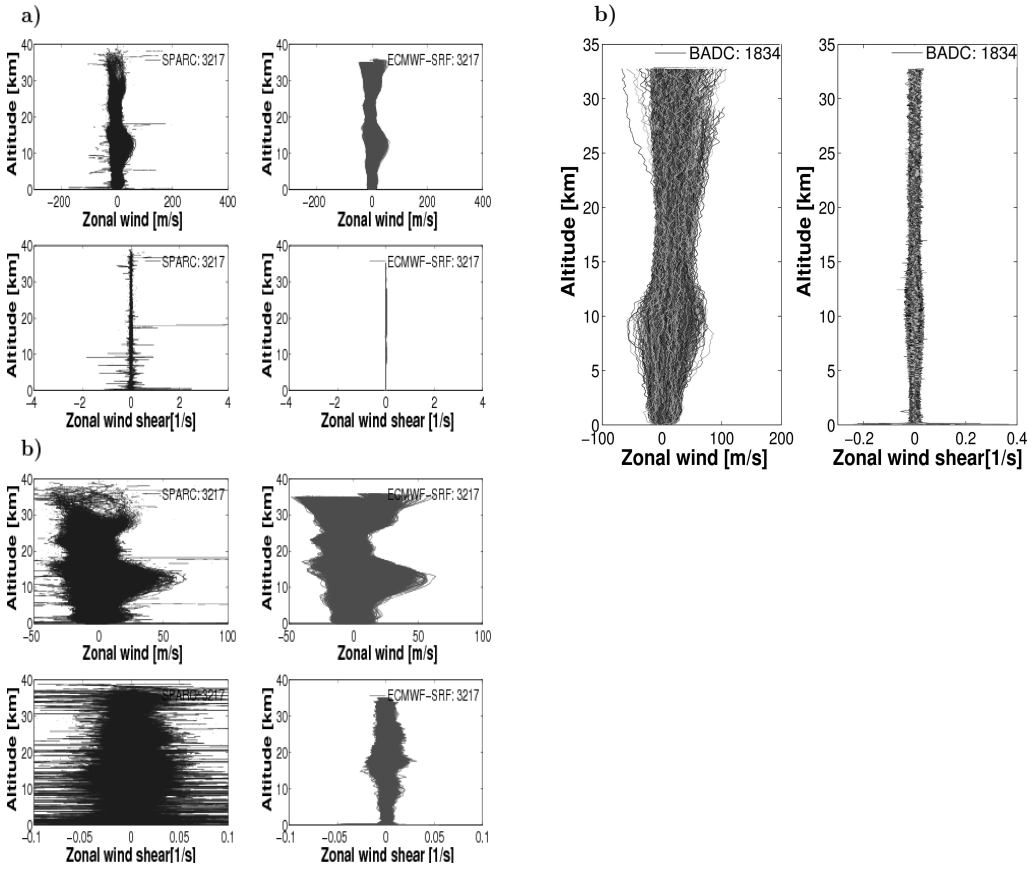


Figure 3.3: a) Tropical zonal wind and wind shear profiles for 11 SPARC radiosonde stations over 2006 at vertical resolution of 6 s (~ 30 m) (left) and the collocated ECMWF 12-hour Short Range Forecast (right). The number of collocated radiosonde profiles used is ~ 80 % (3217/11*365), see the legend inside the plots. Notice that ECMWF model resolution is interpolated to the radiosonde resolution. Below the same plots as a) are shown, but horizontally zoomed; b) Similar plots for 6 BADC midlatitude stations over 2006 at a vertical resolution of 2 s (~ 10 m). The number of available profiles represents more than 83% ($1834/6*365$), as indicated by the legend inside the plots. Notice the very low number of extreme outliers as compared to figure 3.3a.

shows a profile set of wind (top) and wind shear (bottom) for the complete year of 2006 at the tropical stations. The corresponding collocated ECMWF model data are also shown alongside for comparison; the model data are the 12-hour forecast model fields interpolated to the radiosonde launch location. The ECMWF forecast field is preferred to the analysis to avoid incestuous comparison, since the latter contains the current radiosonde data. Large spikes in SPARC data denote the presence of

outliers, which should be removed before generating any statistics. Below figure 3.3a, we can see a zoomed-in version. These plots demonstrate more clearly the global shape of the wind and wind shear profiles for both SPARC Radiotheodolite-radiosonde and ECMWF model, which look pretty similar in some aspects for the wind profiles. However, the profile set of wind shear is substantially broader for the radiosondes. This is not due to the outliers, but suggesting that vertical wind variations are underestimated in the model (Houchi *et al.*, 2010). The figure also provides a rough estimate of reasonable thresholds necessary for wind and wind-shear quality controls.

Figure 3.3b is similar to figure 3.3a and its zoomed-in version, but now for modern WFSs (LORAN-C and GPS) midlatitude BADC data in 2006. This dataset shows very few outliers in wind and shear values; only few values are probably unrealistic near the surface and in the upper troposphere. This suggests that LORAN-C and GPS WFSs are of much better quality than the Radiotheodolite WFS. This is also seen in individual profiles based on LORAN-C and GPS WFSs, e.g., for AMMA and De Bilt (not shown).

3.4 Statistical Quality Control Procedure

3.4.1 Preamble

Although a strict QC is applied on SPARC Radiotheodolite data at UCAR/JOSS (see section 3.2), a visual check of the data shows several points which appear to be unrepresentative and which we suspect to be outliers. Some extreme values of wind and shear exceed the tolerance limit as defined by WMO (DiMego *et al.*, 1985) and some other values have unrealistic changes within a 30-m vertical interval. The outliers cause generally large spikes in the profile distributions as can be seen in figure 3.3a and its zoomed-in version, figure 3.4, figure 3.5a and figure 3.7a. In order to segregate and remove these outliers from the PDF, a SQC procedure, based on percentiles (Eq. 3.4) and mean values of the PDF at each vertical level, has been developed. The percentile method sorts in ascending order the values of the observed variables collected at each 1-km vertical bin. If $X\{x_i\}$ is a vector of n sorted values, then the values can be specified by percentile ranks $P\{p_i\}$, using Eq. 3.3. There are various definitions for percentiles in the literature; in this chapter we adopt the definition as used in the MATLAB software (MathWorks, 2010):

$$p_i = \frac{100}{n}(i - 0.5) \tag{3.3}$$

with $i = 1, ...n$

For any given percentage p between 50/n % and 100-50/n %, the percentile value x_p is defined using a linear interpolation between closest ranks as follows:

$$x_p = x_i + \frac{(p - p_i)}{(p_{i+1} - p_i)}(x_{i+1} - x_i) \tag{3.4}$$

knowing that the minimum and maximum values in X correspond to the minimum

and maximum percentile, resp. $p = 50/\text{n}$ % and $p = 100\text{-}50/\text{n}$ %.

The mean profiles are also calculated, since they can serve as a double check on the median profiles to represent the average state of the analyzed quantity. This may also give an indication on the persistence of the outliers, since the mean is more sensitive to outliers than the median. Globally, the SQC consists in determining thresholds for the observed variables to isolate and eliminate the outliers, i.e., those observations falling outside the threshold values in the PDF. The threshold values are determined statistically at each 1 km vertical bin of the atmosphere using test criteria (see figure 3.6a). The different steps of the SQC method are detailed and illustrated with graphics in the next section.

3.4.2 Description of the Statistical QC Algorithm

Here we describe the most important steps of the SQC algorithm, which are summarized in the flowchart in figure 3.6a. Initially, some very extreme outliers are removed by using the tolerance limits as defined by WMO and cited previously. Following the flowchart, the first step in the SQC procedure is to collect a sufficient data amount (ideally one year or more) at each 1-km range bin and then compute and plot the percentiles (Eq. 3.4) and means. This is done for both wind components u (zonal) and v (meridional), their corresponding shear values and the ascent height increments (dz). The bin size may be modified/adapted to get the desired amount of data at each vertical bin for the QC statistics over a number of stations in a set time period, but here we keep it fixed to $1km$. This is based on our observation that the number of data is almost constant throughout the troposphere and decreases only at the highest levels, because of the difference in the balloon bursting height from one sounding to another (not shown).

Before performing the SQC on the wind profiles, initial wind shear profiles are approximated as a gradient of the horizontal wind over a constant vertical interval (dz), in order to avoid the digitization issue seen in the Radiotheodolite WFS. dz is fixed here to 30 m, which corresponds to ~ $6s$ time sampling and a mean ascent rate of 5 m s^{-1}. So, this fixed height at each vertical level scales all shear values and percentile shear values in the same way and thus does not affect the SQC method based on percentile analysis. The exact wind shear is then recomputed afterwards, using the obtained QC-ed products for wind components u and v and mean dz. For example, QC-ed u is shown in figure 3.6b. In addition, when wind shear values are rejected, the corresponding underlying wind values are also rejected and vice versa. Once the final QC-ed wind profiles are obtained, they are used to establish the statistics, e.g., shown in the right-bottom plots of figure 3.4 and figure 3.5a for wind shear, successively, for 9 tropical and 37 mid-latitude stations.

The density of the number of percentile profiles is enhanced at the extremes when approaching the 0% and 100% percentiles, for each of the analyzed quantities, e.g. u,v,dz, etc. The subsequent percentiles are 17 discrete values in our case, as shown in table.3.1.

Table 3.1: Percentile values of profiles referenced successively from 1 to 17 here, and from left to right in the plots.

$N°$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
$p_i(\%)$	0.0	0.1	0.2	0.3	0.5	1	2	25	50	75	98	99	99.5	99.7	99.8	99.9	100

So, the percentiles from 1 to 9 (from 0 - 50%) increase in value and cover the left side of distribution, while the percentiles from 17 to 9 (from 100 % to 50 %) decrease in value and cover the right side of distribution; see the legends of figure 3.5 or figure 3.7 for instance. The comparison of subsequent differences of percentile values will allow to determine thresholds N at each vertical level, e.g., $x_{100} - x_{99.9} > N.(x_{99.9} - x_{99.8})$ and $x_{0.1} - x_0 > N.(x_{0.2} - x_{0.1})$, respectively for the left and right values of the distribution, defined as condition (1). Condition (2) and so forth, will be the next subsequent differences, as shown in the flowchart in figure 3.6a. N denotes the tolerated magnitude of the relative jump between 2 subsequent percentiles at the edges of the PDF. N is statistically dependent on the data volume (time period and number of stations) present at each level, thus on the width of the percentiles. But, it is more critically dependent on the outlier likelihood, existing in each used data set or subset.

The set value of N determines the trade-off between the rejection of representative extreme values (false alarm rate) and of unrepresentative outliers (outlier detection probability). These two characteristics are clearly illustrated in figures 3.4 and 3.5a. An increase in the number of real outliers, increases the magnitude of the jump in value between the extreme percentiles. But since the number of outliers is arbitrarily different from one data set/subset to another, even for the same WFS, N has to be evaluated heuristically (involving trial and error and visual checks) at least once for each processed dataset. Once N is set for each dataset, the procedure of determining the percentile-threshold and removing outliers at each atmospheric vertical level is automatic. We note however, that the estimation of N should be reconsidered if the adopted bin size is different than the 1 km used here, since the underlying wind and outlier PDFs change with bin size. For one station we have 365 profiles per year and about 12,000 points over a 1 km bin. For the tropical and polar stations, we get good results for $N = 2$, as can be seen in figure 3.6b. On the other hand, for other data sets with multiple stations, the proportion of data in each percentile profile is higher and outliers stand out more clearly at the extreme values, as seen in figure 3.5a as compared to figure 3.4. In these cases, condition (1) is sufficient to eliminate the outliers, and the factor N should be larger to restrict the passing from condition (1) to condition (2) in the QC procedure and avoid the elimination of realistic observations. So, for data sets where we have more than 8 stations, the tests show that for $N = 10$, we remove only the points which are most probably outliers.

By setting pQC as the percentile that defines the threshold, it denotes the level

of rejection at both sides of the PDF distribution. Since outliers appear on both sides of the distributions in similar amounts, we considered a symmetric QC for the determination of the percentile threshold values at each vertical level. As such, we relate pQC to a left and right PDF percentile counter, resp. mm and pp, defined as follows, $mm = 1 + pQC$ (for the left percentile) and $pp = 17 - pQC$ (for the right side percentile). So, for a given pQC, mm sets the percentile position number in ascending threshold value order (minimum accepted PDF value), while pp in descending threshold order (maximum accepted PDF value). For instance, when pQC is initiated at $pQC = 0$, the corresponding percentile position numbers are 1 (0 %) and 17 (100%) and all values in the PDF are accepted. So, pQC is either equal to 0 (no outliers), 1 (0.1% outliers on each side of the PDF, i.e., 1/500 points removed), 2 (2/500 points removed), or 3 and so forth. Its value depends on the fulfillment or not of the conditions defined in the flowchart in figure 3.6a. In our tests, we observed that from $pQC > 2$, we start to remove many realistic observations at some range bins. So, the extreme three percentile levels; i.e., accepting all, percentiles 0.1% to 99.9% and 0.2% to 99.8% , are sufficient to reach our goal to remove the most harmful outliers from the statistics. So, in our datasets a maximum threshold-level of $pQC = 2$ (i.e., 0.2% outliers on both sides) was sufficient to eliminate the outliers causing large and unreasonable jumps in the distributions. The distribution of the QC-ed products is plotted in the same figure as the raw data. The SQC method is applied for both SPARC and non-SPARC datasets, but results in a very small amount of outliers in the latter. Overall QC results are illustrated graphically and discussed below in more detail for the winds, shears and for dz.

3.5 SQC Results

In this section, the SQC applied to the horizontal wind components, their vertical shear and ascent height increments (dz), is graphically illustrated and discussed accordingly. The SQC method can be run for different vertical aggregations. To reduce random noise and for a better comparison, most of the results are established with a smoothed 12 s (~ 60 m) vertical resolution rather than 6 s (~ 30 m) for each analyzed quantity. However, the QC-ed u, v and dz, can be used to compute other important quantities, such as, the drift, the vertical air motion, etc. The final QC-ed products are used in this chapter to derive more reliable wind shear characteristics, including estimation of the vertical motion from the ascent rate statistics (section 3.6).

3.5.0.1 Zonal (u) and Meridional (v) Winds

Figure 3.4 illustrates the SQC method applied to the 2006 tropical wind and shear data of 9 SPARC radiosounding stations. We can see from this figure that the outer percentiles of 0.1% and (99,9%) isolate in most cases the values which present a very large jump (outliers). The 0.2% and (99,8%) percentiles threshold are rarely used for rejection in this case, and generally only near the surface and lower stratosphere. Similarly, figure 3.5a,b illustrates the SQC of wind and shear data respectively for 37 SPARC subtropical and 6 BADC mid-latitude stations in 2006. Except in non-SPARC (modern WFSs) data shown in figure 3.5b, all SPARC Radiotheodolite often

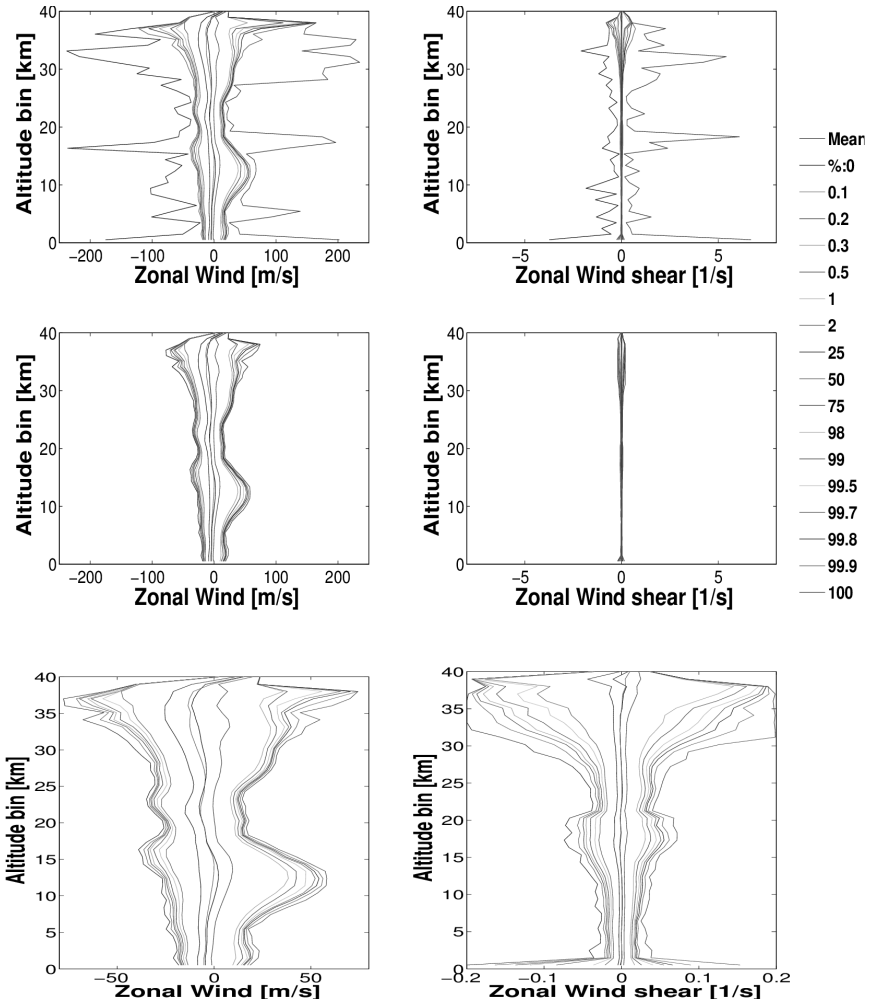


Figure 3.4: Percentile profiles of Zonal wind (left) and wind shear (right) for high resolution radiosondes before (top) and after the automated QC (middle), and a zoomed-in version after QC (bottom). This is based Radiotheodolite data, which contain 9 SPARC tropical stations as shown in figure 3.1. The percentiles are monotonically distributed from 0% to 100%, respectively as indicated in the legend (starting from left curve profile to right). In each plot, only the added mean profile (very black) that can overlap with the median 9^{th} profile (gray) at certain altitudes, and these two latest compare well after the QC. Notice that at each 1 km vertical bin, only 0.1% to 0.2% of data at maximum, is removed by the SQC method.

exhibited a large increase in values between subsequent percentile profiles at the edges of the percentile distributions between (0%, and 0.1%, and 99.9% and 100%). This is not due to the increase in the number of profiles, thus the increased amount of data

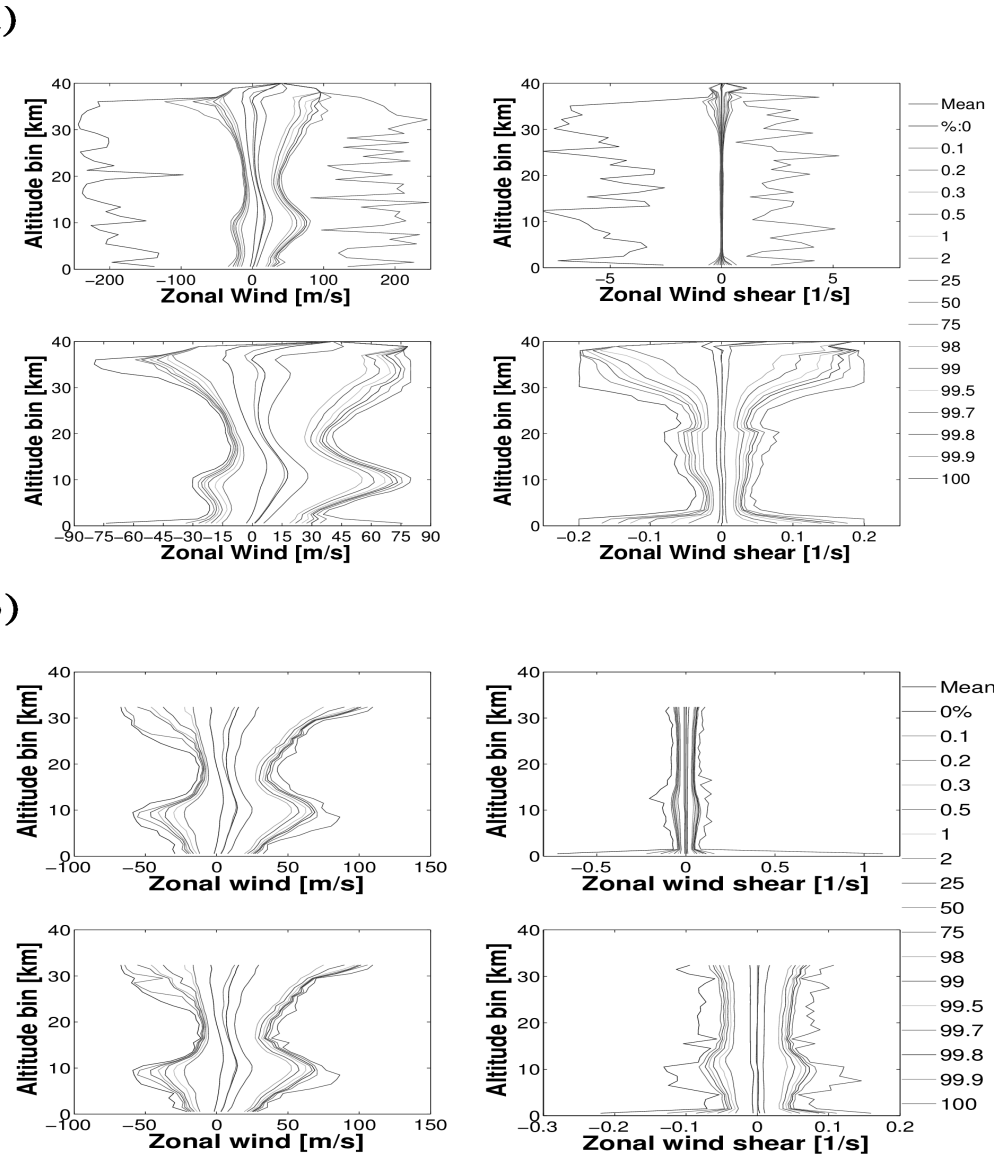


Figure 3.5: a) Same plots as figure 3.4 but now for 37 SPARC (Radiotheodolite) subtropical stations (see figure 3.1) and where only the zoomed-in plots are shown in the bottom panel after SQC. Notice the increased number of outliers compared to figure 3.4; b) Same plots as a) but now for mixed profiles of LORAN-C and GPS WFS, from 6 BADC mid-latitude stations.

collected at each $1km$ vertical range bin, since this would better sample the percentile and presumable result in a smooth vertical line. Near the edges, the large jump be-

tween subsequent percentile profiles is rather due to the frequency of occurrence of extreme outliers, which is highly dependent on the data quality. This can be seen from figure 3.4 and figure 3.5a for SPARC Radiotheodolite data as compared to figure 3.5b for modern WFSs. Indeed, according to these figures SPARC Radiotheodolite data show notable and typical high occurrence frequency of outliers (big spikes). We found this to differ from one station to another, and from one climate region to another, etc., even for the same data type.

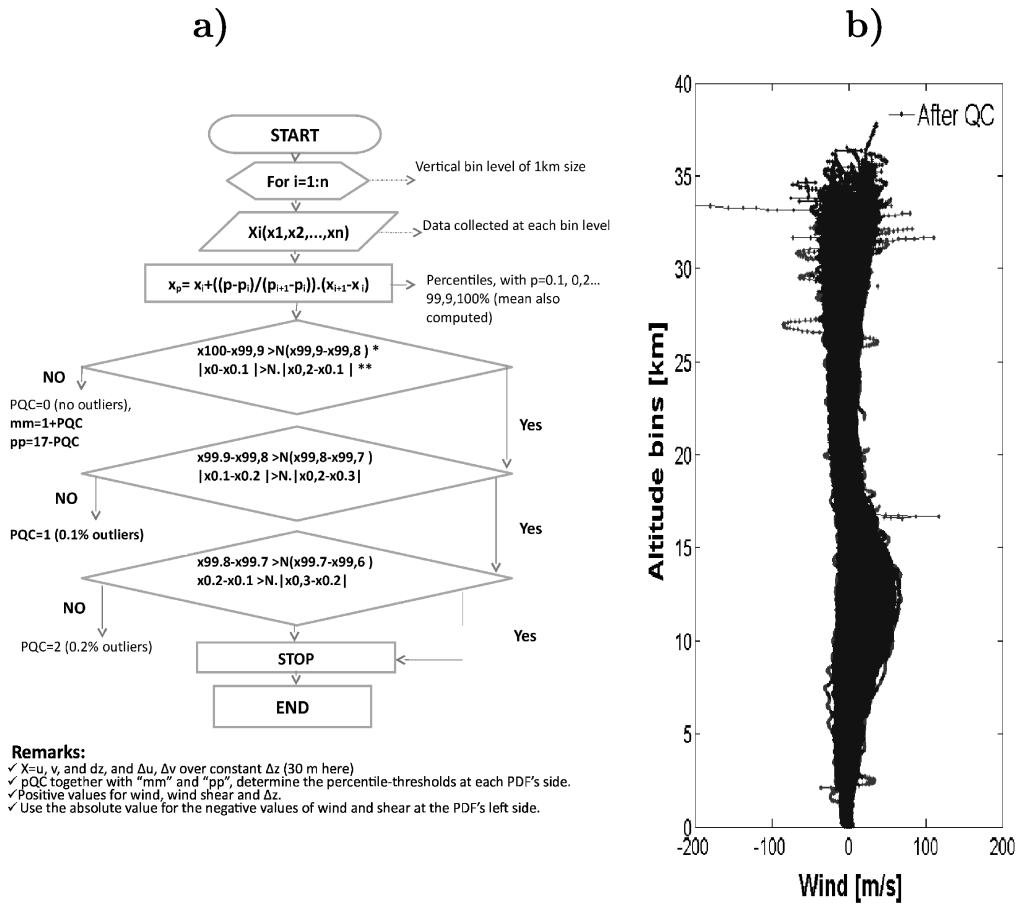


Figure 3.6: a) Flowchart of the SQC Algorithm; b) Illustration of the SQC on a set of zonal wind profiles for the tropical stations in 2006. Gray points at the edges of the different distributions with the altitude are the rejected wind values while the black are the accepted points. Dots are used in the profiles in order to highlight the number of removed data points by the SQC.

Moreover, figure 3.4 and figure 3.5a show not only that the size of the dataset used (number of profiles, thus data amount collected at each vertical bin) affects the

determination of the threshold values in the SQC procedure, but it also affects the magnitude of the percentile jumps at the extremes of the PDF and the associated SQC parameter, N (defined previously). This is due to the fact that the fraction of outliers does not increase systematically with the number of data at each vertical bin. Note that we presume that the PDF exists of a unimodal part with real values and a more wide (symmetric) PDF of outliers. As more data are used, both PDFs become better defined. Therefore, the percentile profiles generally become smoother and less erratic as more input data is used. However, note that big jumps between two outer percentiles are not only related to the amount of data, but also due to the number of cumulative outliers at each level bin. So, for large data sets the N factor is mainly dependent on the data quality (measurement accuracy), which can also differ from one station to another and from one level to another for the same dataset. This justifies the necessity to determine this factor heuristically for each processed dataset, as already mentioned. Figure 3.5b shows the SQC for non-SPARC data (mixed LORAN-C and GPS) for 6 BADC mid-latitude stations of the BADC dataset. In comparison with SPARC data, these results do not show a large jump between the very closely separated outer percentiles at most atmospheric vertical levels. Consequently, it is seen during the SQC of the non-SPARC data that the percentile thresholds pQC remain mostly equal to zero, thereby indicating the good data quality and negligible amount of outliers. Figure 3.6b shows an example of the final QC-ed product for zonal wind profiles shown in black, excluding the removed outliers (in gray) from the raw data (black and gray). Globally, the results show that most outliers fall in the outer 0.1% (99,9%) percentiles profile for the East (West) wind and shear values at most vertical levels, with (N) set to 2 here.

3.5.0.2 Ascent Height Increments dz

We illustrate here the results of the SQC on the ascent height increments (dz) and on the ascent height (z). These results are used in section 3.6, to estimate more reliably the wind shear, ascent rate and vertical motion. Figure 3.7 shows the percentile statistics of dz before and after SQC at each 1 km vertical bin, for SPARC, De Bilt and AMMA data. 37 subtropical stations are used over 2006 for the SPARC dz and z plots, while only one station each for De Bilt and AMMA is used. The statistics of Radiotheodolite radiosonde data (SPARC) show many unrepresentative values (outliers) compared to LORAN-C and GPS radiosonde data (De Bilt and AMMA), in particular near the surface. Moreover, the digitization issue is observed at high altitude in SPARC data, and that originates from pressure profiles as stated in section 3.33.3.1. Although the SQC method is effective in removing outliers at most atmospheric levels in dz profiles, above about 22 km the effect of the pressure digitization; which is amplified by the gravity waves occurrence, persists. In order to reduce this digitization problem and lower the random noise and the gravity wave effects, SPARC data is degraded from 6 s (~ 30 m) to 12 s (~ 60 m) vertical resolution, as shown in figure 3.8a for the statistics of dz and z.

In these last plots, we still see some residual perturbations in SPARC z (thus in dz) above ~ 22 km even after smoothing. This indicates once more, that the system-

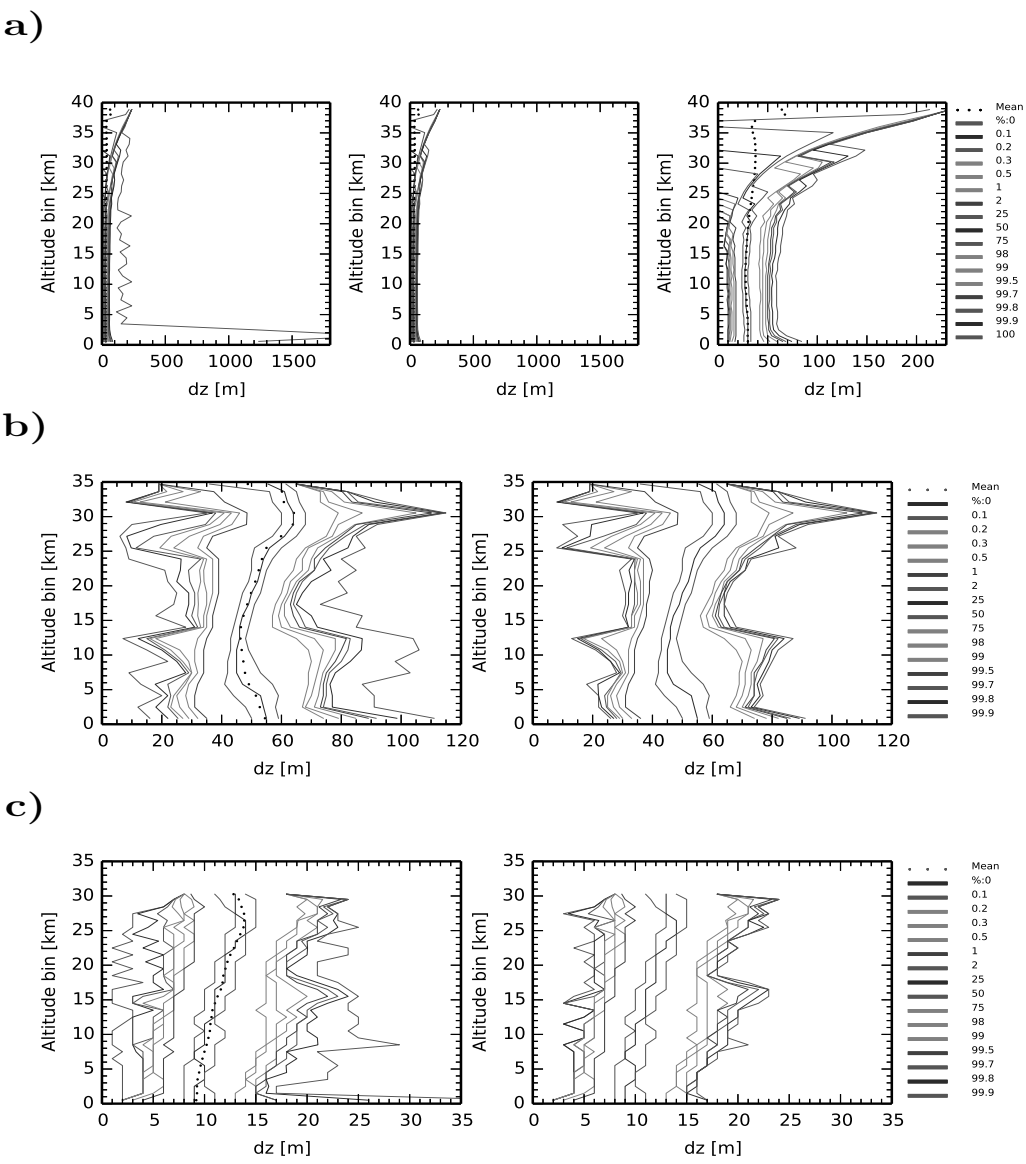


Figure 3.7: a) *dz* profiles before the SQC (left) and after (middle) and the latter zoomed-in (right) at each $1km$ bin (as show in dots for mean profiles) for Radiotheodolite SPARC profiles of 6 s (~ 30 m); b) and c) are the same as a), but successively for De Bilt (10 s ~ 50 m) and AMMA (2 s ~ 10 m), both based on LORAN-C and GPS WFSs. Notice the same legend as in figure 3.5 and figure 3.6b for the percentile and mean profiles. Notice also the very small number of outliers in De Bilt and AMMA data, indicating the excellent data quality.

atic digitization effect is not easy to remove, thus these remaining large oscillations in the percentile statistics in dz plots. Also, some imperfections are observed close to the surface. These are likely due to the known Radiotheodolite WFS limitation for elevation angles below 17 degrees (Vaisala, 2002). However, similar dz and z analyses for the 2 other WFS datasets of De Bilt and AMMA (figures 3.8b and c show a quasi-linear trend in all the percentiles and mean profiles; suggesting a better measurement quality. Apart from the above-mentioned situations, we can see at most atmospheric levels that after the SQC, Radiotheodolite SPARC dz PDFs compare well with the dz PDFs of the modern WFS (LORAN-C and GPS) used in AMMA and De Bilt.

3.6 Vertical Motion and Wind Shear Estimation

In this section, we propose a scheme to estimate wind-derived quantities more reliably; the ascent rate, the vertical air motion and wind shear.

3.6.1 Ascent Rate and Vertical Air Motion

Following our discussion in section 3.53.5.0.2 on the statistics of dz and z shown in figure 3.8, we propose here an approach to approximate vertical air motion based on the ascent rate statistics. The percentile statistics plots for ascent height increments dz, may be seen as the balloon ascent rate statistics, since the latter is defined as the dz over a fixed time sampling Δt, e.g., 12 sec for SPARC data. dz depends on the vertical air motion w, the balloon ascent rate in still air AR and the time sampling Δt: $dz = (w + AR)\Delta t$, such that:

$$w = \frac{dz}{\Delta t} - AR \tag{3.5}$$

This is similar to the model proposed by Wang *et al.* (2009). Their model enables the extraction of the air vertical velocity from radiosonde data, by considering the balloon ascent rate as a result of two contributions, one representing the balloon ascent in still air, and the second the vertical air motion. So, if the balloon ascent rate in the absence of vertical winds may be assumed constant, then, air vertical velocity is obtained by subtracting the ascent rate in still air from the actual ascent rate.

So, the spread in the dz PDF around the median is proportional to the spread in the vertical motion PDF. The AMMA and De Bilt radiosondes show small spread in the stratosphere between 15 and 20 km height. Since vertical air motion at these levels is limited, this spread may be essentially due to variations in balloon lift speed. In turn, such variations at one particular station are typically caused by differences in Helium (He) amounts, as the balloon type and thus launch mass is generally fixed. Variations in He amount cause similar variations in speed at all levels, while the mean ascent speed varies smoothly by a few 10th of a m/s with height. The typical spread (1 σ value) in the dz PDFs represented in figure 3.8 is about 3 m, corresponding to a vertical motion spread of 0.25 m s^{-1}. Note, however, that the 90% percentile in the troposphere deviates relatively much from the other percentiles, particularly

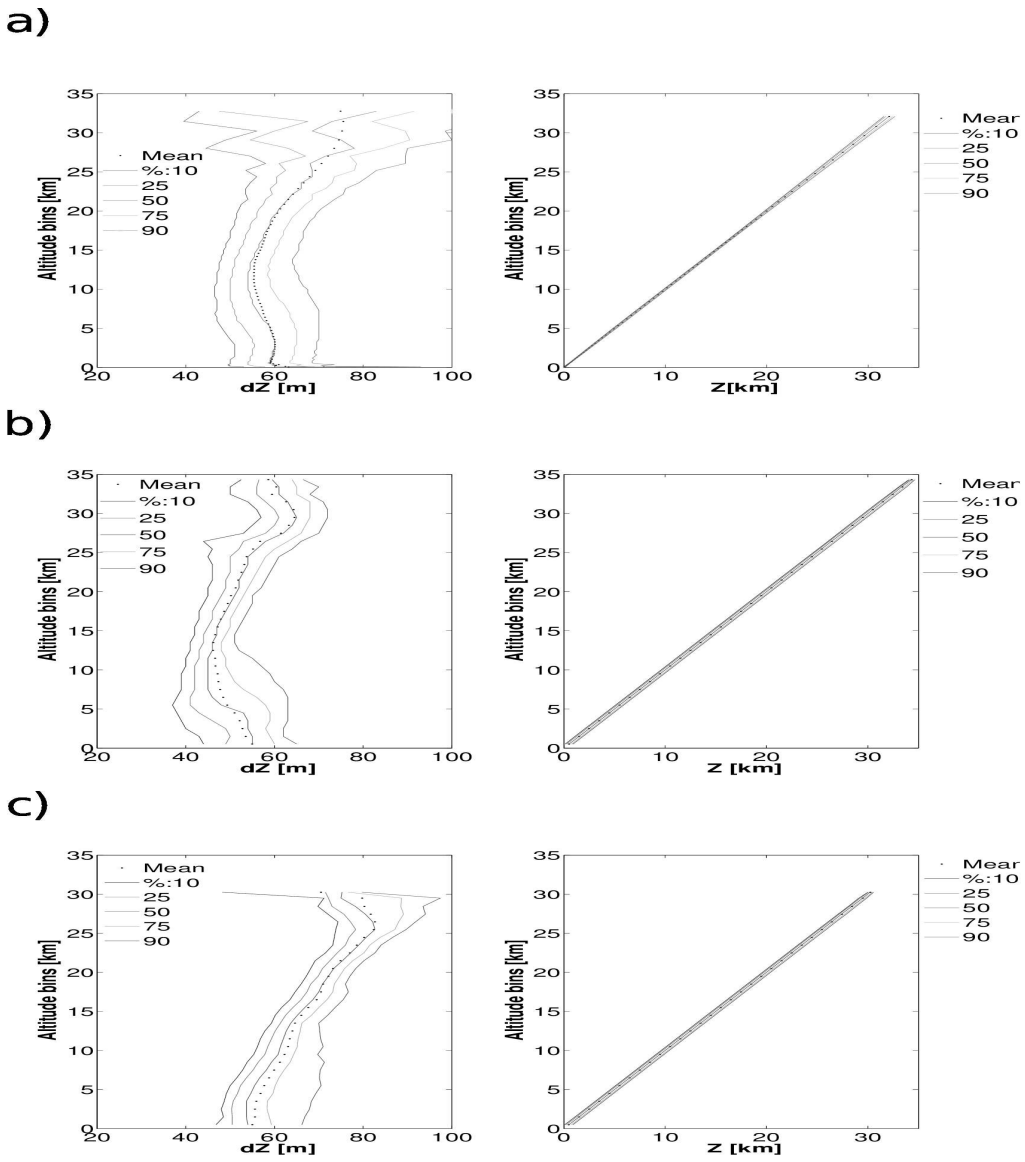


Figure 3.8: a), b) and c) show the percentile and mean profiles for the QC-ed dz (left) and z (right), successively for SPARC Radiotheodolite, mixed LORAN-C & GPS datasets at De Bilt and AMMA GPS datasets; these statistics are established over each 1 km vertical bin. For a better comparison, profiles of each WFS dataset are re-sampled to about similar vertical resolution, i.e., De Bilt at 10 s (~ 50 m), SPARC at 12 s (~ 60 m) and AMMA at 12 s (~ 60 m). Notice the residual perturbations in SPARC data above ~ 22 km, even after smoothing, because of the digitization effect.

for the tropical AMMA data. This PDF skewness towards larger dz and thus larger upward motion is likely associated with the diabatic heating of ascending air due to condensation. The tropical tropospheric deviation of the 90% percentile of dz from the median is about 10 m, corresponding with an upward air motion of roughly 0.8 m s^{-1}. Most other known lift effects, either broaden the PDF on both sides, such as turbulence, or slow down the balloon due to precipitation loading, and do not explain a upward-only motion anomaly. Following this argument, tropospheric PDF broadening due to turbulence and moist convection may be visible in the LORAN-C and GPS dz PDFs. The Radiotheodolite PDFs show more variation in the stratosphere, indicating larger variation in balloon mass and/or He. Therefore, diabatic effects appear less asymmetric and less pronounced in the SPARC climatology.

3.6.2 Wind Shear

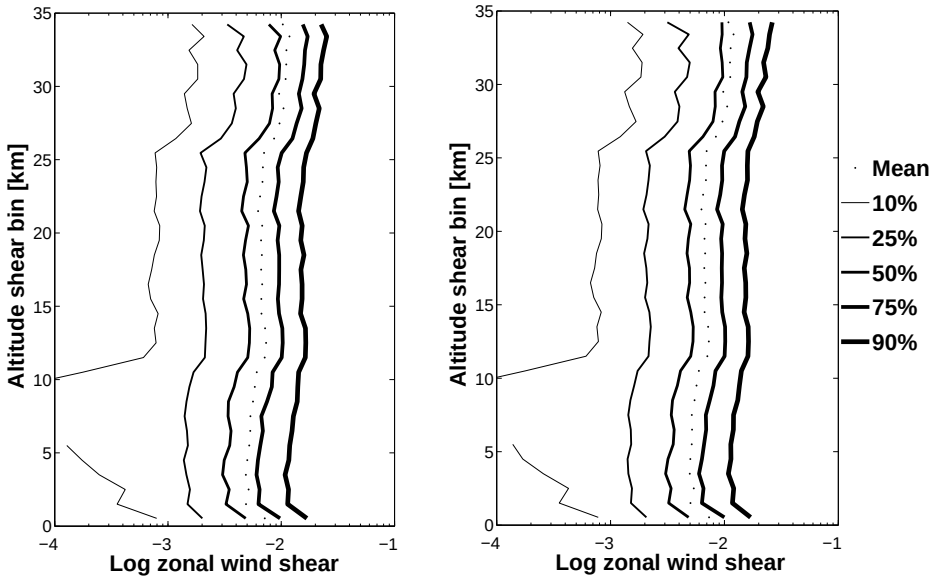


Figure 3.9: Wind shear statistics (percentiles and mean) profiles for De Bilt radiosonde 10 s (~ 50 m) resolution computed with 2 methods, according to the standard definition, Eq. 3.2 (left plot) and using the climatological mean of the QC-ed dz at each 1 km bin (right plot). Notice the logarithmic scales along x-axis)

The direct computation of wind shear using the standard definition (Eq. 3.2) is not reliable for SPARC Radiotheodolite data as mentioned previously, even after the SQC of winds and dz. So, we propose here an approach for the vertical wind shear computation for the Radiotheodolite data, by using the climatological mean dz, rather than an instantaneous dz profile. This method is verified with modern

WFS data (LORAN and GPS), since for these latter the standard calculation of wind shear appears accurate. The comparison of the two methods is tested and shown for De Bilt data over 2008 (LORAN-C and GPS data) in figure 3.9. Absolute shear values are considered for the statistics in this figure and plotted in decadic logarithmic scales (along the horizontal axis) as a function of shear height $(z_i + z_{i+1})/2$. It is clear that this approximation ignores variability in dz due to vertical motion (Eq. 3.5) of the balloon, which corresponds to about 5 %. Moreover, diabatic effects may occasionally further enhance the deviation from a nominal balloon lift. However, the results show that wind shear computed with the two methods leads to very similar results, supporting therefore, our approach for wind shear estimation for Radiotheodolite data.

3.7 Conclusions

In this chapter we present an efficient Statistical QC method to isolate and eliminate outliers from a large distribution of "realistic" profiles of wind, wind shear and ascent height. The SQC method uses fine percentiles, i.e., very small percentage intervals, at the extremities of the PDFs of each analyzed quantity at each atmospheric level (1 km vertical bin size is used in all of our analyses). The SQC method is developed for radiosondes based on the Radiotheodolite WFS, but has also been applied to LORAN and GPS data. For the two latter WFSs, very few outliers are detected, due to the good quality of these measurements. Globally, the SQC method shows a beneficial impact on the different analyses (on individual profiles and statistics) applied on the meteorological variables as studied in this chapter. After QC, the climatology analyses for Radiotheodolite WFSs and more modern WFSs data appear to be comparable. However, above 25 km altitude comparison is more complex, since the data amount can decrease substantially due to the different heights reached by the individual soundings and the increased presence of gravity waves.

A statistical approach is introduced in this chapter to correct Radiotheodolite WFS data. In terms of weather, radiosondes based on Radiotheodolite may be difficult to interpret, as seen for instance here for the ascent rate, vertical air motion and wind shear. This is due to the poor reporting accuracy of such WFS, particularly in the ascent height increments due to the pressure measurement digitization. We estimate the vertical air motion from the ascent rate climatology and in addition derive wind shear profiles and climatology. In all cases the climatological statistics of the QC-ed ascent height increments are used. The posteriori height increment estimation method using climatology has been validated with data from the more accurate WFSs, i.e., GPS and LORAN-C. With this approach, the climatologies of the ascent rate, the horizontal wind components and vertical shear for Radiotheodolite WFSs data become comparable with LORAN-C and GPS WFS results. However, a skewness of the tropospheric vertical motion PDF, probably due to latent heating, is clearly visible in GPS and LORAN-C, but much less so in the Radiotheodolite WFS data. The SQC approach may also be beneficial in case of individual profiles; for instance to characterize physical phenomena, such as turbulence, gravity waves and so forth. A similar approach may also be applied to the other meteorological variables

provided by radiosonde instruments, i.e., pressure, temperature, and humidity. But new threshold criteria have to be derived for these variables. In terms of applications, these variables are of particular interest for high vertical resolution cloud analysis of radiosonde measurements. This would for example be useful for preparing the ESA ADM-Aeolus Doppler Wind Lidar mission, which wind profile quality will depend on combined wind and cloud structure occurrence.

Comparison of wind and wind shear climatologies derived from high-resolution radiosondes and the ECMWF model

CHAPTER 4

Climatology of Wind and Wind Shear

Abstract

The climatology of atmospheric horizontal wind and its vertical gradient, i.e. windshear, is characterized as a function of climate region. For a better representation of the average atmospheric wind and shear and their variabilities, high-resolution radiosonde wind profiles up to about 30 km altitude are compared with collocated operational ECMWF-model Short-Range Forecast winds. Statistics of zonal and meridional winds are established from both datasets. The results show mainly similarity in the probability distributions of the modeled and observed horizontal winds, practically at all levels of the atmosphere, while at the same time the vertical shear of the wind is substantially underestimated in the model. The comparison of shear statistics of radiosonde and ECMWF model winds, shows that model wind shear mean and variability are on average a factor of 2.5 (zonal) and 3 (meridional) smaller than for radiosondes in the free troposphere, while in the stratosphere and the PBL results are more variable. By applying vertical averaging to the radiosonde data, it is found that the effective vertical resolution of the ECMWF model is typically 1.7 km. Moreover, it is found for individually collocated radiosonde-model wind and shear profiles that the model wind may lack in some cases variability larger than 5 m s^{-1} and 0.015 s^{-1}, respectively, due mainly to the effect of lacking vertical resolution, in particular near the jets. Besides the general importance of this study in highlighting the difference in the representation of the atmospheric wind shear by model and observations, it is more specifically relevant for the future Atmospheric Dynamics Mission (ADM-Aeolus) of the European Space Agency, due for launch in 2017. The results presented here are used to generate a realistic global atmospheric database, that is necessary to conduct simulations of the Aeolus Doppler Wind Lidar (DWL) in order optimize its vertical sampling and processing.

4.1 Introduction

With the increase of interest in high-resolution modeling in Numerical Weather Prediction (NWP) and climate research, a more detailed description of the atmosphere dynamics and optical properties is highly needed. Progress is possible because of the exponential development in high-performance computing and advances in instrumentation and measurement techniques, in particular at high resolution. Over the last decade various and important projects exploiting high-resolution observations were accomplished, including SPARC (Stratospheric Processes And their Role in Climate), FASTEX (Fronts and Atlantic Storm Track Experiment), etc. These are motivated by the necessity of understanding and resolving many atmospheric processes occurring generally at meso- and smaller scales and which most current weather models fail to resolve, such as convection, cloud development, gravity waves, turbulence, etc. These effects are generally parameterized using the local mean wind and vertical wind shear as input, among other things. Recall that in this chapter, the vertical gradient of the horizontal wind is referred to as vertical wind shear. For a definition see section 4.2.3. To develop these parametrization schemes, detailed knowledge of the spatial scales of the dynamical processes represented in NWP or climate models is indispensable. The determination of the mean state and variability of the atmosphere can be based on measurements or on models. Håkansson (2001) described global wind statistics utilizing 31 ECMWF (European Centre for Medium-Range Weather Forecasts) model levels of analysis fields and low-resolution radiosonde observations, reported only at standard and significant levels (OFCM, 2006). He shows that except near the surface, there is a large similarity in the wind and shear statistics for both collocated datasets, despite the limited vertical representation in the ECMWF model. One may conclude that the representation of the atmospheric wind by standard and significant radiosonde levels and the ECMWF model is similarly poor. The aim of this study is to statistically describe the climatological wind and wind shear characteristics of the first 30 km above the earth surface at higher vertical detail. This is done by collocating high-resolution radiosonde observations with the ECMWF Short Range Forecast (SRF) model, such that the results may be compared to Håkansson (2001). The effect of vertical smoothing in the radiosonde on both wind and shear variabilities is investigated, first on individual collocated radiosondes-ECMWF profiles, then statistically. The ECMWF model (see section 4.2.1 for the model description) offers a good quality atmospheric wind with a global coverage and with relatively high resolution. On the other hand, radiosonde balloons are currently the only observing system providing continuous vertical wind profiles at very high resolution from the surface to high altitudes (including a large part of the stratosphere), however limited to mainly the Northern Hemisphere continents. The typical balloon ascent rate is about 5 m. s^{-1} (Brock & Richardson, 2001) with only 10-15 % variation in ascent speed along the trajectory, though in the presence of gravity waves the variation may be occasionally higher (Shutts *et al.* (1988), Kitchen & Shutts (1990)) and Shutts *et al.* (1994), as well as in strong convection.

An overview and detailed description of both the model and radiosonde observations are given in section 4.2. However, 3D wind profile observations are one of the most important and lacking meteorological quantities in the current Global Observing

System (WMO, 2008c). Deficiencies, notably in the temporal and spatial coverage of wind observations in the current GOS, are hampering rapid progress in operational weather forecasting and climate-related studies. The future mission of the European Space Agency (ESA), AEOLUS Atmospheric Dynamics Mission (ADM), which in fact motivated this study, will provide more homogeneous global wind profile coverage. This mission, due for launch in 2017, aims to retrieve global wind profiles of the lowermost 30 km of the atmosphere using a Doppler Wind Lidar (DWL) operating in the ultra-violet (at 355 nm) part of the electromagnetic spectrum (Stoffelen *et al.*, 2005). Several studies (Stoffelen *et al.* (2006), Žagar *et al.* (2008), Tan *et al.* (2007a), Marseille *et al.* (2008c), Marseille *et al.* (2008b) and Marseille *et al.* (2008a)) demonstrated using different techniques, that Aeolus measurements would have a significant impact on NWP and climate models. In addition, further understanding of atmospheric dynamics and climate processes are also expected from this mission, see ESAs website:`http://www.esa.int/esaLP/ESAEPG2VMOC_LPadmaeolus_0.html`. ADM-Aeolus vertical range-bin resolution is limited by 24 bins spread over 30 km height. The distribution of these vertical bins is being optimized by conducting simulations of Aeolus-DWL vertical sampling in realistic global atmospheric conditions, i.e., by considering the complex optical and dynamical heterogeneities of the atmosphere. Since the ECMWF model fields are continuous in space and (almost) in time, they may be combined with any high-resolution optical dataset, such as the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO). On the other hand, collocations between, e.g., CALIPSO and high-resolution radiosonde data are rare. However, to use ECMWF collocations for ADM simulation we need to investigate what wind scales are represented by the ECMWF model. By characterizing the wind and wind shear climatology from the ECMWF model and high-resolution radiosonde observations, one may develop an ability to build a realistic global atmospheric data base, needed for Aeolus-DWL simulations.

To this end, available world-wide high-resolution radiosonde observations were collected. A specific year, 2006, with the most abundant SPARC high-resolution radiosonde data has been collocated with Short-Range Forecasts of the ECMWF model (ECMWF-SRF), as described in section 4.2.1. This section also includes the available global data coverage and the definition of the climate regions as used here. This is followed by a discussion about the presence of outliers in radiosonde wind measurements and the difference in accuracy of the various wind-finding systems used to collect the data: radiotheodolite, LORAN and GPS. In section 4.3, an example of collocated radiosonde-ECMWF wind and shear profiles is shown and discussed. It is followed by the statistics of zonal and meridional wind profiles and the resulting shear for different climate regions. The radiosonde horizontal drift from its launch position was also characterized for the different climate regions, in order to verify the validity and consistency of the comparison between the ECMWF model and radiosonde climatologies, since we simplified the spatial collocation to the radiosonde ground-location, i.e., not following the radiosonde ascent trajectory. In addition, to verify the consistency of the 2006 SPARC-collocated statistics, we established similar statistics from new generation and more accurate radiosoundings (BADC, AMMA and De Bilt; described in section 4.2, and by processing 9 years of SPARC data in addition to 2006. The effect of vertical smoothing (resolution) on wind and wind shear

variability is investigated, first on individual collocated profiles, then on the statistics of wind and shear. Finally, by comparing the wind shear statistics obtained from the ECMWF model and the radiosonde observations at different resolutions we could determine the effective vertical resolution of the ECMWF model. In the concluding section, the major results are summarized.

4.2 Data and Method

4.2.1 Data and Collocation

4.2.1.1 ECMWF Model versus High-resolution Radiosondes

The ECMWF model provides a good quality and global atmospheric wind at relatively high resolution. It is important to recall that on February 2, 2006 the ECMWF model moved from 60 to 91 (p_L60 to p_L91) vertical levels, reaching an altitude of about 80 km in both versions. Figure 4.1 shows an enhancement in the number of levels for the L91 model version, particularly in the first 15 km. The L60 model has a horizontal spectral truncation of T511 (Riddaway, 2002) which corresponds to a horizontal mesh size of 40 km, while for the L91 model it is T799 (~ 25 km). But the effective horizontal resolution is larger as shown in Stoffelen *et al.* (2008). The ECMWF horizontal wind power spectra drop substantially for a wavelength below about 250 km, thus denoting the effective horizontal resolution of the model. In line with this, Skamarock (2004) showed for different grid sizes (22, 10, and 4 km configurations of the Weather Research and Forecast (WRF) model) that the effective horizontal resolution is generally around 7 times the grid resolution. Additional new parameterizations and their effect on the atmospheric variability of the ECMWF model are reported by Bechtold *et al.* (2008). But tests with collocated De Bilt radiosoundings and the ECMWF model, over the year 2008, did not show substantial changes of the results as reported in this chapter (see figure 4.7c. On the other hand, radiosonde balloons are the only observing system which provides vertical wind-profiles at very high resolution (up to 1 second) from the surface to high altitudes, continuously and covering a large part of the stratosphere. However, they are limited mainly to the Northern Hemisphere continents and with very sparse coverage over ocean, tropical and Southern Hemisphere areas. In addition, these radio-soundings were for a long time devoted only to weather forecasting, therefore, they were archived only at standard and significant levels which are required to be sent to operational weather centers via the Global Telecommunication System (GTS) for real-time use (Hamilton & Vincent, 1995). Worldwide, there are more than 1500 stations with varying temporal coverage records (Durre *et al.*, 2006). In the early days, not all radiosonde data contained wind measurements and wind quality is being improved progressively following advances in the wind-finding systems, starting from optical radiotheodolite to LORAN (LOng RAnge Navigation), and the new generation of modern GPS (Global Positioning System) systems. For reference, a useful description of radiosonde instruments and data interpretation, including history, development and future prospects, may be found in the two articles of Brettle & Galvin (2003) and Galvin (2003).

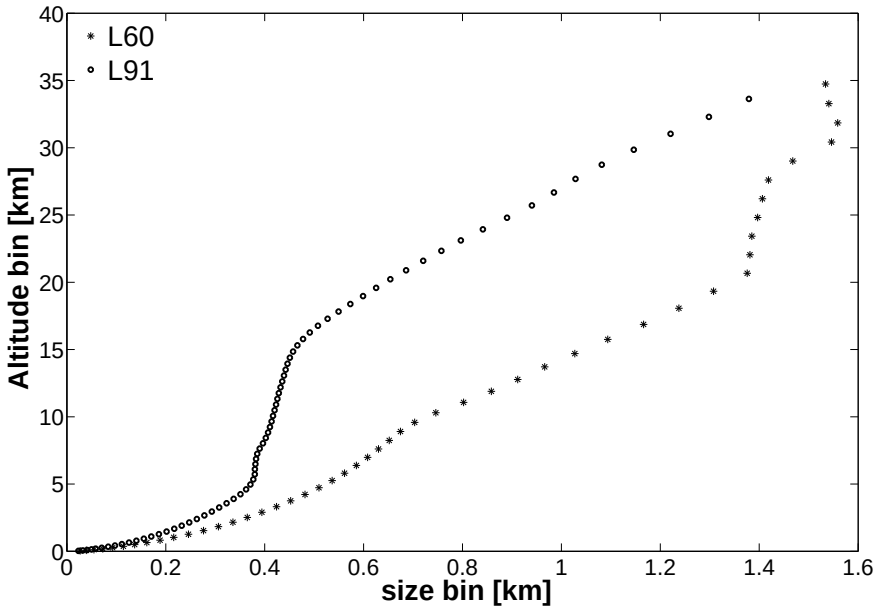


Figure 4.1: Increasing vertical separation of ECMWF model levels with altitude for the two model versions, L60 (blue) and L91 (red). Both models have irregular size bins between one level and another. Notice that the number of levels is enhanced particularly in the PBL and free troposphere for the L91 model.

4.2.1.2 Data Coverage and Collocation Procedure

Here, 92 stations with high-resolution (6 seconds) radiosonde data over the 10 year period from 1998 to 2007, available through the SPARC project website `ftp://atmos.sparc.sunysb.edu/pub/sparc/hres`, were fully exploited in this study. The most continuous in time of the 2006 radiosonde data from the British Atmospheric Data Centre (BADC, see UK-MetOffice (2006-2009), the African Multidisciplinary Monsoon Analysis (AMMA) and the De Bilt radiosondes at the Royal Netherlands Meteorological Institute (KNMI, the Netherlands.) are further selected for the analysis. We briefly note that SPARC observations are intended for the study of gravity waves (Allen & Vincent, 1995) and AMMA for the African monsoon, while BADC and De Bilt are raw data of standard-resolution radiosondes dedicated for weather forecasting. These data cover a large part of the North Hemisphere with highest density in the US; however, only 2 stations are available in the Southern Hemisphere (Falkland Islands and Saint-Helena). It is important to mention that we noticed a difference in the accuracy between SPARC and the rest of the datasets, due to the difference in wind-finding system used to collect each data set. BADC, AMMA and De Bilt

data sets, based on combined LORAN and GPS system, have a better accuracy of wind and ascent height measurements than the SPARC radiotheodolite-based data. Therefore, they are analyzed separately from SPARC to avoid misinterpretation of the statistics. The focus was in particular on the ascent height increment (dz) of the SPARC radiosondes which has uncertainty that degrades the computation of wind shear, du/dz. The 2006 SPARC radiosondes, totaling 85 stations for both 12 UTC and 00 UTC, are collocated with ECMWF Short Range Forecast (ECMWF-SRF) fields for a first comparison between model and observations. The spatial collocation is performed according to the radiosonde launch ground-location, i.e., model wind fields at the different model levels are extracted from the ECMWF archive and interpolated to the ground location (lat, lon) of the radiosonde launch (not following the radiosonde trajectory). The temporal collocation is done with the 12-hour SRF, i.e., a radiosonde launched for instance at 12 UTC (00 UTC) is thus compared with a SRF initiated at 00 UTC the same day (12 UTC the day before). The main reason for using the forecast model rather than analyses is to avoid what is called generally incestuous comparison between model and observations, since the analyses model fields may already contain the comparison radiosonde observations. Also, the difference in the number of levels between the p_L60 and p_L91 model versions is taken into account during the analysis, by interpolating both to 60 m vertical resolution. However, since we focus only on the first 30 km of the atmospheric winds, which generally also correspond to the maximum altitude reached by the radiosonde balloons, only 76 vertical model levels are used, which cover this part of the atmosphere.

Before performing the analysis, both ECMWF model and radiosondes datasets were distributed over 7 climate zones which we define as follows: Northern/Southern Hemisphere polar (70-90), Northern/Southern Hemisphere midlatitude (40-70), Northern/Southern Hemisphere subtropics (20-40) and tropics (20S-20N). The global coverage of the available and analyzed datasets, including their distribution over the defined climate regions, is shown in the map of figure 3.1. This is also summarized in Table 1. We note that the BADC and AMMA have a time resolution of 2 seconds while it is 10 seconds for De Bilt. Radiosondes with time resolutions of 2, 6 and 10 seconds correspond successively to a vertical height increments (dz) of about 10, 30 and 50 m. We report the results at the same resolution when needed, mainly at about 60 m resolution for comparison with SPARC. We sometimes omit deliberately to mention 12 UTC or 00 UTC since we found similar results at both UTC times for each sub-dataset. This is due to the fact that the stations extend over a large longitudinal band covering thus a large range of time zones, more than 6 hours. So, the data at a particular UTC local time zones (day time and night time) launches. For instance, for the midlatitudes and time contain mixed subtropical stations at 12UTC (00UTC) the data cover a time zone extending from at least 3AM to 9AM (3 PM-9 PM).

4.2.2 Quality Control and Wind-Finding System Characteristics

Before establishing the statistics of wind and wind shear, two major issues are faced. The first issue is related to the amount of unrepresentative wind observations present

periods/ Stations	Total	Tropics	Subtropics	Mid-latitudes	Polar
SPARC 2006	85	9	37	38	1
1998-2007	85+7*				
BADC 2006	9	1 (SH: St Helena)	1 (Gibraltar)	7 (1 SH: Falklands Islands)	-
AMMA 2006	6	5	1 (Nouadhibou)	-	-
De Bilt 2006	1	-	-	1	-

Figure 4.2: Distribution of radiosoundings over the defined climate regions for stations of SPARC, BADC, AMMA and De Bilt over the years as indicated. (*) 7 stations among a total of 92 are distributed unequally over climate regions from one year to another. We note here that only 2 stations from the Southern Hemisphere midlatitudes (Falkland Islands and Saint Helena) are available from the BADC data.

in the raw SPARC high-resolution data, even though a strong quality control was applied (NCDC, 1998) at the University Corporation for Atmospheric Research and Joint Office of Science Support (UCAR/JOSS). The second issue is related to the limited accuracy of the ascent height intervals, or what we call shear intervals (dz), compromising the wind shear computation. Both issues are related to the radiotheodolite wind-finding system used to collect these SPARC data and which is generally less accurate than the LORAN and GPS systems. To deal with this issue, a statistical quality control (QC) method was developed. The method (see sub-subsection below) cleans the SPARC data from unrepresentative (outlier) atmospheric wind and wind shear observations, including unrealistic shear intervals (dz). More details on the overall issues in this subsection may be found in (Stoffelen *et al.*, 2009).

4.2.2.1 Removal of Outliers

The statistical Quality Control consists first of accumulating the raw information of wind, wind shear and shear intervals (dz) into probability density functions (PDFs) of these variables at different levels of the atmosphere with uniform vertical bins of 1 km interval. Percentiles of these PDFs are subsequently computed with very fine percentiles sampling at the tails, i.e., the percentile ranks are very closely separated. For such small change in the percentile, e.g., 0.1%, at the tails of the PDF, the change in the observed quantities (wind, wind shear and dz) is expected to be very small and regular. However, outliers fall generally far away from the common PDF, implying they are implausible realizations of the natural wind, shear and dz distributions, and thus cause a relatively big jump in the location of subsequent percentiles at the extremities of the PDFs, particularly notable in case of very dense percentile

sampling. By eliminating such large jumps, the maximum number of data points removed at each 1-km vertical bin does not exceed 1/1000 on either PDF tail. In fact, this amount is only reached when the number of data available in a vertical bin is small, as for the NH polar region where we only have one station. For the other regions, most outlier percentages fall below the 0.1% and above the 99.9% percentile profiles. Therefore, only 1/1000 data points on both sides of the tails are generally removed. The statistical quality control was carefully tested and visually checked. And we made sure that the wind values used are not exceeding the tolerance limits of the extreme values (DiMego *et al.*, 1985), as required by WMO. A detailed description of the method is given chapter 4.

4.2.3 Analysis Method and Definitions

After quality control of wind, wind shear and shear interval (dz) of the radiosonde data, the subsequent processing step consists, as for the QC, of accumulating wind and shear information (including the drift of the radiosonde) at different levels of the atmosphere with uniform vertical bins of 1 km intervals. This is done for both observations and model. To establish the observation statistics, we used mainly 12 seconds (~ 60 m) averages rather than 6 s (raw) data in order to reduce the random noise, in particular for the shear interval (dz). Recall that the quoted spatial resolution between parentheses is given as a rough indication, considering that the mean ascent rate of the radiosonde balloon is about 5 m.s^{-1}, which is generally the case (Brock & Richardson, 2001). The model profiles are interpolated to each 60 m vertical level before performing the statistics. The means and percentiles (successively: 10, 25, 50, 75, 90%) of each quantity are computed at each vertical atmospheric bin. The values obtained at each bin constitute thus the mean and percentile profiles, as can be seen in the overall results in section 3. The wind shear and the balloon drift profiles analyzed here are derived from wind profiles as follows: The wind shear vector, $\mathbf{s}$, is defined as the variation of the horizontal wind vector $\mathbf{v} = (u, v)$ with height z (see eq. 3.2 in section 3.1).

with i indicating the level number of wind or wind shear (knowing that each shear level is between 2 wind levels), $i = 1, ..., m-1$ for shear levels, and m the total number of wind vertical levels, which depends on the vertical resolution used, u and v are the zonal and meridional wind components. Notice that wind shear levels are plotted as a function of shear heights, $(z_i + z_{i+1})/2$.

Notice that, absolute values are considered for wind shear statistics in the next sections and plotted in decadic logarithmic scales (along the horizontal axis) as a function of shear height $(z_i + z_{i+1})/2$. We mention here, that in view of the limited accuracy of $dz = z_{i+1}z_i$ for SPARC data, the mean of dz at each vertical bin is used to compute vertical changes in the horizontal wind as a proxy for the shear. So, here (or dz) in Eq 1 is not the dz reported by the individual radiosonde but the mean dz estimated at each vertical bin for each sub-dataset analyzed. However, reported by the individual radiosonde is taken as the zonal /meridional wind variation over dz. This approach of considering the climatological mean of dz inside the bin to compute the shear has been tested with more accurate GPS radiosonde data; i.e. Shear values

obtained by this approach were compared with the nominal GPS radiosondes shear values, and have showed very similar wind shear results (not shown). The radiosonde balloon drift at height level is computed from the successively reported horizontal positions of the balloon with altitude. This is done by accumulating successive horizontal distances traveled by the balloon as given by eq. 4.1.

$$Drift(z_i) = \sqrt{(\sum_{j=1}^{i} dx_j)^2 + (\sum_{j=1}^{i} dx_j)^2} \tag{4.1}$$

Where and are the zonal and meridional distances traveled by the radiosonde balloon from one atmospheric layer level j to another j+1. are the level heights, where i indicates the level number at which the drift is computed. Because of the two limitations mentioned above in subsection 2.2 for SPARC data and to render the statistical interpretation easier, the radiosonde data from BADC, AMMA and De Bilt (KNMI) based on more recent and accurate wind-finding systems (combined LORAN and GPS) are analyzed separately. The most useful and regular data over 2006 are selected for analysis. This is done by following the same procedure as for SPARC data, i.e., these data were subjected to the same QC, then distributed and segregated according to their location over the defined climate zones. The statistics from the different data sets; based on different wind-finding systems, have been compared, including the results of the Q-Controlled wind, wind shear and shear intervals. Furthermore, to check the validity and the consistency of the statistics obtained over 2006, the remaining 9 years of SPARC data for the period 1998 to 2007 were processed.

4.3 Results and Discussion

4.3.1 Collocated Radiosondes and Model Wind and Shear Profiles

4.3 shows an example of individual wind and wind shear profiles of a high-resolution radiosonde at 6 seconds (~ 30 m) collocated with an independent ECMWF short-range forecast at 90.1 W 32.3 N on 30 December 2006 at 00UTC. The ECMWF profiles broadly compare well in shape to the radiosonde profile, but with clear differences between the profiles in both wind and shear, particularly in the detailed vertical structure. Substantial increases of wind shear are remarkable over up to 1 km extended depths in the lower troposphere, at tropopause height (8 to 15 km) and in the lower stratosphere. Figure 4.10 illustrates the more general observation in our data set that the vertical gradient of the horizontal wind is large for a typical high-resolution radiosonde ascent as compared to the ECMWF-model. The ECMWF wind and shear profiles are very smooth, ignoring thus important vertical structures. The ECMWF model profile shown here is from the L91 model version.

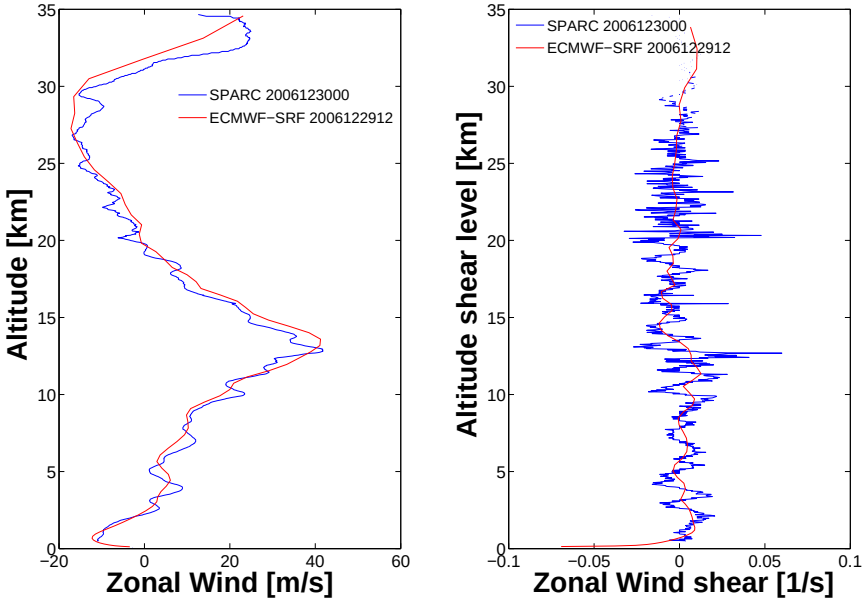


Figure 4.3: Zonal wind (left) and wind shear (right) collocation of the high resolution radiosonde at 90.1 W 32.3 N, from the SPARC dataset (blue) with the ECMWF 12-hour forecast (red). The time resolution of the radiosonde here is 6 s (~ 30 m). 2006122912 means that the 12h forecast was initiated at 12 UTC on 29 December 2006 and verification time is 00 UTC on 30 December 2006.

4.3.2 Zonal/Meridional Wind and Shear Climate Statistics

To investigate the atmospheric wind dynamics and its vertical gradient (shear), the analysis method described in subsection 4.2.3 is applied. The statistics for the horizontal wind and wind shear from collocated ECMWF model and high-resolution radiosonde observations have been established for a one-year data-set, 2006. This was performed successively for zonal and meridional winds and over the defined climate regions. The results are presented here as percentile and mean profiles. The statistics of radiosonde balloon drifts are also provided. This is important to verify how valid the comparison between the model and observation climates is, since the collocation of the model wind is done only at the ground-location of the radiosonde observations, i.e., not following the balloon trajectory ascent. Note that over short distances, i.e. less than hundred kilometers, the climate is not expected to change much, so exact collocation appears less relevant for the comparison of climate data sets of radiosondes and ECMWF model. The results of the radiosonde balloon drift are first summarized for the different climate regions with the mean profiles, and then only the subtropical case is shown with mean and percentile profiles.

4.3.2.1 Zonal Wind and Wind Shear

Figure 4.4 shows the wind and wind shear statistics (means and percentiles) for high-resolution radiosondes and their corresponding model counterpart. Notice that the absolute wind-shear values are plotted in decadic logarithm scale. The radiosonde and model wind statistics show a clear resemblance while the wind shear statistics are different. This is observed in all climate regions (Tropics, Subtopics, Midlatitudes and Polar). The difference in the shear statistics is due in particular to the limited effective vertical resolution of the ECWMF model which does not capture the meso- and small-scale dynamical structures of the atmosphere. As mentioned in subsection 2.1, the ECMWF horizontal wind power spectra drop substantially for a wavelength below 250 km, leaving thus most atmospheric processes occurring below this scale (e.g., turbulence and convection) unresolved. However, Hkansson [2001] found similarity in wind and shear statistics between the ECMWF model and low-resolution radiosonde observations despite the limited vertical representation in the ECMWF model of 31 levels, except near the surface. This indicates that the low-resolution radiosondes lack substantial wind shear in the vertical, similar to the ECMWF model. Near the surface, the large wind shear found in the ECMWF model by Hkansson [2001], exceeding values of 0.1 s^{-1}, is attributed to the misrepresentation of the pressure gradients at the lowest model levels (in turn arising from the misrepresentation of the topography). This includes the (erroneous) horizontal interpolation to a regular latitude/longitude grid. In our results, one may note differences in the values from one climate region to another in wind and shear probability distributions, where these are generally higher in the subtropics and midlatitudes. The median and mean wind profiles are mostly overlapping, while this is not the case for the wind shear. This is mainly due to the fact that we considered the absolute values of the shear, thus mapping negative shear values to the positive side such that the probability distribution becomes skewed. Consequently, this causes a shift of the median profiles away from the mean profiles. The highest averages of wind shear values, given by the median and the mean, are found in the subtropics, respectively 0.008 and 0.01s^{-1} for the radiosondes and respectively 0.004 and 0.005 s^{-1} for the ECMWF model. Thus, radiosondes clearly observe more wind shear than modeled by ECMWF. These high values occur mainly around the tropopause (from 9 to 15 km) near the jet stream, which is associated with high wind values exceeding 55 m s^{-1}, and in the stratosphere. High wind shear values near the surface are apparent, which point to the presence of low-level jets as seen in the raw data (not shown). But also, it may be due to the inaccuracies of the wind or/and height measurements, since the number of rejected points during the quality control is much higher at these low levels and known tracking artifacts exist in the SPARC radiotheodolite data set, in particular for elevation angles below 17 (Vaisala, 2002). The presence of a large number of extreme values near the surface which appear as outliers has also been observed in the raw data (not shown). However, above the boundary layer the quality of the SPARC wind measurements is improved. The midlatitude results show similarity with the subtropics, but the magnitude of the median/mean wind-shear values is smaller, as seen from both radiosonde (0.006/0.008 s^{-1}) and ECMWF (0.0025/0.0035 s^{-1}). This is due to the slowing and quickening of the jet stream, respectively as it

moves northward (towards the midlatitudes) during the warm season (late spring and summer) and southward (towards the subtropics) during the cold season (autumn and winter) (Holton, 1992). In the polar region one may see relatively high median/mean values of wind shear, 0.0053/0.0065 s^{-1} from radiosonde and 0.0021/0.0030 s^{-1} from the ECMWF model, around the tropopause associated with high wind values of more than 25 m s^{-1}. In the Tropics (mainly easterly wind) the highest median/mean shear values are 0.008/0.010 s^{-1} for the radiosondes against 0.005/0.006 s^{-1} for ECMWF. These maximum average values are found in the lower stratosphere, between 15 and 20 km, where the wind changes direction after a short transition around the tropopause where the wind blows westerly. Due to the lack of strong jet flows in the data, the average values of wind shear are relatively small in the polar region as compared to the other regions. This may in turn be caused by the fact that we have only one station in the polar region. From about 25 km and up, gravity wave activity may contribute strongly to the wind and wind shear values. Some studies (Shutts *et al.* (1988) and Kitchen & Shutts (1990)) show that large temperature fluctuations, associated with quasi-stationary gravity waves may lead to a strong wind shear in the horizontal wind and large variations in the balloon ascent rate. Cadet & Teitelbaum (1979) showed that internal inertia-gravity waves can accelerate the mean flow in the altitude range 20-25 km, which may explain the increase of the shear at this level and further up. Note that orographic gravity waves are parametrized in the ECMWF model and do not contribute to the statistics presented in this chapter.

4.3.2.2 Meridional Wind and Wind Shear

Whereas Figure 4 shows results for zonal wind, Figure 5 shows similar plots for the meridional wind and wind shear. One may notice in particular the different behavior of the two horizontal wind components, which can be seen from their variation with altitude. While the mean zonal component is generally large the mean meridional wind tends to be around zero. This is the case, for instance, for the meridional wind in the mid-latitudes and the subtropics, where the zonal wind is generally strong and dominant. The results, particularly from these two last regions where the mean/median of the meridional wind is close to zero at almost all levels of the atmosphere, demonstrate the different character of the zonal and meridional winds. Despite these differences, the meridional wind also produces a strong wind shear, but with magnitudes slightly smaller than the zonal wind shear. In the subtropics for example, the mean/median values observed are 0.007/0.009 s^{-1} in the radiosondes against 0.0027/0.0035 s^{-1} for ECMWF, as opposed to the values seen in the zonal wind shear 0.008/0.010 s^{-1} for radiosondes (see also Figure 7a,b,c of BADC and De Bilt data) and 0.004/0.005 s^{-1} for ECMWF. It is well known that zonal winds dominate the subtropics and the midlatitudes. According to these values, the ECMWF model underestimates the meridional shear by two-thirds and the zonal shear by half. In the tropics, the median/mean meridional wind (mainly northerly for these Northern Hemisphere stations) increases slightly with altitude, consequently increasing the wind shear in particular around the tropopause. In the Polar region, these average values of wind are increasing southerly up to the tropopause, then northerly in the stratosphere. For wind shear, substantial values are seen near the polar jet around the tropopause and in the stratosphere, due

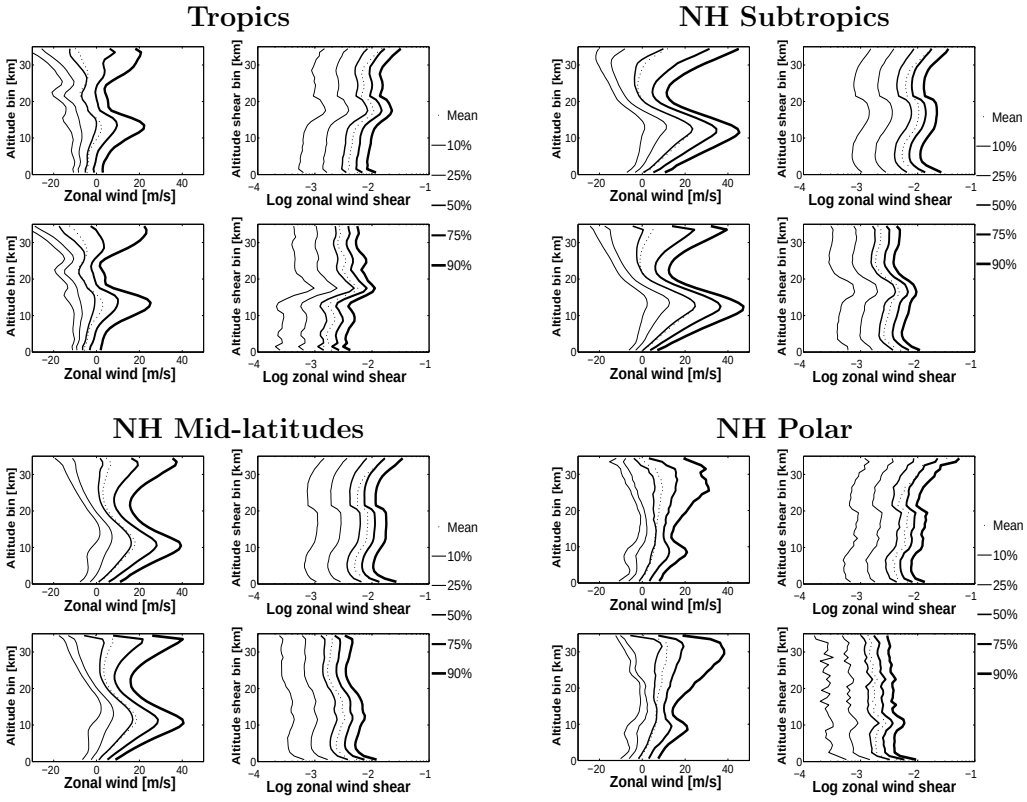


Figure 4.4: Zonal wind and absolute wind shear (in decadic logarithm scale) statistics for different climate regions based on high-resolution 12 s (~ 60 m) SPARC radiosondes (top) collocated with ECMWF SRF (bottom): mean (dots) and percentiles (successively from left to right: 10, 25, 50, 75, 90%), see legend at the right of each panel. The statistics are performed at each 1-km vertical bin for one year, 2006, for both data sets. The stations are distributed over the climate zones as follows: 9 tropical, 37 subtropical, 38 midlatitudes and 1 polar station.

to the frequent occurrence of gravity waves, as seen at this single polar station.

4.3.2.3 Radiosonde Drift

It was mentioned previously that the collocation of the high-resolution radiosonde wind profiles with the ECMWF model is done only according to the radiosonde launch ground location (not following the trajectory of the balloon ascent). Therefore, we compute the radiosonde balloon drifts to check that these are negligible with respect to the spatial scale on which the climatological wind and shear pdfs change. The statistics obtained for the drift have been summarized by mean and median profiles, but only the mean profiles are shown in Figure 6a, as the results are comparable. The

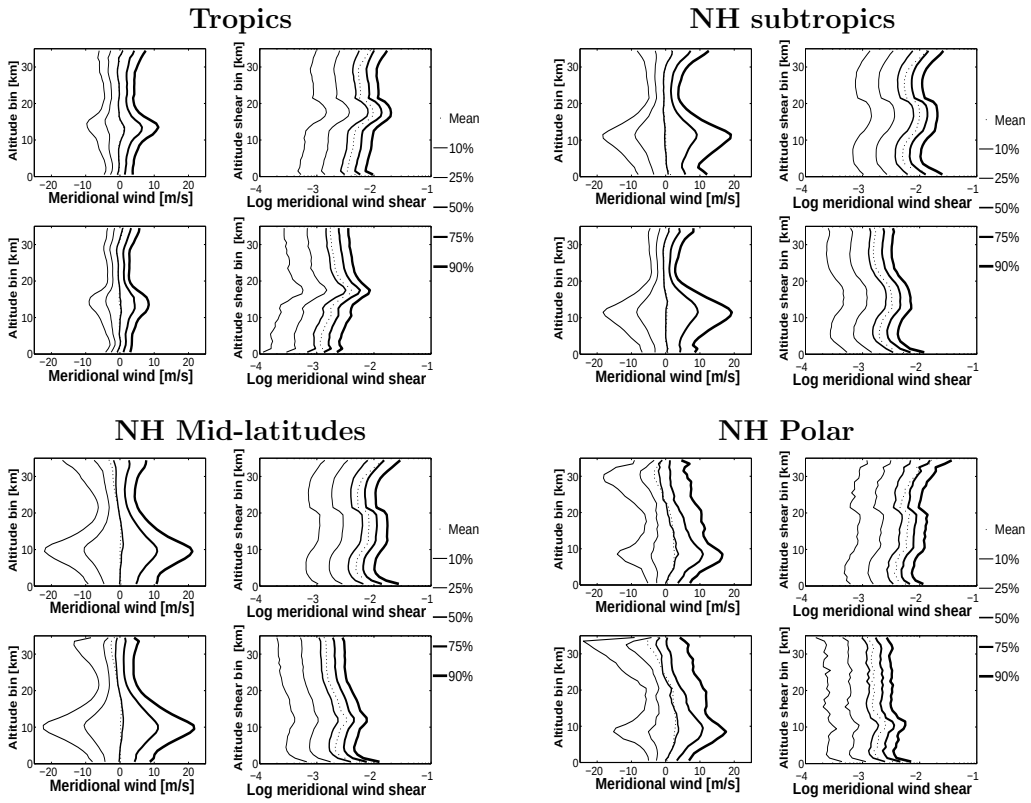


Figure 4.5: Same as in Figure 4 for meridional wind. Note the different wind axis with respect to Figure 4.

maximum drift is recorded in the subtropics where wind magnitudes are higher than in the other regions (see Figure 4). The median/mean maximum values are generally reached at the end of the ascent with values 68/71 km, while it is successively 63/66 km, 43/46 km and 34/35 km in the Midlatitudes, Polar and Tropics. Except for the tropical region, the largest drifts are occurring during the tropopause crossing, near the jets (jet stream and polar jet), and then decreasing notably at higher altitudes in the subtropics and midlatitudes. In the polar region and tropics, the drift increases quasi-linearly with the ascent height. Figure 6b (left-side plot) shows percentile statistics for the radiosonde balloon drift for the maximum mean drift in the subtropics (37 stations over 2006) and where the 90% percentile reaches about 140 km in the stratosphere. From these results, we conclude that the balloon drift is usually below 100 km, which is below the ECMWF model effective horizontal resolution (~ 250 km). Considering these results on balloon drift, we conclude that the comparison between the radiosonde and model climates, just by applying a simple collocation according to the ground-location, is valid and consistent. In addition, Figure 6b (right-side plot) shows the number of data collected at each 1-km vertical bin for the statistics in this

subtropical case. This suggests that in the first kilometer and from 30 km and up, the statistics may be not very representative of the true atmospheric dynamics, because of the rejected data at these levels.

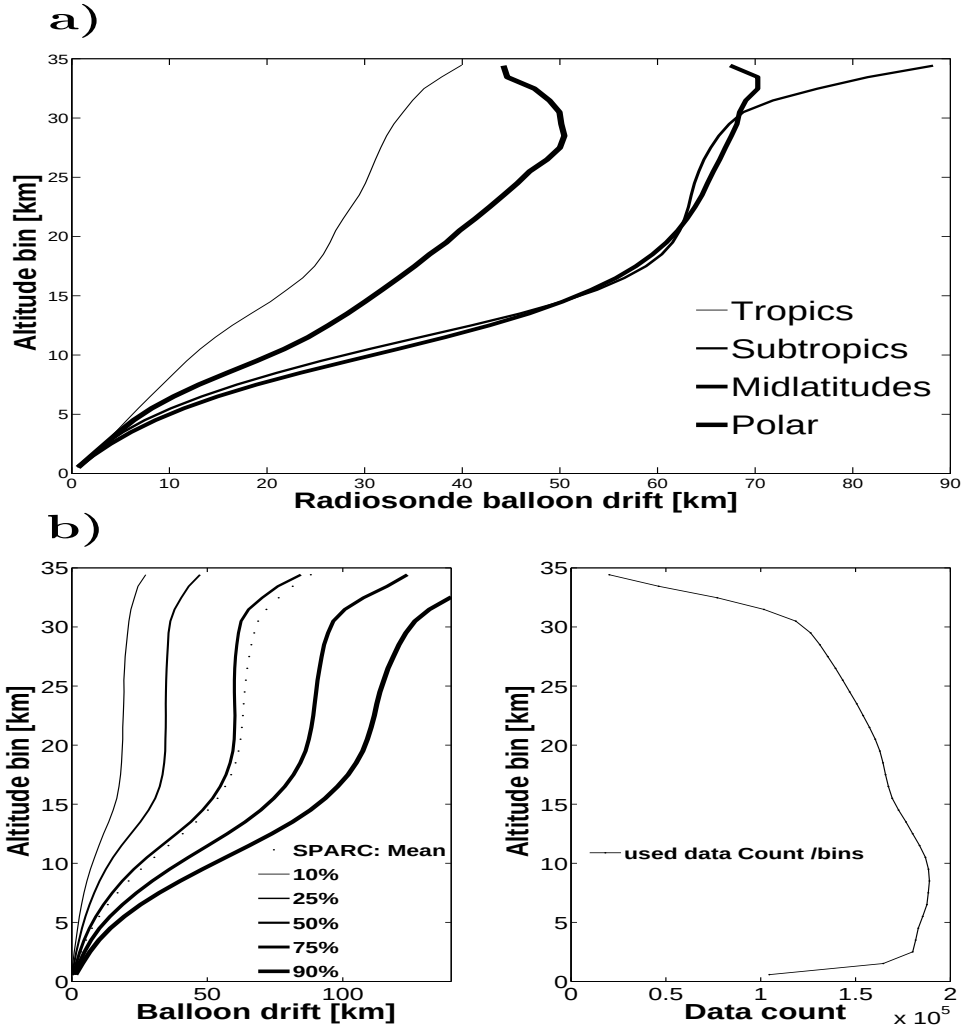


Figure 4.6: (a) Mean radiosonde drift over the different climate regions as shown in the legend, established from one year of SPARC data, 2006; (b) percentile profiles as in figure 4, but for the drift (left) and the amount of data collected and used at each 1-km vertical bin (right), both latter in the subtropics.

4.3.3 Consistency of Climate Statistics

To verify how representative and consistent the statistics obtained for the 2006 SPARC data are, they are compared with similar statistics of multiyear climate established from the remaining 9 years of SPARC data for the period 1998-2007. Since the SPARC statistics are based on radiotheodolite wind-finding systems, they are also compared with other radiosonde data statistics (AMMA, BADC and De Bilt) which are based on new generation wind-finding systems (combined LORAN and GPS). The main objective of this verification is to check on the two major issues faced in this study, as described in subsection 2.2, i.e., the presence of outliers in the wind observations and the limited accuracy in the shear interval (dz). These two issues are both related to the radiotheodolite wind-finding system used to collect the data. First, by following the same method as for SPARC 2006, similar statistics of the radiosonde data based on the more accurate and recent wind-finding systems, such as LORAN and GPS, are established. Figure 7 shows three examples of percentile and mean wind and shear statistics from mixed LORAN and GPS soundings for the Northern (6 UK stations) and Southern midlatitudes (Falkland Islands) BADC stations over the year 2006, but over 2008 for De Bilt station and collocated with the ECMWF model. The wind-shear PDFs are generally similar to the northern midlatitude SPARC data; though a difference in profile shape appears due to increasing shear at the tropopause and in the low stratosphere in particular in the SPARC data. The bump in the shear at these levels in the SPARC data has been seen in the stations over and close to the extended Rocky Mountains, but also with a smaller magnitude, in the stations which are far downstream. This may suggest a long-range effect of such long mountain chain. At the summit of such barriers the flow speeds up, with monthly wind velocity of 12-15 $m.s^{-1}$ (Barry & Chorley, 1992) and thus high wind shear may persist. One may notice also the resemblance of statistics between the Southern and Northern Hemisphere and De Bilt stations, but with remarkably higher extreme wind values in the stratosphere of the Southern Hemisphere station. Also, results from AMMA and the other BADC datasets (not shown), show similar characteristics to SPARC in the horizontal wind and shear statistics for a given climate region. Secondly, the remaining years of the period 1998-2007 of the SPARC data are processed, in addition to the 2006 data. Figure 8 shows the results for two selected cases, tropics and subtropics. One may see clearly that all the wind and shear profiles for the ten years of the SPARC period resemble each other closely. Apart from the remarkable temporal variability associated with the Quasi-biannual oscillation (QBO) in the tropics (Baldwin *et al.*, 2001); all values remain very similar for all profiles over this period.

4.3.4 Effect of Vertical Resolution on Wind and Shear

4.3.4.1 Wind and Wind Shear Variability

According to Hamilton (2006), there have been few systematic studies of the effects of vertical scaling on simulated tropospheric circulation. The goal of this subsection is to investigate the effect of reducing the vertical resolution on the variability of the wind and wind-shear profiles of high-resolution radiosonde observations. The ECMWF profiles are taken as reference. This is achieved by applying a running mean, to

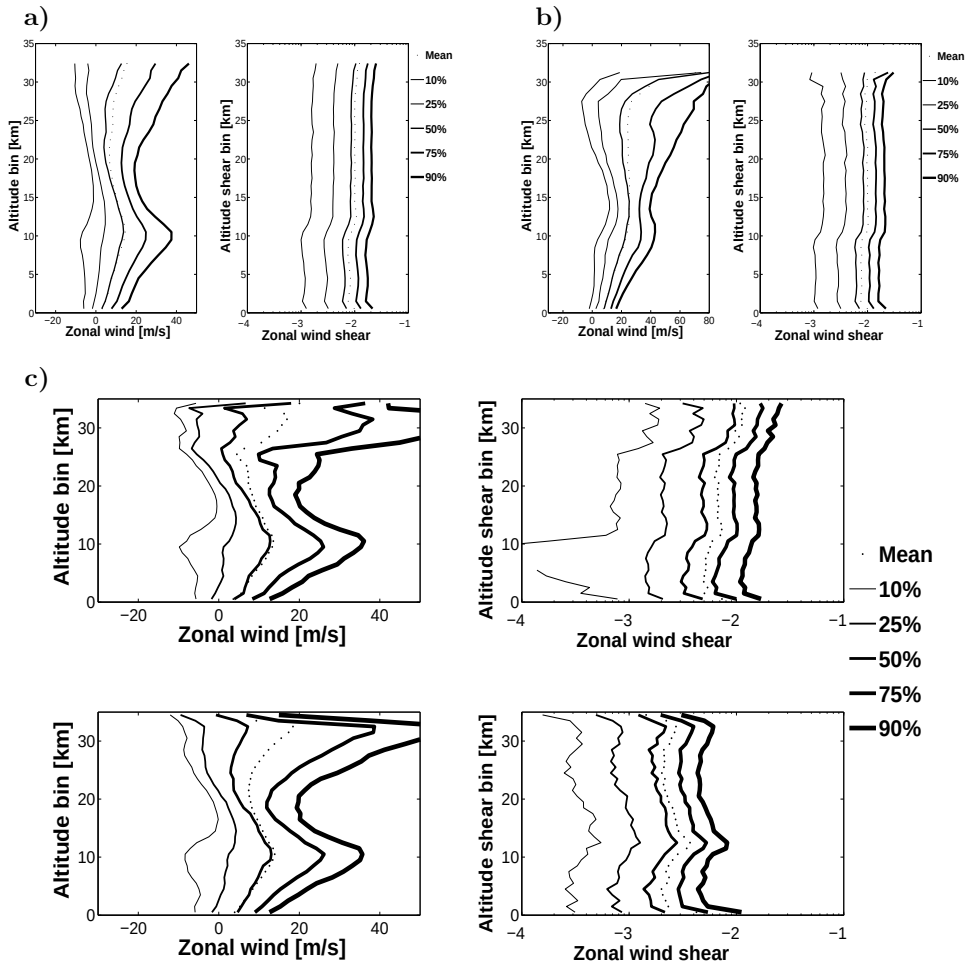


Figure 4.7: (a) and (b) Zonal wind and shear statistics at 12 s (~ 60 m) resolution, successively for Northern (6 UK stations) and Southern (Falkland Islands) Hemispheres midlatitude BADC stations over the year 2006; (c) Same as for (a) and (b), but for De Bilt station, at about 50 m resolution and over the year 2008 (top) collocated with the ECMWF model (bottom). Notice that for De Bilt station only a few radiosoundings go higher than about 25 km, hence the statistics are less significant. Notice also the difference of the horizontal axis for wind in (b).

smooth the raw (30 m) radiosonde wind and shear profiles for successively degraded time (space) resolutions. In the wind profile analysis shown in Figure 9 the means are computed over independent time samples with lengths: 6 (raw), 24, 198 (3 min 18 s) and 396 (6 min 36 s) seconds in order to have independent statistics. These time samples correspond approximately to vertical spatial box sizes of respectively: 30, m

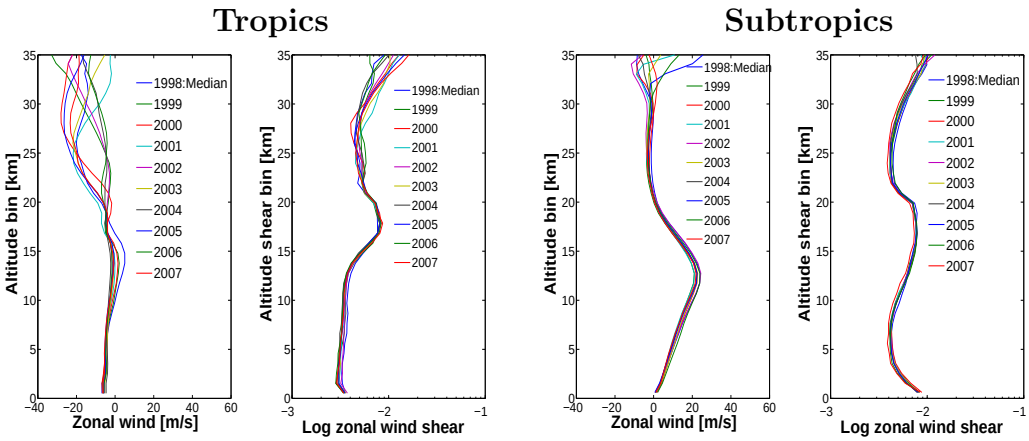


Figure 4.8: Interannual variability of wind and wind shear, shown by median profiles, from SPARC data for the ten year period 1998-2007 in the tropics (left panel) and subtropics (right panel).

(raw), 120 m, 1 km and 2 km. In addition to the mean, the standard deviation (SD) of the raw values within the box is also computed. A typical result for a subtropical station in the US, at 90.10W 32.3N for the 4 selected box sizes is shown in Figure 9. For reference, the 12-h ECMWF forecast at its original vertical separations (see Figure 1) is given at the same ground location and time. Clearly, the lower the resolution, the better the radiosonde profile resembles the smooth ECMWF profile and the more wind variability is lost, since the SD increases. This is highlighted over the profile altitude ranges 9-14 km (Zoom 1) and 19-21 km (Zoom 2) added to the left plot of Figure 9. Note that the smoothness of the ECMWF profile best resembles the 2-km averaging kernel, implying a lack of wind variability in the vertical of about 2-4 ms1 as indicated by the black line (SD). Occasionally it may exceed 5 m s^{-1} near the jets and the surface, and as shown here for this subtropical case at the tropopause and in the stratosphere. Compared to the ECMWF profile, the 2-km (black) profile is smoother and the 1-km (green) profile is less smooth, suggesting an effective vertical resolution of the ECMWF model which is between 1 and 2 km. One may also note that the wind profiles at 30 m and 120 m contexts are very similar and the lost variability (SD) when going from 30 m to 120 m is generally small, i.e., less than 0.2 m s^{-1} for the profile in Figure 9, indicating low random measurement noise in the radiosonde. In other words, both profiles at resolutions of 30 m and 120 m show similar statistics when compared to 1-km and 2-km smoothed profiles and most wind shear variance is present on scales larger than 120 m. This is in line with da Silveira *et al.* (2001) and Nash *et al.* (2005). Nash *et al.* (2005) estimated in particular the random errors; using the standard deviations of wind differences of various radiosondes types, to be between typically 0.2 and 0.4 m s^{-1} in the troposphere and between 0.3 and 0.5 m s^{-1} elsewhere.

The effect of the vertical resolution on the shear of the zonal wind (seen in Figure

9) is more explicitly shown, in Figure 10. As in the analysis of Figure 9, in this figure 10 and also in figure 11, smooth wind profiles are first derived using running mean, but the re-sampling of the raw radiosonde is done for sample sizes (resolutions) of 60, 90, 120, 150, 180, 210, 330, 1000 and 2000 meters. Notice that only the raw and 210 m profiles are plotted in Figure 10. In order to obtain smooth profiles, we applied the running mean using dependent samples, i.e., the shift forward between one sample to another in the profile is only by half the sample size. Wind shear profiles are computed from these smoothed wind profiles, then mean values (Figure 10, left) and standard deviations (Figure 10, right) are reported at the model levels (middle of the layers) and at intermediate levels with layer thickness from one model level to the next, for a fair comparison with the model. Notice that the intermediate levels are added in order to have smooth profiles when computing the mean and SD values by overlapping wind shear values over half layers (Nyquist sampling). Notice also that the model wind shear profile is derived from interpolated wind profiles at 30 m to match the raw radiosonde data. From the left-side plot we conclude that the ECMWF model vertical wind-shear is smoother than the raw and smoothed 210 m radiosonde profiles. And from the left plot, considering the 210 m wind shear profile as a reference, the smoothness of the model implies a missing variability of the wind shear with values between about 0.005 and 0.01 s^{-1} on average. Differences in the wind shear variability between model and radiosonde are more pronounced at certain levels of the atmosphere, e.g., near the jets and the stratosphere, where values may exceed even 0.015 s^{-1}. We can also see that the 210-m smoothed profile closely follows the raw profile, implying a generally small loss of variability by smoothing on this scale. In the next paragraph, the change in the variability of wind and wind shear by vertical smoothing is investigated in more detail.

4.3.4.2 Variability and Noise

Figures 9 and 10 show rather uniform statistics in the vertical. To further investigate the noise and variability properties of radiosondes and ECMWF model, we process wind and shear values in successively increased vertical box sizes, i.e., degraded resolutions, of 60, 90, 120, 150, 180, 210, 330, 1000 and 2000 meters and average the obtained statistics over the full profile. We use the following definition eq. 4.2:

$$RMS(SD) = \sqrt{\frac{1}{N_s}\sum_{j=1}^{N_s}\frac{1}{N_p-1}\sum_{i=1}^{N_p}(x_i - x_m)^2} \tag{4.2}$$

While x_i are the values of the profile time series of zonal wind and wind shear; Np and x_m are successively the number of x_i values and the sample mean in each chosen re-sampling box of size p; and Ns is the number of samples in a profile. N_s can either be independent samples or overlapping samples. For a radiosonde profile of size N (typically 1700 for the raw data), $Ns = floor(N/Np)$ independent boxes exist; for overlapping by a factor two (Nyquist sampling), $Ns = floor(2(N - Np)/Np)$. The square root of the mean variance of dependent samples of wind (left) and wind shear (right) are shown in 4.11. Note that white instrument noise would show up as a constant level of variance (SD), but here we observe increasing variance in increasing

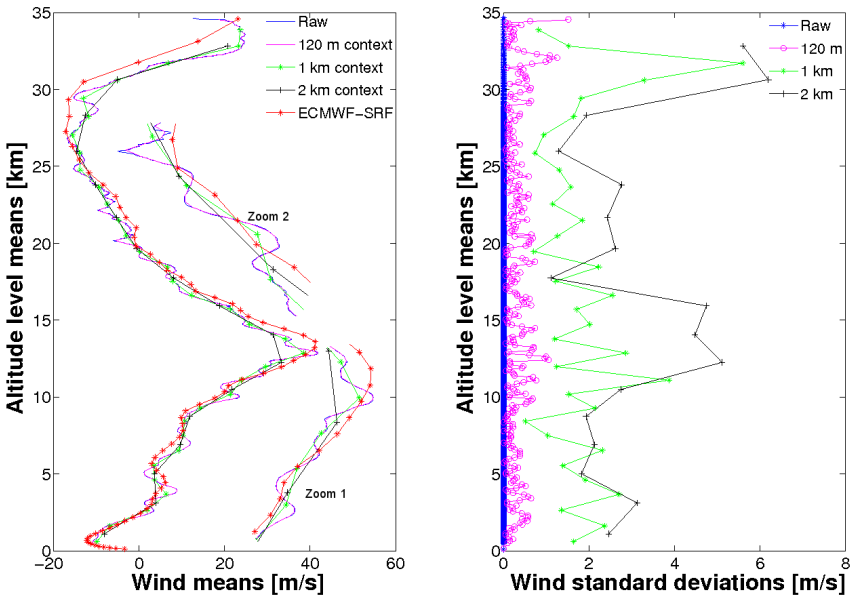


Figure 4.9: Effect of reducing the time (spatial) resolution of raw radiosonde zonal wind: (left) Wind profiles of radiosonde ascent for the raw 6 s (~ 30 m) SPARC data (blue), and for successively degraded radiosonde resolutions with moving time averages over independent samples 24, 198 (3 min 18 s) and 396 (6 min 36 s) seconds, These time samples spatially represent vertical boxes of about 120 m (magenta), 1 km (green) and 2 km (black). Notice the zooms over the altitude ranges 9-14 km (Zoom 1) and 19-21 km (Zoom 2). For comparison, the collocated ECMWF model wind profile (red) is shown on the same plot; (right) Plot showing successively, in the same colors for each box size, the standard deviation (SD) of the raw wind values in the moving box. Note that the SDs of the raw data (blue) are nullified

box sizes, indicative of vertically coherent wind or shear structures. We extrapolate the RMS variance curves to a box of size one and obtain an estimate of the random error level of respectively about 0.1 m s^{-1} and 0.005 s^{-1} for wind and shear. Given the large number of independent samples in a profile, the estimated standard error is well below 0.01 m s^{-1} and 0.0005 s^{-1} for wind and shear respectively. Moreover, for shear, white noise dominates the variance in the radiosonde data for box sizes smaller than about 150 m. For larger box sizes, the natural shear variability dominates the shear variance

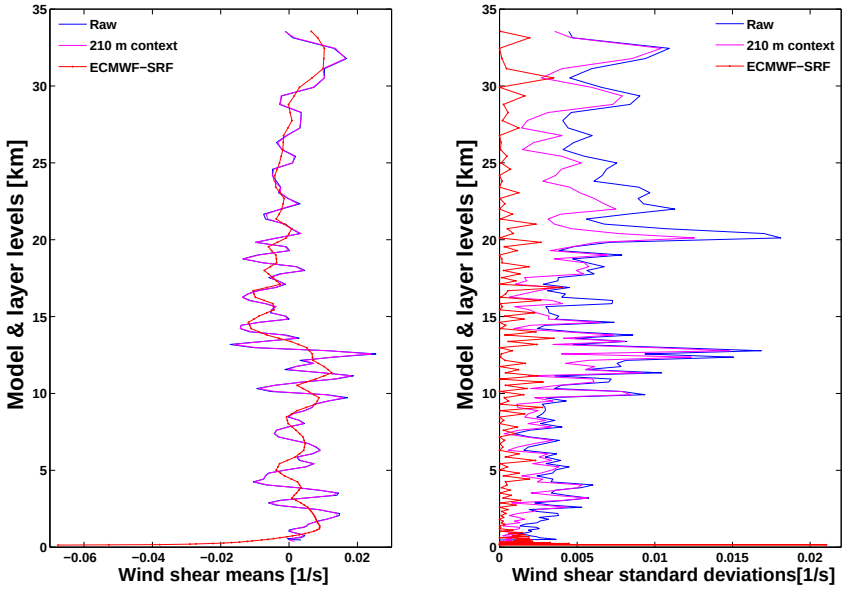


Figure 4.10: Mean (left) and standard deviation (right) at model&layer level altitudes of zonal wind shear profiles derived for two resolutions, 30 m (raw) and 210 m. The smoothed profile at 210 m is derived using a running mean. As a reference, the mean and standard deviation of the model wind shear profiles are also plotted. Notice that model wind shear profile is derived from an interpolated wind profile at 30 m as the raw radiosonde data.

4.3.5 Model Effective Vertical Resolution: Scaling Effect

It is shown previously that the vertical gradient of the horizontal wind (shear) is underestimated in the ECMWF model as compared to the radiosonde observations. To evaluate the degree of difference in the wind shear distributions and examine the effect of the vertical scaling on the wind shear statistics, we compare the shear statistics obtained at each 1-km vertical bin for different radiosonde vertical box sizes with the ECMWF model. This is done by computing the ratio of observation and model quantities (eq. 4.3). Note that the mean and median profiles of zonal and meridional wind shear were computed, but only the median ratio profiles are shown in 4.12.

$$R_{dz}(z) = \frac{|Vsh_{dz}(z)|_{RS}}{|Vsh_{dz}(z)|_{EC}} \tag{4.3}$$

where Vsh is the mean or the median of the absolute zonal (u) or meridional (v) wind-shear, z denotes the center altitude of the bin and dz the successively degraded

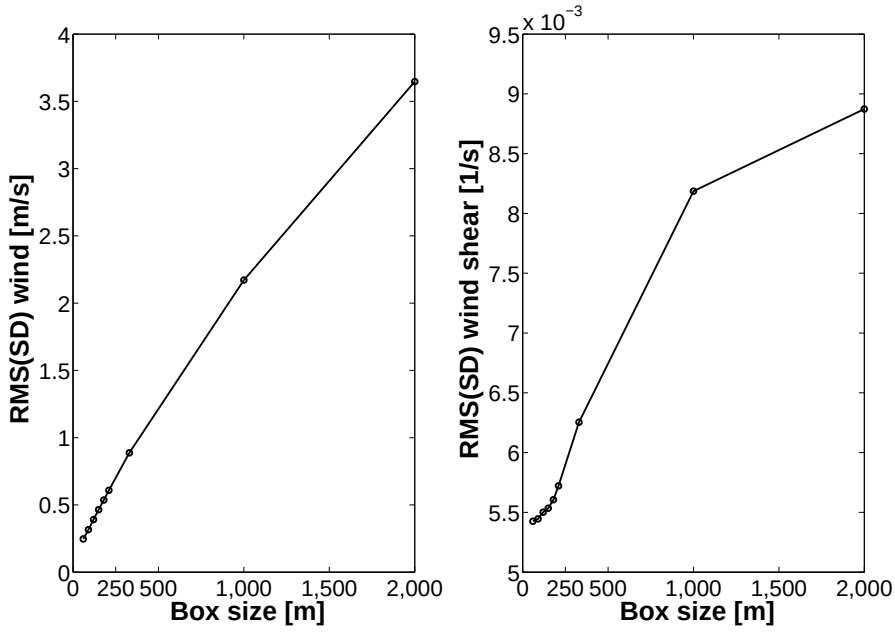


Figure 4.11: Square root of the mean variance of dependent samples of wind (left) and wind shear (right) in successively increased vertical box sizes, i.e., degraded resolutions, of 60, 90, 120, 150, 180, 210, 330, 1000 and 2000 meters (circles), averaged over the full profile.

vertical resolutions (box sizes) of the radiosonde data: 2*6 s (~ 60 m), 16*6 s (~ 480 m), 30*6 s (~ 900 m), 44*6 s (~ 1320 m), 58*6 s (~ 1740 m) and 72*6 s (~ 2160 m). Again, the spatial resolution values between brackets are given as guidance by assuming a mean ascent rate of the balloon of 5 m s^{-1}.

The effect of vertical scaling on the wind-shear statistics is first investigated by reducing the vertical resolution of the SPARC radiosonde data. Then, the ratios of the wind-shear median/mean profiles of these radiosondes at different resolutions and the ECMWF are computed following eq. 4.3. This is done for both zonal and meridional absolute wind-shear. The results in Figure 12 (only median shown) show a proportional decrease in the wind shear ratios when reducing the radiosonde vertical resolution at all different level of the atmosphere. In line with this, Essenwanger & Reiter (1969) demonstrated by using a military wind rocket, the existence of a power law between the vertical wind-shear and shear interval (dz) which explains the dependence of the wind shear on the spatial vertical resolution. Figure 12 also shows a decrease in the average horizontal wind shear until it reaches the amount of average shear seen in the ECMWF model ($R_{dz} = 1$). Notice that this decrease in the average wind shear is also seen when ratios of standard deviation or ratios of inter-quartile shear distances are used, rather than median or mean ratios. The SPARC and ECMWF

model median ratio profiles, for both zonal and meridional winds, are close to one for a radiosonde resolution 1740 m. This indicates a typical effective vertical resolution of the ECMWF model of 1.7 km, at least in the free troposphere and in particular for the meridional wind shear. In the upper stratosphere the effective vertical resolution of the ECMWF model seems poor (¿2 km) mainly because of the coarse vertical model levels, but also the imperfect sub-grid scale gravity-wave parametrization in this part of the atmosphere. Koshyk *et al.* (1999) showed by comparing spectra from high- and low-resolution versions of the same model, that sub-grid scale parameterizations are not representing adequately the effects of the unresolved scales in middle atmosphere. The absence of non-orographic gravity waves in the ECMWF model, for example, will amplify the model wind errors.

In the planetary boundary layer (PBL) where the vertical model levels of the ECMWF model are very dense, the effective vertical resolution does not appear much refined, where one may particularly notice that the wind-shear ratios remain above one for the profiles with radiosonde resolution of 1.32 km. Since the model level separation is small, this draws attention to sub-grid scale parameterization rather than to model level separation. Palmer (2001) suggests that some of the remaining errors in weather and climate prediction models may have their origin in neglecting some sub-grid-scale variability in current parameterization schemes. Above the jet stream (around 17-20 km), the model effective resolution seems to be improved to about 1 km. This is due to the overestimation of the jet level wind and wind shear in the model near the jet, as can be seen from the subtropical statistics of Figure 4 and Figure 5.

4.4 Conclusions

In this study we describe the atmospheric climate wind dynamics using collocated high-resolution radiosonde observations and the ECMWF model Short-Range Forecast (SRF). The results for the horizontal wind from both datasets are consistent since they reproduce pretty similar averages (mean and median) and variability at different levels of the atmosphere and over the various climate regions, as defined in this study. In fact, these results are as expected, since it is seen in most collocated model-radiosonde profiles, that the smooth ECMWF model wind profiles compare generally well in shape with the high-resolution profiles. However, with respect to the radiosondes, important small-scale vertical structures with high vertical wind gradients are lacking in the ECMWF profiles. Consequently, it is found that the average and climate variability of the wind shear is largely underestimated in the ECMWF model as revealed in the statistics. By comparing the statistics of successively in resolution degraded radiosonde profiles with the ECMWF model, the degree of difference in wind shear appears to be a factor of about 2.5 for the zonal wind and a factor of 3 for the meridional wind. Consequently, the effective vertical resolution of the ECMWF model is determined to be typically 1.7 km. It is moreover found that the radiosonde balloon drift is generally smaller than 100 km. Following the observation that the effective horizontal ECMWF model resolution is larger with a value of about 250 km, we conclude that the comparison between the climates of the

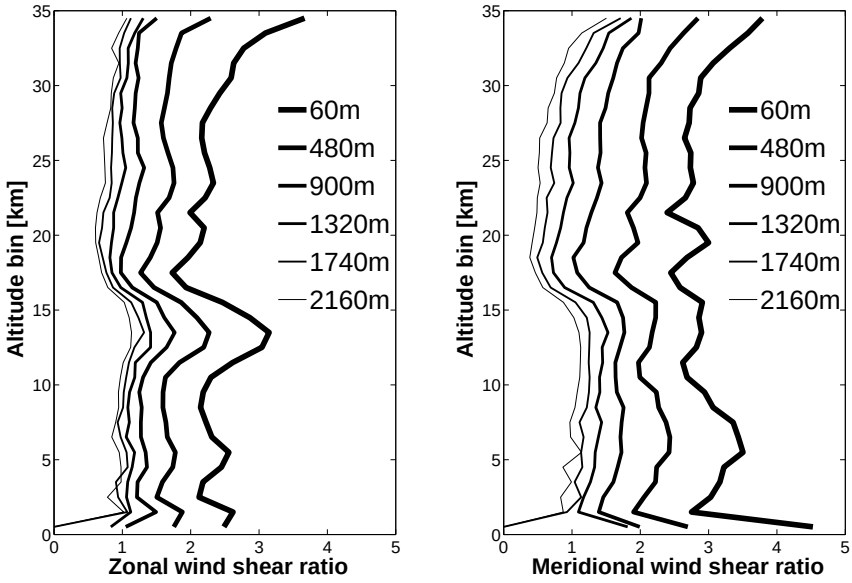


Figure 4.12: Zonal (left) and meridional (right) wind-shear ratios of series of successively degraded radiosonde resolutions and the ECMWF model. The two plots are based on median profiles. This example is from subtropics, and similar underestimation of wind shear has been found in the other climate regions.

ECMWF model and the radiosonde observations in this study is quite valid and consistent, even though the collocation of the radiosonde profiles is performed according to the radiosonde ground location. In addition, it is verified that the wind and shear climate statistics from radiotheodolite wind-finding systems (SPARC data) and from more recent wind-finding systems (LORAN and GPS) are comparable. The climate statistics obtained for 2006 from the SPARC data thus appear valid. The interannual variability computed from 10 years of SPARC data shows a large consistency of the annual climate, as well as some variability due to the QBO. This study highlights the difference in the representation of the atmospheric wind by the ECMWF model and radiosonde observations. We demonstrate in particular that on the one hand the ECMWF model is well capable of simulating the horizontal wind climate and its variability, but on the other hand it is deficient in the wind shear climate and its variability. The effect of vertical smoothing on individual collocated radiosonde-model wind and wind-shear profiles is investigated using four reduced resolutions. This shows mainly that the wind and wind-shear variabilities that are on average lost by smoothing to the ECMWF vertical resolution appears to be 2-4 m s^{-1} for wind and 0.01- 0.015 s^{-1} for wind shear, respectively, when reducing radiosonde resolution from about 30 m to 2 km vertical resolution. But these values may exceed 5 m s^{-1} (wind)

and 0.015 s^{-1} (shear) near the jets. The lack of vertical wind shear in the ECMWF model may have implications for the parameterizations of turbulence, gravity wave drag and convection which, ideally, should be resolution dependent. Besides the importance of this study for NWP and climate modeling, it is used as an immediate application in the framework of ESAs ADM-Aeolus, to investigate the optimal vertical sampling of the Aeolus Doppler Wind Lidar, planned for launch in 2017. The Aeolus DWL has a limited number of vertical range-bins (24) that need to be distributed in an optimal way such that the maximum information content on wind and shear may be obtained from the atmosphere by the mission. This study is thus exploited to build a global and dynamically and optically realistic atmospheric database needed for the simulation of the Aeolus DWL, where adjustments are made to ECWMF winds to obtain a wind and shear climate compatible with the radiosonde database described in this chapter. For more details, see Stoffelen *et al.* (2009).

Temporal variability of wind and wind shear

CHAPTER 5

Temporal Variability of Wind and Wind Shear

Abstract

Temporal variability of wind and wind shear are investigated here from both radiosondes observations and ECMWF model. Monthly, seasonally and interannual statistics are established from the two datasets and over different climate regions. Prior to the temporal variability analysis, we first quantify the spatial variability differences of wind shear between ECMWF model and radiosonde observations in different climate zones, as introduced in chapter 4. Wind shear from a Cloud Resolving Model (CRM) is also compared with the ECMWF model. This is performed by introducing the notion of the Error Multiplier Profile (EMP) that is used also to quantify seasonal and monthly variabilities. Similar notions as EMP have been introduced to quantify the differences between model and observations over land and over sea, but also to quantify land-sea, dawn-dusk differences in both modelled and observed datasets. The results show similar wind shear underestimation factors over land and over sea at all atmospheric levels, apart a the Planetary Boundary Layer (PBL). However, the variability differences between land and sea and dawn and dusk show a factor extending from 2 to 3 for wind shear and 1.5 to 2 for wind within the PBL, while it is close to 1 above. The temporal variability results show mainly that the ECMWF model reproduces globally similar temporal variability for wind and wind shear profiles as compared to radiosonde observations in terms of distribution shapes. However, in terms of magnitude, wind shear is underestimated in the model in line with what is found in chapter 4. The quantification of the variability differences for the different climate regions over the year 2006, over different seasons, and over different months, shows a generally similar underestimation factor of 2-3 in ECMWF model wind shear in the free troposphere and more than 3 above. But the largest contrast in the variability is observed in monthly winter periods and monthly summer periods. Furthermore, the CRM model produces about the same amount of wind shear in the free troposphere while it is larger at the PBL and lower at the

stratosphere.

5.1 Introduction

Seasonal or diurnal variations in wind profile variability may imply changes in the Aeolus-DWL observing strategy (see section 2.4.2 of chapter 2). For building a representative atmospheric database at high resolution of dynamically and optically heterogeneous conditions, the occurrence of wind shear anomalies over time may be relevant too (Marseille *et al.*, 2011). Moreover, differences in the interseasonal, monthly and diurnal variability may guide the adaptation of NWP model winds to better follow observed wind profiles. In chapter 4, wind and wind shear at different atmospheric levels (profile) over four climate regions (tropics, subtropics, midlatitudes and polar) are characterized. it is mainly found that wind distributions in ECMWF model and radiosonde observations are quite similar, while the average and variance of wind shear are grossly underestimated (Houchi *et al.*, 2010).

In this chapter, prior to the temporal analysis, we first describe in section 5.2 the general methodology used to quantify the various differences in data set types which are found in chapter 4: ECMWF model and radiosonde observations over the different climate regions, over land and sea and dawn-dusk differences and also ECWMF and CRM model differences. Then the results of these different quantifications are shown in section 5.3. In section 5.4, we present the analysis of the temporal variability (over months, seasons and within a 10 year period) of wind and shear for the four climate regions, including the quantification of the differences between model and observations based on the notion of the Error Multiplier Profile (EMP). Similarities and differences in the representation of wind and wind shear climate are highlighted and discussed.

5.2 Analysis Method and Definitions

In chapter 4 the statistics of wind and shear have been established for zonal and meridional wind and wind shear as function of height (1-km vertical bin) over various climate regions. We are mainly interested in the temporal variability in this chapter, so similar analysis over months, seasons and a 10-year period have been established. In order to quantify the differences in time (temporal variability) of the vertical wind and shear variations in the radiosonde and the ECMWF model data sets, some notions and definitions based on profile ratios are introduced: first, based on the mean and percentile profiles of wind and wind shear for the two data sets, we define what we call an Error Multiplier Profile (EMP), which is the ratio of the difference between the third quartile (75%) and the median (50%) for the two data sets, i.e., radiosonde (Rs) and ECMWF (Ec), as provided in equation 5.1. The EMP is computed for different seasons and months and over various geographical regions, and it can be seen as a quantification of the difference in variance between model and radiosonde data sets.

$$EMP(z) = \frac{|p_{75}(z) - p_{50}(z)|_{Rs}}{|p_{75}(z) - p_{50}(z)|_{Ec}} \tag{5.1}$$

The above equation is also applied within a dataset, but classified in land and sea to investigate whether the differences between model and observations are much higher or lower over sea than over land.

An example of the segregation between sea and land stations for the subtropics dataset is shown in figure 5.1a. Similar to eq. 5.1, an Error Multiplier type metric can be defined to quantify the differences in wind and wind shear variation between land and sea profiles (LSP) and dawn ($6AM$) and dusk ($6PM$) profiles (DDP) as defined successively by equations eq. 5.2 and eq. 5.3. These two last equations may be applied to both radiosondes and the ECMWF model independently, to compare wind and shear differences between land and sea and between dawn and dusk in both datasets.

$$LSP(z) = \frac{|p_{75}(z) - p_{50}(z)|_{land}}{|p_{75}(z) - p_{50}(z)|_{sea}} \tag{5.2}$$

$$DDP(z) = \frac{|p_{75}(z) - p_{50}(z)|_{dawn}}{|p_{75}(z) - p_{50}(z)|_{dusk}} \tag{5.3}$$

In addition, the differences of the mean and median wind and shear at different levels of the atmosphere is quantified by defining the ratio of mean and median profiles (eq. 5.4) between radiosondes and ECMWF model, between land and sea (for both datasets), or between dawn and dusk (for both datasets). The profile ratios for means or medians are defined successively as follows:

$$Rp(z) = \frac{m(z)_{Rs}}{m(z)_{Ec}} \tag{5.4}$$

Similar ratios are used to quantify the wind and shear variability differences from one year to another by fixing the first year of 1998 as a reference climate year.

$$R_{iYear}(z) = \frac{m(z)_{year}}{m(z)_{1998}} \tag{5.5}$$

with $iYear = 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007$

Remarks:
Note that the word ratio is used as a generic label in some plots. But the legends (or titles for some plots) indicate which equation (or quantities) have been used in each plot. Moreover, the legends indicate also which data have been used, whether radiosondes (Rs) or ECMWF (Ec), land/sea, 6AM/6PM, CRM etc. The title may indicate also whether mean or median ratio profiles have been used. Notice also the decadic logarithm scale along the horizontal axis for wind shear plots

Wind variability matching

In the framework of the VAMP project, establishing a new type of EMP based on the standard deviation ratios of model and radiosonde wind shears (EMP_σ) was necessary to build an atmospheric database for Aeolus-DWL simulations (Marseille *et al.*,

2011). In addition, results of wind shear statistics from CRM have been evaluated against the ECWMF model, thus this more general equation 5.6

$$EMP_{\sigma}(z) = \frac{\sigma_{Rs/CRM}(z)}{\sigma_{Ec}(z)} \tag{5.6}$$

The new (EMP_{σ}) profiles established for the different climate regions are utilized to complement the missing model wind variability to a level of variability compatible with radiosondes.

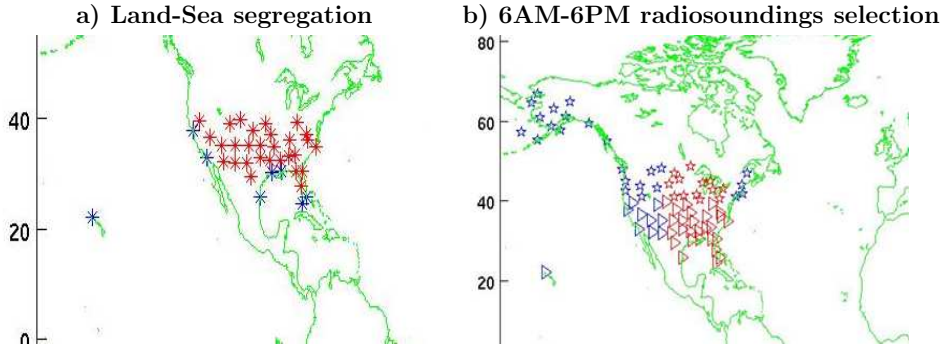


Figure 5.1: a) Segregation of sea (blue) and land (red) radiosonde stations in the subtropics. Notice that the segregation is based on topography (below mean sea level is sea and above is land); b) illustration of the selection of the radiosounding sites with launch time falling within one hour of 6AM (12UTC) and 6PM (00UTC) local times, here shown for 12UTC in red among the whole stations blue. This is shown for both subtropical (triangles) and mid-latitude (stars) stations.

5.3 Quantification of Variability

In this section we first quantify the differences between the modelled and the observed wind shear variability over the various climate regions in 2006 for SPARC collocated data, seen in figure 4.4 of chapter 4. We use for that the notion of the Error Multiplier Profile (eq. 5.1). Notice that the Error Multiplier Profile (EMP) can be seen as an estimation of the difference (using the ratio) in the variances between ECMWF model and observations. We then present similar quantification for separated sea and land data sets. Furthermore, we quantify the differences in wind and shear variability between land and sea, then between dawn and dusk from both model and observations, using the notions of Land Sea Profile (eq. 5.2) and dawn dusk Profile (eq. 5.3). Later on in section 5.4, we use similar definitions to estimate the differences in model and observation variability over seasons and months.

5.3.1 Annual Variability

The difference in wind-shear variability between SPARC radiosondes in 2006 and ECMWF is quantified as shown in the two left plots of figure 5.2, which show from left to right the ratio of the median profiles and the Error Multiplier Profile (eq. 5.1) for the different climate zones. The result shows that the variability in the radiosonde shear is more than 2 times higher than in ECMWF throughout the troposphere. This means that the ECMWF model resolves less than half of the radiosonde wind shear variability. From the tropopause and up, the ECMWF variability is even more reduced which is probably due to the reduction in vertical resolution at these levels which is progressively and rapidly decreased as can be seen in figure 4.1. However, at these high altitudes the results remain uncertain because of the possible effect of gravity waves, which are not taken into account in the ECMWF model (Koshyk *et al.*, 1999) and which may increase errors in wind measurements (Barat & Cot, 1992), especially for the less accurate wind-finding systems (Radio-theodolite here). One notes as well the persistence of this effect for all climate regions.

The profiles of median ratios look very similar to the profiles of the ratio of variance differences (EMP) at all atmospheric levels. This indicates about similar underestimation in the average and variance of wind shear in the ECMWF model as compared to radiosondes. We notice however that the factor (> 2) is slightly smaller than what we found with the method used in figure 4.12. This is due to the fact that the shape of the ECMWF and observed PDFs is not identical. When distributions have similar shape, then they may be described by a single scaling parameter. For example, for a normal distribution the ratio of the interquartile range ($75\% - 50\%$), standard deviation is a fixed number and independent of scaling. So, if both the radiosonde and ECMWF wind shear PDFs would be normal, or more generally, both PDFs have the same scalable shape, a ratio based on percentiles would provide the same number as a ratio based on variances. Seemingly, the shape of the radiosonde and ECMWF model PDFs are different and not mutually scalable, resulting in different percentile and SD ratios. Notice however that results with another second EMP method based on sigmas (EMP_σ, see eq. 5.6) of observed and modelled wind shear; and which gives similar factors as in figure 4.12 is presented in section 5.5 of this chapter.

The right plot shows the quantification of CRM model and ECMWF model wind shear variances. In the free troposphere the ratio is close to 1, while it is higher below and lower above this value. This indicates that this particular CRM model produces more wind shear in the PBL, where generally global circulation models, such as ECMWFs, fail to represent turbulence well, and above about 14 km the CRM variance is reduced due to additional damping towards the stratosphere. Since we have shown that the ECMWF model lacks vertical shear in the free troposphere and is rather smooth vertically, it is surprising indeed that this CRM has similar smoothness.

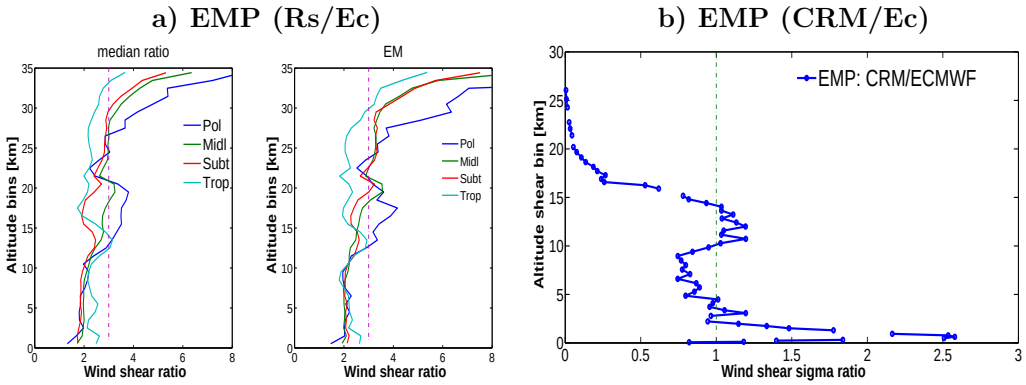


Figure 5.2: a) ratio of the median and Error Multiplier (eq. 5.1) profiles of radiosonde and ECMWF model over the different climate zones, Polar, Midlatitude, subtropics and tropics; b) ratio of interquartiles (EMP) of Cloud Resolving Model (CRM) and ECMWF model (see text).

5.3.2 Variability over Sea and Land

Based on the topography (i.e., at sea-level is sea and above is land), sea and land SPARC data have been segregated for 2006, but only for the subtropics and midlatitude stations because of limited coverage in the other climate regions. Figure 5.1a shows an example of the segregation in the subtropics. Land stations close to the coast, near sea-level, and prevailing wind from the oceans are selected as sea stations. Statistical analysis as in section 4.2 of chapter 4 has been performed over land and over sea, then similar statistics plots (not shown) as in figures 4.4 and 4.5 have established. But the idea here is to quantify directly the differences between radiosonde and ECMWF variability of wind and wind shear over the two earth surface types, using the ratio of mean values and EMP profiles as shown in figure 5.3. However, we notice from the global statistics that wind and wind shear variability are generally larger over land than over sea, as indicated by the following median/mean maximum values for instance for the subtropical region around the tropopause: 20/21 ms-1 (zonal wind), 0.007/0.008 s-1 (zonal shear) and 60/66 km (drift) for coastal stations (sea) and 24/25 ms-1 (zonal wind), 0.0075/0.010 s-1 (zonal shear) and 71/77 km (drift) for land stations.

As can be seen in figure 5.3, the difference in zonal wind and shear between model and observations is similar than that we saw from figure 5.2a, but we see in addition here that the magnitude difference Rs-EC wind and shear is larger over land than over sea throughout the troposphere. The maximum values are seen in the lowest kilometers (PBL), with a wind/shear factor below 1.5/2.5 for sea and exceeding a factor 2/3 over land. This may indicate that beside the underestimation of wind shear as seen previously, the model has more challenges to reproduce a with the observations compatible wind and wind shear in the PBL with its complex dynamics. Furthermore, we notice an overestimation over land of wind shear in the model in the

lower stratosphere, as indicated by the means-ratio below one for wind.

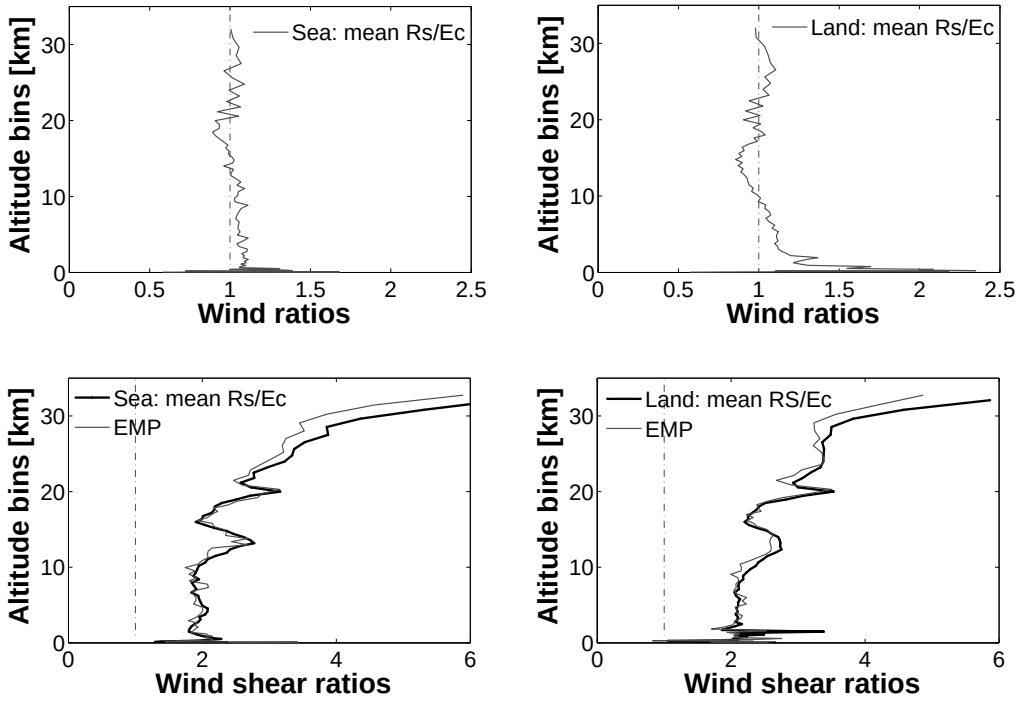


Figure 5.3: Collocated SPARC radiosonde observations and ECMWF model differences in zonal wind and shear statistics (means-ratio and EMP profiles) as seen over sea and land in 2006.

5.3.3 Land-Sea Comparisons

The segregated datasets of land and sea are used here to quantify the effect of the earth's surface type (land and sea) on wind and wind shear variability throughout the atmosphere, from both ECMWF model and radiosonde observations. Previously we have seen the differences in wind and shear variability between radiosondes and model over the 2 earth surface types, but here we rather quantify the differences in wind and shear variability between the 2 surface types from both datasets (radiosonde and model). This is given in figure 5.4 by the means-ratio profile (left) and the variances-ratio profile (LSP) (right). The two profiles look similar, this indicates that the underestimation in the mean and variance of wind and shear is about the same at each given atmospheric level. We clearly see the difference in zonal wind and shear between land and sea in the lowest kilometers (PBL), as illustrated independently from radiosondes and the model, with a Rs/EC factor of about 2/3 for radiosondes and 1.5/3 for ECMWF model. In fact, due to topography over land where, e.g., wind blowing over hills get slowed down close to the surface and an area of "speed-up" is

formed further up, variations in surface roughness up-stream create several so-called internal boundary layers where the wind profile is affected by various upstream surface roughness elements at various vertical levels in the PBL (Stull, 1988). Moreover, the nocturnal boundary layer is stable over land, leading to shallow low-level jets. On top of the stabilized PBL sea-breeze also forms shallow "jets", topography creates shallow low-level channeling of flows, etc.

The results found for the mid-latitudes (not shown here) confirm the differences in sea and land wind/shear variability, as we have seen in the subtropics. Also, similar statistics for a drift as in figure 4.6 which are not shown here, indicated larger radiosonde balloon drift over land than over sea in the subtropical and midlatitude regions. This is proportional to the wind and shear differences seen over the two earth surfaces; since the computation of the drift profile is based on zonal and meridional wind speed components 4.1..

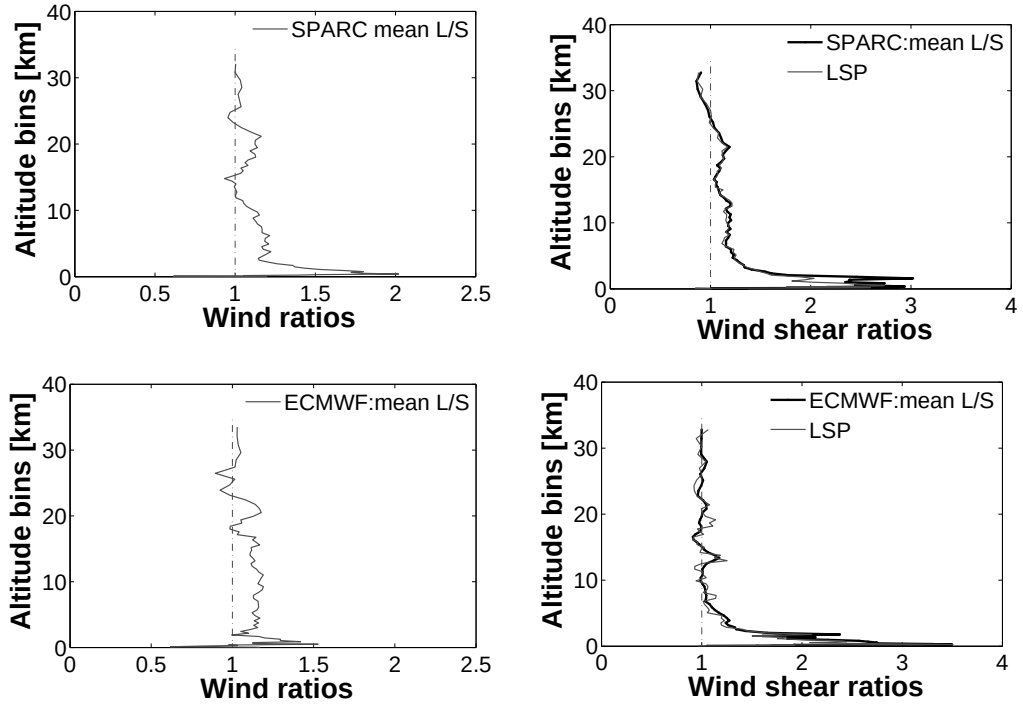


Figure 5.4: Land and sea differences in zonal wind and shear statistics (means-ratio and LSP profiles) as obtained from collocated SPARC radiosonde (top) and ECMWF model(bottom) datasets over 2006.

5.3.4 Dawn-Dusk Comparisons

A number of radiosondes in the SPARC data set can be used to study the atmospheric conditions at dawn and dusk which are the crossing local times of the Aeolus satellite, i.e., the local time of the DWL observations. Radiosondes corresponding to 6AM/6PM local time have been segregated, see map in figure 5.1b. Radiosondes launched between 5AM/5PM and 7AM/7PM local time are used here. This corresponds to all soundings situated in the band of -75 and -105 longitude. The 12UTC radiosonde locations situated between -75 and -105 longitude in the US correspond therefore to the 6AM local time soundings and 00UTC to 6PM local time. Since the longitude band and therefore the radiosonde stations are identical for dawn and dusk, typical dawn/dusk differences may be observed by studying these stations.

As for the analysis of the land/sea comparison, the differences in wind and shear variability of the mean and variance is given by the means-ratio and DDP profiles, as shown in figure 5.5. We see that, apart from the PBL, results of zonal wind and shear for 6AM and 6PM show very similar values, with the means-ratio and DDP profiles closely oscillating around 1 for all levels in the atmosphere. However, in the PBL the difference in wind and shear variability is more pronounced. The dawn/dusk difference appears to be smaller for the ECMWF model with wind/shear factors of about 1.4/2.0 in radiosonde data against 1.3/1.4 for the ECMWF model. Though these findings are similar for both radiosonde observations and ECMWF model, we notice a slight difference in the magnitude, especially in the mean-ratio profile. The same remark can be made for the drift (not shown), which is a little higher at 6AM comparing to 6PM with a few meters difference on average. These findings for dawn-dusk comparisons in the subtropics are valid also for the mid-latitudes (not shown). Notice that similar results are found from various CRM models and CALIPSO satellite data in the frame of VAMP project. This difference in dawn and dusk variability may be explained by the diurnal cycle of boundary layer. Indeed, apart may be in summer period, the nocturnal stable boundary layer with large wind gradients may remain stable at around 6AM, since the sun is may be not rising yet (depending on the altitude). After the sunrise a turbulent boundary layer that breaks verticals wind gradients start developing until it reaches the maximum air-mixing and height level around noon. At dusk, thermal effect that keeps turbulence disappears after sunset, but residual mixing air remains (Shaowu, 1994). This explains the less wind shear at dusk than at dawn. So for Aeolus DWL, we may conclude that the most challenging dynamical scenes in the lower part of the atmosphere are found at dawn.

5.4 Temporal Variability

In the following section, we present wind and wind shear climate over monthly, seasonal and interannual periods, successively. The results are obtained from collocated high resolution radiosondes and ECMWF model data sets and for different climate regions, i.e., tropics, subtropics, midlatitudes and polar. We also present various quantifications of the differences in variability of the modelled and observed wind and shear, based on the method and definitions detailed previously in section 5.2.

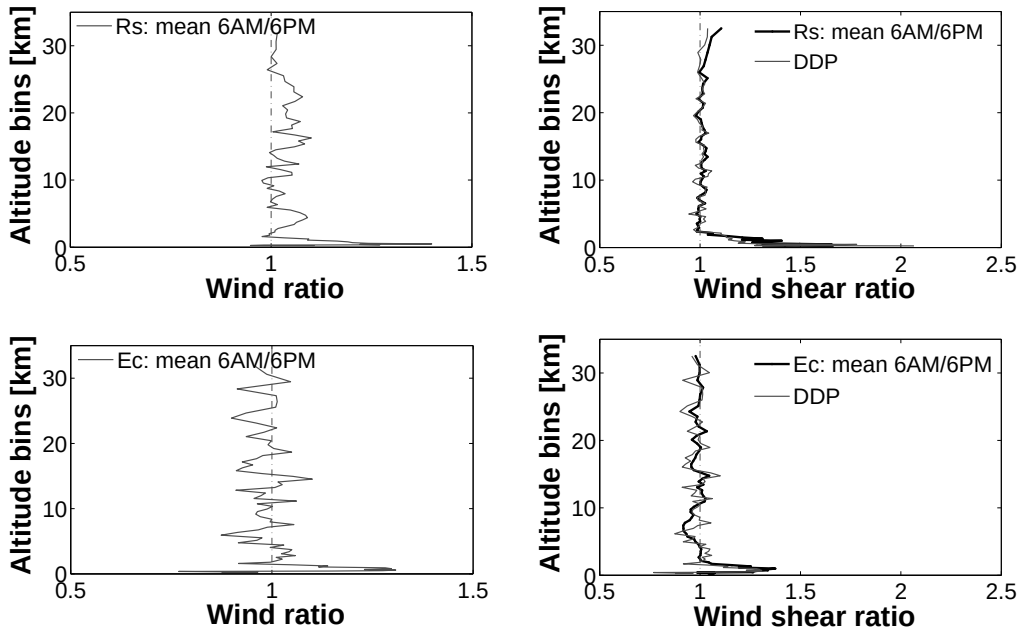


Figure 5.5: Dawn and dusk (at 6AM and 6PM) differences in zonal wind and shear statistics (means-ratio and DDP profiles) as seen from collocated SPARC radiosonde (top) and ECMWF model (bottom) data sets over 2006. Notice that for wind only the mean-ratio profile is given.

5.4.1 Seasonal Variability

We present and discuss here the results obtained for successively zonal and meridional wind and wind shear over the four climate regions as defined previously (see map in figure 3.1).

5.4.1.1 Zonal Wind

Figure 5.6 shows the seasonal statistics (mean profiles), successively for zonal wind and shear at different levels of the atmosphere, as seen from collocated SPARC radiosonde observations and ECMWF model and over different climate regions in 2006. One of the remarkable results is that the ECMWF model reproduces similar seasonal shape distributions in terms of mean profile for wind and wind shear. However, as found in chapter 4, while magnitude wind values in radiosondes and model are similar, wind shear values are underestimated in the model as seen for all seasons (summer, autumn, winter and spring). We also see that, globally, wind shear is well correlated with wind over the seasons, particularly between summer and winter periods. We see that more clearly for instance from the subtropical regions of figure 5.6. In summer for instance, wind and wind shear have the lowest values while they are the highest in winter. Autumn and spring periods show variability with mean values which are

globally between mean values of summer and winter periods, but with slightly larger values in Autumn than spring. For the midlatitude climate region, we show the results separately for sea and land stations. The seasonal variability trends of wind and shear profiles look similar over land and sea but with clearly larger values over land, as seen from both modelled and observed datasets. In addition, the mean seasonal profile values for all the climate regions are in coherence with the mean yearly profiles values of SPARC data over 2006 presented in figure 4.4 of chapter 4. Note also the much higher PBL shear over land, probably due to higher friction.

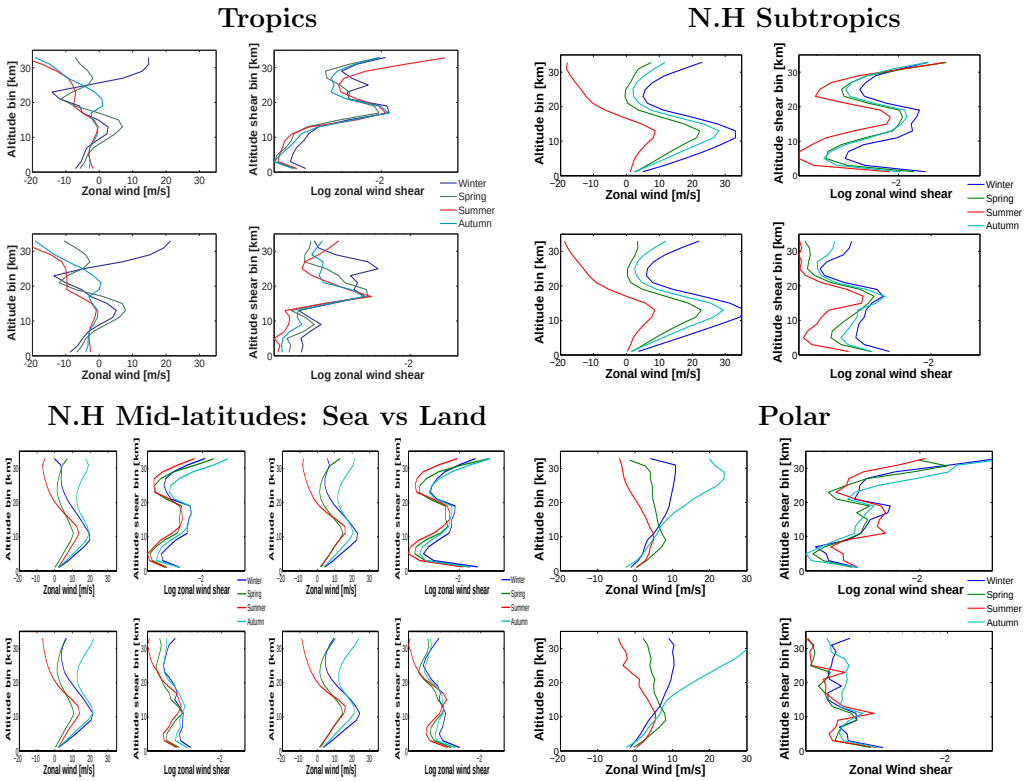


Figure 5.6: Seasonal statistics of zonal wind and shear from collocated SPARC radiosonde observations from 2006 (top) and the ECMWF model (bottom) over different climate regions. Notice that for the midlatitude regions the statistics are presented separately for sea and land successively.

5.4.1.2 Meridional Wind

As for zonal wind and wind shear in figure 5.6, similar statistical analysis is performed for the meridional wind and shear as shown by the mean profile plots in figure 5.7. The zonal wind dominance in the subtropics in autumn, winter and spring provides large interseasonal variability of mean wind, while this is lacking for the meridional wind.

The interseasonal meridional shear variability is also somewhat smaller. For zonal wind, the ECMWF model reproduces quite well the seasonal distributions seen in radiosonde observations. However, for the tropical region, the radiosonde meridional wind appears clearly more pronounced close to the tropopause than ECMWF. Though the magnitude of the meridional wind remains generally smaller than that of the zonal wind for all climate regions.

The large interseasonal contrast is between summer and winter as seen in all climate regions. Again as for zonal wind from midlatitudes plots, we notice similarity in seasonal variability trends of wind and shear profiles over sea and land, but with clearly larger wind shear values over land as seen from both modelled and observed datasets. The lack of wind shear in the ECMWF stratosphere with respect to the radiosondes appears even more pronounced for the Meridional wind than for the zonal wind generally. Furthermore, the general behavior of wind and shear profiles in the different climate zones, as seen in figure 4.5 of chapter 4, are in coherence with these seasonal statistics. For example, the general lack of shear in the ECMWF model as compared to radiosondes occurs in all seasons.

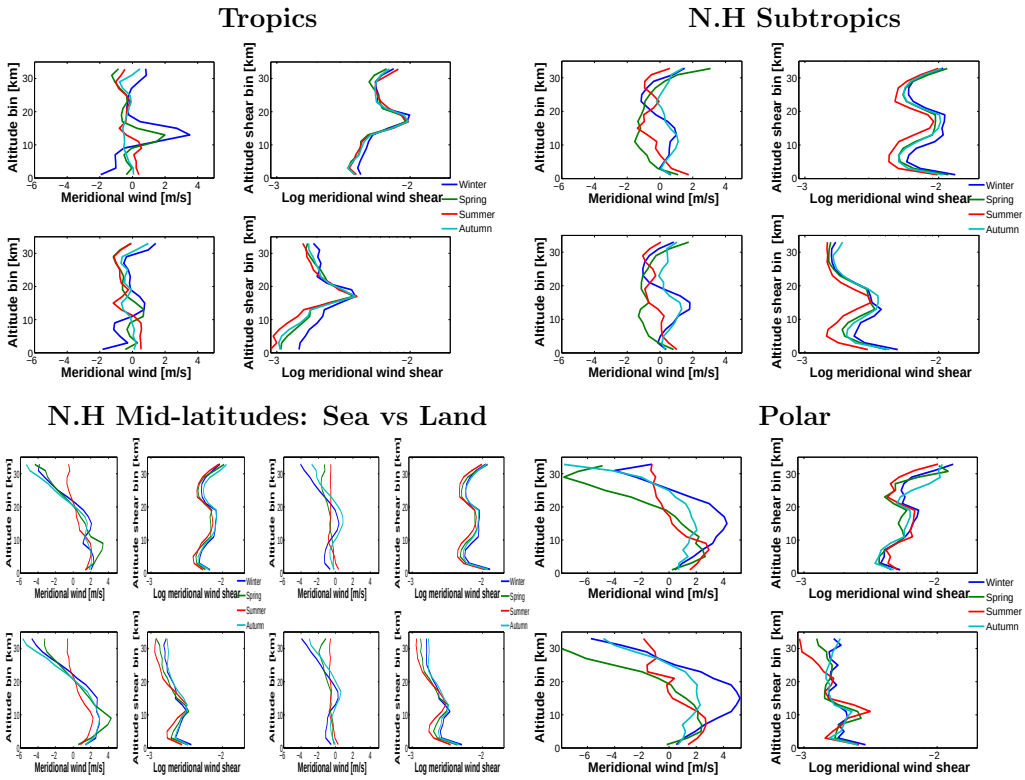


Figure 5.7: Similar plots as in figure 5.6, but now for meridional wind and shear.

5.4.1.3 Mean and Variance Ratios

As in the right plots of figure 5.2, we quantify the difference in zonal wind-shear variability between radiosondes and ECMWF but here over the different seasons (summer, Autumn, winter and spring), as shown in figure 5.8. The estimation of the differences of the average and variance is given successively by the median-ratio profiles (5.4) and the Error Multiplier Profiles (EMP) (eq. 5.1). Globally, the results show similar factors of underestimation of wind-shear mean and variance as the yearly median-ratio and EMP profiles, shown in figure 5.2a for the different climate zones. However, we may notice that a big contrast in wind and wind shear underestimation by ECMWF is between summer and winter periods, and this mainly at the stratospheric levels in the Mid-latitude and tropical regions with gross underestimation of zonal wind and shear in summer by ECMWF.

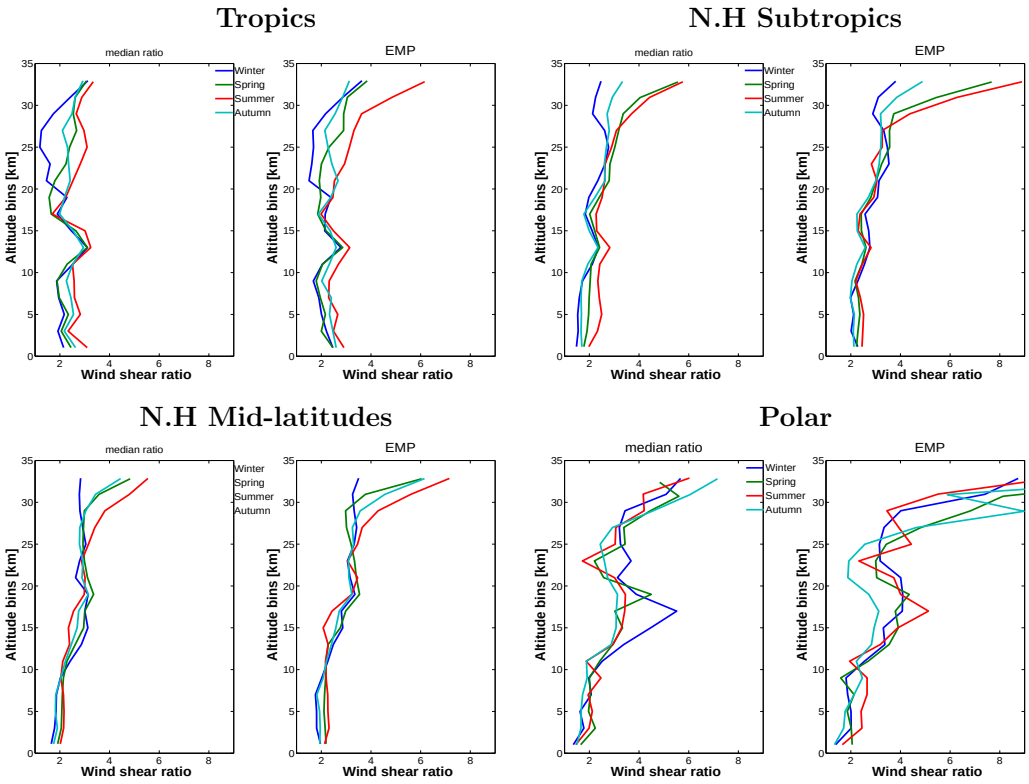


Figure 5.8: Seasonal median-ratio profiles (left) and Error Multiplier Profile (EMP; right) over the different climate regions as indicated. Notice that 12s ($\sim$ 60 m) resolution radiosonde data were used here.

5.4.2 Monthly Variability

Figure 5.9 shows the monthly statistics of zonal and meridional wind and shear over the tropical and polar regions, where the changes in meridional wind are generally larger than in the subtropical region and midlatitudes. This is mainly due to the dominance of zonal wind shear in the two latest regions. In fact, we see large monthly wind and shear variabilities in in the polar region above the tropopause for the zonal wind and throughout the whole atmosphere for the meridional wind. The largest mean wind change is seen between November and December. In the tropics, the largest wind and shear changes are seen in the stratosphere. Generally the results from radiosonde observations and ECMWF model appear similar, apart from the already mentioned substantial underestimation of the wind shear in the ECMWF model as compared to radiosondes.

Figure 5.10 shows the quantification of the differences in the monthly average and variance (given successively by the median-ratio profile and EMP profiles) between model and radiosonde wind shears, over different climate regions. The results show globally similar factors of underestimation in the average and variance of wind shear in 2006 over the whole atmosphere and in all months, successively in figure 5.2a and in figure 5.8. Moreover, the results highlight the largest contrast between model and radiosonde monthly wind shear variabilities in the polar and, to a lesser extent, the tropical regions, in particular for the winter month of January at the north pole. This might explain the difficulty of the ECMWF model to reproduce the complex dynamics that characterizes the polar region associated with the drastic changes in variability that occur over months and seasons, as seen from the global statistics in figures 5.9, 5.6 and 5.6. In the tropical region, the ECMWF model also misses salient features, e.g., the mean meridional wind near the tropopause. This is in line with one of the objectives of the Aeolus mission to improve wind observation in the tropical region.

5.4.3 Yearly Variability

Figure 5.11 gives a global picture of the interannual wind and wind shear variabilities as seen from high-resolution radiosonde observations for the different climate regions over the ten-year period from 1998 to 2007. The results are given by the mean profiles presented for land and sea separately and for all climate zones, except for the polar station while only one station is available over land. So instead, we present at the right side of the polar plots, the quantification of the differences from one year to another by computing the differences and ratios of the annual mean profiles and the chosen reference profile over 1998 for the subtropical land-region, see equation 5.5. Because only one station is available over the polar region, the climatology results for this region may be not very representative and merely provide a first indication. We notice also that because of missing data in the subtropical sea stations for the year 2007, the mean profile of wind and wind shear look biased. The tropics-land profile over 1999 looks also biased, because of the improper vertical sampling for some radiosoundings for this year.

Generally the results show similar climatology of wind and wind shear from one

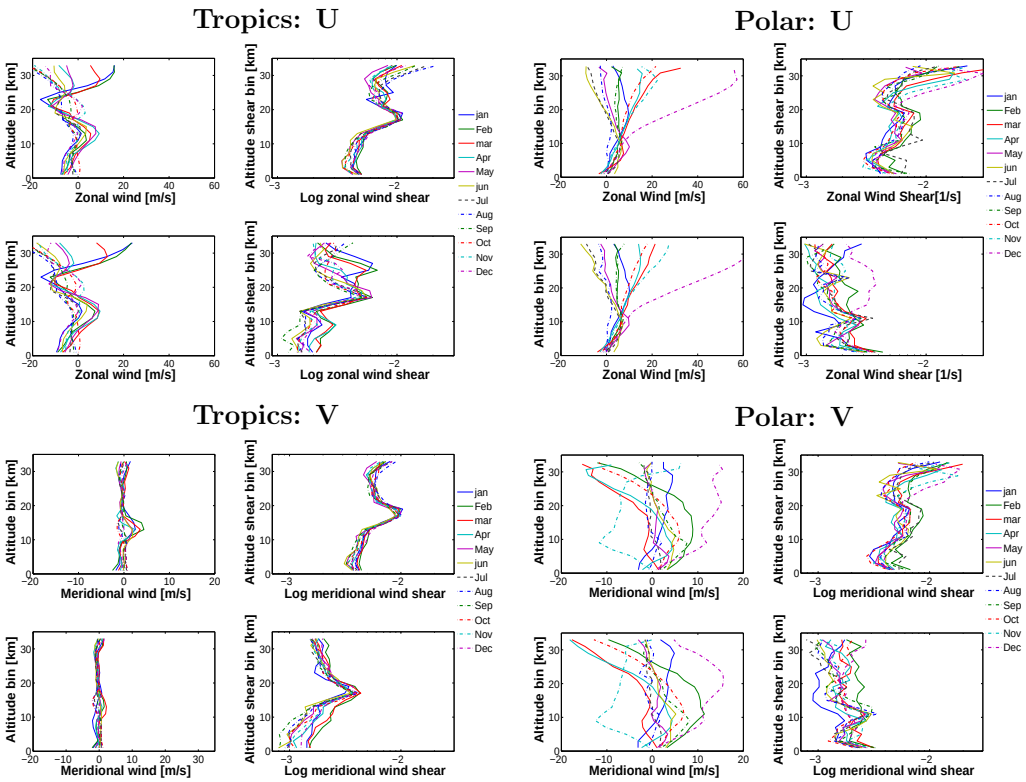


Figure 5.9: Monthly statistics of zonal (upper half) and meridional (lower half) wind and shear from radiosonde observations (top panel) and ECMWF model (bottom panel) over the tropical (left) and polar regions (right).

year to another in all climate regions and almost at all atmospheric levels. At some atmospheric levels, the interannual variability of wind and sometimes wind shear may be significant, as for instance in the tropics at high altitudes (from 15 km upwards). These high fluctuations (seen both over land and sea) are known as the quasi-biennial oscillation (QBO), which is characterized by a periodic wind reversal driven by atmospheric waves emanating from the tropical troposphere and traveling upwards while dissipated in the stratosphere by radiative cooling. The precise nature of the waves responsible for this effect is not clear, but it appears that gravity waves are seen as the major contributor; see Baldwin *et al.* (2001) for more details. The mean profiles of wind and wind shear for the polar station show both important variability over the 10-years period above the troposphere in particular. Up to about 30 km, wind difference from year to another as showed for the subtropical land-region is generally below 5 ms^{-1}, with a maximum value at the troposphere. Above 30 km altitude the values may be higher because mainly of the effect of gravity waves, as mentioned previously.

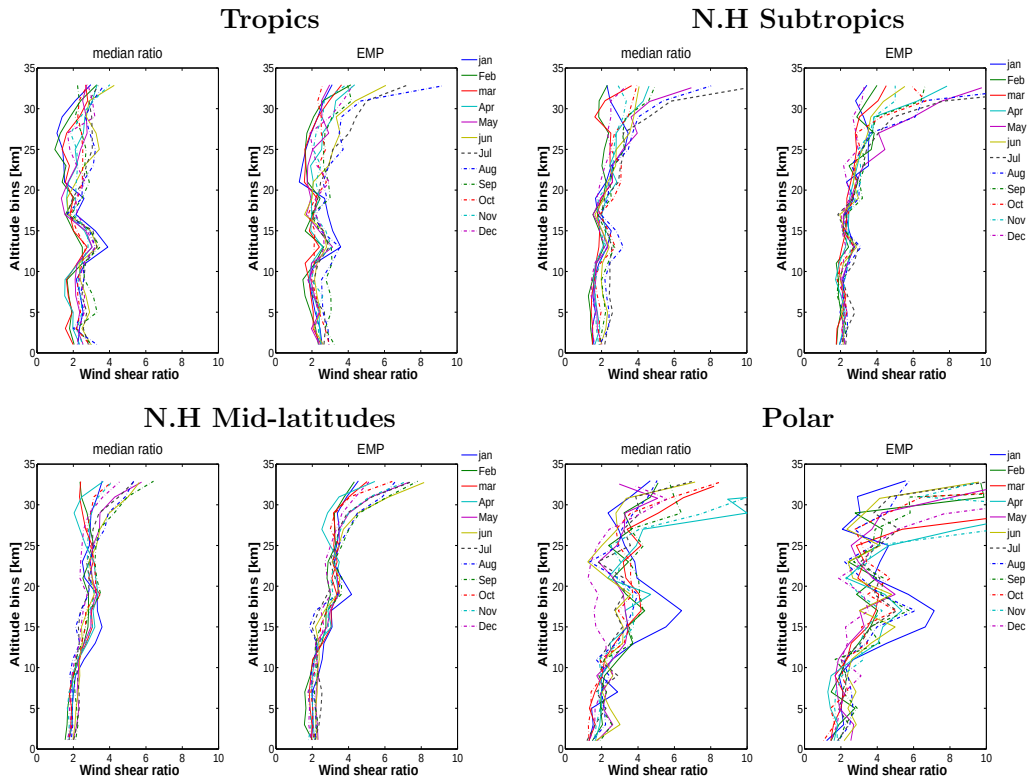


Figure 5.10: Monthly median ratio profiles (left panel) and Error Multiplier Profiles (right panel) for different climate regions as indicated.

5.5 Adaptation of ECMWF Wind Profiles: EMP_σ

As in figure 5.2, figure 5.12 shows a quantification of the differences in the variance of wind shear between radiosonde observations and ECMWF model over four climate regions, using the definition of the EMP_σ (5.6). The results show generally similar factors of underestimation in wind shear in the model with respect to radiosondes, as also found in the method of chapter 4.3. These EMP_σ profiles are used to upscale ECMWF-model wind and shear characteristics to be compatible with those from high-resolution radiosondes. This is necessary to create realistic (including small scales structures) and global wind-shear statistics, that are needed in combination with optical heterogeneities for Aeolus simulations (see chapter 2).

The complete description of the method used to adapt ECMWF model wind and shear profiles is described in a technical report (see Stoffelen *et al.* (2009), in chapter 5). However, an example of the effect on the ECMWF model wind profiles when adding the missing small-scale structures as noted in radiosonde profiles is shown in figure 5.13a. Another method to adapt model winds to radiosonde winds, i.e., render model wind profiles compatible with radiosonde wind profiles, including the

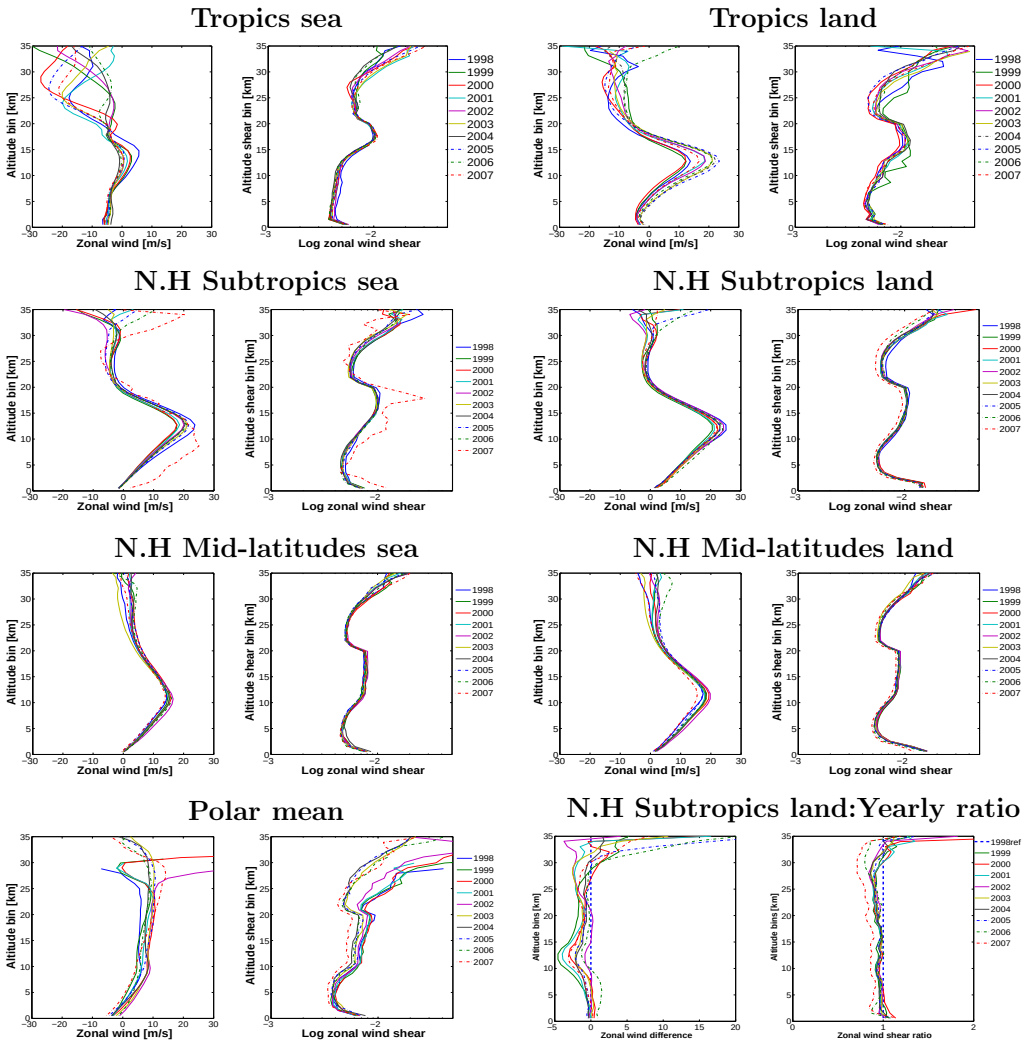


Figure 5.11: Interannual variability of wind and wind shear over land and sea for the ten year period 1998-2007, successively for the four climate regions as defined in the map of figure 3.1. Notice that the last two plots bottom-right show the difference (left) and ratio (right) of the yearly profiles with the 1998 profile for resp. zonal wind and shear.

addition of optical properties from CALIPSO satellite data, is described in details in Marseille *et al.* (2011). This latter method is based on the difference between the observations (O, radiosonde) and background (B, ECMWF model), known as "background departures" or simply O-B in Data Assimilation (DA). See for example figure 5.13b. More precisely, the method is based on an eigenvalue decomposition of the positive definite O-B covariance matrix. For more details, see the last cited

reference.

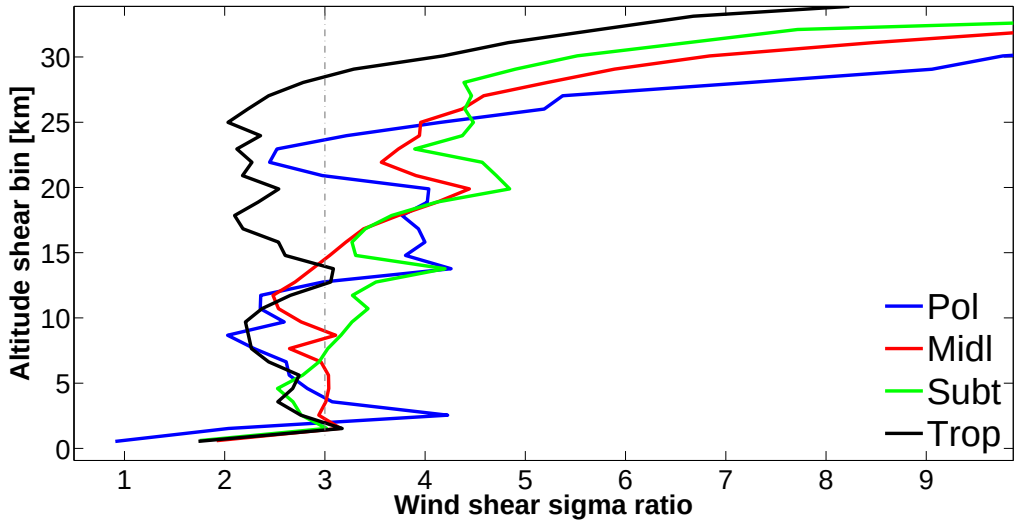


Figure 5.12: Error Multiplier Profile (EMP) based on the standard deviations of radiosonde and ECMWF model, see equation 5.6

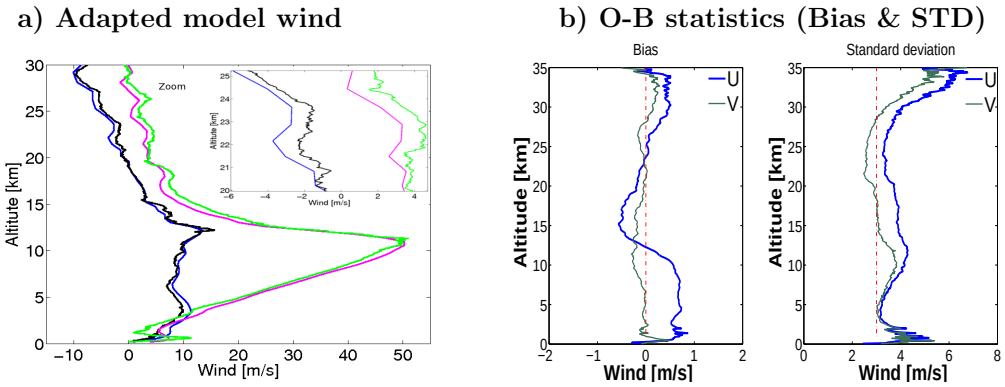


Figure 5.13: a) ECMWF zonal (blue) and meridional (magenta) wind profiles and their corresponding adapted versions in resp. black and green. Notice in the same panel a zoomed pane over about 5 km (between 20 and 25 km) that illustrates more in detail the small-scales structures added to the smooth model winds. b) Observation minus background (O-B) statistics from high-resolution radiosondes and ECMWF model 12-h forecasts; bias (left) and standard deviation (right) of both zonal (blue) and meridional (magenta) wind components.

5.6 Conclusions

In this chapter, first the difference in wind and wind shear variability between ECMWF model and radiosondes is quantified (briefly extended to a CRM), over different climate zones. Also, the differences between land and sea, and dawn and dusk from both modelled and observed datasets have been quantified. The results confirm that radiosondes exhibit more wind shear variability than the ECMWF model, but also more than the CRM model with about similar wind shear underestimation factors (2-3) throughout the free troposphere and more then 3 above, as found earlier with the method (4.3) in chapter 4. Next, except for the Planetary Boundary Layer (PBL), similar results have been obtained over sea and over land at all atmospheric levels. However, in the PBL the variability differences between land and sea and dawn and dusk show factors increasing from 2 to 3 for wind shear and 1.5 to 2 for wind, while for the latter it is close to 1 above the first 2-3 km.

Furthermore, temporal variability (monthly, seasonally, and over a 10-year period) has been investigated through a search for trends in the average (mean and median) and variance (SD and EM) of wind and shear profiles. The analysis is done for modelled and observed data sets allowing a comparison of the results on many aspects. Generally, the results show that the ECMWF model reproduces quite well the monthly and seasonal variations of wind and shear, as seen in radiosonde observations, in terms of trends at all atmospheric levels. But, in terms of magnitude, the monthly and seasonal wind shear is consistently underestimated as found earlier from the yearly EMP computed from collocated SPARC and ECMWF data over 2006 in section 5.3.1 and chapter 4. Only slight differences occur from one season to another, and from one mounth to another. Nevertheless, we notice that the largest contrast in the variability is observed between winter months and summer months . We have also observed from the 10-year period of radiosonde data, that the differences in wind and shear variability from one year to another are generally very small, apart from some specific cases, e.g., the Quasi-biannual oscillation (QBO). Therefore, we basically conclude that EMP is time independent and applicable for any season or year. As such, the analyses on the collocated SPARC and ECMWF data over 2006 are representative for the other years.

The overall results of this chapter on the temporal variability combined with the results on spatial variability of chapter 4 for the different climate regions are very important, specifically in the framework of the VAMP study. They are utilized to adapt the ECMWF model wind and shear variability to match the radiosonde variability. It is in particular concluded that, because of the small differences in temporal variability between ECMWF model and radiosondes, the method developed to adapt the ECMWF model wind shear can be simplified by taking only the spatial variability differences over the climate regions. Therefore, an EMP_σ profile was established for different regions as shown in 5.5, and has been used to adapt the ECMWF model wind shear variability to the more enhanced radiosonde profile variability. The results have been used to build a realistic atmospheric database at high resolution

(Marseille *et al.*, 2011). This optically and dynamically high-resolution database is necessary to perform the Aeolus-DWL simulations.

Cloud vertical structure

CHAPTER 6

Cloud Vertical Structures

Abstract

Cloud Vertical Structure (CVS) are an important information for weather and climate predictions in general. In the context of Aeolus mission studies, CVS are very important because the horizontal and vertical distribution of cloud particles determines the quality of Aeolus winds in cloudy regions. The relative low vertical resolution of Aeolus profiles, with typical bin sizes of 1 km in the free troposphere, implies that the exact location of cloud layers and their particle density inside the bin is inherently unknown without additional information from other sources. Thus, profiles of pressure, temperature and relative humidity (PTU) from high-resolution radiosondes are collocated with ECMWF model profiles in order to derive statistics of CVS, in particular vertical distributions over a fixed 1-km height bin, the frequency of cloud occurrences and ultimately the cloud liquid water and ice content. For the validation of the results, they are compared against CVS statistics derived from the CALIPSO-satellite Vertical Features Mask (VFM). Apart from the gradient profiles of pressure and relative humidity, most of the statistics of the PTU variables from radiosondes and ECMWF model are globally comparable. Indeed, there is a difference between observed and modeled gradient profiles of pressure and relative humidity, which are generally a factor of about 1.5 higher in radiosondes, mainly close to the earth surface. Because of the incomplete information that is provided in each data set, the CVS from the 3 data sets are complementary. Indeed, they show similar trend in the vertical distributions of cloud but with slightly different frequency of occurrences as a function of height. This may be explained by various known reasons, such as: point measurement for radiosonde, limited detection of low cloud from satellite instruments, etc. Finally, we conclude that below 10 km altitude the CVS profiles from radiosondes are consistent with the CVS of other data sets, and therefore may be combined with wind and shear profiles from the same instrument to investigate the effects of cloud on Aeolus DWL measurements. A methodology is proposed and

discussed to simulate Aeolus sampling and performance.

6.1 Introduction

Clouds are among the most fundamental components of the weather and climate systems. In fact, clouds, which represent an important stage of the water cycle in the atmosphere, significantly influence the energy budget of the planet (e.g. Stubenrauch *et al.* (1997)) and the atmospheric circulation (e.g. Wang & Rossow (1998). Clouds are formed at different levels of the atmosphere in humidity-saturated air through two processes which may act together: cooling the air (generally by upward motion) and adding water vapor to it by evaporation. Clouds cool the earth by reflecting the sunlight (short-wave radiation) back to space, but they can also serve as blankets by trapping warmth by absorbing and re-emitting thermal (long-wave) radiation emitted by the earth surface and low atmosphere. See Wielicki & Harrison (1995) for more details on this subject. Also, the role of clouds on the atmospheric circulation has been recognized for many years, in particular at the time scale of climate change. But the interaction of clouds with several atmospheric processes remains not well understood, such as with turbulence, radiation and large-scale circulation. Therefore their integration (parametrization) in climate and weather forecast models is generally associated with uncertainties (e.g. Cess *et al.* (1989)). Since many years, the tendency is to shift from *diagnostic* schemes to represent clouds in circulation models to *prognostic* equations. The latter allow the integration of basic physical phenomena , e.g., such as the hydrological cycle, and takes into account the various feedbacks involved in clouds. However, though this approach is physically more realistic than the diagnostic approach, cloud parametrization remains somewhat inaccurate because of the different uncertainties related to the representation of the involved processes: advective transport of cloud variables, sub-grid scale processes and dynamics, cloud micro-physics and cloud optical properties.

As mentioned above, clouds may occur at various levels of the atmosphere, leading to various types and stacked vertical cloud layers, known also as Cloud Vertical Structures (CVS). In this chapter and in the context of Aeolus space mission we focus on the CVS for the reasons mentioned in chapter 1 and detailed in chapter 2. In fact, we recall here that Aelous DWL measurements may be challenging in unclean atmospheres, i.e., in the presence of clouds and aerosols, since it may attenuate or moderate the Aeolus laser signal, depending on the thickness, height, type, etc., of the cloud, see for details chapter 2.

In particular, because the fact that current Aeolus processing algorithms assume that the cloud is uniformly distributed inside the bin, retrieving CVS and the knowledge of the cloud distribution in Aeolus bins will enable the quantification of the errors in retrieved winds; in case of strong wind gradients inside the cloud bin. Knowing that correlated wind and backscatter variability within an Aeolus bin may result in systematic errors in wind retrieval. Notice that CVS derivation is based on Wang *et al.* (2000) method, but with improved RH_M thresholds as established by Minnis *et al.* (2005), i.e. as function of the temperature and height.

Radiosondes provide wind data at high resolution. In addition, cloud information can be extracted at a similar high resolution and that would give a handle on the occurrence of heterogeneous (both dynamical and optical) atmospheric scenes and enables to quantify the impact on the quality of Aeolus winds in such scenes. In addition such data set can be used to improve the processing algorithms. The main CVS observations are twofold : firstly, those provided by surface weather observers and the derived statistics of the vertical distribution of cloud layers (e.g. Warren *et al.* (1986), Warren *et al.* (1988)) and, secondly, from weather satellites (e.g. Schiffer & Rossow (1983), Rossow & Schiffer (1991), Rossow & Schiffer (1999), Stephens *et al.* (2002)). However these two data sources are incomplete because they can in many cases only identify the lowermost (observer) or uppermost (satellite) cloud layer at each location and time, respectively. Another type of information comes from the analysis of vertical profiles of relative humidity from weather balloons (mainly radiosondes) that does provide a description of the vertical cloud layer structure at each location and time(Wang *et al.* (2000), Wang & Rossow (1995),Rossow *et al.* (2005)). Two limitations here are, firstly, that its geographical coverage is concentrated over land areas in the northern hemisphere, as already mentioned previously in chapter 1, And, secondly, that the balloon humidity measurements are not an accurate indicator of the presence of thin ice clouds, particularly at temperatures well below freezing.

Thus, the combination of high-resolution radiosonde, ECMWF model and CALIPSO satellite data will offer a more consistent and complementary view on the vertical distributions of clouds and the frequency of occurrence. Notice that CALIPSO is part of a constellation of satellites that includes*Aqua*, *Aura*, *CloudSat*, etc. (Stephens *et al.* (2002)). A generalization of the CVS comparison statistics to an extended database of radiosondes, will help to determine the geographical regions and the levels of the atmosphere where the frequency of cloud occurrence is very high and therefore where Aeolus DWL measurements are challenging. Within the limit of accuracy of CVS profiles determined from radiosondes and verified by other datasets, and combined with the wind and and wind shear profiles as studied in the previous chapters, radiosondes allow an assessment of the effect of the CVS on Aeolus DWL wind measurements. This may be done by investigating the coherence in small-scale cloud coverage and wind structures which may cause systematic error in the Aeolus DWL wind processing, in particular in terms of cloud and shear location.

To do so, high-resolution radiosonde profiles of the meteorological variables, denoted PTU (Pressure, Temperature and Humidity), describing the thermodynamic state of the atmosphere up to about 30 km altitude are characterized statistically, to complement the analysis of wind data in the previous chapters. A special focus is given to relative humidity (RH) which is important for the determination of cloud vertical structure (CVS). The ECMWF model and the CALIPSO satellite data are used in combination with radiosondes for the determination of the CVS as presented in section 6.2. In the same section, the following method is presented to analyze the profiles of the PTU meteorological variables and their gradients statistically, both from model and observations (radiosonde and CALIPSO). In sections 6.3 and 6.4, the results of these analysis are presented and discussed. In sections 6.5 a methodology to evaluate the importance of CVS on Aeolus DWL is proposed and discussed.

Such a method has not been fully implemented, because of the difficulty of estimating Cloud Water Liquid Content (CWLC) and Cloud Water Ice Content (CWIC) from radiosonde data. In the last section, the main results are summarized.

6.2 Data and Method

6.2.1 Data

To diagnose the vertical structure of clouds high-resolution radiosondes data are exploited as was done for the wind data in the previous chapters 3, 4 and 5. In this chapter the analysis is extended to other meteorological variables of pressure (P), temperature (T) and humidity (RH) (denoted generally PTU for radio-soundings) . For verification of the climatology of CVS from radiosondes, the results are compared with CVS derived from two other datasets: collocated SRF-ECMWF model data with radiosondes and ECMWF collocated with CALIPSO satellite data. Since close time and space collocations of high-resolution radiosondes and CALIPSO are few and far in between, we used the ECMWF model as a reference for CVS, since ECMWF SRF may be collocated to any observation at any place and any time.

For CALIPSO dataset, CVS are derived from the processed CALIPSO Level 2 Lidar Vertical Feature Mask (VFM) version release 3.1, by the Atmospheric Science Data Center(ASDC) at the NASA Langley Research Center. The detailed procedure of the algorithm procedure is described in Vaughan *et al.* (2009). This latest version of the algorithm (3.1) is improved with respect to earlier data releases, since a substantial overestimate of low cloud fraction is corrected (Vaughan *et al.*, 2010).

6.2.2 Analysis Method

The preliminary analysis consists in verifying the consistency of radiosonde observations of the meteorological variables: pressure (P), temperature (T) and humidity (RH), against the ECMWF model. This is done by establishing (and comparing) the statistics at each 1-km atmospheric bin, as in section 4 for wind and shear profiles, for the PTU variables for the two collocated data sets. In radiosondes data, RH is in particular given with respect to liquid water only. The RH formulation of ECMWF is however with respect to 3 phases: RH_W for liquid water above $0[^\circ C]$, RH_I for ice below -23 $[^\circ C]$ and RH_M for mixed phases (liquid and ice) between -23 and $0[^\circ C]$. For a fair comparison between radiosondes observations and ECMWF model, RH is recalculated according to the ECMWF formulation as summarized below (ECMWF-CY25r1, 2002).

So for a given radiosonde profile data of p, T, relative humidity is defined w.r.t to liquid water only (RH_W) at each atmospheric level Z_j. To compare to ECMWF model data, at certain levels it is rather RH_I or RH_M which is computed depending on temperature as mentioned above. Notice that each vertical atmospheric level referenced with a temperature T_j. and for the simplicity of the equations we used the index j only for temperature, i.e.,T_j. So following the reference above, the general formula for RH_{phase} reads:

$$RH_{phase}(T_j) = \frac{100 \times p.q.R_{w/d}}{[1 + (R_{w/d} - 1).q]es_{phase}(T_j)} \tag{6.1}$$

with *phase* referring to es with respect to water(es_W), with respect to ice (es_I) and mixed phases (es_M). The units for each variable used here are % for RH[%], Pascal for $p[Pa]$ and Kelvin for $T_j[K]$. Notice that the computation of saturation water vapor pressure $es_{W/I/M}$ of atmospheric air includes the individual gas constants for dry air ($Rd = 286.9[J/kgK]$) and water vapor $R_w = 461.5[J/kgK]$. For simplification of the formulas, the ratio of the constants is used ($R_{w/d} = R_w/R_d$). Note that (R_w/R_d) is generally computed.

For radiosonde data, the two unknown that have to be computed are the specific humidity q and the saturation vapor pressure es_{phase} w.r.t to the mentioned different phases. Based on the know radiosonde RH_W, the specific humidity q is then derived as follows (eq.6.2):

$$q \approx 0.622\frac{e(Tj)}{p} \quad with \quad e(Tj) = (RH_W.es_W(Tj))/100 \tag{6.2}$$

Notice that the specific humidity q is related to the pressure p and water vapor pressure e with the simplified approximation formula. while e is in turn derived from RH_W and saturation vapor pressure es_W. We recall here that since RH in radiosonde is defined w.r.t to water only (RH_W), saturation vapor pressure w.r.t to water only (es_W) is necessary to compute the specific humidity (q). Saturation vapor pressure over water es_W is defined as function of temperature Tj by the following formula: $es_W(Tj) = 6.12[exp((17.502 * Tj)./(240.97 + Tj))]$;

Once q is computed, saturation water vapor in case of the presence of the 3 phases es_{phase} is calculated from Tetens's formula (eq.6.3) for *liquid water* , and *ice* phases and from (eq. 6.4) for *mixed* phases, as shown below:

$$es_{W/I}(T_j) = a_1 exp[a_3(T_j - T3)/(T_j - a_4)] \tag{6.3}$$

The parameters are set according to Buck *et al.* (1981) and the AERKI formula of Alduchov & Eskridge (1996), i.e. $a_1 = 611.21hPa$, $a_3 = 17.502$ and $a_4 = 32.19[K]$ over water, and for mixed phases $a_3 = 22.587$ and $a_4 = -0.7[K]$ over ice, with $T_3 = 273.16$. With respect to 3 phases, i.e (liquid water or ice or mixed water-ice), the saturation value over water is taken above $0[°C]$ ($T_j > 273.16K$) and the value over ice is taken for temperatures below $-23[°C]$ ($T_j < 250.16K$). However, for temperatures between -23 and 0 $[°C]$ ($250.16K <= T_j <= 273.16K$ the saturation vapor pressure is computed as a combination of the values over water and ice according to the formula 6.3 with the respective values of its parameters with respect to water and ice phases, thus this formula for mixed phase saturation pressure es_M (eq.6.4):

$$es_M(T_j) = es_I + (es_W - es_I) \times [(T_j - T_i)/(T_3 - T_i)]^2 \tag{6.4}$$

with $T_3 - T_i = 23[K]$. Note that this function provides es_I for T_i and es_W for T_3, with a smooth transition to es_M at both ends.

Finally by considering the values of the different constants defined above for each phase as function of temperature, this allow the computation of the RH_{phase} as given in eq. 6.1. To summarize, RH_{phase} in eq. 6.1 with respect to 3 phases is computed by calculating in eq. 6.2 the specific humidity q from radiosonde (w.r.t to water only) and the saturation water pressure for the different phase (eq.6.3 and eq.6.4)

The results of the application of this method to radiosonde data are presented in figure 6.1. Based on the relative humidity with respect to two phases (water and ice), the CVS statistical analysis is performed based on Wang *et al.* (2000) method that includes detection of cloud vertical distributions and frequency of occurrences at each 1-km altitude bin size.

But contrary to this method where RH_M thresholds are considered constant with temperature over height, we used improved RH_M thresholds for radiosonde observations and ECMWF model,i.e. as function of the temperature and height as established by Minnis *et al.* (2005). However, statistics of CVS from CALIPSO data are established from VFM as mentioned previously in this section. The $Feature_Classification_Flag$ values of CALIPSO VFM data are stored as an 5515 element array for each 5 km "chunk of data" which is constituted of 3, 5 and 15 profiles depending on the altitude range. But for the statistical computations, the number of profiles per 5 km is standardized to 15 by re-sampling the altitude levels that contain only 3 and 5 profiles. Notice that each CALIPSO VFM file is constituted of several times the 5 km "chunk" data (i.e. 15 profiles after standardization), depending on the satellite coverage (see map of figure 6.5). Each array element is a sequence of bit-mapped integers, with one 16-bit integer being recorded for each range resolution element in the *Level 0* lidar data down-linked from the satellite. Decoding the bits in the individual integers yields information on feature type (e.g., cloud, aerosol, or clear air) and sub-type (e.g., water cloud or ice cloud) at each location . The interpretation of the first three bits at each atmospheric bin leads to a construction of 15 cloud profiles from each data element. More technical details about how the VFM are constructed, see ASDC-NASA (2011). Notice that in this study, we focused on cloud only; since this is used for comparison and verification of model and satellite data and it is the unique information that may be derived from radiosondes, based on the profiles of PTU information.

However in the context of the Aeolus mission the methodology proposed; to use the CVS derived from radiosonde to further investigate the coherence between wind dynamics and optical properties of the atmosphere at high resolution, is detailed in a separate section 6.5. This will allow the investigation of the effects of clouds on Aeolus DWL measurements later on.

6.3 Analysis of Observed and Modeled PTU

6.3.1 Humidity Profiles

6.3.1.1 Overview

As mentioned previously, figure 6.1 illustrates the difference between RH profiles of ECMWF model calculated with respect to two phases (Water and ice) with radiosonde

RH of ~$10m$) raw vertical resolution calculated with respect to liquid water only, then when recalculated with respect to 2 phases (liquid water and ice)) as for the ECMWF model. In fact, after the recalculation of radiosonde RH profiles with respect to the two phases about similar result to ECMWF RH profiles are obtained. This approach enables the co-existence of liquid and ice or ice/water alone as a function of temperature as discussed in section 6.2. In addition water vapor over-saturation of the air may occur, implying RH larger than 100 % at certain atmospheric levels as shown in the figure. This standardization of the RH profile calculation is necessary for the derivation of the CVS statistics, since in some conditions no clouds form at saturated pressures at high altitude where temperatures are low and the air remains completely transparent until ice forms, such as in cirrus.

6.3.1.2 Radiosonde RH Noise Estimation

Figure 6.2a gives an overview of RH profiles from radiosondes of 2 sec or ~$10m$ vertical resolution collocated with the ECWMF model for the 28^{th} day of each month from January to August for the year 2007, including their respective vertical gradients. Globally the results show that RH profiles of the ECWMF model follow quite well the trends of radiosonde RH profiles, but are much smoother. This leads to missing RH variability that can exceed 6 % when we smooth significantly (over more than 0.5 km in the vertical). Indeed, figure 6.2b shows the analysis of the effect of a progressive smoothing of radiosonde RH profiles and shows the variability removed by the smoothing. This is done, as in chapter 4 for wind, by degrading successively the vertical resolution of the radiosonde profile from 10 (raw) to 20, 30, 40, 50, 60, 70, 110, 330 and 670 meters. So, we apply a successively increased vertical box size where the standard deviation (SD) of RH and its vertical gradient at each atmospheric box are computed. The RMS of the SDs of RH and grad(RH) are then computed as function of vertical box size, as show in figure 6.2b. The statistics show in particular that above 30 m vertical resolution the standard deviation is clearly increasing. This suggest that above this value the natural variance of RH dominates and white noise appears negligible, since the standard deviation of the white noise should remain constant with box size.

A noticeable difference between the radiosonde and ECMWF profiles is the occurrence of sharp vertical structures. This is most clearly visible in the RH gradients, where radiosonde measurements show spikes of 10-20 percent at various vertical levels, but the ECMWF RH gradients appear relatively smooth. This may be expected from the vertical resolution supported by the discretization of vertical levels in the ECMWF model (see chapter 4). As a consequence, in the box averaging, the variance taken away in the smoothing is almost an order of magnitude higher in the radiosonde profiles then in the ECMWF profiles.

6.3.1.3 RH and Gradients(RH) Statistics

Figure 6.3 shows the statistical comparison of the relative humidity at different levels of the atmosphere (at each 1-km vertical bin) and their gradients, both of radiosonde observations and the collocated ECMWF model profiles. Recall here that the radiosonde vertical resolution of each profile is smoothed with respect to the raw data

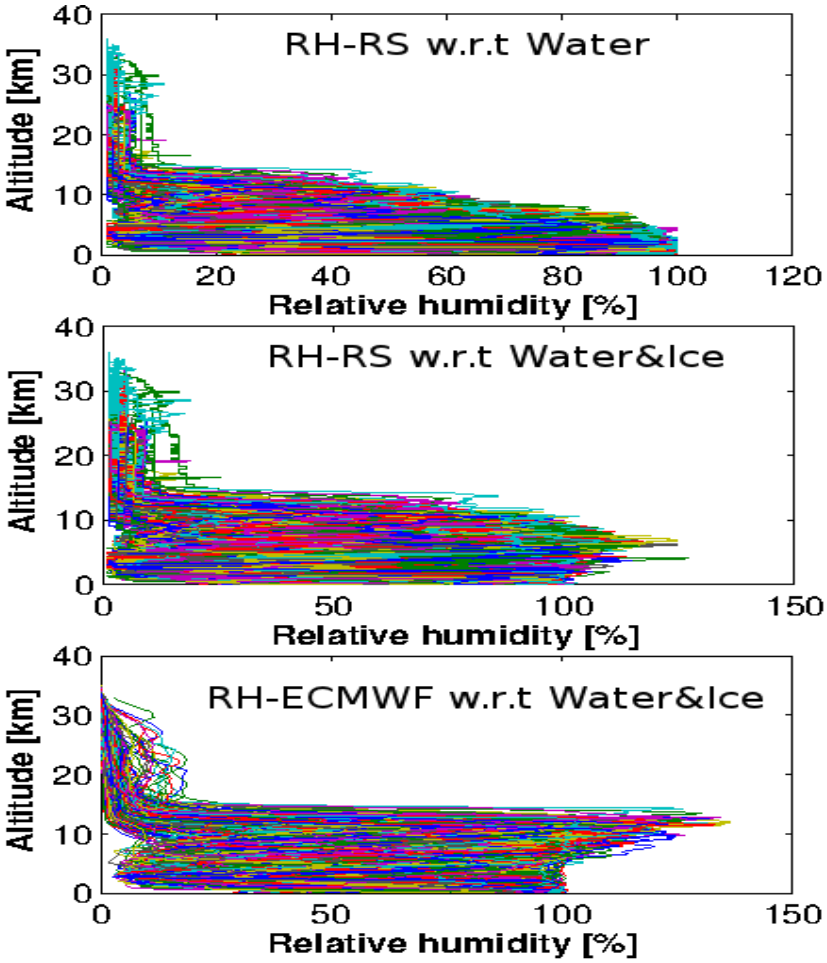


Figure 6.1: Relative humidity of raw resolution (2 sec ~$10m$) vertical profiles for station De Bilt over the year 2007 with successively from bottom to top: ECWMF RH, radiosonde RH recalculated with respect to 2 phases (liquid water and ice) and RH as provided in the radiosonde reports.

of 2 sec (~$10m$) to 6 sec (~$30m$) for the reasons evoked in the previous section in figure 6.2), i.e., to reduce the random noise. For a fair comparison between the modeled and observed RH, ECMWF RH profiles are interpolated to the same vertical resolution as the radiosonde. The statistics of RH and gradient (RH) profiles are given by mean and percentile (0,5,25,50,75,95 and 100%) profiles. The results shows that ECMWF RH statistics are globally similar on average to radiosonde RH in roughly the lower 5 km. Above this altitude, RH is overestimated in the model or may be underestimated by radiosondes.

While the extreme RH gradients recorded for radiosonde profiles can reach more

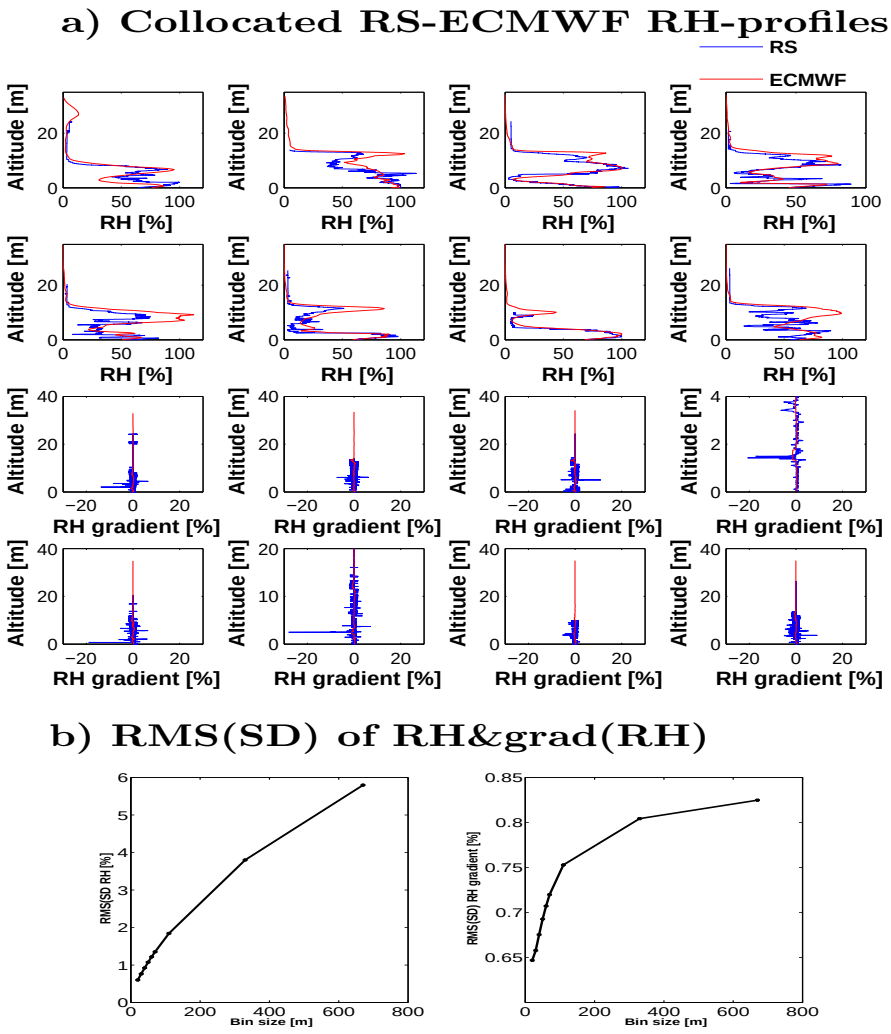


Figure 6.2: a) Relative Humidity profiles from radiosonde (2 sec or ~$10m$) and collocated ECWMF model for the 28^{th} day of each month from January to August in the year 2007, including their respective vertical gradients; b) Square root of the mean variance of box samples of RH (left) and RH gradients (right) in successively increased vertical box sizes, i.e., degraded resolutions, from 10 to 20, 30, 40, 50, 60, 70, 110, 330 and 670 meters(circles), averaged over the full profile.

than 40 %, as can be seen in figure 6.3, the most extreme gradient percentiles of ECMWF are clearly higher around the tropopause, perhaps indicating some RH sensor limitations in cold air; since radiosonde RH measurements are known to be unreliable at cold temperatures (Miloshevich *et al.* (2010), Wang *et al.* (2003)).

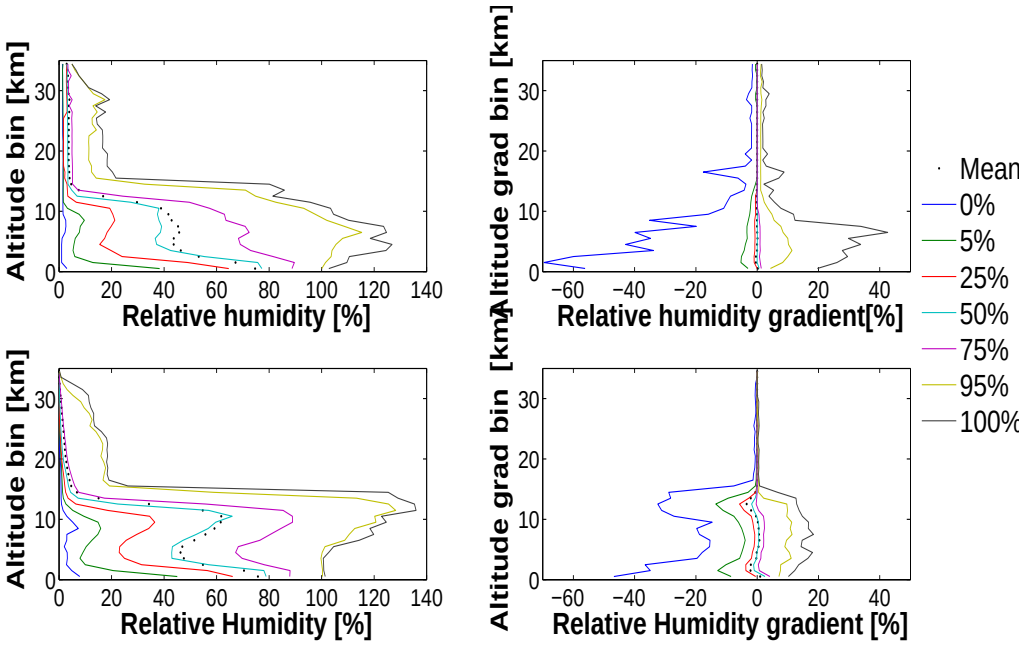


Figure 6.3: Statistics (mean and percentile profiles) of RH and its vertical gradients for the year 2007, as seen from radiosondes (top) and ECMWF model (bottom); radiosonde launched from De Bilt.

6.3.2 Pressure and Temperature and their Gradients Statistics

Pressure and temperature profiles are involved in the RH-profiles computation, so it is important to verify how well the two variables compare between model and observations. The gradients of p and T are also computed since they provide information on the RH variability in the vertical. Figure 6.4 shows similar plots as figure 6.3 of the statistical comparison of radiosonde observations and the ECMWF model for successively pressure and temperature and their gradients. The results show that the ECMWF model reproduces similar statistics of pressure and temperature profiles as radiosonde observations, but not for their gradients. Though the statistics of radiosonde pressure gradients show a similar mean profile to the model, i.e., very small bias as revealed by mean ratios plot (not shown), a little more vertical variability is seen in radiosondes shown by the large spread of the percentiles in the figure. The quantification by percentile ratios between radiosonde and ECMWF model reveals a pressure variance higher by about a factor of 1.5 in radiosonde as compared to the model.

However for the statistics of temperature and their gradients, the results show a close resemblance between radiosonde and ECMWF profiles. This is revealed by the ratio profiles of radiosonde and ECMWF, which show values oscillating around a factor value of 1. describing the similar thermodynamic state of the atmosphere.

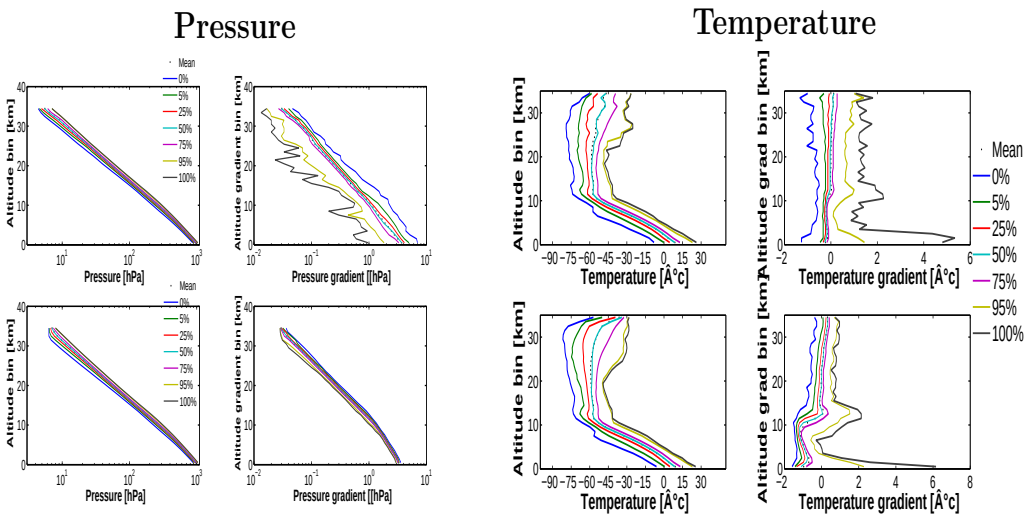


Figure 6.4: Statistics (mean and percentile profiles) of the atmospheric pressure and temperature and their vertical gradients in the year 2007, as seen from radiosondes (top) and ECMWF model (bottom); radiosonde launched from De Bilt.

6.4 Cloud Vertical Structure (CVS)

In this section we present the CVS derived from CALIPSO VFM as function of different radius coverage around De Bilt station, that comprise different representations of the De Bilt CVS climatology and a different number of data files involved in the statistics. We present a comparison with the results of CVS from collocated radiosonde and ECMWF model, by choosing the CVS statistics as function of radius coverage that contains the maximum number of files of CALIPSO data, e.g., for a radius of daily coverage of 1000 km only 284 files are involved in the statistics in 2007.

6.4.1 CVS w.r.t Spatial Coverage

Figure 6.6 shows different CVS profile statistics as function of increasing the radius coverage around De Bilt station (lat=52.10, lon=5.2), from 100 to 1000 km. As can be seen in the legend of the plot, the number of files of CALIPSO data collected in 2007 are reported. The results show mainly that for sufficient number of data files, the vertical distributions and the frequency of occurrences of clouds (over one year data here) are quite similar. In other words, for a radius between 500 and 1000 km, the climate of CVS is quite similar. We recall here that each CALIPSO file is constituted of several 5-km data "chunks" of 15 profiles (see section 6.2.2), depending on the satellite coverage (figure 6.5). For the comparison with collocated radiosonde and ECMWF model profiles, we considered the set of CVS profiles that contains the maximum number of data files, i.e., with a radius coverage of 1000 km around De Bilt as shown in figure 6.5 for the day and night coverage. For all data in 2007 and

a radius coverage of 1000 km, 284 and 280 files are counted in the CVS statistics, successively, for the day and night coverages.

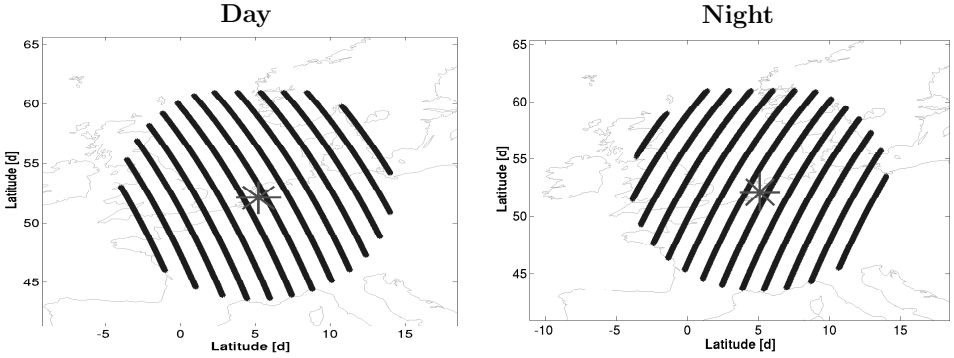


Figure 6.5: Daily (left) and night (right) CALIPSO satellite coverage within 1000 km radius around De Bilt station (lat=52.10, lon=5.2), indicated by a star. For this radius, 283 files with CVS statistics are available from the VFM of CALIPSO data. Each profile in the plot represents the CVS statistics over different area sizes around De Bilt, denoted by the maximum distance from this location.

.

6.4.2 CVS from 3 Data Sources

Figure 6.7 shows the vertical distributions and the frequency of occurrences of cloud for the day and night coverages, as obtained from 3 data sets, successively from radiosonde, collocated ECMWF model and CALIPSO datasets. The time coverage for each dataset is as follows: 12 UTC (day) and 00 UTC (night) for the collocated radiosonde and ECMWF model, and which correspond to about 1 AM and 1 PM local times at De Bilt station. For CALIPSO the time coverage is from about 11 AM up to 1 PM during the day and from 1 AM up to 3 PM during the night.

For the day data coverage for instance, the results show a typical common trend of the vertical distributions of clouds in the three data sets, though the cloud frequency of occurrence slightly differs from data-set to another, depending on the altitude. The most notable difference is above 10 km where radiosondes record less cloud than ECMWF and CALIPSO. This indicates probably the limitation of temperature or humidity sensors of radiosondes at these levels with cold temperatures and sun irradiation, because of the fact that satellite clouds such as from CALIPSO are quite accurate for high level clouds and since high-level night-time radiosonde cloud cover is much higher and comparable to CALIPSO. We notice also that the ECMWF model overestimates the frequency of clouds by 5-7 % with respect to CALIPSO data. Contrary to high level clouds, radiosondes shows a slightly higher frequency of occurrence for middle- and low clouds than ECMWF model, but much higher than CALIPSO, with values of about 7-8 %. As for the day coverage, similar comments may be made

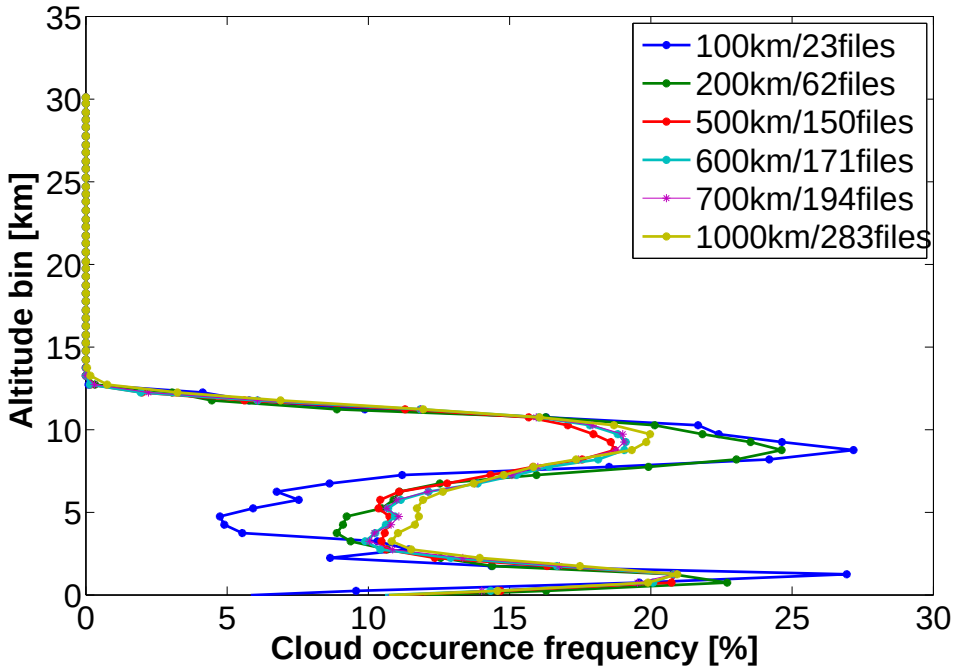


Figure 6.6: Statistics of CVS around De Bilt (lat=52.10, lon=5.2) as translated from the Vertical Features Mask of CALIPSO data. Each profile in the plot represents the CVS statistics over different area sizes around De Bilt location, denoted by the maximum distance from this location. Notice that the number of files collected for each area size is shown at the right-side of the radius in the legend.

for the night statistics. But we notice in addition a much higher frequency of cloud occurrence in the night than during the day. This is mainly remarkable for high-level clouds in the radiosonde data as mentioned above.

At around 5 km height radiosonde and CALIPSO statistics overlap well, with ECMWF providing less clouds. Near the surface, radiosonde clouds are most frequent, since CALIPSO is not always able to detect PBL clouds in the presence of overcast and optically thick clouds. Therefore, ECMWF appears to slightly underestimate PBL clouds as well. From the overall results we can conclude that radiosonde CVS profiles below 10 Km are quite consistent with satellite data and model results, and therefore may be used to further investigate their effects on wind and shear profiles in the context of the Aeolus mission. Notice that a radiosonde is a unique instrument in that it can provide such combined information on optical and dynamical structure of the atmosphere.

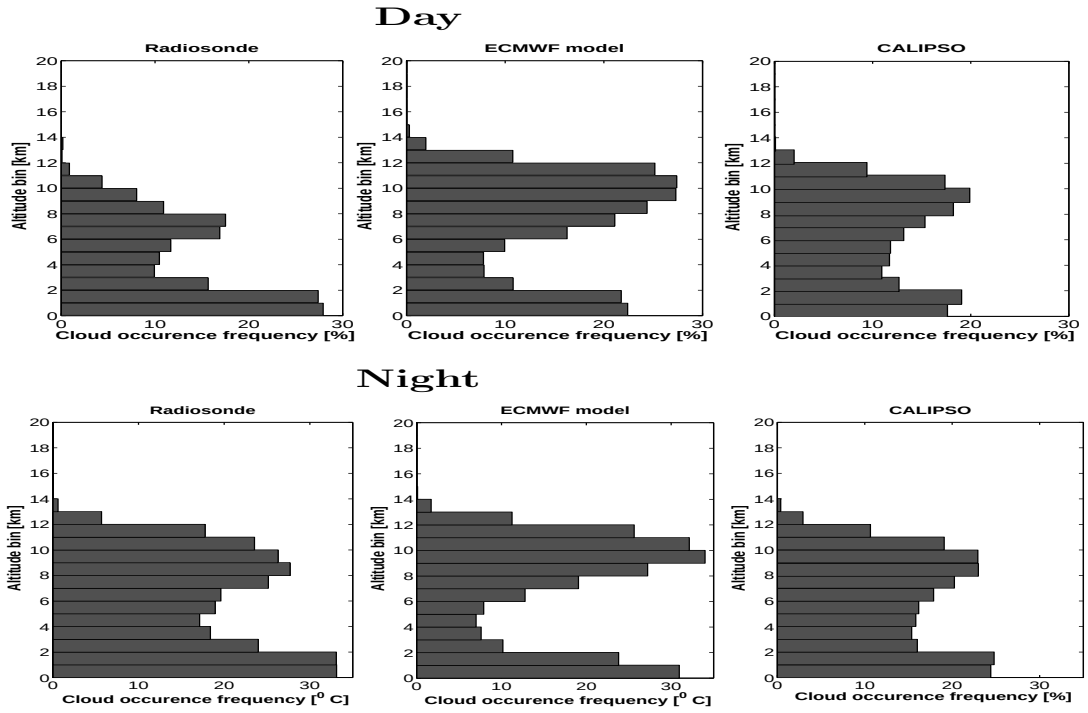


Figure 6.7: Daily (top) and night (bottom) CVS statistics around De Bilt (lat=52.10, lon= 5.2) from 3 data sets: successively from Radiosonde, ECMWF model and the VFM of CALIPSO data. For CALIPSO data a radius of 1000 km around De Bilt is considered to establish this 2007 climatology.

6.5 CVS and Aeolus DWL

In this section we present a methodology to investigate the effects of cloud on wind measurements based on wind and CVS profiles derived from high-resolution radiosonde observations. As mentioned previously in chapter 2, using light-scattering particles in the air (cloud/aerosols and molecules) as tracer, LOS wind speed is determined as a function of height ranges (bins) by integrating over vertical bins of varying depth, see equation eq. 2.10 of chapter 2, and now in slightly different form:

$$V_{m,i} = \frac{\int_{z_i}^{z_{i+1}} S_x(z) V_t(z) dz}{\int_{z_i}^{z_{i+1}} S_x(z) dz}; V_{t,i} = \int_{z_i}^{z_{i+1}} V_t(z) dz \tag{6.5}$$

with $V_{m,i}$ the mean measured wind inside the bin, V_t is the "true" wind, and S_x is total energy of the return signal (in number of photons) as a function of height z; where the subscript x denotes the type of scattering particle, i.e. a for aerosol and

cloud and m for molecular and where $V_{t,i}$ is the mean wind speed in the bin from radiosonde wind profiles.

For a given radiosonde wind profile, wind errors in Aeolus DWL measurement that may be caused at each level bin by the presence of clouds may be written as follows:

$$\epsilon_i = V_{m,i} - V_{t,i} \tag{6.6}$$

which can be computed straightforward. However, $V_{m,i}$, the mean Aeolus-DWL wind speed in bin i has to be simulated using the formula above (eq. 6.5).

Based on this equation, we may estimate what would have been the total energy of the return signal for a given CVS and wind obtained from radiosonde profiles. The peak integrals S_a and S_m relate to the total energy of the return signal (in number of photons) from aerosol/cloud and molecular scattering respectively from an atmospheric layer at a given altitude. They are quantified by Winker *et al.* (1996) as follows:

$$S_x(z) = \frac{E\lambda}{hc}\frac{\pi D^2}{4}\frac{\beta_x(z)}{R(z)^2}\tau(z)^2\Delta z_{LOS}T_{RX}T_{TX} \tag{6.7}$$

where the number of scattering particles in the illuminated volume is denoted by $\beta_x(\lambda, z)$ and the one-way atmospheric transmission by $\tau(\lambda, z)$ (see eq. 6.8). Instrumental properties include laser energy E, the telescope diameter D, the range from the atmospheric layer to the instrument $R(z)$ and the depth of the illuminated layer, $\Delta z_{LOS} = \Delta z/cos(\phi(z))$, with Δz the range gate resolution and $\phi(z)$ the local line-of-sight LOS or laser beam incidence angle. System optics are described with T_{TX} and T_{RX} that denote transmit and receive optics transmission respectively, c is the speed of light and h is Plancks constant. For typical values of the different parameters, see Table 2 in Marseille & Stoffelen (2003). τ is the one-way atmospheric transmission of laser light, between the satellite and the observed atmospheric layer at altitude z. It is derived from aerosol, molecular and cloud extinction. By focusing only on cloud levels as given by the CVS profiles from radiosonde, we can write the simplified equation for τ as follows:

$$\tau(z) = exp[\frac{-1}{cos(\phi(z))}\sum_{i}^{i+1}\{\alpha_m(z) + \alpha_{cld}(z)\}] \tag{6.8}$$

with α_{cld} is the cloud extinction as modeled in (Stephens, 1984).

$$\alpha_{cld}(z) = \frac{3X10^{-6}}{2} \times \frac{CLW(z)\rho(z)}{\rho_1(z)r_e} \quad with \quad r_e = (-3.8X10^{-2}p + 43.8)10^{-6} \tag{6.9}$$

where ρ is the air density, ρ_1 is the density of liquid-water, and p is the air pressure in mb.

This equation states that cloud extinction is inversely proportional to the mean effective radius, r_e, of the size distribution of the cloud droplets. The parametrization on the right-hand side of equation eq. 6.9 may be applied to the ECMWF model or radiosonde profiles. This states that the mean droplet size in low-level clouds is smaller than in high level clouds, making low-level clouds optically denser for a given water content.

Cloud backscatter is modeled by a constant backscatter-to-extinction ratio of $b_{cld} = 0.0552sr^{-1}$ (Vaughan *et al.*, 1998), hence

$$\beta_{cld}(z) = \alpha_{cld}(z) b_{cld} \tag{6.10}$$

while molecular backscatter is modeled by a constant backscatter-to-extinction coefficient $b_m = 3/8\pi 2sr^{-1}$, hence

$$\beta_m(z) = \alpha_m(z) b_m \tag{6.11}$$

Clouds play a crucial role in the performance of a space-borne DWL. Opaque clouds for instance generally hamper the laser beam to reach the underlying atmosphere. But transparent clouds such as thin cirrus may yield a strong return signal and yet transmit a significant part of the laser beam. The cloud liquid-water content (CLW) and the size distribution of cloud droplets quantify the cloud penetrating capability of the laser beam. So the challenge for this methodology is to derive the CWLC and CWIC from radiosonde observations and then combine them with the shear profile. The water amount in the atmosphere, be it liquid or solid, may be estimated depending on p and T, where climatological aspects of the derived backscatter may be compared to the CALIPSO backscatter data (still under investigation) for consistency. This will allow, as mentioned previously, the investigation of the effect of CVS combined with large wind vertical gradients on the return signal of DWL laser and therefore the wind measurement errors.

6.6 Conclusion

A radiosonde is a unique instrument that provides combined information on the mass and wind field, which may be used to derive optical and dynamical properties of the atmosphere at high vertical resolution. In this chapter, profiles of the thermodynamic variables of pressure, temperature and humidity from collocated radiosondes and the ECMWF model are first compared and analyzed, and subsequently used to derive the distribution of cloud vertical structure, or CVS. The CVS statistics are derived from both data sets and compared with CVS derived from the vertical feature mask of the CALIPSO satellite data set. Globally, the statistical profiles of pressure, humidity and temperature show on average about similar results between radiosonde observations and the ECMWF model. But apart from temperature, the statistical profiles of their gradients are different. Indeed, the relative humidity and pressure gradients show more variability in the radiosondes as compared to the ECMWF model, generally

with a factor below 1.5 on average. The value of this factor may be larger for the extremes of the percentile profiles at certain atmospheric levels, such as near the surface where RH and pressure are not always well represented in the model.

The preliminary statistical results of the assessment and comparison of the vertical cloud structures show similarities between the various data sets: collocated high-resolution radiosondes, ECMWF model and CALIPSO data sets at De Bilt station for the daily and night coverages, in terms of vertical distributions and frequency of occurrence of clouds. The similarity in the results from these three data sets is even more consistent for the general trend of the CVS, though the frequency of occurrences may be some times higher or lower at some altitude levels, generally because of known limitations of each measurement instrument and the ECMWF model. So, from the overall results we can conclude that radiosonde CVS profiles below 10 km at daytime are quite consistent with satellite data and model results, and therefore may be used to further investigate cloud effects on wind and shear profiles in the context of the Aeolus mission. Night-time CVS profiles are also usable above 10 km.

In the context of the Aeolus space mission, generalizing the statistical results of CVS to all climate regions (with SPARC data, etc.) will first help to geographically determine the levels of the atmosphere where the frequency of simultaneous cloud and wind shear occurrence is high, challenging the retrieval of good-quality Aeolus DWL observations. The goal is to obtain optically realistic atmospheric scenes with different vertical cloud structure configurations and associated wind dynamics, to simulate such challenging scenes for the Aeolus DWL. The database is being used to optimize the vertical sampling of Aeolus and its wind retrieval performance. Furthermore, a methodology is proposed and described to use combined CVS and wind profiles from radiosondes to allow the investigation of the effect of clouds on Aeolus DWL measurement and estimate the errors that may be caused by the coherence between wind dynamics (wind shear) and optical properties (CVS). The correlation between wind shear and the optical properties of the atmosphere will be used to examine and improve the interpretation of ADM-Aeolus observations. However, we note that a challenge remains in the derivation of the Cloud Water Liquid Content (CWLC) and Cloud Water Ice Content (CWIC), for example as function of temperature and pressure.

The recent update on the subject and the use of this high-resolution database, specifically for De Bilt region (Netherlands) Sun *et al.* (2014), confirmed some of our results presented in this chapter.6, e.g., regarding CVS features and statistics. To insure a good collocation of wind information and optical properties of the atmosphere, Sun *et al.* (2014) also adopted Zhang *et al.* (2010) methodology. The latest consists in detecting clouds along the radiosonde path from measured temperature and Humidity (RH), while the parametrization of cloud and aerosol backscatter and extinction is established and discussed according to the measured RH along the radiosonde path. Additional data from UV lidar (at Cabauw observation site, NL) have been used to validate the parametrization of aerosol backscatter and extinction. it is further used to simulate errors of Aeolus DWL Winds retrieved from

Mie and Rayleigh channel signals in heterogeneous atmospheric conditions. It is mainly found that Rayleigh channel wind errors through an heterogeneous atmosphere is substantially smaller than Mie channel wind errors. it is also showed that wind errors standard deviation grows linearly with increased bin size. More detailed information may found in the above mentioned paper. In the end, all this aims to optimize the vertical sampling of Aeolus and its wind retrieval performance.

CHAPTER 7

Conclusion

In this thesis available worldwide meteorological observations at high resolution mainly from radiosonde instruments are gathered and analyzed. This is done in the framework of ESA'S Aeolus space mission, to study wind and shear and CVS profiles. This unique collection of radiosonde data-sets at high resolution is completed mainly by collocation with the ECMWF model profiles and compared to CALIPSO satellite data. Besides scientific contributions related to weather and climate modeling as summarized below, the present work contributes mainly in the context of the ESA Aeolus mission, e.g., through inputs for the generation of a high resolution atmospheric data base. This data base provides a more realistic (i.e., including small scales structures) representation of the atmosphere than other data sources, which is necessary to conduct detailed simulations of the Aeolus Doppler Wind Lidar (DWL). This will allow to understand and improve the wind measurements performance of the DWL. The data base may also be used for other research studies related to weather and climate. To reach these objectives, the following research questions are addressed:

1. What are the best and most useful observations to investigate the atmospheric dynamics and optics at high resolution, which are relevant for the preparation of ESA's Aeolus space mission ?

2. Which data sets can be used to build a global atmospheric data base at high resolution and that is needed to simulate Aeolus DWL realistically; GCMs, observations or both ?

3. What is the accuracy (bias) and precision (error variance) of the various selected data sets? Is quality control needed ?

4. What is the difference in the representation of the atmosphere dynamics in terms of spatial and temporal variabilities between the observations (i.c., high resolution radiosondes) and a Global Circulation Model (i.c., ECMWF model) ? Similarly, what are the differences in the observed and modeled CVS of radiosondes and GCM with respect to CALIPSO satellite data ?

5. Can we determine reliable CVS profiles from radiosonde data which may be used to investigate CVS effects on the estimation of wind measured with Aeolus DWL at the different vertical height ranges ?

The approach and response to these research questions are provided in this thesis and summarized below. Prior to these investigations, the necessary theoretical knowledge on wind dynamics, its properties, its representation and its importance in the GOS and GCMs and many other domains, and available or potential measurement instruments, is introduced (chapter 1). The ESA Aeolus space mission that guided most of the investigations and initiated this study is presented in chapter 2, with a special focus on the vertical sampling of the DWL that inspired most of this thesis work. This issue is very important because significant errors in wind measurements may occur, due mainly to the limited number (24) of the atmospheric vertical range gates that can be sampled by Aeolus-DWL, including the complex and coherent optical and dynamical properties of the atmosphere targeted by the DWL.

The preparation of a global radiosonde data set at high resolution requires confidence both in accuracy and precision in the different meteorological variables to be analyzed: pressure, temperature and humidity (abbreviated as PTU), as well as wind profiles. Therefore the data set needs to be consistent between the different identified radiosonde types. One of the special focus among the addressed research questions above is to ensure the quality of the data, in particular for wind data which is derived from radiosondes of different Wind-finding Systems (WFS): radiotheodolite, LORAN-C and GPS. Thus, an effective statistical QC method is developed to screen the high-resolution radiosonde data for unrepresentative observations (outliers), to allow the comparison of the various radiosonde WFS types and their accuracies. This is in particular very important for an accurate derivation of vertical profiles of wind and wind shear and subsequent statistics. The main results showed that though the radio-theodolite-based data are most abundant, it is less accurate than data based on the other wind-finding systems. But after QC (chapter 3), radio-theodolite data provide a good statistical representation of horizontal wind and vertical wind shear (i.e., vertical gradients of the horizontal wind) climatologies. Moreover, this allowed a preliminary statistical estimation of the vertical wind by considering the mean ascent of the radiosonde balloon with height.

After ensuring the consistency of the various radiosonde WFS types by applying QC and necessary corrections, the ensemble of data sets are collocated and compared against climatologies of the short-range forecasts of the ECMWF model. This data collocation of radiosondes with the ECMWF model allows at first stage to highlight the differences and resemblances between the observed and modeled representation of wind and shear profiles, while differences are particularly large for wind shear. This is determinant in the generation of a detailed description of the atmospheric dynamics, in particular in terms of vertical wind shear climatologies (chapter 4), for the various defined climate regions (tropics, subtropics, mid latitudes and the poles). The main results showed similar wind distributions at different vertical atmospheric levels, while we notice important discrepancies in wind shear (mean and variability) which are crucial for the DWL measurements as explained in chapter 2. It is found that observed

wind shear is much higher than modeled wind shear by a factor of 2.5 (for zonal) and 3 (for meridional) wind profiles. The systematic vertical sampling effects on wind and wind shear investigated both in single radiosonde profiles and statistically, showed consistent results. Consequently, it is demonstrated that the effective ECMWF vertical model resolution is about 1.7 km, failing thus to well resolve the atmospheric phenomena which may occur below this scale, such as turbulence, convection, gravity waves, etc. High resolution radiosonde observations that represent more realistic (i.e., including small scale structures) atmospheric dynamics are however limited in terms of time and space coverage. Therefore, the statistical results are used to develop a method to complement the relatively smooth but global ECMWF model profiles with the missing shear variability as observed by the radiosondes (chapter 2).

The ensemble of radiosonde data sets is further exploited to extend the analysis of the spatial variability of wind and wind shear profiles to variations over time (chapter 5), i.e., determine the monthly, seasonal and yearly variabilities over different climate zones as seen from both radiosonde observations and the ECMWF model. This includes an assessment of the differences and a quantification of wind and shear variabilities between the two data sets for the various periods, between land and sea and between dawn and dusk. From the various detailed results of this extended analysis of wind and shear variability, it is mainly concluded that the differences in the temporal variability between the observed (radiosonde) and modeled (ECMWF) wind and shear are less significant than the differences in the spatial variability. Consequently, this may be negligible when we generate the high-resolution data base and one may only consider the spatial variability differences; i.e., apply similar corrections in all climate regions. As mentioned above, a method to adapt the variability of modeled wind shear to the level of the observed shear is presented. This will allow the creation of a global atmospheric dynamics data base at high resolution; i.e., including small scale structures. Thus, a global (from ECMWF model) and realistic (from radiosondes observations) wind and shear data base has been created for Aeolus DWL simulations. This is complemented by adding the atmospheric optical properties from CALIPSO data collocated with the ECMWF model dynamical variables in the context of ESA's VAMP project (Marseille *et al.*, 2010).

We also exploited the other radiosonde meteorological variables of pressure, temperature and humidity (PTU) to study the Cloud Vertical Structures (CVS) (chapter 6). We recall that CVS is important in the context of the Aeolus mission, since the DWL laser cannot penetrate thick cloud layers and clouds may be associated with small-scale atmospheric motion, such that uncertainty in wind measurements may occur at these cloud levels. First, the collocation and comparison of the radiosonde PTU data with ECMWF showed an overall similar representation of the thermodynamics of the atmosphere, though some small differences are seen for the vertical gradients in pressure and humidity profiles. The differences are on average well below a factor 1.5 for the two variables, and limited to the lowest 5 km of the atmospheric. The preliminary assessment of the CVS as a function of altitude for De Bilt station showed general consistency between the three data sets (radiosonde, ECMWF model and CALIPSO), in terms vertical distributions and occurrence frequencies of clouds.

We also found a quite similar trend in the vertical cloud distributions for the day and night coverages. However, the frequency of cloud occurrence is quite different at some altitude levels, generally because of known limitations of each measurement instrument and the ECMWF model (e.g., poor detection of low cloud from satellite, etc.), limitations in daytime radiosonde temperature and humidity measurements, and a lack of mid-troposphere model clouds. We conclude from the overall results that radiosonde CVS profiles are quite consistent with satellite data and model results at night time and at day time only below 10 km.

Therefore, it is recommended to further investigate CVS effects on wind and shear profiles in the context of the Aeolus mission, since a radiosonde is the unique instrument that can provide combined information on the optical structure and dynamics of the atmosphere at high resolution. For this purpose, a methodology is proposed to estimate errors in Aeolus DWL measurements in the presence of combined strong wind shear and opaque clouds. However, this methodology is subject to one crucial step which is the derivation of the Cloud Water Liquid Content (CWLC) and Cloud Water Ice Content (CWIC) profiles, based on CVS radiosonde profiles and may be linked to CALIPSO backscatter measurement climatologies. Moreover, generalizing the statistical results of CVS to all climate regions (using the radiosonde data collected during this study) would be interesting to determine where frequent and simultaneous very high optical and dynamical variability occur at different levels in the atmosphere, and that challenge Aeolus DWL measurements. Optically realistic atmospheric scenes with different CVS configurations and associated wind dynamics may be identified and simulated to optimize Aeolus DWL vertical sampling and wind retrieval performance. Recent update on the subject by Sun *et al.* (2014), confirmed some of our results presented in this chapter.6, by using this high-resolution database specifically for De Bilt region (Netherlands). But it is further used to simulate errors of Aeolus DWL Winds retrieved from Mie and Rayleigh channel signals in heterogeneous atmospheric conditions. It is mainly found that Rayleigh channel wind errors through an heterogeneous atmosphere is substantially smaller than Mie channel wind errors. it is also showed that wind errors standard deviation grows linearly with increased bin size. More detailed information may found in the above mentioned paper. In the end, all this aims to optimize the vertical sampling of Aeolus and its wind retrieval performance.

Outlook

The research questions addressed in this thesis lead to new issues that pave the way to further future investigations. While some questions that require more exhaustive studies as suggested below, some studies are limited by the information given by the data and/or lack data coverage. For instance, it will be particularly useful to extend the wind and shear climatologies, by integrating the high-resolution radiosonde data obtained from France and Australia to the database (after this study). More data from other sources needs to be collected from weather centers and observations sites, in particular in the southern hemisphere countries, when it is possible.

The vertical wind in the thesis is estimated only statistically by considering a mean ascent of the radiosonde balloon. An attempt to accurately estimate the vertical wind on single profiles has not been performed with the current data, because of several unknown parameters (e.g., charge load, initial diameter of the balloon at each ascent, etc.) for each radiosonde launch. Indeed, these parameters are essential to estimate the diameter and the drag coefficient of the radiosonde balloon, which are varying during the ascent with height (i.e., with density, turbulence conditions and pressure) and needed to derive the vertical wind. So, by performing systematic radiosonde launches (currently performed at KNMI) and taking care of the above mentioned variables, it may be possible to obtain an improved estimation of the diameter and the drag coefficient, thus a better estimation of the vertical wind, e.g. In case of night flights soundings (Gallice *et al.*, 2011).

It would also be interesting to extend the CVS statistics, obtained from radiosonde data at De Bilt and here validated against ECMWF model and CALIPSO data sets, to radiosonde data over the various climate regions. Generalizing the statistical results of CVS to all climate regions (SPARC data, etc.) would be interesting to determine the geographical regions and the levels of the atmosphere where the frequency of cloud occurrence and associated dynamics is very high, and that challenge Aeolus DWL measurements. These data would be important for the Aeolus DWL simulations and ground processor development. Some recent updates on the simulation of Aeolus performance in heterogeneous atmospheric conditions may found in Sun *et al.* (2014). However, the study of the effects of CVS on Aeolus DWL measurements is hampered by the difficulty of deriving the CWLC and CWIC form radiosonde data and which needs to be solved. It is suggested to estimate the CWLC and CWIC as function of atmospheric conditions, temperature, pressure and relative humidity profiles, as approximated in NWP model diagnostic schemes for clouds parametrization. These days prognostic cloud schemes are used in NWP, that contrary to diagnostic schemes propagate cloud fields in time, such that differences between diagnostic and prognostic cloud schemes would need to be analyzed, e.g., using the ECMWF model. This will help to understand the limitations in radiosonde CVS based on diagnostic schemes. The simulation of Aeolus DWL cloud backscatter based on radiosonde CVS may need an extensive comparison with CALIPSO backscatter measurements to ensure locally consistent PDFs. This extends the analyses of the collocated data as presented in chapter 6.

BIBLIOGRAPHY

Ahrens, C. (2007). *Meteorology today: an introduction to weather, climate, and the environment*. Brooks/Cole Pub Co.

Ahrens, C. (2011). *Essentials of meteorology: an invitation to the atmosphere*. Brooks/Cole Pub Co.

Alduchov, O. & Eskridge, R. (1996). Improved Magnus Form Approximation of Saturation Vapor Pressure. *Journal of Applied Meteorology*, **35**, 601–609.

Allen, S. & Vincent, R. (1995). Gravity wave activity in the lower atmosphere: Seasonal and latitudinal variations. *Journal of Geophysical Research*, **100**, 1327–1350.

Andersson, E., Haseler, J., Undén, P., Courtier, P., Kelly, G., Vasiljevic, D., Brankovic, C., Gaffard, C., Hollingsworth, A., Jakob, C. *et al.* (1998). The ECMWF implementation of three-dimensional variational assimilation (3D-Var). III: Experimental results. *Quarterly Journal of the Royal Meteorological Society*, **124**, 1831–1860.

ASDC-NASA (2011). *CALIPSO Data Products Catalog (Version 3.2)*. Atmospheric Science Data Center(ASDC)-NASA Langley Research Center.

ASTRIUM (2011a). End-to-end simulator detailed processing model. [521801_E2SDPM_3_00.doc, internal astrium document (not public)], 31-Jul-2011.

ASTRIUM (2011b). Level 1b processor detailed processing model. 521800_DPM_v3_00.doc, internal astrium document (not public), 28-Oct-2011.

Baker, W., Emmitt, G., Robertson, F., Atlas, R., Molinari, J., Bowdle, D., Paegle, J., Hardesty, R., Menzies, R., Krishnamurti, T. *et al.* (1995). Lidar-Measured winds from space: A Key Component for Weather and Climate Prediction. *Bulletin of the American Meteorological Society*, **76**, 869–888.

Baldwin, M., Gray, L., Dunkerton, T., Hamilton, K., Haynes, P., Randel, W., Holton, J., Alexander, M., Hirota, I., Horinouchi, T. *et al.* (2001). The quasi-biennial oscillation. *Rev. Geophys*, **39**, 179–229.

Barat, J. & Cot, C. (1992). Wind shear rotary spectra in the atmosphere. *Geophysical Research Letters*, **19**, 103–106.

Barry, R. & Chorley, R. (1992). Atmosphere, weather and climate. 421 pp, Chap. 6(c).

Basker, S. & Williams, P. (2010). Navigating eloran: Challenges and the way forward. In *Proc XVIIth IALA Conference*, [http://www.gla-rrnav.org/file.html?file=652be60e268e09c0d91c5b6a976fbc30].

Bechtold, P., Köhler, M., Jung, T., Doblas-Reyes, F., Leutbecher, M., Rodwell, M., Vitart, F. & Balsamo, G. (2008). Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Quarterly Journal of the Royal Meteorological Society*, **134**, 1337–1351.

Brettle, M. & Galvin, J. (2003). Back to basics: Radiosondes: Part 1-The instrument. *Weather*, **58**, 336–341.

Brock, F.V. & Richardson, S.J. (2001). *Meteorological Measurement Systems*. Oxford University Press.

Buck, A. *et al.* (1981). New equations for computing vapor pressure and enhancement factor. *Journal of Applied Meteorology*, **20**, 1527–1532.

Cadet, D. & Teitelbaum, H. (1979). Observational Evidence of Internal Inertia-Gravity Waves in the Tropical Stratosphere. *Journal of Atmospheric Sciences*, **36**, 892–907.

Cess, R., Potter, G., Blanchet, J., Boer, G., Ghan, S., Kiehl, J., Le Treut, H., Li, Z., Liang, X., Mitchell, J. *et al.* (1989). Interpretation of cloud-climate feedback as produced by 14 atmospheric general circulation models. *Science*, **245**, 513.

Chanin, M.L. (1995). The SPARC programme. *Physics and Chemistry of The Earth*, **20, Issue 1**, 21–32.

CNES (1988). BEST scientific objectives and preliminary definition of a satellite mission dedicated to GEWEX and Global Change. Tech. Rep. pp 58, Centre National d'Etudes Spatiales.

Courtier, P., Gauthier, P., Rabier, F., Flamant, P., Dabas, A., Lieutaud, F. & Renault, H. (1992). Study of preparation for the use of Doppler wind lidar information in meteorological assimilation systems. *ESA Contract Rep*, **3453**, 82.

Courtier, P., Andersson, E., Heckley, W., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., Fisher, M. & Pailleux, J. (1998). The ECMWF implementation of three-dimensional variational assimilation (3D-Var).

I: Formulation. *Quarterly Journal of the Royal Meteorological Society*, **124**, 1783–1807.

da Silveira, R., Fisch, G., Machado, L., Dall'Antonia, A., Sapucci, L., Fernandes, D., Marques, R. & Nash, J. (2001). The WMO intercomparison of GPS radiosondes. *Final report, Alcantra, Brazil*, iOM 90 (TD 1314).

Daley, R. (1993). *Atmospheric Data Analysis*, vol. 2. Cambridge university press.

de Kloe, J. (2011). Test cases for the l2b processor. [ae-tn-knmi-gs-0031a-v1_1]: Tn3.1a, KNMI report, 7-Oct-2011.

de Kloe, J. (2012). Test results for the l2b processo. Ae-tn-knmi-gs-0031b-v1_2: Tn3.1b, KNMI report, 13-Jan-2012.

de Kloe, J., Marseille, G.J. & Stoffelen, A. (2007). Vertical aeolus measurement positioning. Ae-tn-knmi-vamp-001, KNMI report, 28-Mar-2007.

DeFelice, T. (1998). *An introduction to meteorological instrumentation and measurement*. Prentice Hall.

DiMego, J.G., Phoebus, P.A. & McDonnel, J.E. (1985). Data processing and quality control for optimum interpolation analysis at the national meteorological center. NMC Office Note 306, NOAA.

Doppler, J. (1842). Über das farbige Licht der Doppelsterne und einiger anderer Gestirne des Himmels: Abn. Königl. Böhm. *Ges. Wiss*, **2**, 465–482.

Durre, I., Vose, R. & Wuertz, D. (2006). Overview of the integrated global radiosonde archive. *Journal of Climate*, **19**, 53–68.

ECMWF-CY25r1, I. (2002). IFS Documentation CY25R1 (Operational implementation 9 April. Chap.5.

ESA (1989). ALADIN Atmospheric Laser Doppler Instrument, Doppler Lidar Working Group. ESA Technical repot SP-1112, European Space Agency.

ESA (1995a). ESA Doppler Wind Lidar Workshop. ESA Technical repot, European Space Agency, workshop 20-22 September.

ESA (1995b). Mission Operations Concept Document. AE-TN-ESA-GS-006 issue 1, European Space Agency.

ESA (1996). Atmospheric Dynamics Mission. ESA Technical report SP-1196(4), European Space Agency, reports for Assessment.

ESA (2008). ADM-AEOLUS: science report. In *Special Publication*, vol. 1311.

ESA (2011). Adm-aeolus level1b products. Ae-rp-dlr-l1b-001_l1b_atbd_iss_2.0, atbd, internal esa document (not public), 10-May-2006.

Essenwanger, O. & Reiter, E. (1969). Power spectrum, structure function, vertical wind shear, and turbulence in troposphere and stratosphere. *Meteorology and Atmospheric Physics*, **18**, 17–24.

Fang, B.T. (1986). Trilateration and extension to global positioning system navigation. *J.Guidance of Control and Dynamics*, **9**, 715–717.

Fitzgibbon, J., Facundo, J. & Spring, S. (2005). Update on the implementation of the national weather service's radiosonde replacement system. In *9th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere,(IOAS-AOLS)*, 11.5, NOAA/NWS and Maryland, Ogden, UT.

Gallice, A., Wienhold, F., Hoyle, C., Immler, F. & Peter, T. (2011). Modeling the ascent of sounding balloons: derivation of the vertical air motion. *Atmospheric Measurement Techniques*, **4**, 2235–2253.

Galvin, J. (2003). Radiosondes: Part 2-Using and interpreting the data. *Weather*, **58**, 387–395.

Gong, J., Geller, M.A. & Wang, L. (2008). Source spectra information derived from US high-resolution radiosonde data. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **113**.

Håkansson, M. (2001). *Determination of Atmospheric Wind Statistics*. Ph.D. thesis, Stockholm University, Department of Meteorology.

Hamilton, K. (2006). High resolution global modeling of the atmospheric circulation. *Advances in Atmospheric Sciences*, **23**, 842–856.

Hamilton, K. & Vincent, R. (1995). High-resolution radiosonde data offer new prospects for research. *Eos Trans. AGU*, **76**, 497.

Holton, J. (1992). *An introduction to dynamic meteorology*. San Diego, Academic Press, ISBN 0-12-354355-X.

Houchi, K., Stoffelen, A., Marseille, G. & De Kloe, J. (2010). Comparison of wind and wind shear climatologies derived from high-resolution radiosondes and the ecmwf model. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **115**.

Houchi, K., Stoffelen, A., Marseille, G.J. & De Kloe, J. (2015). Statistical quality control of high-resolution winds of different radiosonde types for climatology analysis. *Journal of Atmospheric and Oceanic Technology*, **32**, 1796–1812.

Ide, K. (1997). Unified notation for data assimilation: operational, sequential and variational. *Practice*, **75**, 181–189.

Int. Loran Ass. (2007). Enhanced loran (eloran) definition document. *version 0.1, January*.

Javelle, J. (2000). La météorologie du baromètre au satellite. Mesurer l'atmosphère et prévoir le temps (La bibliothèque du naturaliste).

Keshavamurthy, R. & Rao, M. (1992). *The physics of monsoons*. South Asia Books.

Kitchen, M. & Shutts, G. (1990). Radiosonde observations of large-amplitude gravity waves in the lower and middle stratosphere. *Journal of Geophysical Research*, **95**, 20451.

Körnich, H. (2010). Impact of the Vertical Sampling Scenarios on NWP and Stratospheric Wind Analysis. Tech. rep., Stockholm University, Department of Meteorology.

Koshyk, J., Boville, B., Hamilton, K., Manzini, E. & Shibata, K. (1999). Kinetic energy spectrum of horizontal motions in middle-atmosphere models. *Journal of Geophysical Research*, **104**, 27177.

Lorenc, A., Graham, R., Dharssi, I., Macpherson, B., Ingleby, N. & Lunnon, R. (1992). Preparation for the use of Doppler wind lidar information in meteorological assimilation systems. *ESA Contract Rep*, **3454**, 90.

Marseille, G. & Stoffelen, A. (2003). Simulation of wind profiles from a spaceborne Doppler Wind Lidar. *Quarterly Journal of the Royal Meteorological Society*, **129**, 3079–3098.

Marseille, G., Stoffelen, A. & Barkmeijer, J. (2008a). A cycled sensitivity observing system experiment on simulated Doppler wind lidar data during the 1999 Christmas storm'Martin'. *Tellus A*, **60**, 249–260.

Marseille, G., Stoffelen, A. & Barkmeijer, J. (2008b). Impact assessment of prospective spaceborne Doppler wind lidar observation scenarios. *Tellus A*, **60**, 234–248.

Marseille, G., Stoffelen, A. & Barkmeijer, J. (2008c). Sensitivity Observing System Experiment (SOSE)a new effective NWP-based tool in designing the global observing system. *Tellus A*, **60**, 216–233.

Marseille, G., Stoffelen, A., de Kloe, J., Houchi, K., Schyberg, H. & Körnich, H. (2010). VAMP Executive Summary. Tech. report, [Available through KNMI), Ref: Executive summary, ESA, 21p., or to full report].

Marseille, G., Houchi, K., Kloe, J.d. & Stoffelen, A. (2011). The definition of an atmospheric database for aeolus. *Atmospheric Measurement Techniques*, **4**, 67–88.

MathWorks (2010). MATLAB: Statistics Toolbox User's Guide. Technical report, Chap 3. [Available online at http://www.pi.ingv.it/ longo/CorsoMatlab/OriginalManuals/stats.pdf].

Miloshevich, L., Vömel, H., Paukkunen, A., Heymsfield, A. & Oltmans, S. (2010). Characterization and correction of relative humidity measurements from Vaisala RS80-A radiosondes at cold temperatures.

Minnis, P., Yi, Y., Huang, J. & Ayers, K. (2005). Relationships between radiosonde and RUC-2 meteorological conditions and cloud occurrence determined from ARM data. *Journal of Geophysical Research*, **110**, D23204.

Nash, J., Smout, R., Oakley, T., Pathack, B. & Kurnosenko, S. (2005). The WMO intercomparison of high quality radiosonde systems. *Final report, Vacoas, Mauritius*, **IOM 83**, 2–25.

NCDC (1998). Data and documentation for rawinsonde 6-second dataset. Tech. report, [Available online from NCDC (National Climatic Data Center) at ftp://atmos.sparc.sunysb.edu/pub/sparc/hres/TD6211.pdf].

OFCM (2006). Rawinsonde and pibal observations. Handbook no. 3, chap 5, FCM-H3-1997 Washington DC, [Available from OFCM (Office of the Federal Coordinator for Meteorological services and supporting research) at http://www.ofcm.gov/fmh3/pdf/00-entire-FMH3.pdf].

Palmer, T. (2001). A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quarterly Journal of the Royal Meteorological Society*, **127**, 279–304.

Pope, F. (1888). The American inventors of the telegraph, with special references to the services of Alfred Vail. *The Century: Illustrated Monthly Magazine*.

Rabier, F., McNally, A., Andersson, E., Courtier, P., Undén, P., Eyre, J., Hollingsworth, A. & Bouttier, F. (1998). The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions. *Quarterly Journal of the Royal Meteorological Society*, **124**, 1809–1829.

Rees, G. (2001). *Physical principles of remote sensing*. Cambridge Univ Pr.

Reitebuch, O., Lemmerz, C., Nagel, E., Paffrath, U., Durand, Y., Endemann, M., Fabre, F. & Chaloupy, M. (2010). The airborne demonstrator for the direct-detection Doppler Wind Lidar ALADIN on ADM-Aeolus. Part I: Instrument design and comparison to satellite instrument.

Riddaway, R.W. (2002). Numerical methods. ECMWF document, revised by M. Hortal, March 2001. Chap 6. pp.

Rossow, W. & Schiffer, R. (1991). ISCCP cloud data products. *Bulletin of the American Meteorological Society*, **72**, 2–20.

Rossow, W. & Schiffer, R. (1999). Advances in understanding clouds from ISCCP. *BULLETIN-AMERICAN METEOROLOGICAL SOCIETY*, **80**, 2261–2288.

Rossow, W., Zhang, Y. & Wang, J. (2005). A statistical model of cloud vertical structure based on reconciling cloud layer amounts inferred from satellites and radiosonde humidity profiles. *Journal of Climate*, **18**, 3587–3605.

Schiffer, R. & Rossow, W. (1983). The international satellite cloud climatology project(isccp)- the first project of the world climate research programme. *American Meteorological Society, Bulletin*, **64**, 779–784.

Schwartz, B. & Govett, M. (1992). A hydrostatically consistent north american radiosonde data base at the forecast systems laboratory. NOAA Tech. report ERL FSL-4.

Shaowu, W. (1994). Diagnostic Studies on the Climate Change and Variability for the for the period of 1880-1990 [J]. *Acta Meteorologica Sinica*, **3**.

Shutts, G., Kitchen, M. & Hoare, P. (1988). A large amplitude gravity wave in the lower stratosphere detected by radiosonde. *Quarterly Journal of the Royal Meteorological Society*, **114**, 579–594.

Shutts, G., Healey, P. & Mobbs, S. (1994). A multiple sounding technique for the study of gravity waves. *Quarterly Journal of the Royal Meteorological Society*, **120**, 59–77.

Skamarock, W. (2004). Evaluating mesoscale NWP models using kinetic energy spectra. *Monthly Weather Review*, **132**, 3019–3032.

Stephens, G. (1984). The parameterization of radiation for numerical weather prediction and climate models. *Monthly Weather Review*, **112**, 826–867.

Stephens, G., Vane, D., Boain, R., Mace, G., Sassen, K., Wang, Z., Illingworth, A., O'Connor, E., Rossow, W., Durden, S. *et al.* (2002). The CloudSat mission and the A-Train. *Bulletin of the American Meteorological Society*, **83**, 1771–1790.

Stoffelen, A. & Marseille, G.J. (1998). Study on the utility of the Doppler wind LIDAR data for Numerical Weather Predition and climate. Tech. rep., European Space Agency.

Stoffelen, A., Marseille, G.J., Flamant, P., Hkansson, M., E., K., Pailleux, J. & H. Schyberg, M.V. (2002). MERCI - Measurement Error Correlation Impact on the Atmospheric Dynamics Mission. Tech. rep., KNMI.

Stoffelen, A., Pailleux, J., Källén, E., Vaughan, J., Isaksen, L., Flamant, P., Wergen, W., Andersson, E., Schyberg, H., Culoma, A. *et al.* (2005). The Atmospheric Dynamics Mission for Global Wind Field Measurement. *Bulletin of the American Meteorological Society*, **86**, 73–87.

Stoffelen, A., Marseille, G., Bouttier, F., Vasiljevic, D., De Haan, S. & Cardinali (2006). ADM-Aeolus Doppler Wind Lidar Observing System Simulation Experiment. *Quart. J. Roy. Meteor. Soc.*, **132**, 1927–1947.

Stoffelen, A., Marseille, G.J., de Kloe, J., Dabas, A., Huber, D. & Reitebuch, O. (2008). Comparison of aeolus burst and continuous mode concepts. Tech. report, [Available through KNMI), Ref: AE-TN-KNMI-GS-051, version: 0.8, 6 October 2008].

Stoffelen, A., Körnich, H., Marseille, G., Houchi, K. & Kloe, J.D. (2009). Assessment of optical and dynamical atmospheric heterogeneity. Tech. report, [Available through KNMI),Ref: AE-TN-KNMI-VAMP-002_v7_KNMI_150609, version: 0.8, 15 June 2009].

Stubenrauch, C., Del Genio, A. & Rossow, W. (1997). Implementation of subgrid cloud vertical structure inside a GCM and its effect on the radiation budget. *Journal of climate*, **10**, 273–287.

Stull, R. (1988). *An introduction to boundary layer meteorology*. Springer.

Sun, X., Zhang, R., Marseille, G., Stoffelen, A., Donovan, D., Liu, L. & Zhao, J. (2014). The performance of aeolus in heterogeneous atmospheric conditions using high-resolution radiosonde data. *Atmospheric Measurement Techniques*, **7**, 2695–2717.

Tan, D., Andersson, E., Fisher, M. & Isaksen, L. (2007a). Observing-system impact assessment using a data assimilation ensemble technique: application to the ADM-Aeolus wind profiling mission. *Quarterly Journal of the Royal Meteorological Society*, **133**, 381–390.

Tan, D., Andersson, E., Poli, P., Dabas, A., De Kloe, J., Marseille, G. & Stoffelen, A. (2007b). ADM-AEOLUS Level-2B Algorithm Theoritical Baseline Document (Mathematical description of the AEOLUS Level-2B Processor). *Change*, **2**, 23–12.

Tan, D., Andersson, E., De Kloe, J., Marseille, A., G.J.A.N.and Stoffelen, Poli, P., DENNEULIN, M., Dabas, A., Huber, D., Reitebuch, O. *et al.* (2008). The ADM-Aeolus wind retrieval algorithms. *Tellus A*, **60**, 191–205.

UK-MetOffice (2006-2009). Uk high-resolution radiosonde data. Tech. report, British Atmospheric Data Centre, [Available from http://badc.nerc.ac.uk/data/ukmo-rad-hires/].

Vaisala (2002). Vaisala radiotheodolite rt20. Tech. report, [Available online from http://www.hobeco.net/pdf/rt20_brochure.pdf].

Vaughan, M., Powell, K., Kuehn, R., Young, S., Winker, D., Hostetler, C., Hunt, W., Liu, Z., McGill, M. & Getzewich, B. (2009). Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements. *Journal of Atmospheric and Oceanic Technology*, **26**, 2034.

Vaughan, M., Kuehn, R., Tackett, J., Rogers, R., Liu, Z., Omar, A., Getzewich, B., Powell, K., Hu1, Y., Young, S., Avery, M., Winker, D. & Trepte, C. (2010). Strategies for Improved CALIPSO Aerosol Optical Depth Estimates. In *Proc.25th ILRC, St. Petersburg, Russia.*.

Vaughan, N., Geddes, P., Flamant, C. & Flesia, C. (1998). Establishment of a backscatter coefficient and atmospheric database. *ESA-CR12510*.

Veldman, S., Knobbout, H., Stoffelen, A., Marseille, G. & Fuchs, J. (2000). *Lidar Performance Analysis Simulator-LIPAS*. Citeseer.

Wang, J. & Rossow, W. (1995). Determination of cloud vertical structure from upper-air observations. *Journal of Applied Meteorology*, **34**, 2243–2258.

Wang, J. & Rossow, W. (1998). Effects of Cloud Vertical Structure on atmospheric circulation in the GISS-GCM. *Journal of Climate*, **11**, 3010–3029.

Wang, J., Rossow, W. & Zhang, Y. (2000). Cloud vertical structure and its variations from a 20-yr global rawinsonde dataset. *Journal of Climate*, **13**, 3041–3056.

Wang, J., Carlson, D., Parsons, D., Hock, T., Lauritsen, D., Cole, H., Beierle, K. & Chamberlain, E. (2003). Performance of operational radiosonde humidity sensors in direct comparison with a chilled mirror dew-point hygrometer and its climate implication. *Geophyical Research Letter*, **30**, 1860.

Wang, J., Bian, J., Brown, W.O., Cole, H., Grubišic, V. & Young, K. (2009). Vertical air motion from t-rex radiosonde and dropsonde data. *Journal of Atmospheric and Oceanic Technology*, **26**, 928–942.

Wang, L. & Geller, M.A. (2003). Morphology of gravity-wave energy as observed from 4 years (1998–2001) of high vertical resolution US radiosonde data. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **108**.

Warren, S., Hahn, C., London, J., Chervin, R. & Jenne, R. (1986). Global distribution of total cloud cover and cloud type amounts over land. Tech. rep., Washington Univ., Seattle (USA). Dept. of Atmospheric Sciences; Colorado Univ., Boulder (USA); National Center for Atmospheric Research, Boulder, CO (USA).

Warren, S., Hahn, C., London, J., Chervin, R., Jenne, R. *et al.* (1988). Global distribution of total cloud cover and cloud type amounts over the ocean. Tech. rep., USDOE Office of Energy Research, Washington, DC (USA). Carbon Dioxide Research Div.; National Center for Atmospheric Research, Boulder, CO (USA).

Werner, C. (2005). *Doppler Wind lidar*. Springer.

Wielicki, B. & Harrison, E. (1995). Mission to planet earth: Role of clouds and radiation in climate. *Bulletin of the American Meteorological Society*, **76**.

Winker, D., Couch, R. & McCormick, M. (1996). An overview of LITE: NASA's Lidar in-space technology experiment. *Proceedings of the IEEE*, **84**, 164–180.

WMO (1987). LAWS Laser Atmospheric Wind Sounder. Instrument Panel Report Earth Observing System, Volume IIg, NASA TM 86129, National Aeoronautical and Space Administration.

WMO (2002). Statement of Guidance How Well Satellite Capabilities Meet WMO User Requirements in Several Application Areas. WMO Satellite Reports SAT-26 WMO/TD No. 1052, World Meteorological Organization.

WMO (2003). Affiliate and CEOS Member Requirements. Mission and Instrument Database (CD-ROM). Tech. report, February 2003. Version 2.5, World Meteorological Organization.

WMO (2008a). Measurement of meteorological variables; guide to meteorological instruments and methods of observation. WMO (World Meteorological Organization) CIMO Guide N8, 7th Edition. Part I.

WMO (2008b). Measurement of upper wind; guide to meteorological instruments and methods of observation. WMO (World Meteorological Organization) CIMO Guide, 7th Edition. chap 13.

WMO (2008c). Status of the space-based global observing system. Tech. report, World Meteorological Organization, [http://www.wmo.int/pages/prog/sat/meetings/documents/], [cm8Doc051.pdf].

Žagar, N., Stoffelen, A., Marseille, G., Accadia, C. & Schlüssel, P. (2008). Impact assessment of simulated Doppler wind lidars with a multivariate variational assimilation in the tropics. *Monthly Weather Review*, **136**, 2443–2460.

Zhang, J., Chen, H., Li, Z., Fan, X., Peng, L., Yu, Y. & Cribb, M. (2010). Analysis of cloud layer structure in shouxian, china using rs92 radiosonde aided by 95 ghz cloud radar. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **115**.

APPENDIX A

Appendix

A.1 Inventory of High-resolution Radiosondes

An inventory of the available high-resolution radiosondes which are gathered and analyzed in the different studies of this thesis are summarized in table A.1. These data are mainly from the SPARC and AMMA projects and from the British Atmospheric Data Centre (BADC). Mainly the SPARC data set provided radiosonde stations covering an entire year with all seasons well represented. In this respect only few AMMA and BADC stations remain useful. The Figures A.1, A.1, A.3) and A.4 in Appendix A illustrate distributions over time that allow the selection of the most regular data in time, from which a representative climatology has been made. Additional stations such as from the FASTEX project, stations Cabauw and Paramaribo-STAR (Support for Tropical Atmospheric Research) project are not considered here because of their insufficient time coverage. Station De Bilt however looks more continuous and has been analyzed. The map in figure 3.1 shows the selected and analyzed data that were considered relevant for this study. All these data are obtained either by contacting weather centers on a personal basis, or otherwise through anonymous ftp access. Further high-resolution radiosonde data from France and few from Australia and Belgium were received recently, so they were not integrated in the database used in the analyses underlying this thesis.

A.2 Data Selection

A.2.1 BADC Data

Station(s) / Features	AMMA (Africa)	SPARC (USA)	BADC / Met Office (UK)	De Bilt (NL)
Vertical Resolution [m]	$\sim$ 10	$\sim$ 30	$\sim$ 10	$\sim$ 50 ($\sim$ 10 on request)
Location (lon,lat) [degrees]	18W-10E/0-21N	071W171E/14S-71N	UK and scattered	5.18 E/52.13 N
Number of stations	21	92	21	1
Sampling Periods	2005 to 2007	1998 to 2007	1990 to 2008 (UK) 1999 to 2008 (Others)	1993 to 2008
Number of Ascents a day	00&12 UTC and 06&18 UTC (for some locations)	00&12 UTC	00&12 UTC and 06&18 UTC (for some locations)	00&12 UTC
Wind-finding system	LORAN and GPS	Radio-theodolite	LORAN and GPS	Loran-C and GPS
Data Format	ASCII	ASCII	Mostly Binary	ASCII
Region	North west Africa	USA and pacific Ocean	UK-stations, Gibraltar, St Helena, Falkland	De Bilt, Netherlands

Table A.1: Available high-resolution radiosonde data. SPARC data are devoted to gravity wave studies; AMMA to understanding the African monsoon and BADC and De Bilt dedicated to weather forecasting. The BADC and De Bilt stations are part of the GOS and also available in WMO standard resolution.

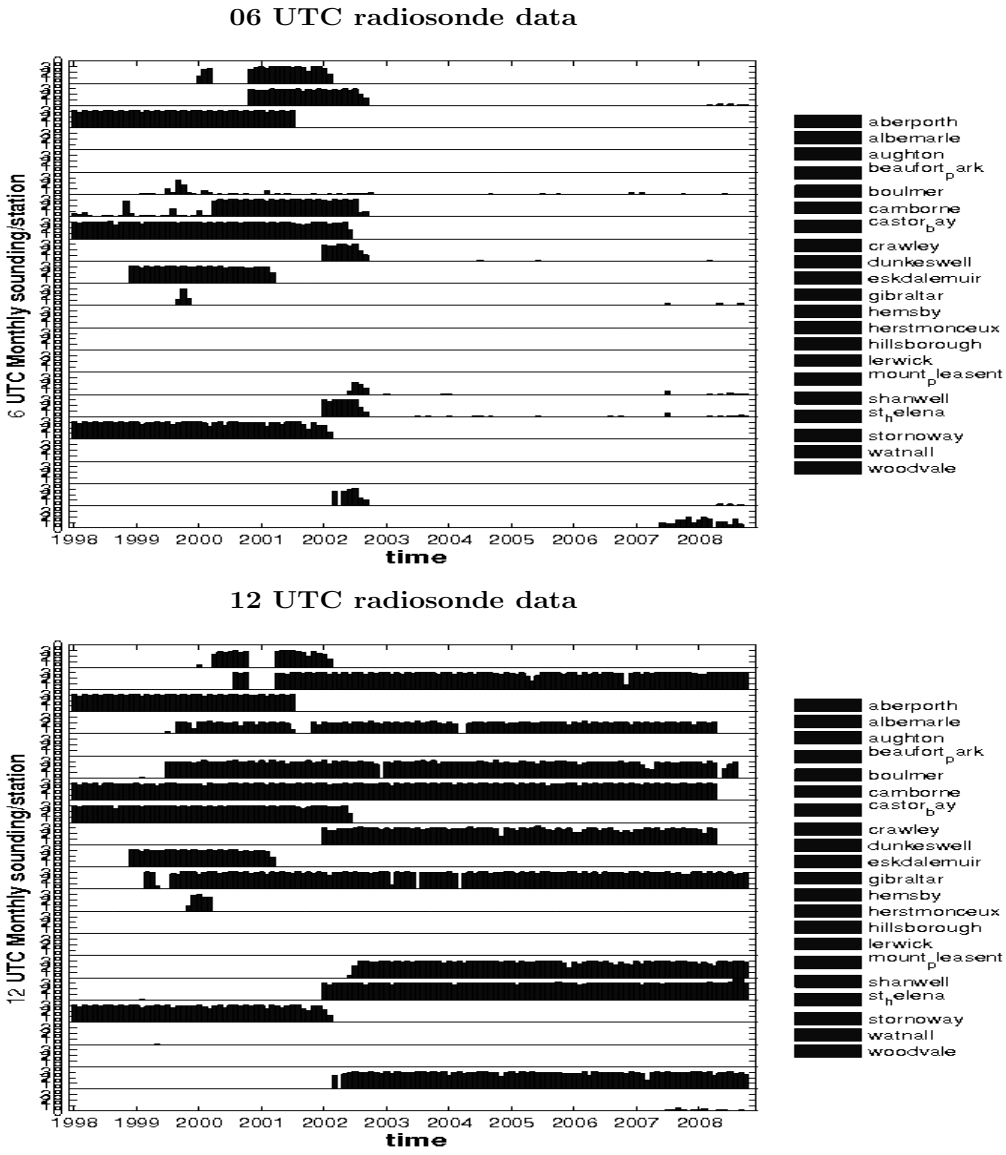


Figure A.1: Day coverage available at 06 UTC (top) and 12 UTC (bottom) of high-resolution stationsfrom the British Atmospheric Data Centre (BADC). Note that the data entries in the plot (from bottom to top) correspond to the station list (from the top to bottom). The vertical bars in each data entry run from 1 to 40 and depicts the number of days with a sounding per month. When the bar exceeds the number of days in a given month, then some additional soundings were available.

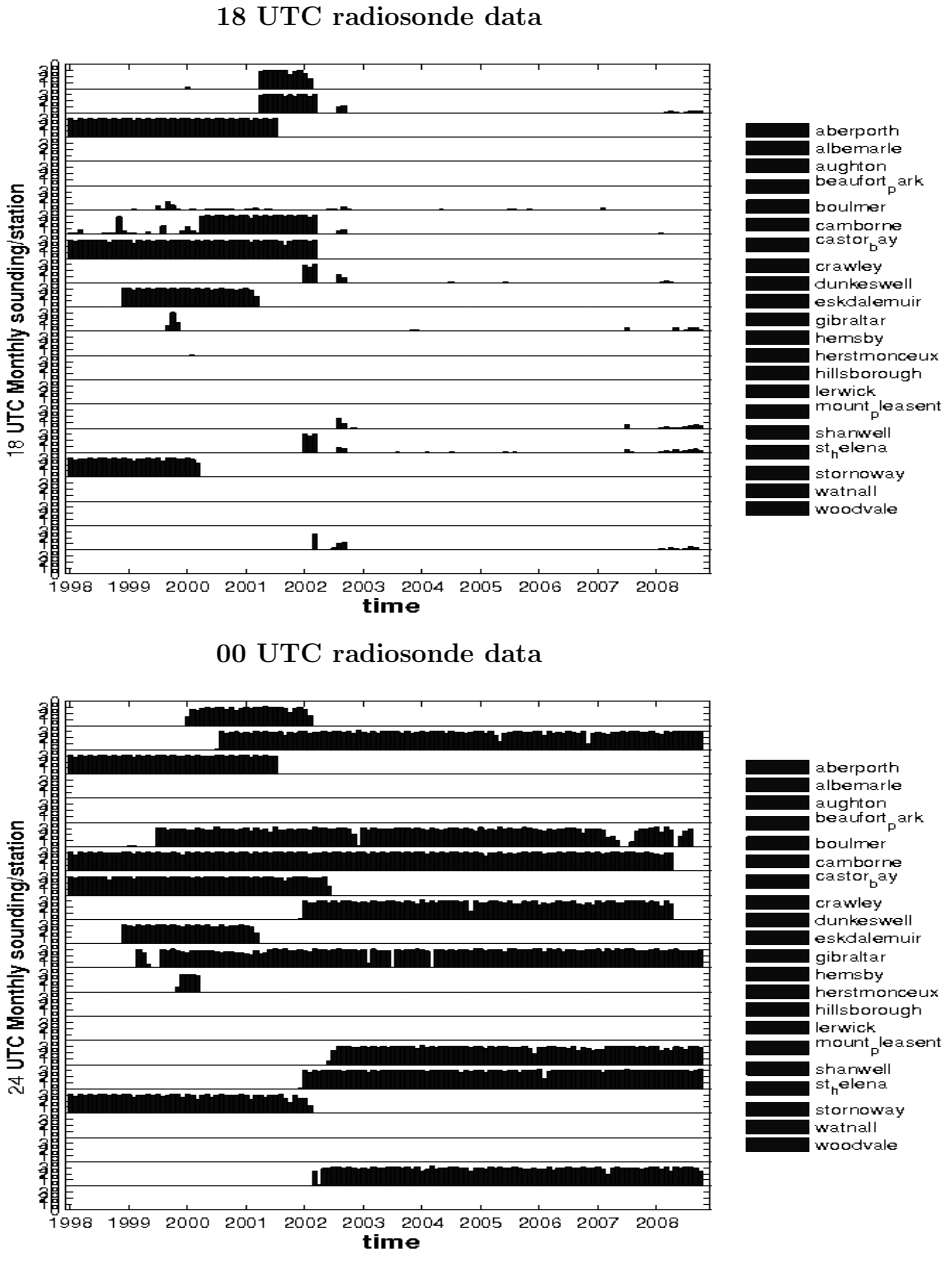


Figure A.2: Same figure as A.1, but now for 18 UTC and 00 UTC.

A.2.2 AMMA Data

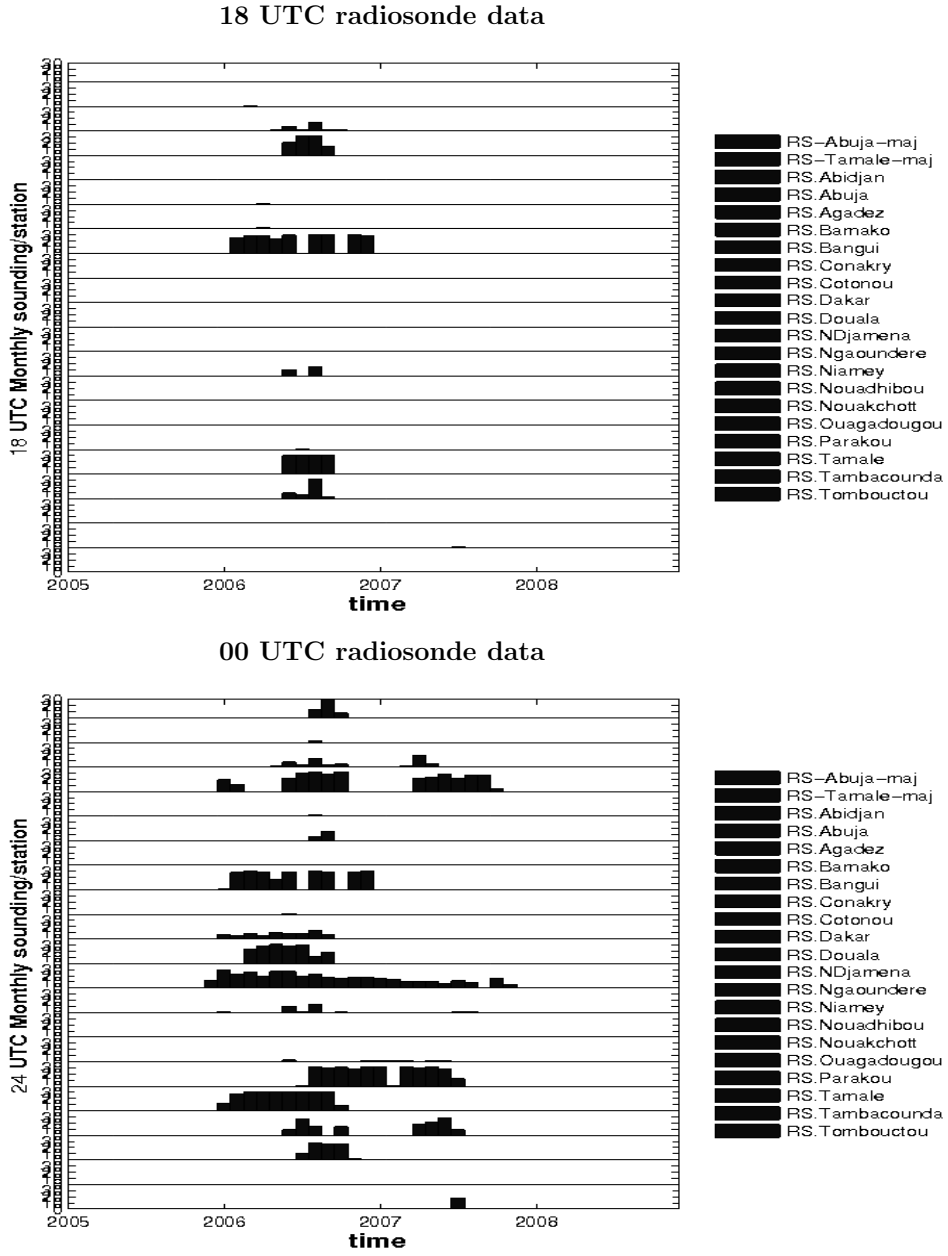


Figure A.3: Same figure as A.1, but now for AMMA data sets (see map 3.1

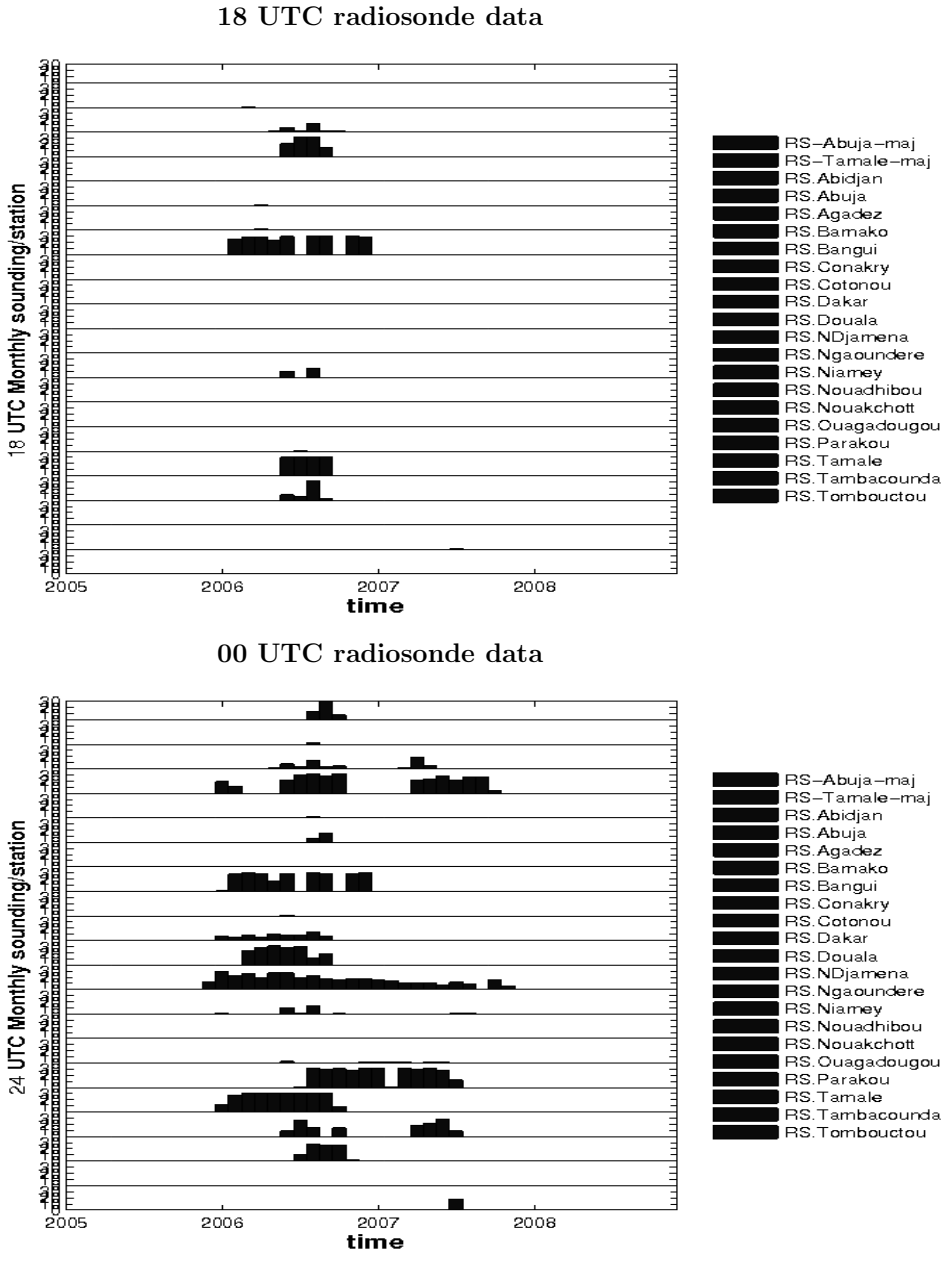


Figure A.4: Same figure as A.3, but now for 18 UTC and 00 UTC.

A.3 Radiosonde Wind Formulas

Most radiosondes contain three sensors, one for pressure, temperature and humidity. These quantities are successively reported in the following units (hecto-Pascals (hPa) (one hP equals one mbar), degrees Kelvin and percent resp.. The conversion from the engineering units to meteorological units will be specific to the radiosonde and its manufacturer for all these variables. The other important meteorological variable measured via balloon radiosonde is wind, by determining the successive balloon positions at regular time intervals (up to 1 second, but 2 seconds more generally). The method utilized to retrieve balloon locations and wind from subsequent locations, is known as a Wind-Finding System (WFS) or a Navigation-Aid (NAVAID)-derived wind system. The WFS is under the control of the manufacturer and may not be always the same for one radiosonde type to another. When a radar-reflector is integrated in a radiosonde, then the balloon radiosonde can be tracked by a radar (known as a radiotheodolite). Other radiosonde types with more modern WFS to derive wind are LORAN and GPS.

Wind information can be recorded in polar form, in degrees from true north and in knots or in a vector form, in meters per second. Winds are computed using earth spherical geometry, ensuring an accurate determination in case of high winds. The calculation of wind from the position of the balloon shall be made using earth spherical coordinates. The position of the balloon should also be used for establishing the position of the information at the levels selected for the rawinsonde message. According to the WMO meteorological convention, wind direction is specified as that from which the wind blows, contrary to oceanographic convention, in which currents are labeled with the direction towards which the current is flowing. A compass oriented clockwise from north is used; thus, a wind blowing from the west is given a 270 degree direction, and from the south a 180 degree wind direction. In meteorological usage the vector velocity component values are calculated from the wind speed and direction. The u-component (blowing from west to east) of the wind is given by:

$$u = speed.sin(direction) \tag{A.1}$$

and the v- component (blowing from south to north) by

$$v = speed.cos(direction) \tag{A.2}$$

Wind speed can be determined from the vector components by

$$speed = \sqrt{u^2 + v^2} \tag{A.3}$$

The mean wind included in the coded message at standard levels is calculated by averaging separately the u- and v-components obtained at high resolution between the surface and the 1525 m (5000 ft) levels and between the 1525 m and the 2048m (10,000 ft) levels. More details on all meteorological variables measured by balloon radiosonde, may be found in OFCM (2006).

A.4 WMO Limit-Values

The table A.4 below summarize the extreme possible values that can be tolerated for the various radiosonde measured quantities, PTU, wind speed and height. This table is important in case of the QC of radiosonde data.

ValidityChecks					
	Height(*m*)		*Temp.*(*C*)		*MaxWindSpeed*(*Kts*)
GeopotentialLevel(*mb*)	*Low*	*High*	*Low*	*High*	
1000	-588	601	-65	60	70
850	634	1853	-50	45	90
700	2101	3473	-50	30	120
500	4505	6121	-57	5	200
400	5870	7791	-66	-10	250
300	7726	9952	-72	-20	300
250	8835	11274	-76	-25	300
200	10260	12699	-78	-30	300
150	12094	14533	-85	-30	200
100	14000	17500	-95	-30	200
70	16496	19596	-95	-25	200
50	18402	21602	-95	-15	200
30	21003	25503	-95	-5	200
20	23501	28001	-95	5	200
10	27003	33003	-95	15	200
<10	N/A	45000	-95	15	200

Table A.2: Tolerance limits for the extreme values of temperature, height and wind speed at each mandatory pressure level, according to WMO (DiMego *et al.*, 1985). Notice the units; 1Kts ~ 0.5 ms-1.

Acronyms and Symbols

Acronyms

ACCD	Accumulation Charge-Coupled Device
ADM	Atmospheric Dynamic Mission
Aeolus	Name of the ESA ADM
ALADIN	Atmospheric LAser Doppler INstrument Lidar
AMMA	African Monsoon Multidisciplinary Analysis
AMV	Atmospheric Motion Vector
ATOVS	Advanced TIROS Operational Vertical Sounder
BADC	British Atmospheric Data Centre
BRC	Basic Repeat Cycle
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation
CLIVAR	Climate and Ocean: Variability, Predictability and Change
CRM	Cloud Resolving Model
CVS	Cloud Vertical Structures
CWIC	Cloud Water Ice Content
CWLC	Cloud Water Liquid Content
DA	Data Assimilation
DLR	Deutschland fũr Luft-und Raumfahrt
DWL	Doppler Wind Lidar
E.m.	Electromagnetic
E2E	End-to-End
E2S	E2E Simulator
ECMWF	European Centre for Medium-range Weather Forecasts
ENSO	El Niño/La Niña Southern Oscillations
EO	Earth Observation
ESA	European Space Agency
FASTEX	Fronts and Atlantic Storm-Track EXperiment
GCM	General Circulation Model
GEWEX	Global Energy and Water Cycle Experiment
GOS	Global Observing System

GPS	Global Positioning System
GWP	Global Wind Patterns
HLOS	Horizontal projection of the LOS wind component
HSRL	High Spectral-Resolution Lidar
IFS	Integrated Forecasting System
IPCC	Intergovernmental Panel on Climate Change
ITCZ	Intertropical Tropical Convergence Zone
JOSS	Joint Office of Science Support
KNMI	Koninklijk Nederlands Meteorologisch Instituut
L1Bp/L2Ap/L2Bp	Level 1B/2A/2B processors
LAWS	Laser Atmospheric Wind Sounder
LIDAR	LIght Detection And Ranging
LIPAS	Lidar Performance Analysis Simulator
LMD	Laboratoire de Meteorologie Dynamique
LORAN	LOng RAnge Navigation
LORAN-C	Type C LORAN
LOS	Line-Of-sight
LST	Local Solar Time
MERCI	Measurement Error and Correlation Impact on ADM
MF	Meteo France
MicroART	Processor ground radiosonde system
MW	Microwave
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCDC	National Climatic Data Center
NE	Northern East
NERC	Natural Environment Research Council
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Production
O-B	Observation minus Background
P	Pressure
PBL	Planetary Boundary Layer
PRF	Pulse Repetition Frequency
PTU	Pressure-Temperature-Humidity
QC	Quality Control
RADAR	RAdio Detection And Ranging
RH	Relative Humidity
Rs	Radiosondes
SE	Southern East
SH	Southern Hemisphere
SNR	Signal-to-Noise Ratio
SODAR	SOund Detection And Ranging
SPARC	Stratospheric Processes And their Role in Climate
SRF	Short-Range Forecast
STAR	Support for Tropical Atmospheric Research
SYNOP	Surface SYNOPtic observation

T	Temperature
TEMP	WMO code for radiosonde message: PTU and wind velocity
TIROS	Television Infra-Red Observation Satellite
U	Humidity
UCAR	University Corporation for Atmospheric Research
UTC	Coordinated Universal Time
UV	Ultra-Violet
VAMP	Vertical Aeolus Measurement Positioning
VFM	Vertical Features Mask
WCRP	World Climate Research Programme
WFS	Wind-Finding System
WMO	Word Meteorological Organization

Symbols

DDP	Dawn-Dusk Profile
EMP_σ	Sigma Error Multiplier Profile
EMP	Error Multiplier Profile
LSP	Land-Sea Profile
$p-QC$	Percentile QC
SD	Standard deviation

Acknowledgments

I am very grateful to Professor Hennie Kelder and Pierre Flamant for their guidance and encouragement, and giving me the opportunity to carry out this research work in the framework of collaboration between Technical University of Eindhoven and the "École Polytechnique". I would like to express my deepest gratitude to my mentor Ad Stoffelen for his guidance during the thesis studies, in particular for his advice which were invaluable for the successful completion of this research work. I thank him again for his strenuous job of proof reading parts of the manuscript, doing his best in the short time available, and for his excellent, professional, and extensive advice.

My special thank goes to my office mate Gert-Jan Marseille for his inspiring discussions, his tireless help, and patience in providing me with the details about Aeolus mission and more. similarly I thank Jos de Kloe for his support, help and willing to share invaluable knowledge, in computing sciences in particular.

I would like to thank the people who helped us in the long-winded task of gathering the high-resolution radiosonde data and who provided support: Stefan Liess from the SPARC project, Guillaume Brissebra from AMMA, David Hooper and his team from BADC (UK Met-Office), and Richard Rhoe, Marc Allaart, Jitze van der Meulen, Henk Klein Baltink, and Ge Verver from KNMI who made De Bilt high-resolution radiosonde data available and provide data access and documentations.

I would like to thank Anne Grete and Olivier Le Rille from ESA-ESTEC (Noordwijk) for their constructive comments, suggestions and advice, during regular VAMP project meetings related to my thesis work.

I thank all KNMI colleagues for the pleasant working atmosphere, in particular those who supported me in any respect during the completion of the project: Maria Belmonte, Davit Harutyunyan, Jan-Otto Hooghoudt, Marcos Portabellla, Niels Zweers, Tim Vlemmix, Bas Mijling, Aart Overeem, Anton Verhoef, Roeland Stappers, Jeroen Verspeek, Jur Vogelzang, John de Vries, Kees Kok, Rob Roebeling, Maurice Schmeits, Hans de Vries. Without forgetting the football and the running teams: Wim de Rooij, Sander Tijm,, Frank Selten, Wim Verkley, Cisco de Bruijn, etc. Without the generous help of these individuals, this work would not have been possible, and sorry for the forgotten names.

Lastly, I thank my parents and my lovely wife, Dhya, for her unwavering support and revision of some parts of this thesis.

Curriculum Vitae

Karim Houchi was born in July 1976 in Great Kabylia (Tizi Ouzou), Algeria. He started his university studies in Aeronautic Engineering at the Aeronautical Institute of Blida the University in Algeria. He moved to France in 2001 to pursue his curriculum at the University of Paris XII, where he obtained successively a "Licence" and "Matrise" degrees on Mechanical Technology in 2002 and 2003. He received a French government grant to follow a Master degree on "Fluid mechanics & Energetics" at the University of Pierre Marie Curie (Paris VI) that he obtained in 2005 (with distinction). The subject of the Master internship was about investigations on marine turbulence by performing in situ measurements with an Acoustic Doppler Current Profiler (ADCP) and using multi-scales analysis methods. This work was done at the Marine Station of Wimereux (Boulogne sur- Mer, France) under the supervisory of Dr. Francois Schmitt, currently director of the Marine Station. In 2006 the author followed courses and an internship related to climate and environment research at the LSCE (Saclay, France). The focus of this internship was the study of the role of the anthropogenic nitrogen in the response of the biosphere to climate change and CO2 increase, by means of vegetation model, ORCHIDEE-N. In January 2007, he joined the team of Dr. Ad Stoffelen at the Royal Netherlands meteorological Institute (KNMI) to prepare a PhD in the framework of Aeolus ESA's space mission, which was sponsored by the Dutch Ministry of Transport, Public Works and Water Management (Ministerie van Verkeer en Waterstaat). The focus of the research during the PhD was about high-resolution wind and wind shear profiles and cloud vertical structures, which are mainly relevant for the preparation of the Aeolus mission. During the PhD period, he attended various conferences/workshops/progress meetings, where he presented his work. Just after the PhD, he joined the CRP-GL (Centre de Recherche Public Gabriel Lippmann) for a PostDoc position in Luxembourg, during 10 months. The research topic was the assessment of regional climate change impacts on the water cycle in Luxembourg and neighboring regions. Since the end of the Postdoc, he has been teaching Physics and Maths at the European School in Luxembourg-I.

.

Publications

In this thesis Chapter 3 is based on the second paper listed below and Chapter 4 on the first paper.

Publications

Houchi, K., Stoffelen, A., Marseille, G.J., de Kloe, J. (2010). Comparison of wind and wind shear climatologies derived from high-resolution radiosonde and ECMWF model. Journal of Geophysical Research, 115, D22123.

Houchi, K., A. Stoffelen, G.J. Marseille, J. De Kloe (2015). Statistical Quality Control of High-Resolution Winds of Different Radiosonde Types for Climatology Analysis. Journal of Atmospheric and Oceanic Technology, 32, 1796-1812.

Marseille, G.J., Houchi, K., Stoffelen, A. & de Kloe, J. (2011). The definition of an atmospheric database for ADM-Aeolus. Atmospheric Measurement Techniques, 4, 67-88.

Houchi, K., A. Stoffelen, G.J. Marseille, J. De Kloe (2009), Characterization of wind and wind-shear profiles from high-resolution radiosondes and ECMWF model. Extended Abstarct, conference proceedings of the 8th International Symposium on Tropospheric Profiling: Integration of Needs, Technologies and Applications (ISTP8).

Stoffelen, A., H. Körnich, G.J. Marseille, K. Houchi, and J. De Kloe (2009), Assessment of Optical and Dynamical Atmospheric Heterogeneity, available through KNMI, ESA's Report, Ref: AE-TNKNMI-VAMP-002_v7_KNMI_150609, version: 0.8, 15 June 2009.