Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN FISICA
Ciclo XXIV

COORDINATORE Prof. Filippo Frontera

Multi-sensor Satellite Precipitation Estimate for Hydrogeological Hazard Mitigation

Settore Scientifico Disciplinare FIS/06

Dottoranda:
Dott. Lisa Milani

Tutore:
Prof. Federico Porcú

Anni 2009–2011

To my parents

Contents

List of Acronyms **v**

Introduction **1**

1 Nowcasting and flood risk mitigation **3**

1.1 High-impact meteorological events 4

1.1.1 Global, continental and local relevance of the problem 4

1.1.2 Research activities in Europe 10

1.2 The research in Italy and the PROSA System 18

1.2.1 The PROSA structure 19

1.2.2 PROSA Products 20

1.2.2.1 PROSA Precipitation Products 22

1.2.2.2 PROSA soil products 26

1.2.2.3 PROSA Numerical Weather Prediction products . . 27

1.2.3 PROSA validation and demonstration activities 28

1.3 Nowcasting systems 28

1.3.1 UK Met Office Nowcasting Systems 28

1.3.1.1 The Nimrod System 28

1.3.1.2 The GANDOLF System 29

1.3.1.3 STEPS System 31

1.3.2 The TAMORA algorithm 31

1.3.3 PERSIANN Systems 33

1.3.4 Satellite Application Facility to support Nowcasting (SAFNWC) 34

2 Instruments for precipitation estimate **37**

2.1 Ground based instruments 37

2.1.1 Rain-gauges . . . 37
2.1.2 Meteorological Radar . . . 38
2.2 Satellite Remote sensing . . . 44
2.2.1 Basic definitions . . . 46
2.2.2 Visible . . . 50
2.2.3 3.9 μm . . . 52
2.2.4 Infrared: atmospheric window wavelengths . . . 57
2.2.5 Infrared: absorption wavelengths . . . 59
2.2.6 Microwave . . . 61
2.3 Space-borne sensors . . . 63
2.3.1 SEVIRI sensor . . . 64
2.3.2 AMSU sensors . . . 64

3 Artificial Neural Networks 69
3.1 Neural Network in geophysics and remote sensing . . . 69
3.2 ANN introduction . . . 70
3.3 Single Layer Perceptron . . . 72
3.4 Multi-Layer Perceptron . . . 74
3.5 ANN building phases . . . 76
3.5.1 Data set building phase . . . 76
3.5.2 Training phase . . . 76
3.5.3 Testing phase . . . 77
3.6 Interpretation of network outputs . . . 78
3.6.1 Regression problems . . . 78
3.6.2 Classification problems . . . 80

4 Satellite precipitation estimate algorithms 83
4.1 The original algorithm . . . 83
4.2 PROSA Product 1 . . . 89
4.3 ANN training over Italy . . . 90
4.4 Night-time ANN . . . 95
4.5 Classes of precipitation intensity . . . 98
4.6 PROSA Product 2 . . . 104
4.7 Precipitation rate as a continuous variable . . . 106

4.8 Microwave data for ANN online calibration 110
4.9 PROSA Product 3 115

Conclusions **119**

A Satellite estimates validation **121**
A.1 Dichotomic parameters 122
A.2 Multicategorical parameters 124
A.3 Continuous variables 125

Acknowledgments **127**

List of Acronyms

ALOS Advanced Land Observing Satellite

AMV Atmospheric Motion Vectors

AMSR-E Advanced Microwave Scanning Radiometer for EOS (onboard Aqua)

AMSU Advanced Microwave Sounding Unit

ANN Artificial Neural Network

ASAR Advanced Synthetic Aperture Radar (onboard ENVISAT)

ASI Italian Space Agency

AVI Aree Vulnerate Italiane (Italian areas affected by mass-movements and floods)

BB Binary Bias

CDRD Cloud Dynamics & Radiation Database

CETEMPS Centro di Eccellenza per l'Integrazione di Tecniche di Telerilevamento e Modellistica Numerica per la Previsione di Eventi Meteorologici Severi - Excellence Centre for Remote Sensing Technique Integration and Numerical Modeling for Severe Meteorological Events Forecasting

CMORPH CPC MORPHing method

CNR Consiglio Nazionale delle Ricerche (National Council of Research of Italy)

COST COoperation in Science and Technology

CPC Climate Prediction Center

CRM Cloud Resolving Model

CSI Critical Success Index

DPC Dipatimento della Protezione Civile (Department of Civil Protection of Italy)

DTM Digital Terrain Model

ENVISAT European ENVIronment SATellite

EO Earth Observation

ETS Equitable Threat Score

EUMETNET Network of European Meteorological Services

EUMETSAT European Organization for the Exploitation of Meteorological Satellites

FAR False Alarm Ratio

FP Framework Programme

FPn Framework Programme number n

FSE Fractional Standard Error

GANDOLF Generating Advanced Nowcasts for Deployment in Operational Land-based Flood forecasts

GEO Geostationary Earth Orbit

GNDCI Gruppo Nazionale per la Difesa dalle Catastrofi Idrogeologiche (National group for the prevention of hydrogeological disasters of Italy)

GOES Geosynchronous Operational Environmental Satellite

GPM Global Precipitation Measurement

HK Hanssen-Kuipers Skill Score

HRV High Resolution Visible

H-SAF EUMETSAT Satellite Application Facility on Support to Operational Hydrology and Water Management

HSS Heidke Skill Score

IFAC Istituto di Fisica Applicata 'Nello Carrara' - Applied Physics Istitute 'Nello Carrara'

IFOV Instantaneous Field Of View

IPWG International Precipitation Working Group

IR Infrared

ISAC Istituto di Scienze dell'Atmosfera e del Clima Institute of Atmospheric Sciences and Climate (CNR)

LAMPOS Lotta alle Alluvioni Mediante Prodotti di Osservazione Satellitare

LEO Low Earth Orbit

LINET Lightning Detection Network

MHS Microwave Humidity Sounder

MLP Multi-Layer Perceptron

MODIS Moderate Resolution Imaging Spectroradiometer

MSG Meteosat Second Generation

MVIRI Meteosat Visible and InfraRed Imager

MW Microwave

NBIAS Normalized Bias

NIR Near infrared

NOAA National Oceanic and Atmospheric Administration (USA)

NRLT Naval Research Laboratory Technique

NWP Numerical Weather Prediction

OOM Object Oriented Model

OPERA Operational Programme for the Exchange of weather RAdar information

PALSAR Phased Array type L-band Synthetic Aperture Radar

PDF Probability Density Function

PERSIANN Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks

PERSIANN-CCS PERSIANN Cloud Classification System

PERSIANN-GT PERSIANN GOES-IR-TRMM TMI

PMW Passive Microwaves

POD Probability Of Detection

PoP Probability of Precipitation

PROSA Prodotti di Osservazione Satellitare per Allerta Meteorologica

QPE Quantitative Precipitation Estimation

QPF Quantitative Precipitation Forecast

RADAR RAdio Detection and Ranging

RMSE Root Mean Square Error

RR Rain-rate

SAF Satellite Application Facility

SAFNWC Satellite Application Facility to support Nowcasting

SAR Syntetic Aperture Radar

SEVIRI Spinnig Enhanced Visible and Infrared Imager

SICI Sistema Informativo Catastrofi Idrogeologiche (Information System on Hydrogeological Catastrophes)

SPOT Systme Pour LObservation de la Terre

SSM/I Special Sensor Microwave/Imager

SSMIS Special Sensor Microwave Imager/Sounder

SSP Sub Satellite Point

STEPS Short Term Ensemble Prediction System

TAMORA TAMSAT, Met Office Rainfall for Africa

TB Brightness Temperature

TAMSAT Tropical Applications of Meteorology using SATellite data and ground-based observations

TBV Brightness Temperature (vertical polarization)

TBH Brightness Temperature (horizontal polarization)

TRMM Tropical Rainfall Measuring Mission

TS Threat Score

UNISDR United Nations International Strategy for Disaster Reduction

UTC Coordinated Universal Time

UW-NMS University of Wisconsin - Nonhydrostatic Modeling System.

VIS Visible

WGS84 World Geodetic System 84

WMO World Meteorological Organization

WV Water Vapor

Introduction

High-impact meteorological events have in the last decade received increasing interest for the relevant negative consequences on both social and economical matters. Considerable efforts have been made to mitigate the effects of floods and landslides on human life and economy, on one hand trying to mitigate ground effects with progresses in civil engineering, environmental protection and in the strategies to optimize the reaction of the population to the predicted hazard, and on the other hand improving the monitoring and forecasting of meteorological forcing.
Many projects addressing different aspects of the risk mitigation strategy have been financed from Governmental Institutions, focusing on meteorological, hydrological, management and social aspects. Among others, one of the last project is PROSA (Prodotti di Osservazione Satellitare per l'Allerta Meteorologica - Satellite products for meteorological alert), funded by the Italian Space Agency (ASI) as a pilot study designed to provide institutional end users, such as the Italian Civil Protection Agency, with advanced nowcasting tools. With the term nowcasting are indicated a class of algorithms and techniques that aim to provide the detailed description of the current weather and the state of the atmosphere at a local scale and the extrapolation of the current state in the forthcoming hours. PROSA represents the Italian attempt to solve the meteorological side of the hazard mitigation scheme and it is oriented to design, develop, test and demonstrate a prototype system dedicated to the innovative dynamic characterization of meteorological parameters at the ground by means of satellite data.
This Thesis work arises from PROSA and finds its development as implementation and optimization of a satellite precipitation estimation algorithm. Started from a similar algorithm implemented over British Isles, the algorithm uses Artificial Neural Networks to correlate multi-sensors satellite data, in the Visible, Infrared and Microwave bands, and precipitation rate at ground.

The whole work is divided in three main steps, following the PROSA structure: the first version of the algorithm gives a qualitative precipitation estimate divided in rain and no-rain classes, with seasonal and day-time characterization of the precipitation maps. The second version gives a quantitative estimate, dividing the rain-rate in five precipitation classes. Finally, the last version provides precipitation maps with quantitative values expressed in $\mathrm{mm\,h^{-1}}$.
The study carried out and the results obtained are reported here in four chapters, followed by a final concluding section.

Chapter 1 gives an overview of the meteorological hazard problem with a quantification of human and economic losses and the possible solutions in order to mitigate the risk. In this chapter many projects addressed to mitigation strategies are described and in particular the Italian PROSA Project with all its aspects and algorithms developed.

Chapter 2 illustrates precipitation retrieval instruments and satellite sensors, giving a description of the spectral bands used for this work and of the physical principles on which the use of satellite measured reflectance and radiance to infer precipitation is possible.

Chapter 3 describes the theoretical basis of Artificial Neural Networks, mentioning some examples of the use of Neural Network for precipitation estimates.

Chapter 4 gives a detailed description of the algorithms developed in this work. Main characteristics of the used technique are explained and major problems are illustrated. Finally, validation results are reported and the performances of the technique are discussed, with insights on possible future work.

Chapter 1

Nowcasting and flood risk mitigation

Over recent years we are experiencing destructive precipitation events, many of which generate floods with consequent relevant effects at the ground, i.e. a great number of casualties and economic losses worldwide. This is due not only to an increasing of high-impact weather events, but also to the vulnerability created by various human activities. Considerable efforts have been made to mitigate the effects of floods on human life and economy, on one hand trying to mitigate ground effects, and on the other hand improving the monitoring and forecasting meteorological aspects.

With the term Nowcasting, after Scofield (1977) [1], is intended the detailed description of the current weather and the state of the atmosphere at a local scale and the extrapolation of the current state in the forthcoming hours. The ability to make frequent and detailed observations of the atmosphere is fundamental to any attempt to nowcast the evolution of mesoscale processes, such as, for example, high-impact weather events.

This chapter gives an overview of the research efforts undertaken in the last years to mitigate the meteorological risk, and presents a review of different nowcasting approaches, illustrating some example of successful operational systems. Finally, a description of the system designed under the PROSA project is presented.

1.1 High-impact meteorological events

Heavy precipitation is the most dangerous meteorological hazard affecting the European and Mediterranean countries, followed by windstorms and hail. Generally speaking, the impact of a meteorological event on the society mainly depends on four different factors: the severity of the meteorological phenomena, the geomorphological characteristics of the ground, the population density and assets value and the readiness of the ground structures (and population) to react to the hazard [2]. Following the United Nations International Strategy for Disaster Reduction terminology (UNISDR, 2009) the term *hazard* can be used to mention the first two factors, while the term *vulnerability* refers to the other two. Thus, the actual damage is determined by the complex interactions of the four basic factors: in order to mitigate the effects of high-impact meteorological events, a strategic action on any of them is strongly required. While improving numerical weather predictions, with an emphasis on extreme events, is recognized to be of primary importance, an effective losses reduction can only be achieved if the improvement in forecast is accompanied by progresses in civil engineering and in the strategies to optimize the reaction of the population to the predicted hazard. Quite often, however, meteorology could be regarded as the dominant factor in case of very high rain intensities, especially if produced by rapidly growing cloud systems (such as thunderstorms). Convective cloud systems at all scales, from isolated, single-cell thunderstorm to large storm cluster, are often responsible for high rain-rates and their initiation and development are difficult to forecast with conventional numerical weather prediction tools. For these cases, a forecasting approach based on detailed observation and simple evolution models, such as nowcasting, can be considered a valuable operational choice for a more effective prevention strategy [3][4].

1.1.1 Global, continental and local relevance of the problem

The collection of detailed and complete data about floods and hydro-meteorological hazards at regional to continental scales is made difficult because of the need of information of different nature, from a quantitative description of the meteorological settings to the estimates of human and economic losses.
Local authorities and hydrological institutes often collect flood data at basin, regional or national scale: a valuable option to create a flood episodes database is the

search of newspaper records. As an example, the Swiss Federal Institute for Forest, Snow and Landscape Research has been systematically collecting informations on flood and mass movement damage in a database since 1972, estimated direct financial damage as well as fatalities and injured people have been documented using press articles as the main source of information [5]. This strategy allowed the AVI (Aree Vulnerate Italiane - Italian Areas affected by mass-movements and floods) and SICI (Sistema Informativo Catastrofi Idrogeologiche - Information System on Hydrogeological Catastrophes) Italian Projects [6] to create a database on flood and landslide in Italy from 1918 onwards, and similar efforts are carried on for Catalonia (Spain) to record local flood episodes since 1982 [7]. Moreover, the press, and more generally mass-media, can play a relevant role in increase the population readiness to the hazard and significantly influence how the population and the government view, perceive, and respond to hazards and disasters [8].

A further source of information for flood episodes are records of insurance companies, that usually collects data for internal purposes. Porcú et al. (2003) [9] used at European scale the data of the reinsurance company Munich RE to create a short term climatology of flood episodes: the Munich RE archive was searched for relevant damages and meteorological archives (of both conventional and satellite data) to complete the classification of flood episodes. At smaller scale, Llasat et al. (2008) [8] used the record of the Consorcio de Compensacion de Seguros, a Spanish insurance company, to evaluate damages produced by floods in Catalonia and to evaluate the trend.

The Munich RE Natural Catastrophes archive (MRNatCat) issues periodical reports with some statistics about natural catastrophes occurrence and distribution. In fig.1.1 are shown the number of natural catastrophes recorded from 1980 to 2010 in the world, distributed in four categories: geophysical (earthquakes and volcano eruptions) in red, meteorological (storms, wind, hail) in green, hydrological (floods and landslides) in blue and climatological (heat waves, droughts, wildfires) in orange.

Meteorological and hydrological events represent the large majority of the catastrophic events in the considered decades and, while earthquakes and climatological catastrophes account for the largest number of human losses (75% of total casualties), meteorological and hydrological events causes the 62% of the total estimated losses (2500 billions of US$) [10].

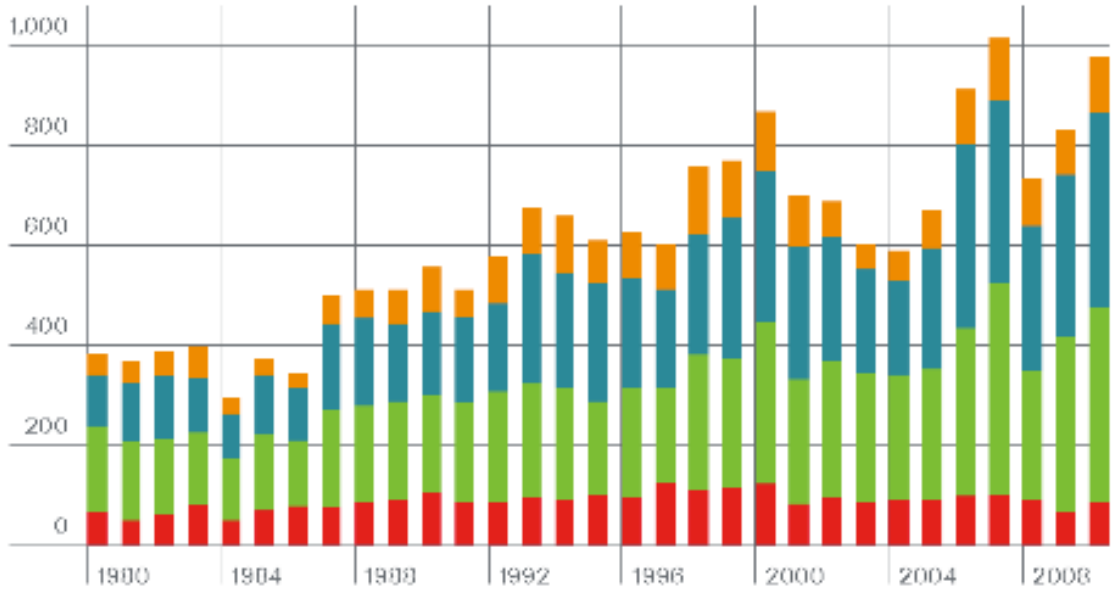


Figure 1.1: Number of significant natural catastrophes over the world from 1980 to 2010. Red bars indicate *geophysical events*, blue bars indicate *hydrological events*, green bars indicate *meteorological events*, and orange bars indicate *climatological events*. See text for definitions.

A deeper analysis of the he MRCatNat archive shows that hydro-meteorological hazards in Europe are characterized by elevated economic losses and limited number of fatalities if compared with events occurred in Asia or Southern and Central America. In 2009-2010 considering the most catastrophic hydro-meteorological events, in Europe were estimated economic losses for 16800 millions of US$ and were recorded 109 fatalities. For the same period, in Asia the economic losses were less than the double as in Europe (27875 millions of US$), but the number of fatalities was as high as 5902. This indicates that in Europe the hydro-meteorological catastrophes prevention and mitigation strategies developed in the latest decades are effective in protecting population but the areas most prone to frequent severe episodes are among the economic heartlands of the European Union, rich in assets and productive capacity, and difficult to defend effectively.
Considering the decade 1987-1996, out of the events reported in MRNatCat archive, Porcú et al (2003) [9] estimated damages for about 27000 millions of US$, and 974

fatalities in Europe due to hydro-meteorological events. From a rough comparison with 2009-2010 figures reported above, it seems that the overall number of fatalities has decreased for a factor of 2, while the economic losses has increased of a factor of 3. The spatial distribution of catastrophic events in the 1987-1996 decade reported in fig.1.2, points out that a number of European areas are particularly prone to hydro-meteorological catastrophes: Mediterranean coasts of Spain, France and Italy, Alpine region, Po Valley in Italy and Central Germany collect a large part of the total number of events occurred in Europe.

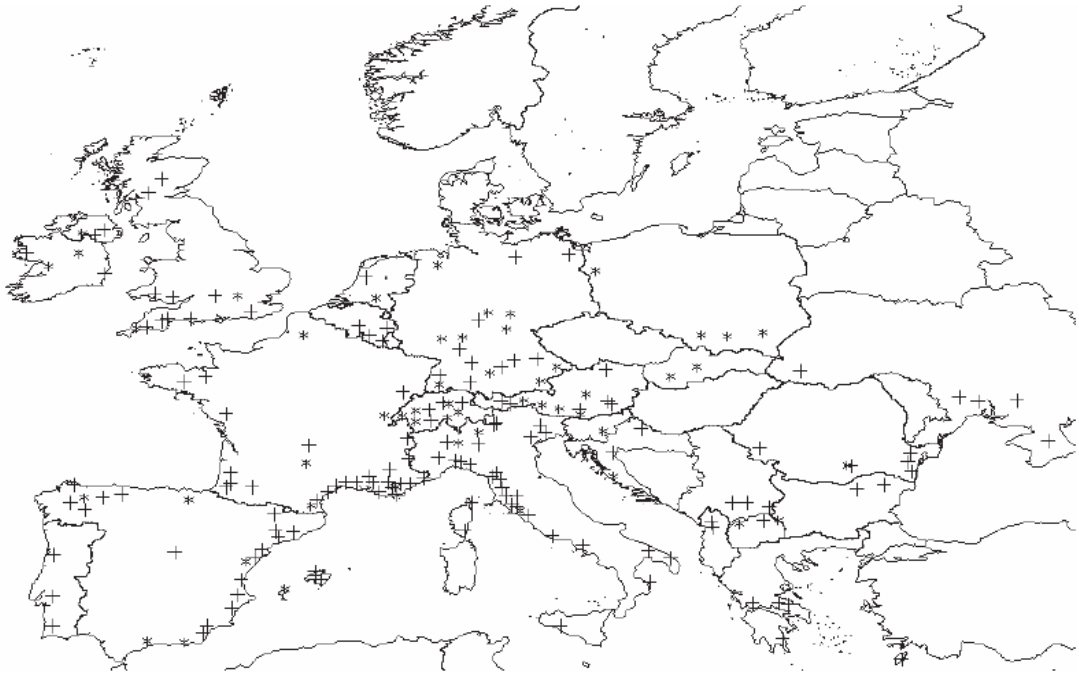


Figure 1.2: Spatial distribution events with recorded damages in Europe in the decade 1987-1996. Crosses indicate events with human and economic losses reported; asterisks indicate only economic losses [9].

Italy is one of the European regions most exposed to hydro-meteorological hazards. From the meteorological point of view, the surrounding sea provide a source of moist and unstable air, that in warm month is likely to form convective clouds, while autumn is dominated by cyclonic systems with frontal development, that usually lead to long lasting precipitation [11]. On the other side, the peculiar geomorphology of the Italian Peninsula is favorable to the triggering and enhancement of

precipitation episodes, given the presence of steep reliefs close to the coasts (such as in Liguria, Toscana, Campania and Calabria), the Alps [12][13][14][15] and the general short hydrological response time of the drainage basins. Moreover, the risk is also strongly influenced by human activities: increasing population density, illegal building practices, continuous deforestation, use of farming methods with little respect for the environment and failure to keep the hill and mountain sides and water courses in safe condition, have certainly increased the vulnerability and evidenced the fragility of the Italian territory to a greater degree.
A basic study to understand the relevance of the problem in Italy resulted in the implementation of AVI archive, that collected the Italian areas affected by mass-movements and floods since 1918 [16], and the follow-on project SICI [6]. The AVI Project, based on newspaper records, collected more than 15000 landslide and flood episodes in Italy until 1994, and the SICI continued the archive, adding also more detailed informations coming from other datasets. The SICI archive is online at the website `sici.irpi.cnr.it` and can be freely accessed to compile plots and maps. As an example, the occurrence maps for Italy for floods and landslides are reported from Guzzetti and Tonelli (2004) [6] in fig. 1.3a and 1.3b, respectively.

These maps confirm the presence of the already mentioned high-risk areas: coastal areas, Central Italy Apennines, different sectors of the Alpine chain, Po Valley, Calabria. The "patchy" pattern of flood-prone areas in Italy has been more deeply addressed by Salvati et al. (2010) [17] evaluating at the regional scale the societal risk, defined as the relationship between frequency and the number of people suffering from a specified level of harm in a given population from the realization of specified hazards. The geographical distribution of sites where landslides and floods have caused direct damage to the population shows that harmful landslide and flood events are not distributed homogeneously in Italy. To rank the 20 Italian Regions on their societal landslide and flood risk levels, the frequency of the harmful events versus the intensity of the events are computed, revealing that societal landslide risk in Italy is highest in Trentino-Alto Adige and Campania Regions and lowest in the Emilia-Romagna Region, and that societal flood risk is largest in the Piemonte and in Sicilia Regions, and lowest in the Umbria and Basilicata Regions.
Given the impact of meteo-hydrological hazards on the European Countries, The European Union funded in the last decades a number of research and application Projects and Programmes aiming to an improvement of the risk mitigation strategies

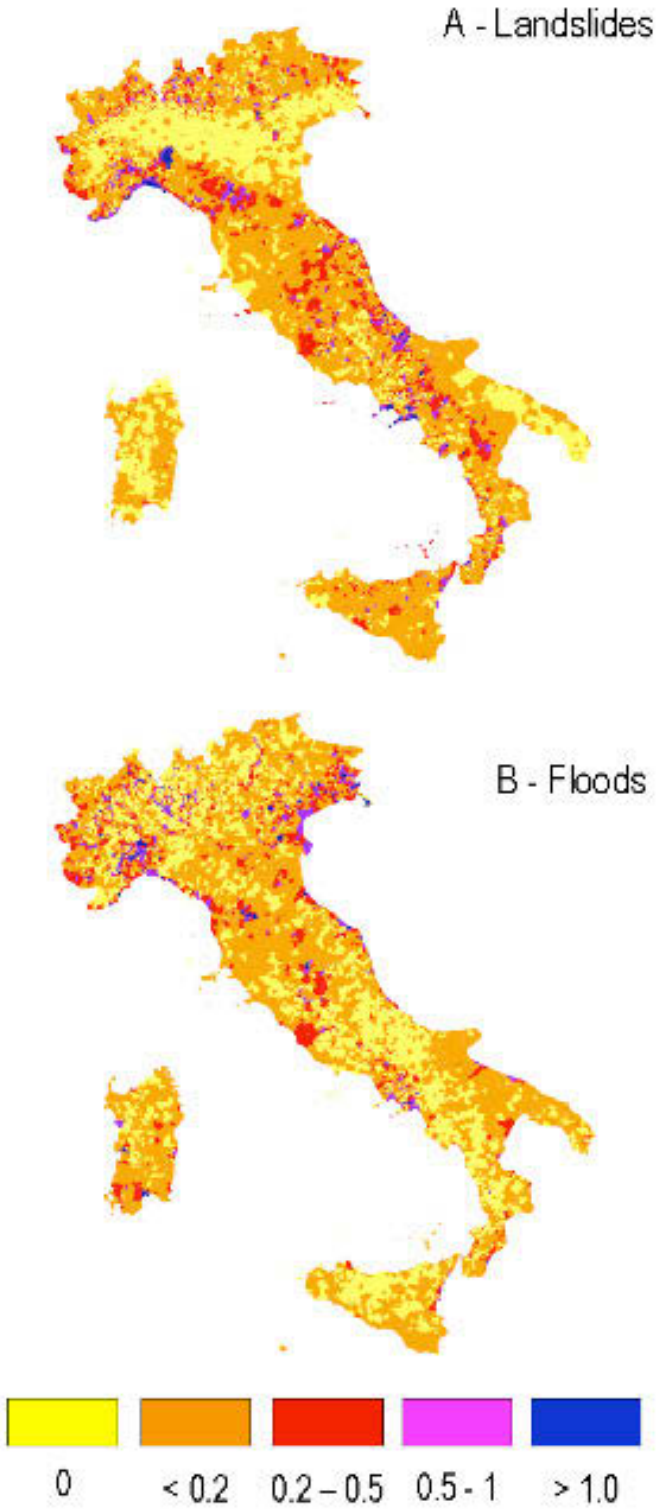


Figure 1.3: Density of landslide (A) and flood (B) events sampled on a 10×10 km^2 resolution.

in the continent. In the next section an overview of these efforts is presented.

1.1.2 Research activities in Europe

European research focused on flood and hydro-meteorological risk since early '80s, financing under Framework Programmes a number of Projects addressing different aspects of the risk mitigation strategy. In tab.1.1, a list of EU-funded Projects is reported from the newest to the oldest, attesting these activities: project full name, acronym, total project costs and end year are reported. In the third column is also shown the main focus of the Project in terms of Meteorological (Me) or hydrological (H) aspects, management (Ma) and social (S) issues. Projects with focus on Me and H developed or optimized meteorological and/or hydrological models, respectively, for flood simulation, and/or worked on data analysis for flood study and monitoring. Project labeled with Ma deals with the study of optimal risk reduction strategies, risk management, and often the hydro-meteorological risk is treated as part of a more general risk ensemble, including earthquake, fires, volcanism, drought and others. Projects labeled with S focus on the social aspects, including economical effects and populations reactions and post event organization by civil protection agencies.
Other type of projects, dealing with aspects such as the broad impact of floods on ecosystems or on urban structures are not included in this list. After a total cost of more than 56 millions of euro, several achievements have been reached and some example, relevant for this thesis work, will be discussed below.

Table 1.1: EU-founded projects of risk mitigation.

Full project name	**Acronym**	**Focus**	**Cost M€**	**End**
Organizing disaster	OD	S	1.18	2014
Improving preparedness and risk management for flash floods and debris flow events	IMPRINTS	Ma, S	3.28	2012

Table 1.1: continued on next page ...

Table 1.1: ... continued from previous page

Full project name	**Acronym**	**Focus**	**Cost M€**	**End**
Integrated advanced distributed system for hydro-meteorological monitoring and forecasting using low-cost high-performance X-band mini-radar and cellular network infrastructures	HYDRORAD	Me, H	1.44	2011
Observations, Analysis and Modeling of Lightning Activity in Thunderstorms, for use in Short Term Forecasting of Flash Floods	FLASH	Me	1.64	2010
Hydrometeorological data resources and technologies for effective flash flood forecasting	HYDRATE	H, Ma	3.42	2010
Integrated flood risk analysis and management methodologies	FLOODSITE	Ma, S	14.02	2009
Flood and drought risk assessment tools using modeling and Earth observation for early warning systems in the Nile basin	FATE-EWS-NILE	H, Ma, S	1.7	2008
Earthquakes, tsunamis and landslides in the Corinth rift, Greece A multidisciplinary approach for measuring, modeling, and predicting their triggering mode and their effects	3HAZ-CORINTH	Ma	2.01	2007
Mediterranean grid of multi-risk data and models	MEDIGRID	Ma	1.36	2006
Achieving technological innovation in flood forecasting	ACTIF	Me, H, Ma	0.27	2006

Table 1.1: continued on next page ...

Table 1.1: ... continued from previous page

Full project name	**Acronym**	**Focus**	**Cost M€**	**End**
Real-time flood decision support system integrating hydrological, meteorological and radar technologies	FLOODRELIEF	Me, H, Ma	2.5	2006
Near real-time flood forecasting, warning and management system based on satellite radar images, hydrological and hydraulic models and in-situ data	FLOODMAN	Me, H	3.35	2006
Towards natural flood reduction strategies	ECOFLOOD	H	0.35	2005
Multi-sensor precipitation measurements integration, calibration and flood forecasting	MUSIC	Me, H	1.9	2004
Investigation of extreme flood processes and uncertainty	IMPACT	H	2.6	2004
Decision support system for Risk Assessment and Management of FLOODs	RAMFLOOD	H, Ma	1.8	2004
Earth Observation Linking SMES To face real time natural disaster management	EOLES	S, Ma	1.8	2004
Mitigation of climate induced natural hazards	MITCH	H, Ma	0.198	2003
European satellite rainfall analysis and monitoring at the geostationary scale	EURAINSAT	Me	2.43	2003
An European flood forecasting system	EFFS	Me, H	3.05	2003

Table 1.1: continued on next page ...

Table 1.1: ... continued from previous page

Full project name	**Acronym**	**Focus**	**Cost M€**	**End**
Systematic, palaeoflood and historical data for the improvement of flood risk estimation	SPHERE	H	1.52	2003
Cluster Initiative for Flood and Fire emergencies	CLIFF	Ma	0.15	2002
Data Fusion for Flood Analysis and Decision Support	ANFAS	Ma	3.8	2002
Runoff and atmospheric processes for flood hazard forecasting and control	RAPHAEL	Me, H		2000
Satellite and combined satellite-radar techniques in meteorological forecasting for flood events	MEFFE	Me		1999
River Basin Modeling, Management and Flood Mitigation	RIBAMOD	H		1999
The development of active on-Line hydrological and meteorological models to minimize impact of flooding	HYDROMET	H, Me, Ma		1999
Heavy precipitation in the alpine area	HERA	Me		1999
Storms, Floods and Radar Hydrology		H, Me		1995
Flood hazard control by multisensors storm tracking in Mediterranean areas		Me		1994
A comprehensive forecasting system for flood risk mitigation and control	APHORISM	H		1994

Table 1.1: continued on next page ...

Table 1.1: ... continued from previous page

Full project name	**Acronym**	**Focus**	**Cost M€**	**End**
Flood hazard assessment		Me, H		1994
Weather radar and storm and flood hazard		Me, H		1992
European Flood Study				1983

Table 1.1: ... concluded from previous page

A first task of many Project was to make an inventory of the past flood events over different regions of interest, to increase the knowledge of the event itself and to assess its relevance on the local environment and human activities. As an example, one of the outcome of the project MEFFE, funded under FP4 [18], was to collect flood events over Europe to propose a classification scheme for a European-level flood database, in which severe flood episodes were related to meteorological conditions and mesoscale settings [9]. It was clear that extratropical cyclones are the most common meteorological phenomena related to flood episodes and any attempt to improve meteorological forecast of floods should include better knowledge of extra-tropical cyclones and, thus, better modeling capabilities. These tasks have been pursued by a number of FP funded projects listed in tab.1.1 (HYDROMET, HERA, among others) and also projects funded by other organizations, such as The Observing System Research and Predictability Experiment (THORPEX) [19], funded by WMO.

A number of Project aimed to provide better estimation (or measure) of the precipitation intensity at different scales, especially using remote sensor such as weather radar and radiometers on earth orbiting platforms. Quantitative Precipitation Estimates (QPE) from ground weather radar, due to the large spatial coverage, the high resolution and the availability in real time, provides reliable data for spatially detailed hydrological simulations [20][21]. However, the set up an effective ground radar network for QPE could be difficult and costly in case of complex terrain [22][23][24], and often the QPE has large uncertainties that can impact on both hydrological modeling results and hydrological model parameter estimates [25]. Radar

operational activity in Europe are harmonized under the Operational Programme for the Exchange of weather RAdar information (OPERA 3) a joint effort of the European Network of National Weather Services (EUMETNET). The main tasks of this long lasting Project (OPERA 1 has started in 1999) are: to maintain an inventory on European weather radars, their characteristics, their data, and products derived from them and to foster the exchange and harmonization of weather radar data and products throughout Europe and to maintain the Odissey data center [26][27].

European Commission coordinated cooperation on radar meteorology also by means of COST Actions: COST (European Cooperation in Science and Technology) is one of the longest-running European instruments supporting cooperation among scientists and researchers across Europe. In particular, COST Actions 72 and 73 established first cooperation protocols among member states on radar operational uses and networking [28]. More recently COST Action 717 examined and defined the requirements on European radar data for their use in Numerical Weather Prediction (NWP) and hydrological models, and, in order to increase the level of management of river flow in rural and urban catchments, brought together radar engineers, meteorologists and hydrologists. They initiated a European-wide action to homogenize procedures related to radar data [21]. This effort continues beyond the duration of the Action in the above mentioned EUMETNET‘s OPERA initiative. COST Action 731, focused more on the quality and uncertainty of meteorological observations, along with their impacts on hydro-meteorological outputs from advanced forecast systems. The Action also made a major contribution to the genesis of series of European Conferences on Radar Meteorology, which became one of the main channels of the dissemination of the results [29].

A space view of rainfall pattern and intensity can be obtained by satellite rainfall estimation techniques. Among the first efforts to the use of satellite data it has to be mentioned the FP4 Project MEFFE [18], that used combination of satellite (both microwave and visible-infrared spectra) and detailed ground radar observation to optimize the accuracy of instantaneous rainfall maps. Significant results were reached in studying the clouds and precipitation structures of cyclonic development [30][31][32][18], in proposing new nowcasting-oriented, satellite estimation techniques [33][34] and multi-sensor studies on severe convective systems [35][32][36]. Moreover, an attempt was made to introduce satellite rainfall maps as direct input

in hydrological models [37].
The FP5 Project EURAINSAT focused on the satellite precipitation retrieval techniques development and assessment, exploiting the new capabilities of multispectral satellite sensors, cloud resolving numerical models and weather radar [38]. A large part of the results are collected in the book Measuring Precipitation from Space: Eurainsat and the Future [39]. The main tasks of the Project were: 1) contribute to improving the knowledge of clouds and precipitation formation processes using meteorological satellite sensors, and 2) make available new precipitation products for weather analysis and forecasting. The first task has been accomplished by studying the impact of aerosols in determining the structure of convective clouds [40][41], and the relationship between lightning cloud activity and convection for the discrimination of convective and stratiform regimes [42]. As for the second task, several approaches have been pursued, publishing a number of new techniques based on microwave data or on combination of data from different platforms [43][32][44][45].
The European Commission funded projects related to the prevention and mitigation of hydro-meteorological hazards also under the Interreg programmes (now in the IV stage) designed to promote the cooperations among European Regions, financed by the European Fund for Region Developing. In particular, Interreg III funded the Risk Advanced Weather forecast system to Advise on Risk Events and Management (Risk-Aware) with two general objectives: the prevention of geo-hydrological natural disaster at regional, national and trans-national level forced by meteorological situation, and the design and implementation of regional, national and trans-national programmes for geo-hydrological hazards assessment and risk management. A part of the efforts under Risk-Aware aimed to increase the knowledge of precipitating systems by means of new multi-sensor techniques [46] based on polarimetric radar and multispectral satellite observation to infer internal structure of relevant precipitating cloud systems. New techniques were also introduced to improve the quantitative precipitation estimates from weather radar [47] and satellite data [48][49], and also to optimize the use of raingauges data [50]. More oriented to operational activities, Verdecchia et al. (2008)[51] introduced a flood forecasting platform that uses based on coupled mesoscale atmospheric and a newly developed distributed hydrological model with in-situ and remote sensing data integration. The focus is on small-catchment flood forecast in complex topography in Central Italy, with emphasis on the integration of numerical models and retrieval algorithms with aim to provide

an overview of an objective system for hydro-meteorological alert-map emission. A fully model-based approach has been proposed by Diomede et al. (2008)[52], based on a hydrological ensemble forecasting approach that uses multiple precipitation scenarios provided by different high-resolution NWP models, driving the same hydrological model. In this way, the uncertainty associated with the meteorological forecasts can propagate into the hydrological models and be used in warnings and decision making procedures relying upon a probabilistic approach. A strategy based on geostationary satellite, weather radar and lightning network data, has been introduced by Tafferner et al. (2008)[53] to track the development of deep convection in Central Europe for nowcasting purposes. The same objective has also been pursued by Zinner et al. (2008)[54] only relaying on Meteosat 7 SEVIRI 6.2 μm and 10.8 μm channels data and tropopause temperature obtained from ECMWF analysis.
Under Interreg III has been also funded the Project Weather Risk Reduction in the Central and Eastern Mediterranean (RISKMED) with the aim to create an Early Warning System (EWS) for the eastern Mediterranean [55]. A further Interreg III funded project was HYDROPTIMET, with the main goal to improve of the knowledge of severe events phenomena, the optimization of the meteorological and hydrological aspects for the prevention of natural hazards, the experimentation of new tools (such as numerical models) to be used operationally for the Quantitative Precipitation Forecast (QPF) and the improvement of the collaboration between the partners (exchange of data, methodologies, information). The territory affected by the activities of the project includes the following regions: the Italian side of the western part of the Alps, subject to severe events and to strong vulnerability due to the complex orography; the regions along the Apennines and some region of South Italy, where the interaction with the sea is more pronounced; the northeast part of Spain, where severe events are less frequent, but produce intense damages due to the vicinity of the sea with the mountains; the south-eastern France is also prone to heavy precipitating events, specially during fall. The results of the project after the three years of activity, can be subdivided into four main categories [56]: general description of the test cases, results of the meteorological applications, meteo-hydrological chain, and connection with the end-users.
To complete the overview of the European coordinated efforts for the mitigation of meteo-hydrological risks it has to be mentioned the set up in 2005 of the Satellite Application Facility on Support to Operational Hydrology and Water Management

(H-SAF), funded by the European Agency for the Exploitation of Meteorological Satellites (EUMETSAT) in the frame of the European network of Satellite Application Facilities. The aim of this effort (expected to end in 2015) is to provide new satellite-derived products from existing and future satellites with sufficient time and space resolution to satisfy the needs of operational hydrology. In particular, European users will benefit of the following identified products: precipitation (liquid, solid, rate, accumulated); soil moisture (at large-scale, at local-scale, at surface, in the roots region); snow parameters (detection, cover, melting conditions, water equivalent); Moreover, H-SAF performs independent validation of the usefulness of the new products for fighting against floods, landslides, avalanches, and evaluating water resources. These activity includes: downscaling/upscaling modeling from observed/predicted fields to basin level; fusion of satellite-derived measurements with data from radar and rain-gauge networks; assimilation of satellite-derived products in hydrological models; assessment of the impact of the new satellite-derived products on hydrological applications.

1.2 The research in Italy and the PROSA System

Italian researchers are very active in the European and global efforts to mitigate natural risks, participating in most of the initiatives outlined in the previous section. Nevertheless, significant activities are carried on since 80s funded and coordinated by Italian Ministries and Governmental Institutions devoted to develop and make applicable techniques on the Italian territory. First activities were gathered under the National Group for Hydro-Geological Disaster Prevention (Gruppo Nazionale per la Difesa dalle Catastrofi Idrogeologiche, GNDCI), a research network of the Italian National Research Council (CNR). The group is composed of research organizations from Universities, from State Administration, and from private and public organizations. Since 1986, the GNDCI has been the operating group of the Italian Civil Protection Department (Dipartimento della Protezione Civile, DPC) for issues related to hydro-geological disasters, with the following main tasks:

> Coordination of interdisciplinary activities to improve the scientific knowledge in the fields of civil defense against flood, landslides, aquifer deterioration and coastal erosion;

Provision of scientific and technical consulting in the field of hydro-geological risk, to the national, regional and local authorities, with particular attention to the civil protection authorities;

Management of the scientific community in support to the hydro-geological disasters mitigation activities of the Italian Civil Protection Department.

In parallel, the Italian Space Agency (ASI) started financing national earth observation programs with focus on precipitation systems. The project 'Hydrological Cycle from Satellite: Clouds and Precipitation' lasted for 5 years (1995-2000) to work on case studies of severe events over Italy using combined infrared-passive microwave techniques [57]. As a follow on project, in 2001 MeditRain (Use of newly-available advanced space-borne instruments for prototype generation of precipitation information over the Mediterranean area) has been funded by ASI for 2 years to continue on satellite estimation techniques development.
The 2003-2005 National Space Plan of ASI assumed as main mission of the Earth Observation branch the improvement of the capability of prevention natural disasters in Italy by means of satellite products. To set up the research structure and to identify potential users, the Pilot Project LAMPOS (Lotta alle Alluvioni Mediante Prodotti di Osservazione Satellitare - Flood mitigation by means of satellite observations) was funded and completed the preliminary study, suggesting the guidelines for the implementation of a National Project to provide operational satellite tools for the prevention of meteo-hydrological catastrophes.
The Project PROSA (Prodotti di Osservazione Satellitare per l'Allerta Meteorologica - Satellite products for meteorological alert) started in 2007 to: design, develop, test and demonstrate a prototype system dedicated to the innovative dynamic characterization of meteorological parameters at the ground. This project will be illustrated in detail in the next section.

1.2.1 The PROSA structure

PROSA is based on combined use of EO satellites data, "in situ" data and traditional technologies. PROSA is aimed at supporting the Italian DPC in managing the risks associated with floods and hydro-geological hazardous events. The project,

started on October 2007 is managed by Carlo Gavazzi Space SpA whereas the scientific coordination is carried out by CNR-ISAC Bologna (Prof. Franco Prodi). PROSA system had been developed in three years time; three incremental versions of PROSA had been developed, along with the gradual integration in the system of the technical innovations coming from research activities of the scientific institutes (CNR-ISAC, CNR-IFAC and CETEMPS) and universities (University of Rome "La Sapienza", University of Ferrara and Politecnico di Bari). PROSA system had been developed in collaboration with ASI and the institutional end-user, the Italian DPC. The PROSA system can be divided into three main sub-systems: 1) instantaneous and cumulate precipitation estimate and observation; 2) numerical weather predictions; 3) surface products (snow parameters and soil moisture). A scheme of the whole system is shown in fig.1.4
PROSA is a prototype system developed to be used in a demonstration mode and designed to receive near-real-time data from satellite sensors (SSM/I, SSMI/S, AMSR-E, AMSU-A, AMSU-B, MHS in the MW band, MSG in VIS-IR bands, that will be introduced in the next chapter) and LINET data for lightening. Algorithms based on MSG data can produce final products every 15 minutes, just some minutes after the availability of data. MW algorithms, instead, are usually intermediate products, used in others algorithms as calibrating data. This type of products, depending on sun-synchronous satellite, can be available just in presence of a satellite track passing on the target area. The project had been developed in three years time span with 41 products available for operational application as part of the PROSA operational chain. Many of these products are different versions of the same algorithm, released in three versions during the project.

1.2.2 PROSA Products

Before entering in the detail of Ferrara University work-packages, we want to mention the different algorithms developed for PROSA. They are 20 EPP products (Precipitation Products Elaboration), 11 EMN products (Elaboration of Numerical Modeling prediction products) and 10 EPS products (Surface Products Elaboration).

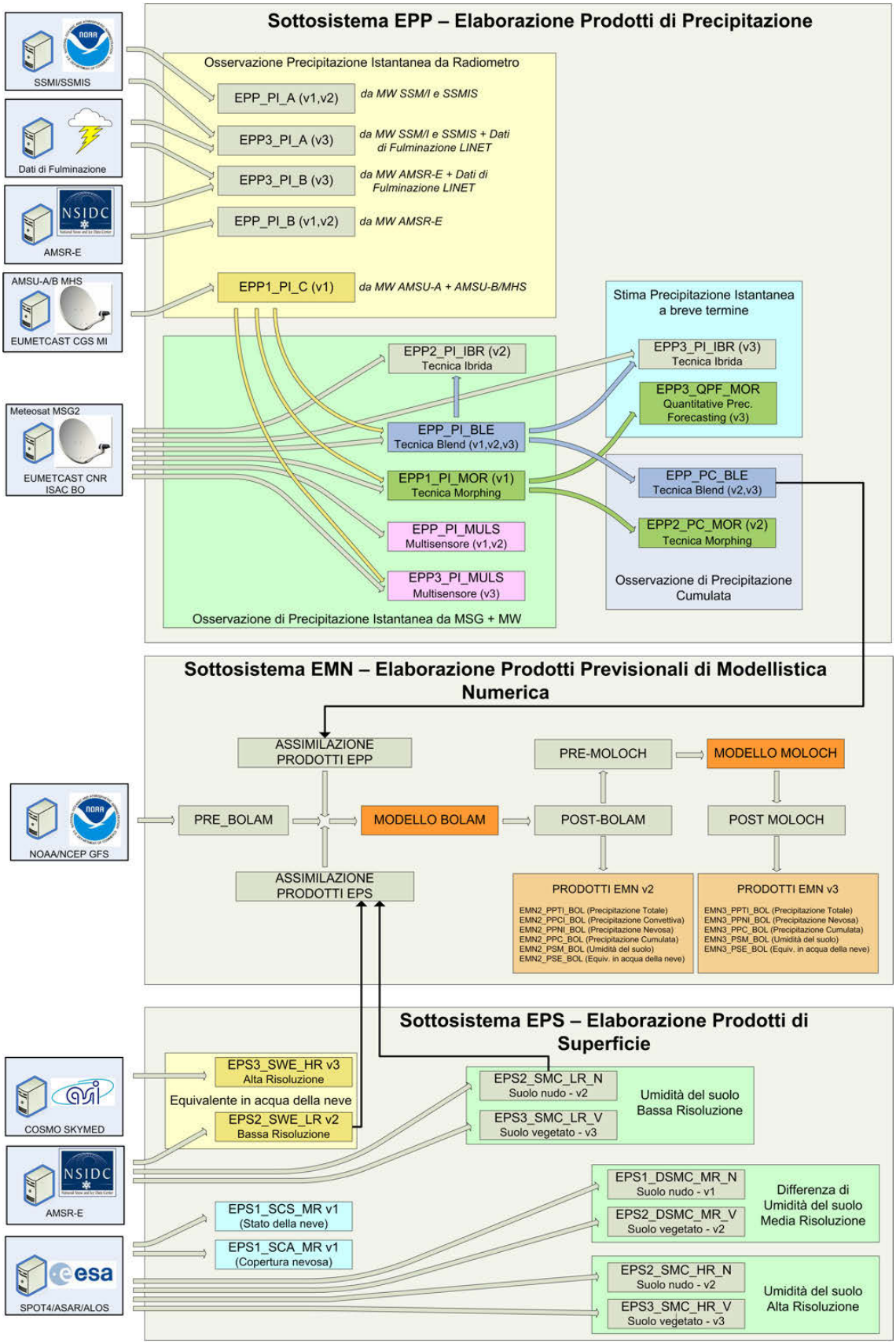


Figure 1.4: PROSA system scheme.

1.2.2.1 PROSA Precipitation Products

The PROSA structure provides several precipitation products, obtained with different techniques and satellite data in order to provide the uses a wide range of possible options to have precipitation maps [58]. The validation program undertaken under PROSA and the operational use of the different products is expected to give a trade-off of potential and drawbacks of each technique, and to drive the user to choose the more suitable product to describe at best the current weather situation. The techniques used in PROSA will be briefly described below. The EPP MULS technique, developed within this PhD thesis work will be only briefly introduced because it will be the subject of chapter **??**.

Instantaneous precipitation estimates by PMW conical scanning radiometers This technique provides ground instantaneous precipitation (in $\mathrm{mm\,h^{-1}}$) observed in the MW (Microwave) band (with liquid or solid phase flag) from SSM/I (Special Sensor Microwave/Imager) - SSMI/S (Special Sensor Microwave Imager/-Sounder) sensors [59][60][61]. The algorithm uses as input the TB of PMW channels and a Digital Terrain Model (DTM). The first part of the algorithm, that works offline, aims to solve the direct problem: a database is constructed (Cloud Dynamics and Radiation Database CDRD) where the TBs as would be measured by a satellite sensor are associated with synthetic cloud structures generated by a Cloud Resolving Model (CRM). In the CDRD each cloud microphysical vertical profile is associated the set of simulated TBs, the terrain structure, the dynamic and thermodynamic characteristics, and the lightning activity as measured by the LIghtning NETwork (LINET). The second part solves the inverse problem, computing the actual rain-rate. For each satellite Instantaneous Field of View (IFOV), the CDRD is searched to find the set of simulated TBs that are closes, in terms of Bayesian analysis, to the set of actual TBs, as measured by the sensor. The precipitation rate is assigned to the IFOV as computed from the assigned cloud profile microphysical characteristics. The CDRD includes 23 different simulations (with about 10^6 cloud profiles) carried on by the University of Wisconsin Nonhydrostatic Modeling System (UW-NMS, Tripoli, 1992), on precipitation events occurred in the Mediterranean area, focusing on heavy precipitation cloud systems. Finally, a screening procedure is adopted in a

data pre-processing phase, to separately consider coastal IFOV, where the retrieval is made difficult by the different emissivity of land and sea. The algorithm applies on data from conical scanners, such as SSM/I, SSMIS on board the Low Earth Orbit (LEO) satellites of the US Defense Meteorological Satellite Program (DMSP), and the Advanced Microwave Scanning Radiometer for EOS (AMSR-E) on board the Earth Observing System (EOS) Aqua LEO spacecraft. Specific CDRDs are needed for the different sensors, since the number and the central frequency of the channels varies, and the radiative transfer simulations have to be performed accordingly. These algorithms are used to retrieve the products labeled EPPn_PI-A (for SSM/I and SSMIS) and EPPn_PI-B (for AMSR-E), where n indicates the number of the product release in the PROSA system. Precipitation maps are delivered from 4 to 8 times/day (depending on the LEO satellite overpasses) at a ground resolution of 12×12 km^2 over Italy.

Instantaneous precipitation estimates by PMW cross-track scanning radiometers This algorithm computes instantaneous precipitation rates by means of data from the PMW cross-track scanning sensors Advanced Microwave Sounding Unit (AMSU), on LEO satellites NOAA, EOS-Aqua, MetOp, and Microwave Humidity Sounder (MHS) on MetOp and NOAA [62].
The first phase of the algorithm aims to reduce the IFOVs of the different channels (ranging from 16 to 52 km at nadir) at the same resolution, the one of the final product, i.e. 16 km at nadir; moreover, in this phase a correction for the slant path of the radiation through the atmosphere in case of larger viewing angle (the so called limb broadening effect). In the second phase, a screening is applied to the TBs, in order to eliminate contaminate IFOV, for example due to snow at the ground. In the third phase the retrieval of precipitation rate at the ground is performed, by an Artificial Neural Network (ANN) algorithm. The training of the ANN has been performed on a supervised dataset using 38 NOAA15 orbits and the corresponding precipitation data as computed by the US weather radar network NEXRAD. Input data to the ANN are for each valid IFOV: the actual sensor scanning angle, 3 TBs in the water vapor absorption band, the first 3 principal components of the TBs in the oxygen absorption band and the first 2 principal components of the TBs in the window channels. The Principal Component Analysis has been applied in parallel to the ANN training aiming to select the TBs with more information related to the

precipitation layers.
This product is labeled EPP1_PI-C in the PROSA system, and it is delivered in real time just 30 minutes after the reception of the satellite data. The has a ground resolution at nadir of 16 km, to be intended as the diameter of the circular IFOV, that become elliptical and larger close to the edge of the swath (ellipsis axes around 23 and 52 for the last IFOV in the swath). The product is available depending on the satellite orbits and availability, from 4 to 8 times a day.
Finally, it has to be mentioned that this EPP1_PI-C is the only PMW product available in near real time, and thus it is used as input for other PROSA products.

Multi sensor precipitation estimates from VIS-IR and PMW data The products of these algorithms are labeled by the PROSA system as EPPn_PI-MULS. Three version of the ANN-based algorithm are released: the first algorithm computes the probability of precipitation from geostationary satellite Visual-Near Infrared-Infrared (VIS-NIR-IR) data, the second one computes precipitation rates divided in 5 intensity classes, using VIS-NIR-IR data, the last makes use of PWM (Passive Microwave) precipitation products combined to VIS-IR-NIR data and the output is given as rain-rate in $\mathrm{mm\,h^{-1}}$.
Since the development and implementation of this algorithms is the main subject of this PhD Thesis, they will be described in details in chapter **??**.

Precipitation estimate by means of a blended technique This algorithm makes a synergistic use of satellite IR and PMW data to retrieve instantaneous precipitation rate. It is based on the Naval Research Laboratory Technique (NRLT) [45] and uses co-located SEVIRI IR TB (at 10.8 μm) and rain-rate from PMW algorithm (in this case EPP1_PI-C product). Statistical relationships between TB and rain-rate are created based on the dataset, and updated every time a PMW sensor overpasses the region of interest. The matching is performed on wide boxes of 2.5×2.5 degrees. The SEVIRI images preprocessing includes a screening to exclude certainly non precipitating pixel based on thresholds applied to different SEVIRI channels: considered channels are: 0.65 μm (during daytime to screen out thin clouds, 1.6 μm (during daytime to screen out snow at the ground), 3.9 μm (to screen out partially cloudy pixels), while 6.2 μm and 12.0 μm are used in combina-

tion to flag overshooting top of convective clouds. The products is at the resolution of SEVIR sensor, i.e. 3.5×4.5 km^2 over Italy.
The 15-minutes output of this algorithm is also used to construct cumulated precipitation products (labeled EPP3_PC-BLE) at 3, 6, 12 and 24 hours.

Precipitation estimates by hybrid techniques This technique aims to provide a forward extrapolation of the rainfall fields produced by other algorithms of the PROSA system. In particular, SEVIRI images and the EPP2_PI-BLE product are entered as input, and the algorithm provides the precipitation map for one hour later.
The approach used here is the Steady State Displacement [63]: two successive SEVIRI IR images are processed and the displacement of the cloud entities in the images is computed as motion vector. Given the short time lag between SEVIRI images (15 minutes), it is assumed that the image characteristics and cloud shapes do not change significantly: the *nowcasted* image will be the last received image, translated accordingly to the computed motion vector. The motion vector is computed by a cross-correlation algorithm between subsequent SEVIRI images.
This product is labeled and EPPn_PI-IBR in the PROSA system, has the same ground resolution of SEVIRI (around 3.5×3.5 km^2 over Italy) and it is delivered for Italy and surrounding area every 15 minutes.

Instantaneous and cumulated precipitation estimates by means of a morphing algorithm The morphing algorithm is derived from the CPC MORPHing (CMORPH) technique developed for global precipitation estimates [64] and it is tailored for Italian region for the PROSA system, and new features are also added. The technique ingests the last PMW precipitation product (in this system EPP1_PI-C product is used), and the nearly simultaneous IR SEVIRI TB image (at time t_0), and the rain-rate is re-sampled over the IR TB image, identifying in the IR image the precipitating areas (as defined by the PWM algorithm) by a pattern matching technique. At time t_0+15 min a new IR image is available, and the rain areas are searched on this image, by means of a cross correlation algorithm, and new precipitating areas are identified in the t_0+15min image by translation and deformation (morphing) of the areas at time t_0. Actual rain-rates at time t_0 are transferred to

the new areas and the estimate at time t_0+15min is produced. This procedure goes on iteratively at every 15 min step, for each new SEVIRI TB map, until a new PMW precipitation map becomes available, and the procedure starts from the beginning.

1.2.2.2 PROSA soil products

The final design of PROSA structure includes the delivery of six soil products derived from satellite observations. These products are listed below.

Low resolution soil moisture content (bare and vegetated soil) This product is derived by AMSR-E data at 6.9 GHz. The algorithm is based on an ANN procedure trained with archive data: it takes as input TB6.9V (sensitive to soil moisture), polarization indices at 10.65 and 18.7 GHz (sensitive to vegetation coverage) and TB36.5 to estimate the soil temperature. The output is given as percentage on a 30×30 km^2 grid on a daily basis.

High resolution soil moisture (bare and vegetated soil) This product is derived by SAR (PALSAR on board ALOS and ASAR on ENVISAT) data and it based on a Bayesian inversion technique. Optical data and land use maps are also used to classify the land cover (bare/vegetated). The output is as percentage of water over dry soil in the surface layer, it is delivered on a grid of 100×100 m^2, and can be obtained over regional scale areas. The revisiting time spans between 15 to 35 days, depending on SAR data availability.

Medium resolution snow water equivalent and snow cover The algorithm works on ENVISAT-ASAR data and optical data from the SPOT4 satellite. It is based on neural networks, the output is on a 500×500 m^2 grid, and the revisiting time is between 15 to 35 days.

Medium resolution snow surface condition (wet/dry) This algorithm uses ENVISAT-ASAR data and classify the areas as wet snow/dry on a grid 500×500 m^2 over small areas. The revisiting time is between 15 to 35 days.

Low resolution snow water equivalent This product is derived by AMSR-E data at 10.65, 18.7 and 36.5 GHz by means an ANN algorithm trained on archive

data and validated on field campaigns over basin-scale target areas. The output is delivered daily on a 20×20 km^2 grid, as meters of water.

High resolution snow water equivalent and snow cover The algorithm works on SAR data (ALOS-PALSAR, ENVISAT-ASAR, COSMO-SkyMed-X-SAR) and optical satellite data (SPOT4). Provides snow water equivalent in meters of water and a flag for wet snow. The output is available on a regional scale over a grid with variable (depending on the sensor used) between 100 and 200 m. The revisiting time is between 15 and 35 day.

1.2.2.3 PROSA Numerical Weather Prediction products

In addition to the remote sensing products, PROSA provides also NWP products, despite they are usually not considered in nowcasting systems, given the relatively longer leading time of the forecast (12 hours). Two series of products are operatively disseminated, based on two Limited Area Models: BOLAM and MOLOCH.
BOLAM [65][66] is an hydrostatic model and it is designed to operate within PROSA system on a grid of 15×15 km^2. The MOLOCH model [67][68] is a non-hydrostatic model operating at an horizontal resolution of 0.0225×0.0225 degrees (about 2.5 km). The operational chain takes in input low resolution analysis for BOLAM initialization (GFS from NCEP on a 0.5×0.5 degree grid) and geographical data, and starts the BOLAM processing. BOLAM outputs are delivered after post-processing as a medium resolution products, and also are used to initialize MOLOCH that runs nested on the BOLAM outputs. MOLOCH outputs are post-processed and disseminated as high resolution forecasts. Data assimilation of satellite products (EPP2_PI_BLE) is performed during the BOLAM runs by means of a nudging scheme [66].
NWP product are: high and medium resolution precipitation rate forecast in $mm\,h^{-1}$; medium resolution convective precipitation forecast; medium and high resolution snow rate forecast in $mm\,h^{-1}$; medium and high resolution cumulated precipitation forecast at 3, 6, 12 and 24 hours in mm; medium and high resolution soil moisture forecast as fraction of water mass respect to dry soil mass in m^3m^{-3}; medium and high resolution snow amount forecast in m of water equivalent. All the products are updated every 12 hours on grids covering whole Italy.

1.2.3 PROSA validation and demonstration activities

Within the PROSA project efforts have been devoted to validate the satellite products with respect to independent rainfall measurement, that are assumed as ground truth. Validation of satellite products with other datasets, of different nature, is a thorny issue not yet resolved by the scientific community at any spatial and temporal scale ([69], [70], [71], among others). From one side, the user community needs a detailed knowledge of the error structure of the products to better exploit their capability and get aware of their limitations; from the other side, the algorithm developers claim that validation of satellite products by matching with rainfall data measured by instruments operating with different physical principles. The latter is particularly true when we validate satellite instantaneous, areal integrated rain-rate with time-integrated, point-like rain-gauges measurements.

1.3 Nowcasting systems

Before to describe in a detailed way the algorithm developed for this Thesis, it is useful to introduce few examples of current nowcasting systems developed or in operational use, with a particular attention to precipitation retrievals. It is important to stress that nowcasting systems have the relevant role of forecasting meteorological events in the few minutes after the availability of the data: a powerful and rapid computational system is fundamental to deliver products in near-real-time. Also an user-friendly interface is very important to facilitate public users to interpret products and, in case, to quickly disseminate warnings. Nowcasting systems developers made great efforts to take into account all these characteristics and implement complete and rapid systems [3][4].

1.3.1 UK Met Office Nowcasting Systems

1.3.1.1 The Nimrod System

Nimrod is the first operationally used system of the UK Met Office. Developed during the early 1990s, it integrated nowcasting techniques with NWP model products to provide forecasts over the UK and surrounding waters up to six hours ahead. With this system it was possible to analyze and forecast the main weather variables,

such as precipitation, clouds and visibility. The visibility and precipitation products were generated on a 5 km grid, while cloud and most of the derived products were produced on a 15 km grid. Cloud and visibility analyses and forecasts were updated hourly, with forecasts generated half-hourly to six hours ahead. The precipitation component, however, used a cycle of 15 minutes with forecasts also generated every 15 minutes. The NWP component used by Nimrod was the Mesoscale Model version of the Met Office's Unified Model and its input was updated every 6 hours at present, the new data becoming available about 3 hours after data time.
The system was divided into five major components: observation processing, NWP, data blending, merged forecasts and product generation.
The main source of observation were satellite imagery, radar imagery and surface synoptic reports. Upper air information was obtained indirectly through the NWP assimilation procedure. Satellite imagery was received from both geostationary (Meteosat) and polar orbiting (NOAA) platforms. Radar imagery were pre-processed at the radar site to calibrate the observed reflectivity, to remove fixed clutter and to re-map to a common 5km grid.
The NWP used a version of Unified Model with a grid spacing of about 17 km for short range forecasting. Forecasts to at least 18 hours ahead were updated four times a day. For data blending two basics algorithms were used. The visibility and precipitation rate analyses used a two-dimensional variational scheme based on a standard descent algorithm, while the cloud analysis used a two-dimensional recursive filter algorithm applied to each level.
The forecast procedure consisted of two main steps: computation of an extrapolation from recent values or trends and optimal merging of this with other independent forecast estimates [72].

1.3.1.2 The GANDOLF System

A disadvantage of Nimrod's precipitation forecasting was represented by its limited ability to generate new areas of precipitation. In the short-range prediction of frontal precipitation, the modeling of atmospheric disturbances with time scale of many hours or days is not required, and it is more effective to extrapolate the current weather and its recent trends in such a way to preserve the most significant features. In the nowcasting of showers associated with outbreaks of air mass convection this

sort of approach is much less successful. With this assumption, the Environment Agency requested to develop an operative system to make the Nimrod system more complete.
GANDOLF (Generating Advanced Nowcasts for Deployment in Operational Land-based Flood forecasts) system had been designed to address this problem. Developed in the mid-1990s, this system is able to forecast the development of 'daughter cells' associated with existing convective cells.
System data inputs were received from various operational and database computer systems at Met.Office HQ. These included near-real time multi-site radar network data with a 5 km spatial resolution and 15 minutes time step and Satellite data from geostationary satellite (Meteosat) sampled every 30 minutes with a spatial resolution of 5 km in the IR and 2.5 km in the VIS band at the sub satellite point. Mesoscale Model data from Nimrod system were also used for forecast with a 15 km spatial resolution and updated four times daily.
The precipitation model implemented in GANDOLF was the Object-Oriented conceptual Model of convection (OOM): it could benefit of 2 km spatial resolution and 10 minutes temporal step. This means that a GANDOLF product could be received within 10 minutes of Data Time. The OOM incorporated a conceptual model of the life cycle of a shower cloud. Five cell growth stages were recognized, among which the developing, mature (young and fully mature), early dissipating and dissipating phases in the life of a convective cloud. Cell stages were distinguished on the basis of vertical rain rate profiles, derived from multi-beam, single-site radar data. The development potential of a cell was determined by comparing the current development stage with that classified 10 minutes and 20 minutes earlier [73].
The OOM was specifically designed to run during periods of non-frontal air mass convection. On occasions when precipitation was frontal in origin, or was associated with convection embedded in a frontal feature, the life cycle model was not applicable. The accurate identification of meteorological conditions suitable for running the OOM was therefore critical to the successful operation of the GANDOLF system. The main method used operationally for identifying the types of rainfall pixels (convective or non-convective) used a neural network cloud classifier. In the GANDOLF system was then implemented a Neural Network Cloud Classifier capable of distinguish four cloud classes. [74].

1.3.1.3 STEPS System

Since 2008, the UK Post-Processing system includes the Short Term Ensemble Prediction System (STEPS). The goal of this algorithm, developed in collaboration with the Australian Bureau of Meteorology, is to generate products to replace those previously provided by the Nimrod and Gandolf precipitation nowcasting systems. STEPS is an ensemble-based probabilistic precipitation forecasting scheme that blends an extrapolation nowcast with a downscaled NWP forecast. The use of ensembles allows the scheme to be used for applications that require forecasts of the probability density function of areal and temporal averages of precipitation.
STEPS derives forecast data from two separate sources, radar-based rain analyses and NWP model forecasts. These data are merged with uncertainty estimates, in order to give the right weight to every contribution (extrapolation forecast, NWP model forecast and noise). The output from NWP forecast model is downscaled so that the small scales not represented accurately by the model are injected into the forecast using stochastic noise. As the forecast progresses, the extrapolation component will become less skillful, especially at the smallest scales. The loss of skill is modeled to allow precipitation features to be smoothed out trough a decreasing in power of the extrapolation forecast cascade. The power lost at the small scales can be replaced by noise. This allows the scheme to better represent the distribution of precipitation rate at spatial scales finer than those adequately resolved by operational NWP. Finally, in order to produce nowcasts with a useful range of at least 6 hours, it is necessary to merge these nowcasts with NWP forecasts. This merging will allow the resultant precipitation nowcasts to reflect the influence of the large-scale dynamical evolution of the atmosphere on the precipitation field [75].

1.3.2 The TAMORA algorithm

A multi-spectral rainfall estimation algorithm has been developed for the Sahel region of West Africa with the purpose of producing accumulated rainfall estimates for drought monitoring and food security. Several Satellite-based techniques have been designed for rainfall estimation over Africa, including TAMSAT [76], an algorithm using only a single IR channel from the SEVIRI instrument on board MSG. After that work, much work has been done on utilizing multiple channel data from SEVIRI

and other similar instruments for cloud classification and rainfall identification.
The starting point is the UK Met Office SatPrecip algorithm which uses a real-time radar calibration to produce rainfall estimates over Europe and the North Atlantic. This algorithm, reconfigured and recalibrated to take account of West-African conditions, is named TAMORA (TAMSAT, Met Office Rainfall for Africa) algorithm.
The most useful combinations of SEVIRI channels for rainfall identification used in both UK Met Office SatPrecip and TAMORA algorithms are the 0.8, 1.6, 3.9 and 10.8 μm channels during daytime and the 3.9, 10.8 and 12.0 μm channels at night. The data from each SEVIRI channel are binned into a number of discrete classes and combining these multi-channel binned data into a contingency table, a number of satellite pixel data classes are obtained. Taking the radar data as "truth", coincident radar and satellite pixels are compared and a probability of rainfall for each satellite data class is computed. This process is performed for each rain-rate threshold used by the algorithm, and in this case "probability of rainfall" is the probability of a rain-rate greater than the rain-rate threshold. The probability threshold is chosen for each rain-rate by the "minimum percent" method, whereby the difference between the number of observed and diagnosed rainy pixels is minimized by an iterative procedure.
The differences between SatPrecip and TAMORA start with a different calibration of the two algorithms, the first from the Uk and European networks and the latter from Niamey (Niger) radar. As well as using different calibration data, several other changes are made in order to optimize TAMORA for use in Africa. The relationship between SEVIRI brightness temperature/reflectance and probability of rainfall vary greatly between European and Niamey data, so binning of radiances in the contingency tables is altered increasing bin resolution for areas where the probability of rainfall is relatively high for Niamey. An other big difference between the two algorithms is in the long timescale accumulation: SatPrecip simply estimates instantaneous rain-rate every hour and accumulates this directly to longer timescales, TAMORA uses an half-hour estimate obtained averaging over 4 scans (one every 10 minutes) of radar data and maps them to the equivalent half-hourly mean by histogram matching. This equivalent half-hourly rain-rate is used as output of TAMORA and accumulated to decadal timescale [77].

1.3.3 PERSIANN Systems

PERSIANN (Precipitation Estimation from Remotely Sensed Information using an Artificial Neural Network) are satellite-based rainfall estimation algorithms. The PERSIANN-GT (PERSIANN GOES-IR-TRMM TMI) algorithm, developed from University of Arizona, uses an Artificial Neural Network model that estimates rainfall rates using geostationary satellite imagery (GOES-8, GOES-10, GMS-5, Metsat-6, and Metsat-7) in the IR and VIS band and rainfall estimates from low-orbital satellites, including TRMM, NOAA-15, -16, -17, DMSP F13, F14, F15. An adaptive procedure is used to recursively update the network parameters when low-orbital satellite estimates are available. This system can generate a global rainfall estimation with a spatial resolution of $0.25° \times 0.25°$ and a temporal scale of 30 minutes to 6 hour accumulated rainfall [78][79][80].
The PERSIANN-CCS (PERSIANN Cloud Classification System) is the new version of the PERSIANN algorithm: it uses computer image processing and pattern recognition techniques to develop a patch-based cloud classification and rainfall estimation system based on satellite infrared images. This system enables the categorization of cloud-patch features based on cloud height, areal extent, and variability of texture estimated from satellite imagery. At the heart of PERSIANN-CCS is the variable threshold cloud segmentation algorithm. In contrast with the traditional constant threshold approach, the variable threshold enables the identification and separation of individual patches of clouds. The individual patches can then be classified based on texture, geometric properties, dynamic evolution, and cloud top height. These classifications help in assigning rainfall values to pixels within each cloud based on a specific curve describing the relationship between rain-rate and brightness temperature. Precipitation intensity and distribution of classified cloud patch is initially trained using ground radar and TRMM observations. The PERSIANN-CCS enables recursive (in space and time) data assimilation and system training, allowing for flexibility in the adjustment of the cloud-rain distribution curves as new ground or space-based radar measurements become available [81][82].

1.3.4 Satellite Application Facility to support Nowcasting (SAFNWC)

The SAFs for Support to Nowcasting and for Very Short-range Forecasting provides product extraction from the Meteosat Second Generation (MSG) data for near-real-time needs. Activities started in 1997 by a consortium of meteorological institutes, hosted by Spain, integrating Austria, France and Sweden, and it is designed to be completed in 2012 [83]. The general objective of the SAFNWC is to provide operational services (i. e. satellite products) to ensure the optimum use of meteorological satellite data in Nowcasting and Very Short Range Forecasting by targeted users. This is applicable to the MSG and the polar satellite systems over European area. To achieve this goal, the SAFNWC is responsible for the development and maintenance of appropriate SW Packages, as well as of all related tasks for user's support. SAFNWC is delivering a number of algorithms to operationally derive atmospheric parameters useful for nowcasting procedures. A first product family is labeled "cloud products" that includes four parameters: cloud mask, cloud top temperature, pressure and height, cloud type, and precipitation rate (in 5 intensity classes). The first three products are derived by means of thresholds in different channels of both LEO and GEO sensors, while the precipitation product is derived from PWM data (AMSU/MHS sensor). A GEO satellite product (labelled "precipitating clouds") deals with precipitation probability (0 to 1): a linear combination of SEVIRI radiances, calibrated with rain-gauges rain-rate measurements, is used, with also NWP products (e.g. temperature at the ground) to infer this parameter. A further GEO-scale precipitation product is called "Convective rainfall rate", and it is computed by using linear combinations of SEVIRI radiances, taking also into account the time evolution (growth and decay) of the convective system, calibrated with ground radar network rainfall maps. The "rapid development thunderstorm" is also dealing with convection: based on IR channels SEVIRI sequences, it allows the identification, monitoring and tracking of convective and rapid developing cloud systems.
The product "high resolution winds" computes Atmospheric Motion Winds, by using a cross-correlation approach between 15 min apart SEVIRI images in the two VIS channels, two WV channels and two IR window channel (10.8 and 12 μm). The product includes pressure level information (10 levels), and a quality control flagging giving some indication of its error in probabilistic terms, with auxiliary indicators

about how the product was determined.
The "air mass analysis" is currently comprised of 4 sub-products: 1) classification of air masses (polar, tropical, arctic and equatorial); 2) analysis of dark stripes in WV6.2; 3) analysis of the ridge lines of the equivalent-potential temperature field; 4) analysis of gradient zones of the equivalent-potential temperature field in cloudfree areas. All the product corresponds to well established conceptual models widely used for operational forecast and present in the literature. All these products are mainly derived by SEVIRI absorption channels and NWP products.
The "SEVIRI physical retrieval" product is based to the retrieval of vertical profiles of tropospheric temperature and humidity. The main parameter computed from IR window and absorption channels, and NWP forecast are: Total Precipitable Water, Layer precipitable Water (corresponding to water vapor in the pressure layer between surface and 850 hPa), middle layer water vapor (between 850 hPa and 500 hPa), high layer water vapor (above 500 hPa), and three stability indexes (Lifted Index, Showalter Index (SHW) and K-index).
Finally, the product labelled as "automatic satellite image interpretation", makes use of SEVIRI data, NWP forecasts, and some other SAFNWC products to help the user in recognizing features in the images that can be interpreted as conceptual models such as: cold fromt, warm front, occlusion, frontal wave, dry intrusion, MCS, comma clouds, and many others.
All documentation about SAFNWC, from where the above reported information are derived, is available at the website nwcsaf.org.

Chapter 2

Instruments for precipitation estimate

In this chapter the instruments commonly used for operational precipitation estimates are reviewed: in the first section the ground based instruments, such as rain-gauges and weather radar, are briefly introduced. In the second one the rain detection capabilities of the two sensors used in this work are addressed with more detail

2.1 Ground based instruments

2.1.1 Rain-gauges

Rain-gauge, or pluviometer, is the only instrument able to measure directly the amount of precipitation at the ground, collecting hydrometeors in a funnel and sampling them over a given cumulation time (usually spanning from a minute to a day, one hour for most of the cases). Rain-gauges network data are widely used as precipitation reference measure for most of the operational application, from the direct precipitation monitoring, for weather radar and satellite algorithms calibration and validation, for NWP model forecast verification and as input to hydrological models. Nevertheless, it is well known that the rain-gauge measure is affected by several sources of uncertainty difficult to mitigate, especially in case of tipping bucket instruments, the most common device worldwide used.

First of all, very light rain-rates (2 $\mathrm{mm\,h^{-1}}$ and less) can be incorrectly estimated due to the long time it takes the rain to fill the bucket [84]. On the other side, high rain-rates (above 50 $\mathrm{mm\,h^{-1}}$) are usually underestimated due to the loss of water during the tips of the buckets [85]. Drifting wind can also greatly reduce the size of the effective catching area, if rain does not fall vertically, resulting in a rain-rate underestimation quantitatively assessed in about 15% for an average event [86]. Further errors occur in case of solid precipitation (snow or hail), when frozen particles are collected by the funnel but not measured by the buckets, resulting in a temporal shift of the measurements since the melting (and the measure) can take place several hours (or days, depending on the environmental conditions) after the precipitation event [87][88]. This error can be mitigated by an heating system that melts the particles as soon as are collected by the funnel, but, in our case, Italian rain-gauges are not equipped with.
All these errors can be mitigated and reduced, but in general not eliminated, by a careful maintenance of the instrument and data postprocessing [89].
Other disadvantages concern the matching between rain-gauge data and satellite or radars data. First of all, the relatively small section of the funnel (about 20 cm of diameter) allow to obtain just a point-like measure of precipitation. Secondly, the hourly temporal sampling gives a measure of accumulated rainfall, missing any information about the time evolution of the meteorological event within the cumulation time.
In this study the gauges of the Italian network collected by the Italian DPC are used as *true* data for the Artificial Neural Networks training and validation. They are irregularly distributed over the surface (see fig.2.1), and have a cumulation time of one hour, and the number of operational instruments varies from 1500 to 1700. More detailed information on the rain-gauge network used in this work will be given in the next chapter.

2.1.2 Meteorological Radar

The Radar is an instrument that operates emitting radiation at a well defined wavelength and measuring the radiation fraction backscattered by a target. It is composed by a transmitter which transmits electromagnetic waves at a fixed wavelength

Figure 2.1: Italian rain-gauges network, where the 1512 rain-gauges used in this work are positioned (grey dots).

and a receiver, which receives the electromagnetic signal after its interaction with the target.

The weather radar is designed to measure the radiation backscattered by hydrometeors that, under a number of conditions, can be related to the precipitation rate it is expected from the scanned volume. Wavelengths used for meteorological radar are the ones that can interact with atmospheric objects, such as rain drops, ice aggregates and hail. At centimetric wavelengths, the radiation emitted from the radar can be absorbed or scattered from hydrometeors. The backscattered signal is the information that reaches the receiver and can be analyzed to understand intensity and motion of precipitation. The radiation is emitted and received on a narrow conical beam (see scheme in fig.2.2), usually around 1 degree width (β angle), and can be measured the signal on a maximum range between 100 and 200 km, depending on the radar power and wavelength. The parabolic antenna usually rotates, in the operational surveillance mode, around a vertical axis in order to acquire a series of conical scans, at different elevation angles (θ angle). The ratio between transmitted and received signal (after backscattering from hydrometeors) depends on a number of radar characteristics (such as wavelength, antenna pattern and gain) and the characteristics of the target, i.e. the hydrometeors.

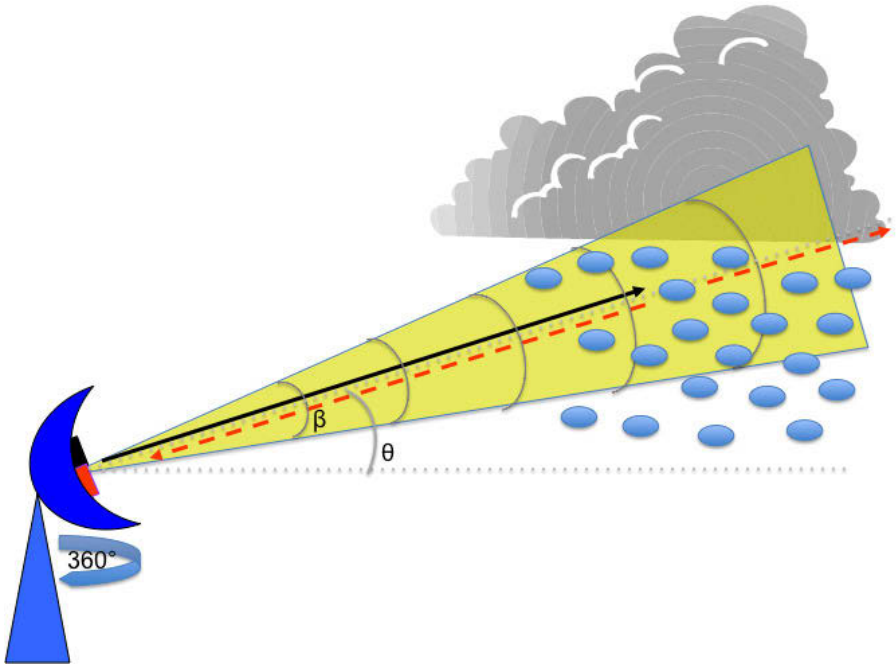


Figure 2.2: Schematic representation of radar system.

The spectral bands used by meteorological radars define the radar capabilities, and are:

L-Band: $\lambda = 0.3$ m
S-Band: $\lambda = 7$ - 15 cm
C-Band: $\lambda = 4$ - 7 cm
X-Band: $\lambda = 2$ - 4 cm
Ku-Band: $\lambda = 1$ - 2 cm
Ka-Band: $\lambda = 8.5$ mm
W-Band: $\lambda = 3$ mm

Shorter is the wavelength, smaller are the objects which interacts with the radiation. Usually, for rain drop sizes and quite long distances, the S or C-band radars are used. X-band are short distances precipitation radars, whereas Ku, Ka and W-band are mainly clouds radars.
In case of Rayleigh scattering approximation the hydrometeor characteristics entering the radar equation are expressed by the reflectivity factor, defined as:

$$Z = \int_0^\infty D^6 N(D) dD \tag{2.1}$$

where $N(D)$ is the hydrometeor size distribution and D is the hydrometeor diameter: $N(D)dD$ represents the number of hydrometeors for unit of volume with diameter in the interval $[D, D + dD]$, and Z is measured in $\mathrm{mm}^6\,\mathrm{m}^{-3}$. The actual rainfall rate is also related to the function $N(D)$ as follows:

$$R = 0.6\,\pi\,10^{-3} \int_0^\infty D^3 v(D) N(D) dD \tag{2.2}$$

where $v(D)$ is the hydrometeor terminal velocity, depending on the diameter. The exponential relationship between Z and R, for operational uses, is experimentally fitted as:

$$Z = aR^{\,b} \tag{2.3}$$

where a and b are parameters that are to be experimentally determined and depend on the radar characteristics, the wavelength and, more important, on the characteristics of the precipitation. A number of a and b sets are proposed in the literature and operational practice, and usually each radar site computes its own parameter set, depending on the rainfall type. The radar conical output are remapped onto a regular cartesian volume composed by $1\times1\times1$ km^3 elementary volumes, while the precipitation product is usually released on a horizontal, constant altitude, map. Precipitation maps can be obtained at rather high temporal frequency, varying from few minutes (in case of severe precipitation events) to half an hour.
The weather radar certainly is the principal instrument for operational precipitation monitoring and precipitation nowcasting, and it is a critical component of any flood prevention forecasting system ([20] [90], among others): it provides high resolution rainfall maps in real time, with an acceptable level of quantitative accuracy. However, two classes of error affect the radar precipitation estimation reliability. The first class groups the error due to the possible fail of assumptions made to apply the Rayleigh theory to the radar signal: hydrometeors in the unit volume analyzed are assumed to be homogeneous dielectric spheres small with respect to the wavelength, the elementary volume is homogeneously filled by random scattered hydrometeors, multiple scattering is negligible, the hydrometeors size distribution is known.

Other error family is related to the interaction of the radar beam with atmosphere and hydrometeor, and are summarized as follows, with reference to figure 2.3:

1. precipitation below the radar beam colud be underestimated;
2. evaporation and strong winds below the cloud base may results in overestimation of rainfall rate;
3. orographic enhancement effects can be underestimated;
4. melting layer signal usually results in a rainrate overestimation (bright band);
5. radar beam attenuation behind strong rainshaft;
6. anomalous propagation and secondary echoes;
7. beam blocking on fixed obstacles.

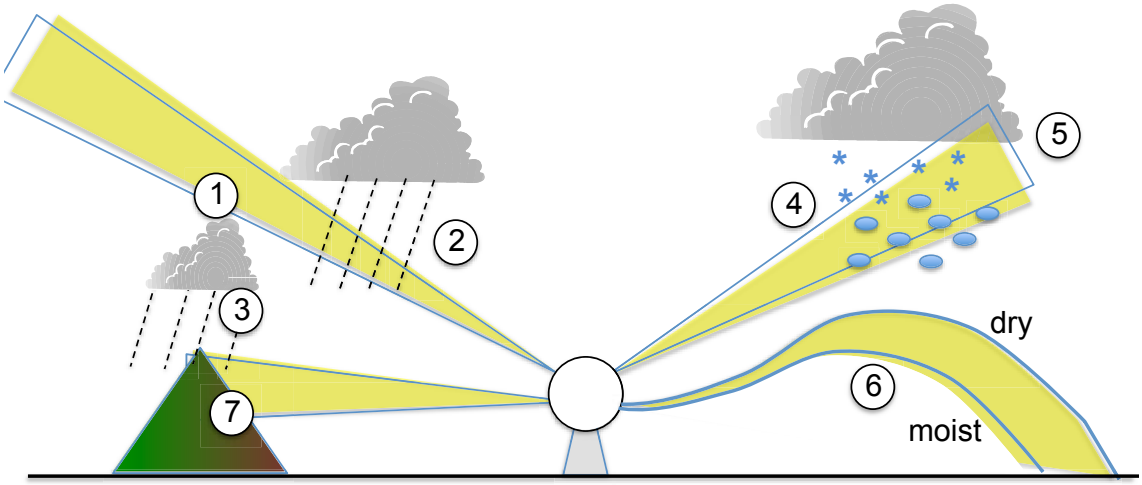


Figure 2.3: Schematic representation of radar errors due to interaction of radar beam with atmosphere and hydrometeors.

These errors can be mitigated by using specific algorithms, e.g. the bright band ([91], among others), but the more effective strategy for radar errors reduction is to create a radar network. A number of radar are distributed over a region with

partially overlapping areas, where to apply mutual correction algorithms and intercalibrate: an example of the reduction of the attenuation problem by using two partially overlapping radars is presented in Celano et al. (2006) [92].
A new class of weather radar is used mainly for research purposes but the extension to operational use is envisaged, the polarimetric radar. The polarimetric radar transmits both horizontal and vertical polarization electromagnetic wave pulses and receive the backscattered signals in co-polar and cross-polar components. The polarization depends on hydrometeors physical parameters (i.e. shape, orientation, phase etc.) therefore the polarimetric capability gives a more direct information on the nature of the scatterers. Some of the fundamental variables measured by polarimetric radars are introduced below.
The differential reflectivity (ZDR) is the ratio between the reflected horizontal and vertical power returns. It depends on the asymmetry of the shape and it is a good indicator if the particle is oblate or prolate. ZDR is positive for oblate raindrops, near zero for hail and graupel.
The linear depolarization ratio is a ratio of a vertical power return from an horizontal pulse or an horizontal power return from a vertical pulse. Depolarization of a polarized pulse is caused by asymmetric particles with their major or minor axis not aligned nor orthogonal to the electric field. Depolarization is high for melting snow and hail precipitation.
The specific differential phase shift (K_{dp}) is a comparison of the returned phase difference between the horizontal and vertical pulses. This phase difference is caused by the difference along the propagation path for horizontal and vertically polarized waves, because rain drops are oriented and their larger dimension is horizontal. The specific differential phase is a "propagation effect", it is independent by the radar calibration and it results a very good estimator of rainfall.
The use of polarimetric variables (especially K_{dp}) reduces the uncertainties on hydrometeor size distribution and properties, allowing higher quality rainfall estimation [93]. In figure 2.4 are reported the results of an intercomparison among algorithms using different polarimetric variables [94], using the Root Mean Square Error (RMSE) between hourly radar estimate and rain-gauge measures as quality indicator. Results show the great impact of the use K_{dp} and other parameters in reducing the error for both point-like and areal integrated values. The diffusion of the operational use of polarimetric radar networks is expected to increase the accu-

racy of precipitation measurements and the overall quality of nowcasting and very short term forecasts [95].

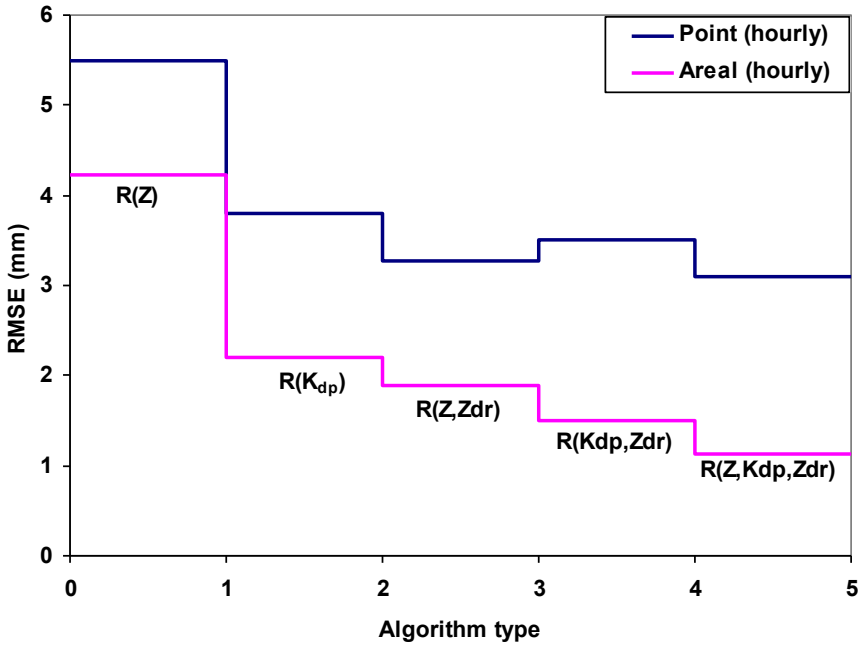


Figure 2.4: RMS errors of point and areal hourly rainfall accumulations for various algorithms (from Ryzhkov et al., 2005).

2.2 Satellite Remote sensing

As described in the previous chapter, satellite observations represent a valuable solution for a wide scale monitoring of atmospheric parameters. Remote sensing is based on the understanding of the relationship between satellite measured radiation and the physical and chemical characteristics of atmospheric gases, particles, clouds and Earth's surface.

We know that the radiation we deal with interacts with atmospheric objects and can be absorbed, emitted or scattered, depending on its wavelength and on objects dimension and nature. Radiation at different wavelengths can then bring to the

sensor different informations and for this reason the multi-sensor remote sensing is the preferred approach.
Satellite sensors for atmospheric monitoring are designed to work at VIS-IR and MW wavelengths because of the object dimensions they want to observe, and the radiative characteristics of relevant atmospheric component.

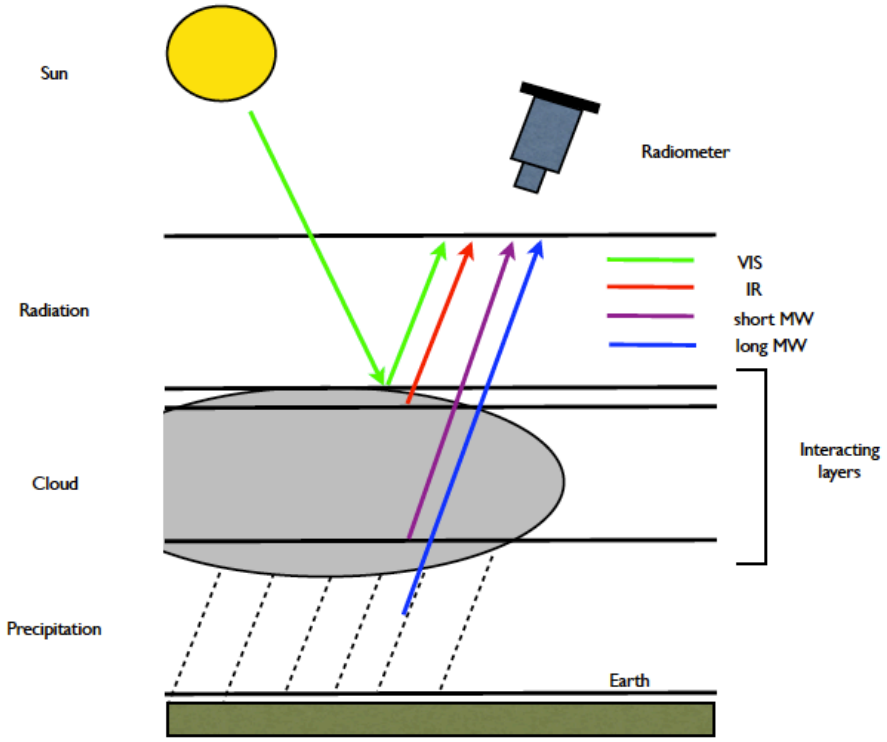


Figure 2.5: Schematical representation of VIS-IR-MW radiation interaction.

If we are interested in clouds, we have to consider the interaction between radiation and hydrometeors. Shorter wavelengths, such as the ones in the VIS band, coming from the sun, weakly interact with atmospheric gases, but strongly interact with clouds particles and are not able to penetrate them. The informations that reach the satellite radiometer come then from the top of clouds (green lines in fig.2.5).
Radiation in the IR band measured by a satellite sensor is emitted by objects and gases: the amount of emitted radiance is a growing function of the object or gas temperature, following the Planck's Law. Focusing on clouds, each hydrometeor

layer absorbs warmer radiation from layers below and emits radiation isotropically, according to its equilibrium temperature: the satellite sensor measures the radiation coming from the upper cloud layer, and no information on the underlying cloud structure is provided (red line in fig.2.5).
MW radiation can be considered the more direct physical information about precipitation layers, because it comes from the inner layers of clouds (blue line in fig.2.5) or even from precipitation layers (violet line in fig.2.5), thanks to its longer wavelength, which weakly interacts with smaller cloud particles.
In the next sections will be explained with more detail the physical properties of the atmospheric radiative transfer at the wavelengths used for this work.

2.2.1 Basic definitions

Electromagnetic radiation is a form of energy (Radiant Energy Q [J]) that travels trough the space from a source without the necessity of transfer medium. The rate of energy transfer by electromagnetic radiation is called Radiant Flux and is denoted by:

$$F = \frac{dQ}{dt} \qquad [J\,s^{-1}\,(W)] \tag{2.4}$$

The radiant flux per unit area is called Irradiance or Radiant Flux Density and can be expressed as:

$$E = \frac{dQ}{dt\,dA} \qquad [W\,m^{-2}] \tag{2.5}$$

The irradiance per unit wavelength (or frequency or wavenumber) interval is called Monochromatic Irradiance:

$$E_\lambda = \frac{dQ}{dt\,dA\,d\lambda} \qquad [W\,m^{-2}\mu m^{-1}] \tag{2.6}$$

In general the irradiance is determined by the contribution of energy coming from all directions. The irradiance per unit solid angle is called Radiance and is denoted by:

$$I_\lambda = \frac{dQ}{dt\,dA\,d\lambda\,d\Omega} \qquad [W\,m^{-2}\mu m^{-1} sr^{-1}] \tag{2.7}$$

For a perfect emitter (black body) the spectral distribution of radiant energy is defined by the Planck Law:

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5}\left(exp\left(\frac{hc}{\lambda kT}\right) - 1\right)^{-1} \qquad [W\,m^{-2}\mu m^{-1}sr^{-1}] \tag{2.8}$$

where $h = 6.626 \times 10^{-34} J\,s$ is the Planck constant, $c = 2.99792 \times 10^8\,m\,s^{-1}$ is the speed of light, $k = 1.38054 \times 10^{-23} J\,K^{-1}$ the Boltzmann constant and T is the equilibrium temperature in Kelvin degrees (K).
Alternatively, the Planck Law can be more simply formulated as:

$$B_\lambda(T) = \frac{c_1}{\lambda^5}\left(exp\left(\frac{c_2}{\lambda T}\right) - 1\right)^{-1} \qquad [W\,m^{-2}\mu m^{-1}sr^{-1}] \tag{2.9}$$

where $c_1 = 1.19104\dot{1}0^{-8}\,W\,m^{-2}cm^{-4}sr^{-1}$ and $c_2 = 1.438769\,cm\,K$ are the first and the second radiation constants, respectively.
Blackbody radiation represents the upper limit to the amount of radiation that a real substance may emit at a given temperature. At any given wavelength λ, Emissivity is defined as the ratio of the actual emitted radiance to that from an ideal blackbody (refer to the scheme in fig.2.6 for notation):

$$\epsilon_\lambda({}_\uparrow\hat{\Omega}, T) \equiv \frac{I^e_\lambda(\hat{\Omega})cos\theta d\Omega dA}{B_\lambda(T)cos\theta d\Omega dA} \tag{2.10}$$

where $\uparrow$ denotes the emission direction and in the following equations $\downarrow$ will denote the incident direction.
Considering now an incident beam of radiation (primed symbols) with a particular direction within a solid angle $d\Omega'$, there will be a radiation fraction absorbed by atmosphere. The Absorbance is the ratio of the absorbed radiation to the incident one:

$$\alpha_\lambda({}_\downarrow\hat{\Omega}, T) \equiv \frac{I^{e\downarrow}_\lambda(\hat{\Omega}')cos\theta' d\Omega' dA}{I^{\downarrow}_\lambda(\hat{\Omega}')cos\theta' d\Omega' dA} \tag{2.11}$$

At the same time, Transmittance can be defined as the ratio of the transmitted radiation over a particular layer to the incident radiation:

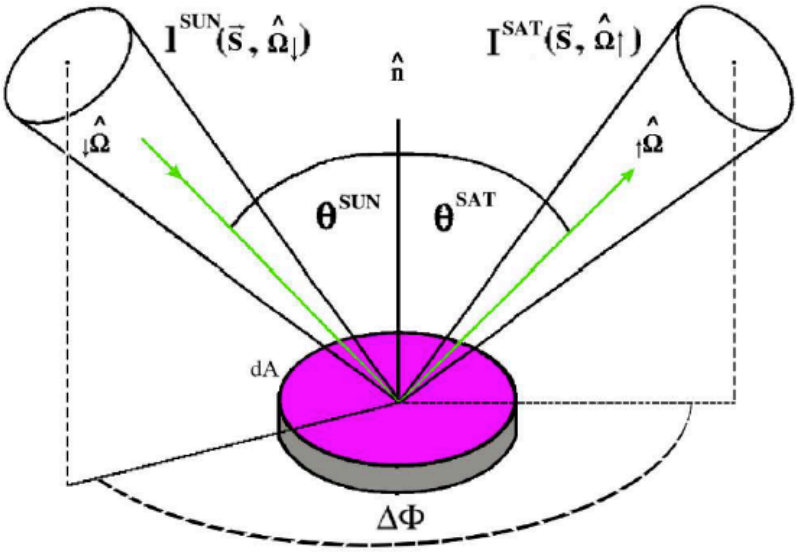


Figure 2.6: Sun-pixel-satellite geometry.

$$\tau_\lambda({}_\downarrow\hat{\Omega}',{}_\downarrow\hat{\Omega}) \equiv \frac{dI_\lambda^{t\downarrow}(\hat{\Omega})}{I_\lambda^{\downarrow}(\hat{\Omega}')cos\theta' d\Omega'} \tag{2.12}$$

And concluding, the reflectance bidirectional distribution is the ratio of reflected radiation in $\hat{\Omega}$ direction to the incident energy over the surface (denoting with surface in our case a cloud layer):

$$\rho_\lambda({}_\downarrow\hat{\Omega}',{}_\uparrow\hat{\Omega}) \equiv \frac{dI_\lambda^{r\uparrow}(\hat{\Omega})}{I_\lambda^{\downarrow}(\hat{\Omega}')cos\theta' d\Omega'} \tag{2.13}$$

When electromagnetic radiation at a given wavelength passes through a medium, the Beer-Lambert Law (or Extinction Law) describes the attenuation of an incident beam because of absorption or scattering with the medium components (in our cases the atmospheric components, particles and gas molecules):

$$dI_\lambda(s,\hat{\Omega}) = k_{ext} I_\lambda(s,\hat{\Omega}) ds \tag{2.14}$$

that measures the decrement of I_λ due to the interaction with the medium along

the infinitesimal path ds, where k_{ext} is the extinction coefficient, linked to the extinction cross section through the definition:

$$\sigma_{ext} \equiv \frac{k_{ext}}{n_a} \tag{2.15}$$

where n_a is the number of absorber-particles in the unit volume.

Equation 2.14 can be integrated along a vertical path from the surface to the sensor to yeld:

$$I_\lambda(sat, \hat{\Omega}) = I_\lambda(0, \hat{\Omega}) exp(-\tau/\mu) \tag{2.16}$$

where $I_\lambda(0, \hat{\Omega})$ is the incident beam intensity (i.e. the radiation upwelling from the surface), $I_\lambda(sat, \hat{\Omega})$ is the radiance reaching the satellite position and $\mu = \cos\theta$, being θ the satellite viewing angle with respect to the vertical.
τ is the atmosphere optical depth (or optical thickness) along the path from the surface to the satellite and is defined as:

$$\tau \equiv \int_0^s k_{ext} ds \tag{2.17}$$

Finally, we introduce a quantity widely used in climate modeling to estimate the radiative impact of clouds on the Earth's energy balance, and in cloud microphysics as a relevant feature of the cloud structure: the hydrometeors effective radius (r_{eff}) defined as follows:

$$r_{\text{eff}} = \frac{\int_0^\infty r^3 n(r) dr}{\int_0^\infty r^2 n(r) dr} \qquad [\mu m] \tag{2.18}$$

where $n(r)dr$ represents the hydrometeor concentration with radius between r and $r + dr$. The meaning of this parameter can be summarized as follows. The scattering process occurs on particle surface, given the fact that the electromagnetic field is zero inside the particle, while the absorption process depends on the mass of the particle: scattering depends on r^2 while absorption on r^3. The net reflectance is thus qualitatively proportional to the ratio between scattering and absorption, i.e. the ratio between the contribution of all the particles surface area and all the particles volume in the cloud layer, which is proportional to r_{eff}^{-1}. Therefore, where absorption is dominant, reflectance is related to the r_{eff} profile: large particles have low reflectance, while small particles are more reflective.

2.2.2 Visible

For wavelengths between 0.4 and about 4 μm, solar radiation is dominant (see in fig.2.7 the solar spectrum in the VIS band). At these wavelengths thick clouds, snow and ice can be easily recognized because of their intense reflection of solar radiation. Comparing images 0.6 an 0.8 μm, it is expected that clouds are better discriminated from clear sky at 0.6 μm, because of the different reflectivity values of leaves and soil in the two wavelengths, low at 0.6 μm and higher at 0.8 μm (in fig.2.7 indicated with orange and green arrows).

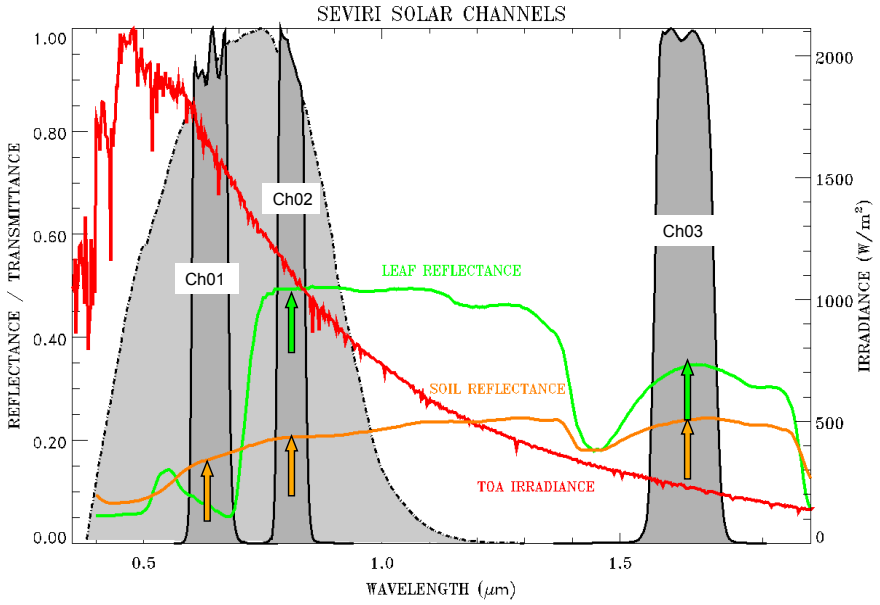
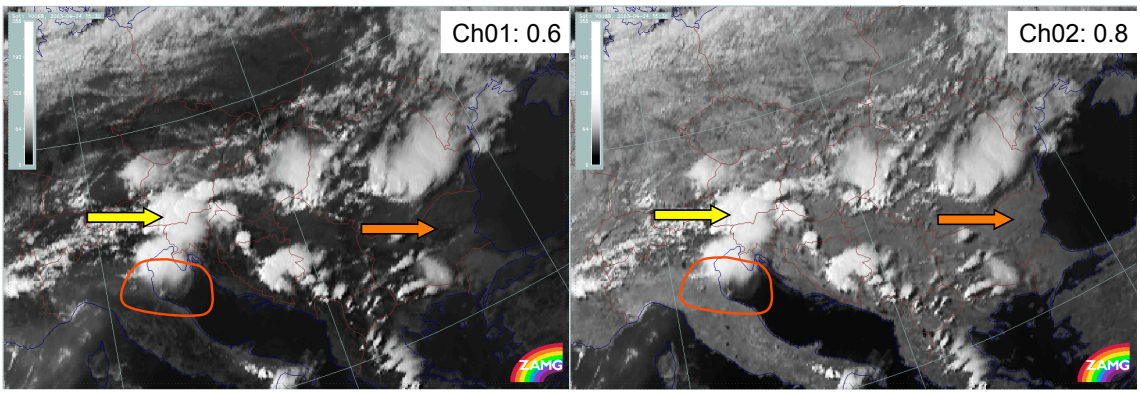


Figure 2.7: Solar energy spectrum in the VIS wavelengths (Copyright 2010 EUMETSAT).

In fig.2.8 it is shown an example of satellite images in the two different wavelengths and it is possible to note that in both cases thick clouds are clearly detected (green arrows), whereas thin clouds appear only in the 0.6 μm because of the higher contrast between land and clouds (red closed line) that is clearly observable comparing the two images (orange arrows). Because of the great separation between leaf and soil reflectance curves at 0.8 μm (in fig.2.7 indicated in orange and green), using this wavelength it is also possible to recognize Earth's surface characteristics (soil or vegetation).



(a) (b)

Figure 2.8: Example of satellite images in the 0.6 μm (a) and the 0.8 μm (b) wavelengths (Copyright 2010 EUMETSAT).

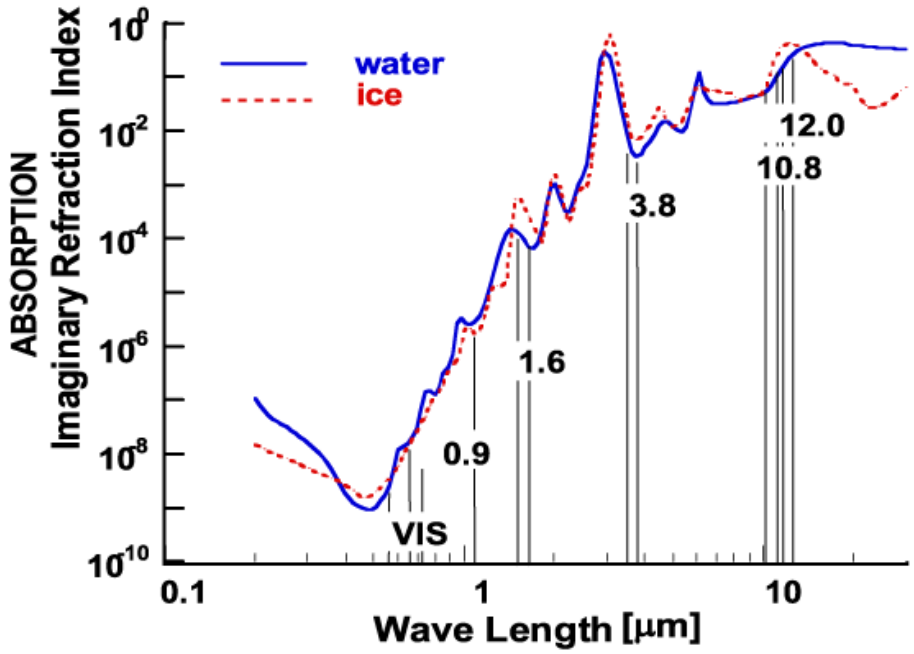


Figure 2.9: Absorption of ice and water cloud.

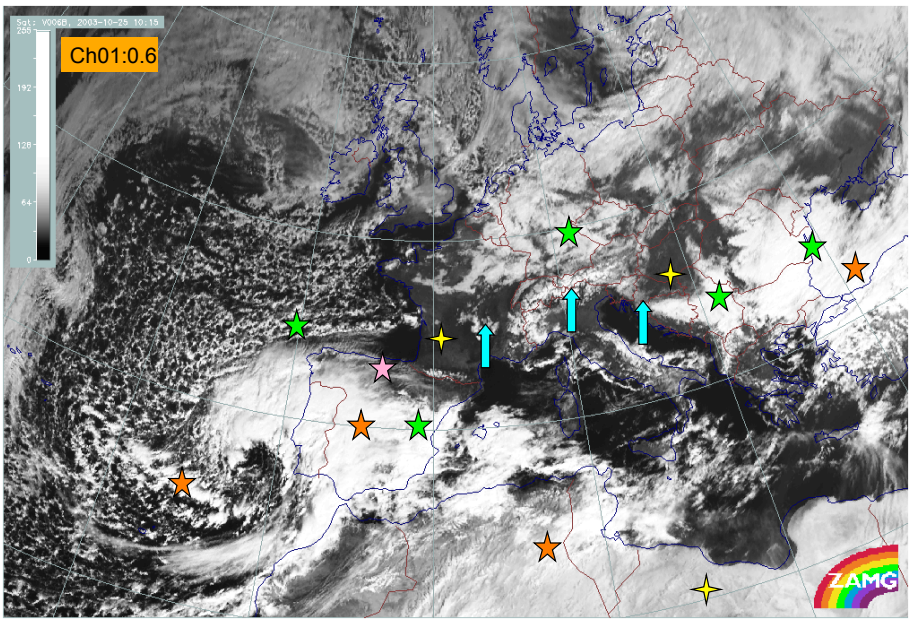
An other important wavelength to be considered is the 1.6 μm, in which it is possible to distinguish between ice and water clouds because of the separated trend between the imaginary part of the refraction index (related to absorption properties) of water (blue solid line in fig.2.9) and ice (red dotted line in fig.2.9). This wavelength, because of this important property, can help also for a correct classification of snow at ground. Fig.2.10 is an example of satellite images comparison in which it is possible to correctly classify water clouds, ice clouds and snow at ground just through a simple comparison between them.
The possibility to discriminate clouds from snow at the ground is due to the different reflectance of snow and clouds at 0.6 μm and 1.6 μm: clouds are good reflectors at both wavelengths, while snow reflects much less at 1.6 μm.
A further application of 1.6 μm wavelength will be presented in the next section discussed in parallel to radiation at 3.9 μm.

2.2.3 3.9 μm

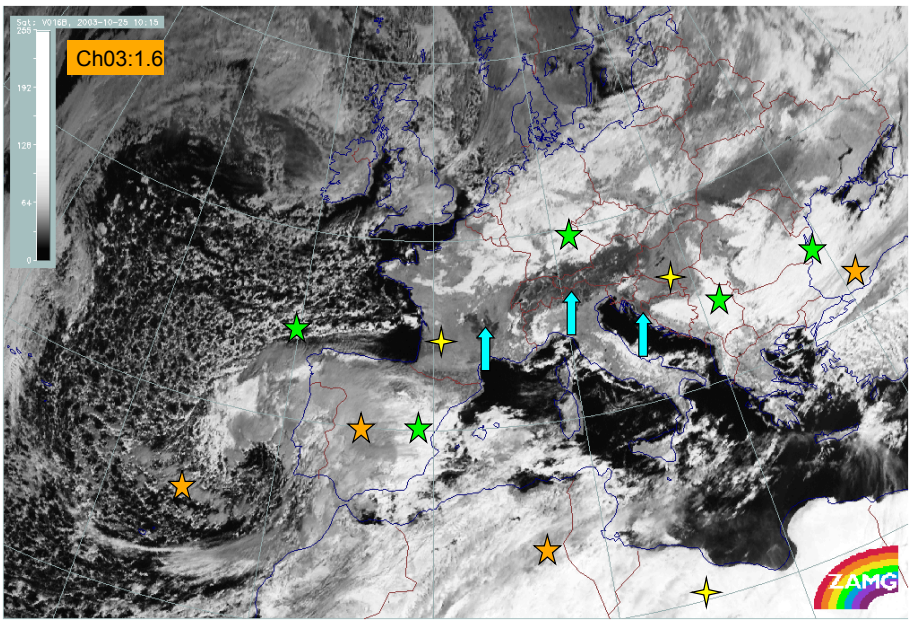
Between VIS and IR spectral bands there is a small part of electromagnetic spectrum for which the measured radiation is composed by scattered solar radiation and Earth system emitted thermal radiation. The behavior of this particular band is different between day and night-time because of the strong dependence on solar zenith angle of the solar component. For this reason, applications and algorithms using this wavelength have to be different for day and night-time.

In fig.2.12 it is highlighted the strong interpretation difference between the two situations (day and night-time) for the same wavelength. In red the actual scene temperature, measured during night-time as a normal IR wavelength and in blue the satellite measured scene temperature, higher with respect of the actual one because of the reflected sunlight contribution.

In fig.2.13 it is reported an example of day and night-time images. During night-time (fig.2.13b) there is only thermal contribution and clouds are brighter (colder) than ocean surfaces, given the fact that IR channels images are rendered in gray scale with an inverse relationship between radiance and pixel intensity. During day-time (fig.2.13a) there are thermal and solar contribution and low clouds appear darker than ocean surfaces: they are colder than the ocean, but reflect much more solar radiation. High, cold-top clouds appear white in both day-time and night-time scenes.



(a)



(b)

Figure 2.10: Satellite images at 0.6 μm (a) and 1.6 μm (b). With blue arrows are indicated snow areas, with orange ice clouds, with green water clouds, with pink an example of transparent cloud and with yellow it is indicated Earth's surface (Copyright 2010 EUMETSAT).

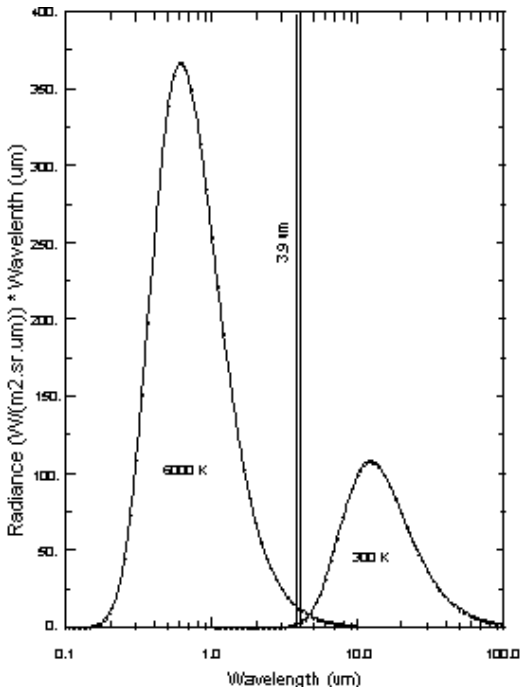


Figure 2.11: 3.9 μm wavelength between Solar and Earth spectra.

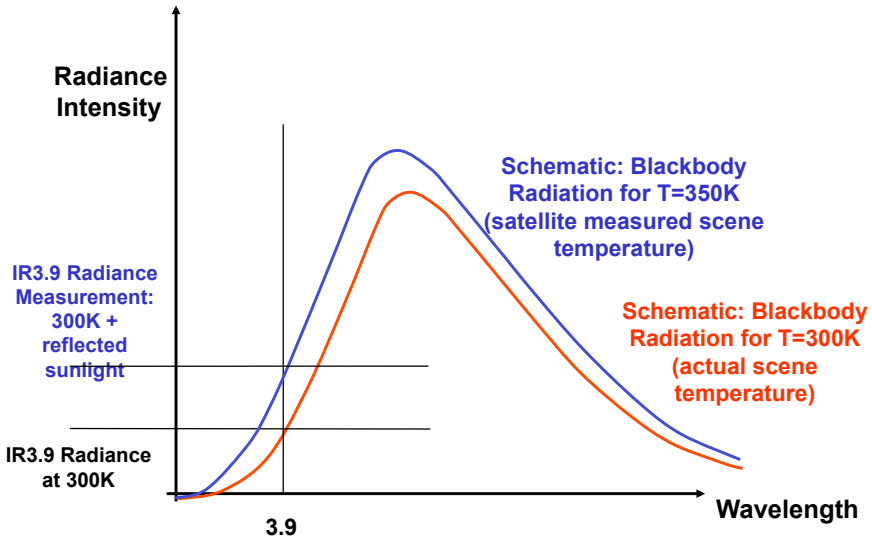
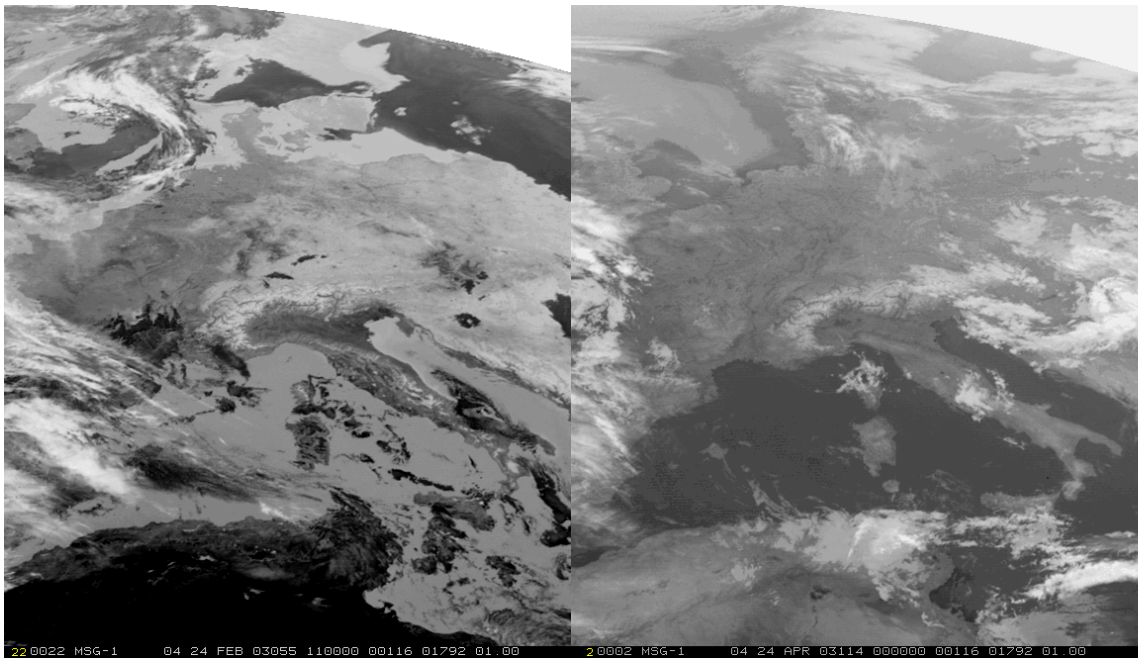


Figure 2.12: Scheme of different blackbody radiation considering or not reflected sunlight (Copyright 2010 EUMETSAT).



(a) (b)

Figure 2.13: Satellite images at 3.9 μm for day-time (a) and night-time (b) (Copyright 2010 EUMETSAT).

This wavelength because of its particular characteristic, can be used for different applications. During both the day and night-time, it is used for the detection of low clouds, thin cirrus, super-cooled clouds and fog. A peculiar application is for night-time images due to the high sensitivity of this wavelength to warm spot in the satellite pixel. The temperature sensitivity α is defined as the percentage change in radiance corresponding to a percentage change in temperature:

$$\frac{\Delta B}{B} = \alpha \frac{\Delta T}{T} \tag{2.19}$$

and substituting the Planck equation 2.9 it can be solved for $\alpha = c_2/\lambda T$. Integrating the equation in B and T between T_{ref} and T and B_{ref} and B we obtain:

$$B = B_{ref}(\frac{T}{T_{ref}})^{\alpha} \tag{2.20}$$

Computing α at 3.9 μm and at 11 μm, it is verified that the temperature sensi-

tivity is much higher for the shorter wavelength. Thus, in case of partially cloudy pixel, the temperature measured at 3.9 μm will be sensitively higher than if measured at 11 μm.
This property poses two prompt applications of the differences between 3.9 μm and 11 μm TB at night: 1) in case of partially cloudy pixel TB3.9 will be higher then TB11, while in case of homogeneous cloudy coverage the two temperatures would be very similar; in case of the presence of a small, very hot spot inside the pixel, as in case of wildfire, again TB3.9 will be higher then TB11. The TB3.9-TB11 difference could be used for the discrimination of sub-pixel cloud and for the early detection of wildfires.
A last feature of the radiation between 1 and 4 μm is that the cloud reflectance depends on the r_{eff} radius of the cloud particles. By means of radiative transfer simulations Capacci [96] computed the reflectance at 1.6 μm and 3.9 μm as function of the cloud layer optical thickness for different values of the cloud particles effective radius, and the results are shown in figure 2.14 for a given triplet of solar and satellite zenith and azimuth angles.

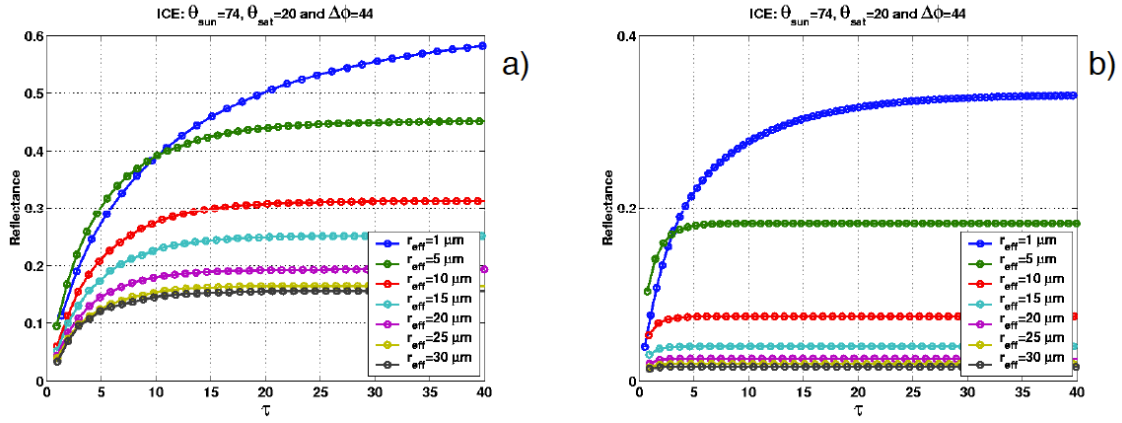


Figure 2.14: Reflectance at 1.6 μm (a) and at 3.9 μm (b) as function of cloud layer optical thickness for different values of the cloud particle effective radius [96].

At 1.6 μm the dependence of the reflectance from r_{eff} becomes negligible for optical thickness higher than 20, indicating that the 1.6 μm radiance can be used to retrieve cloud particle r_{eff} for $\tau > 20$ for effective radius smaller than 25 μm, where the signal saturates. The 3.9 μm plot shows similar behavior with significant

differences. First, the radiance values are lower by a factor around 2; secondly, the reflectance dependence on τ is negligible for optical thichness between 5 and 10, indicating the radiance can be used to estimate r_{eff} in very thin cloud layer. Finally, the reflectance sensitivity to r_{eff} is limited to $r_{\text{eff}} < 25\,\mu$m. This indicates that the reflected component of daytime 3.9 μm radiance, that can be extracted from the total radiance by simple radiative transfer computations, can be used to infer an estimate of cloud top r_{eff}, while by means of the reflectance at 1.6 μm is possible to compute r_{eff} for cloud layers below the cloud top. This approach was used by [97] to infer the cloud upper layers structures by using the four channels between 1 to 4 μm of the Moderate resolution Spectroradiometer MODIS.
Moreover, Lensky and Rosenfeld (1996) [98] made an important study about relation between effective radius and precipitation and they developed an algorithm for precipitation retrieval based on effective radius detection principle. The algorithm first of all detects fully clouded pixels, then it assigns an effective radius of cloud-top particles to these pixels, it divides the area into windows of about 2000 km^2, it assigns to each window a cloud type by the cloud radius parameter and it assigns rain area and rain rate to each of the cloud types. Capacci [96] continued in this direction and used the 3.9 μm wavelength to appreciate the cloud particles effective radius and correlate it with precipitation. In fig.2.15 is shown an example of comparison between satellite image, radar precipitation estimate and effective radius map. It is clear from the graphs (fig.2.16) that effective radius greater than 15 μm is confirmed as being correlated with the precipitation region in case of convective precipitation.

2.2.4 Infrared: atmospheric window wavelengths

In the IR spectral band, where Earth thermal emitted radiation is dominant, in order to investigate clouds and Earth's surface with negligible gases absorption, it is necessary to choose particular wavelengths in which radiation does not interact much with atmospherical gases. To clarify the roles of IR wavelengths it is useful to introduce the weighting functions. The weighting function of a wavelength, defined as the vertical derivative of the atmospheric transmittance, indicates the contribution of the different standard atmosphere layers to the radiation reaching the top of the atmosphere (and thus the satellite sensor) at the given wavelength. It is clear from the weighting functions (fig.2.17, light blue, pink and yellow lines) that at

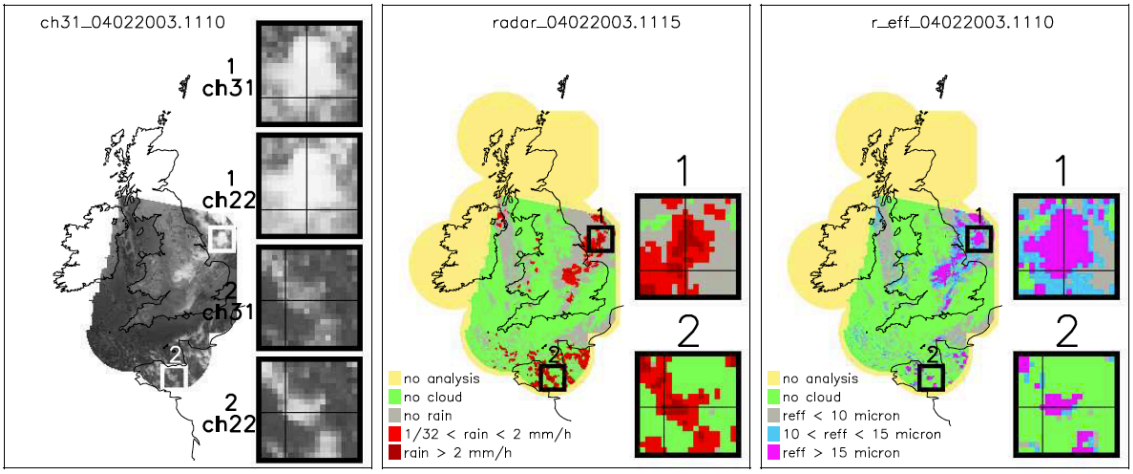


Figure 2.15: IR satellite image (left), precipitation radar map (center) and effective radius map (right). The selected cases in the boxes are convective clouds [96].

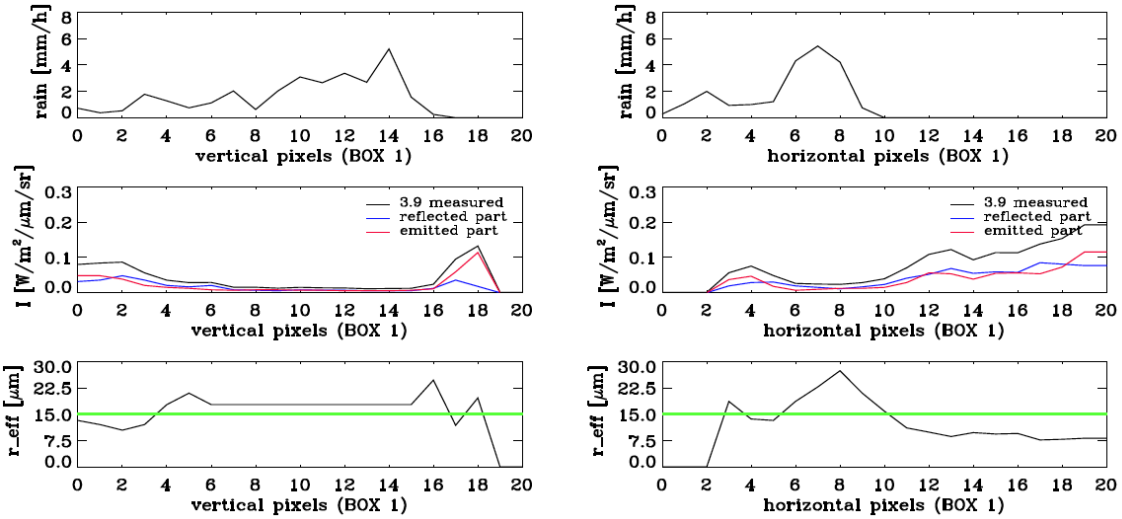


Figure 2.16: Plots of rainfall rate (top), 3.9 μm measurement, with the separated emitted and reflected parts (centre) and effective radius values along the vertical and horizontal lines in the box 1 of fig.2.15 (bottom) [96].

this particular wavelengths, the radiation reaching the satellite sensor, in clear sky conditions, comes effectively from Earth's surface, in fact the normalized weighting function is equal to 1 for surface pressure. These wavelengths, called *atmospheric windows*, are 8.7, 10.8 and 12.0 μm and are indicated in fig.2.18 with red lines.

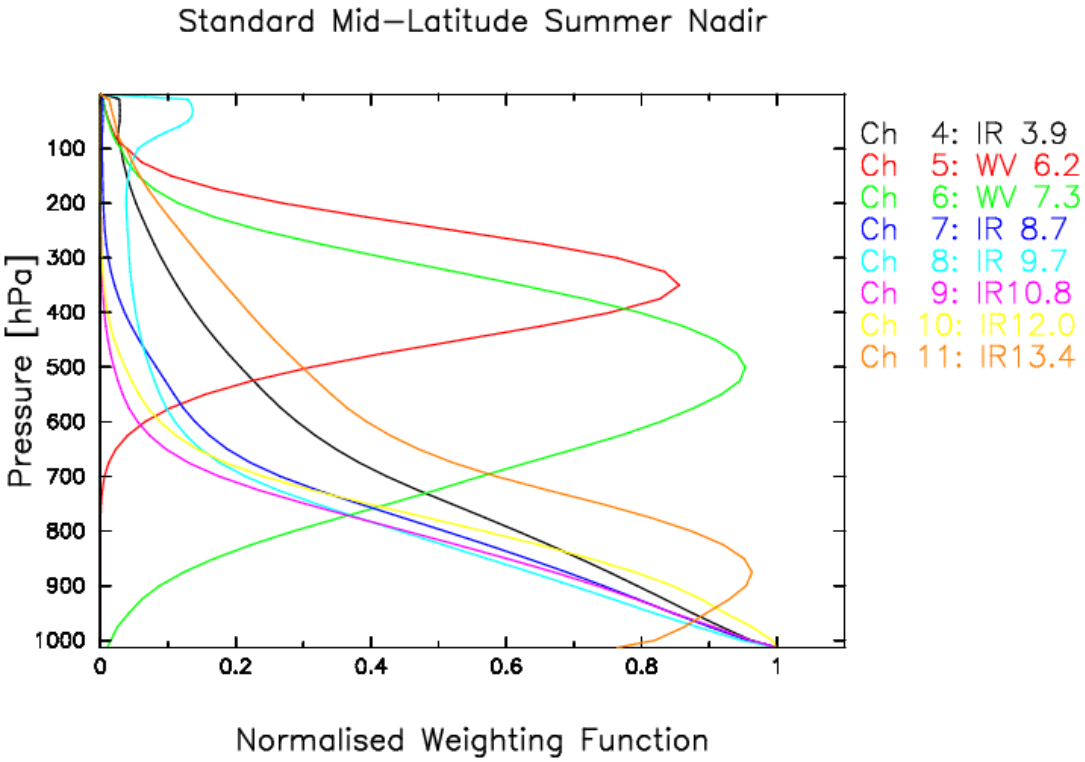


Figure 2.17: IR wavelengths weighting functions (Copyright 2010 EUMETSAT).

These wavelengths are useful for cloud systems recognition thanks to the thermal radiation of clouds and Earth's surface. They are indicative of high thin cirrus clouds (particularly the 12.0 μm) and are a good support to 1.6 μm for the discrimination between ice and water clouds (see in fig.2.9 also for 10.8 and 12.0 μm the significative difference between water and ice imaginary refraction index, responsible for absorption).

2.2.5 Infrared: absorption wavelengths

In the IR band, apart from atmospheric window, it is useful to consider also some wavelengths where the atmosphere transmittance is low (called also *opaque spectral regions*) to understand atmosphere composition. Water Vapor (WV), among absorption gases, can be considered the most important for precipitation retrieval and for air mass classification and tracking. Strong absorption bands at 6.2 and 7.3 μm (see in fig.2.18 indicated with blue lines) are important to estimate the content of

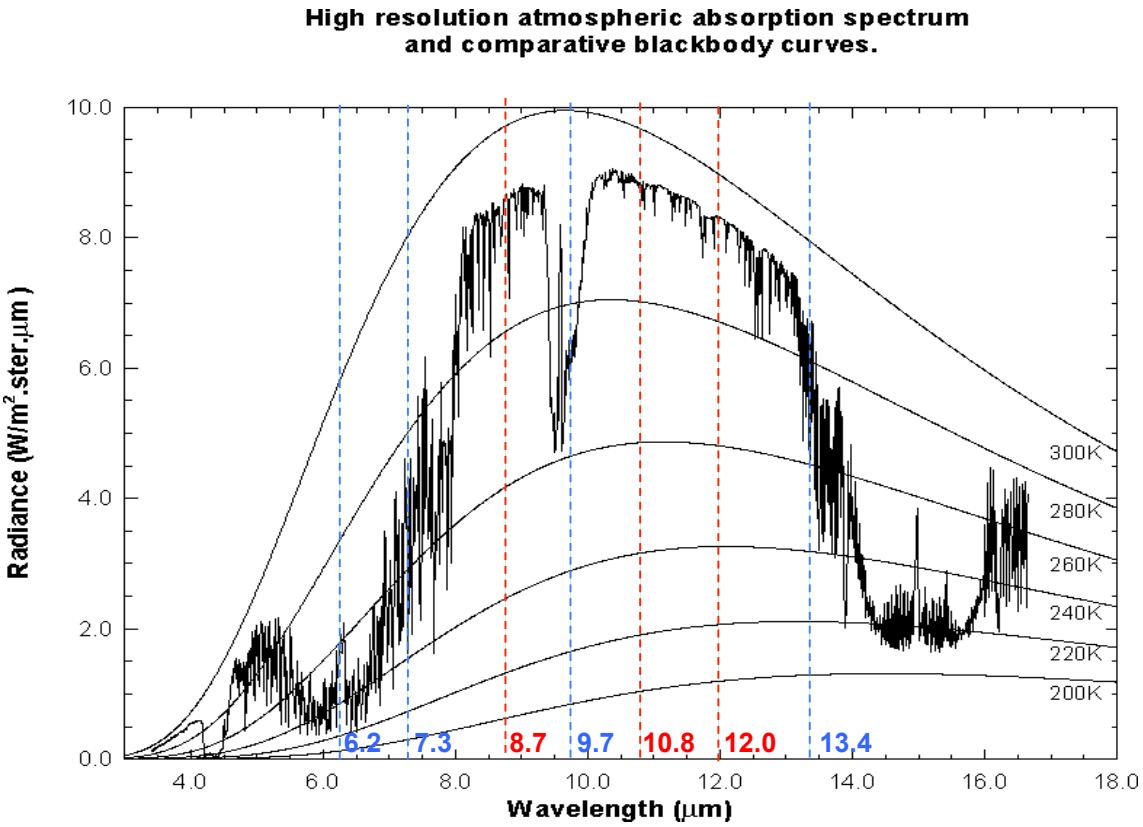


Figure 2.18: Earth emitted spectrum.

WV in two different layers of troposphere (in fig.2.17 red and green lines represent the weighting functions of WV). In particular, the 6.2 μm is more in the center of absorption band with a consequently strong absorption and gives informations about the WV content in the higher tropospheric layers, whereas, the 7.3 μm is more to the wings of absorption band, with less absorption and giving informations about lower layers of the troposphere. These two wavelengths are also useful for the height determination of transparent clouds and for the Atmospheric Motion Vectors (AMV) computation in clear sky regions. Other two absorption bands are retrieved in the IR, but they are not relevant for this work and for this reason they will be just mentioned. They are the 9.7 μm and the 13.4 μm, in the O_3 and in the CO_2 absorption band respectively. The Ozone wavelength gives information about high stratospheric Ozone concentrations which have protruded relatively far down into the troposphere (its weighting function is the light blue one in fig.2.17). The Carbon

Dioxide band is useful for a temperature profile and for informations about winds and instabilities.

2.2.6 Microwave

MW radiation is the most physical information able to reach satellite sensors. It comes from the inner layers of clouds or even from precipitation layers because it weakly interact with clouds droplets and small ice crystals. In fig.2.19 we can note that Mie scattering at MW wavelengths is possible only with objects larger than 10 μm, so the interaction can happen only with large cloud droplets or larger particles (such as rain drops, graupel, snowflakes and hail).

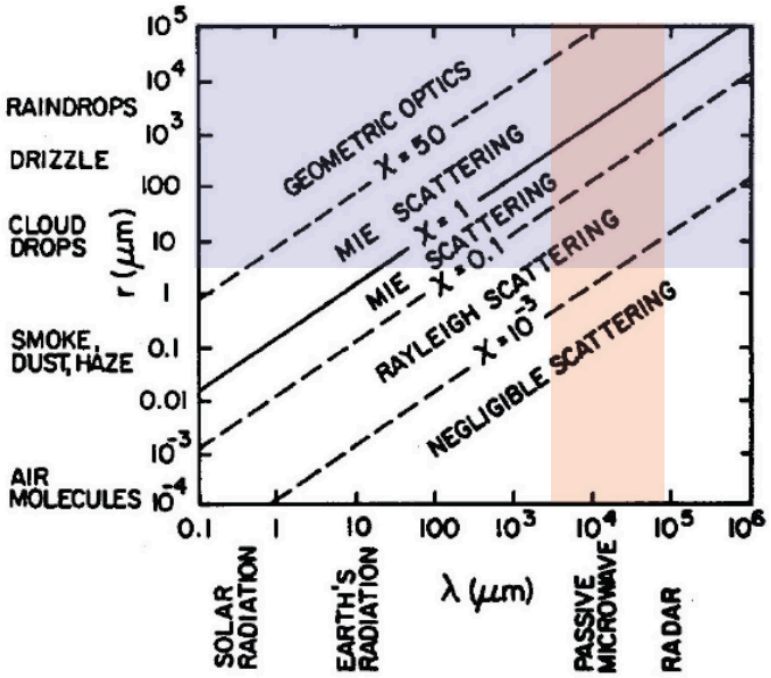


Figure 2.19: Type of scattering for different particles size in the different wavelengths.

Precipitation estimate in this spectral band, physically speaking, can then be considered the most accurate and reliable, but because of technical problems it is difficult to obtain a good time and spatial resolution (this concept will be better explained in sec.2.3). The only physical problem is the different behavior of radiation

depending on background, because of the different and time varying emissivity of land and sea surfaces. To solve this problem, Staelin and Chen (2000)[99] developed an algorithm for MW wavelengths using absorption bands, because they are able to sense precipitation over land without strong surface emissivity effects. This work is used for the MW precipitation product of PROSA Project, described in sec.1.2.2.1. The frequency used are the 54 and 183 GHz and are respectively in O_2 and in H_2O absorption bands (see fig.2.20).

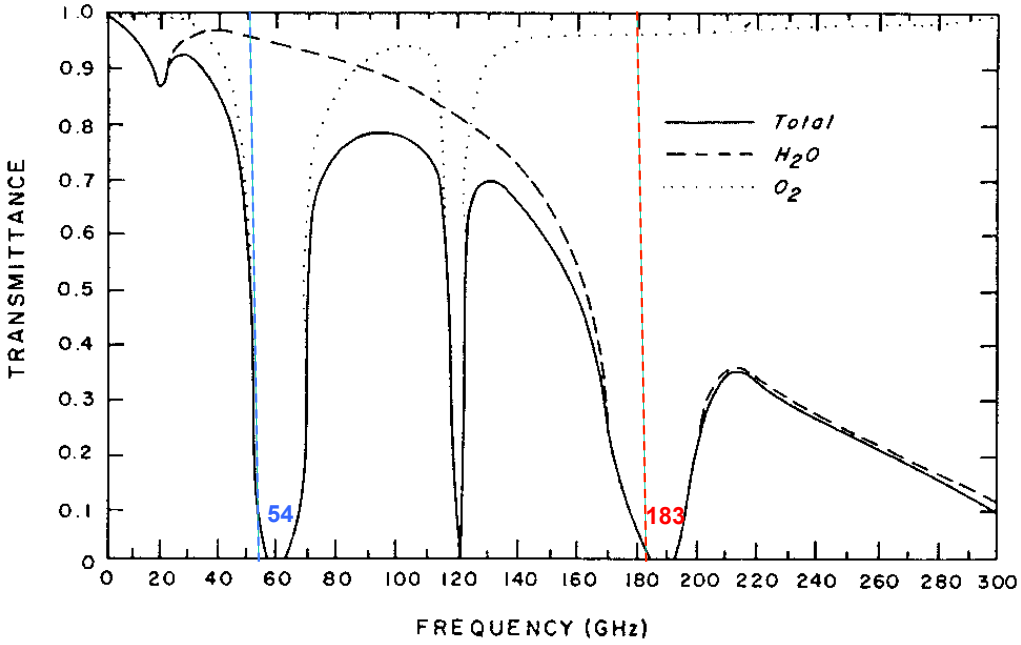


Figure 2.20: MW spectrum.

At 54 GHz, all solid, liquid or mixed-phased hydrometeors absorb and scatter radiation and the ones above about 4 km are generally visible against the warmer opaque atmospheric background below (in contrast with the tendency of window frequencies over land to respond to both liquid hydrometeors and random surface variations). Because of a general dependence of frequency from penetration depth of radiation, low-altitude hydrometeors impact primarily those sensors channels at lower, more transparent frequencies, whereas hydrometeors at higher altitudes produce a cold spot across all tropospheric sensor channels. In this way, 54 GHz signal is sensitive to the altitude of precipitating layers, even when they are hidden under thin cirrus or other light clouds. Moreover, at this frequency it is possible to

establish a close relationship between rain-rate and vertical velocity and absolute humidity of the saturated air of convective cells because this velocity is directly related to the cell top altitude [99].
The 183 GHz frequency is also important for convective cells identification, because they are generally characterized by ascending air masses and wet adiabatic temperature profiles. In absence of hydrometeors, for an idealized adiabatic atmosphere, the relation between water vapor content and temperature at any altitude is fixed and also the penetration depth of 183 GHz frequency is fixed. In this way the idealized observed brightness temperature at this frequency is fixed and independent from surface emission at any latitude. Out from this ideal situation, in presence of hydrometeors scattering and absorbing radiation, the brightness temperature measured from satellite sensors is lower then its nominal value for a saturated atmosphere. This cold precipitation signature is unique because brightness temperature below those corresponding to saturation generally can not be generated otherwise. 183 $\pm$1 GHz nominal brightness temperature corresponds to ice in saturated atmosphere, whereas 183$\pm$7 GHz temperature corresponds to both ice and water. This leads to a strongly response of 183$\pm$7 GHz to stratiform rain and low precipitation rates, contrarily 183 $\pm$1 better responds to strong convection thrusts large ice particles, substantially higher in the atmosphere.

2.3 Space-borne sensors

The lack of data in some regions not covered by rain-gauges or meteorological radars gave the idea of using satellite sensor to monitor precipitation at wide scale. There are two types of meteorological satellites: the Low Earth Orbit (LEO), sun-synchronous, orbiting from 400 to 1500 km from the Earth-surface, and the Geostationary Earth Orbit (GEO) satellites, orbiting at 36000 km at the same angular velocity of the Earth rotation, in a fixed position with respect to the Earth's surface. Most space-borne sensors are passive instruments which collect radiation emitted from Earth-surface or atmosphere. GEO satellites can carry visible/infrared (VIS/IR) sensors, while on LEO spacecrafts there can be also antennas for microwave (MW) signals. The problem of this type of sensors stays in the dimensions of antennas: following the Rayleigh criteria (eq.2.21), it is known that higher orbits require larger telescope dimensions making impossible to carry them on GEO

satellites.

$$r = \frac{h\lambda}{d} \tag{2.21}$$

where r is the minimum resoluble dimension at ground from sensor, h is the satellite orbit altitude, λ is the wavelength of radiation and d is the antenna diameter.

2.3.1 SEVIRI sensor

The Spinning Enhanced Visible and Infrared Imager (SEVIRI) sensor is a radiometer on board MSG (Meteosat Second Generation) satellite. Launched in 2002, since 2004 it produces images of the Earth's disc centered at 0° latitude and 0° longitude every 15 minutes.
Its operating principle is based on collecting radiation from a target area and focusing it on detectors sensitive to 12 different bands of the electromagnetic spectrum by means of a telescope.
The 1 km resolution of the High Resolution Visible (HRV) channel is achieved by using 9 broad-band detection elements. The other channels are sampled at 3 km resolution by using 3 narrow-band detection elements per channel.

In tab.2.1 are resumed the main features of SEVIRI channels.

The full Earth disc image is obtained after 1250 scan line steps (south north direction) of 9 km SSP per line step with 3750 total pixels. The rapid scan (line scan) is performed from East to West thanks to the satellite rotation around the spin axis (spin rate at 100 rotations per minute). The spin axis is perpendicular to the orbital plane and is nominally oriented along the South-North direction.
For this work a SEVIRI image dimension of 444×705 pixels is used, in order to consider only the Italian region.

2.3.2 AMSU sensors

The Advanced Microwave Sounding Unit (AMSU) is a cross-track microwave radiometer on board the NOAA satellite (NOAA-15, NOAA-16, NOAA-17). The instrument is composed by two sensors, AMSU-A observes at 15 frequencies up to 90 GHz with about 50 km spatial resolution and AMSU-B observes at 5 frequencies, 88-191 GHz with about 15 km resolution at nadir with a circle IFOV (Instantaneous

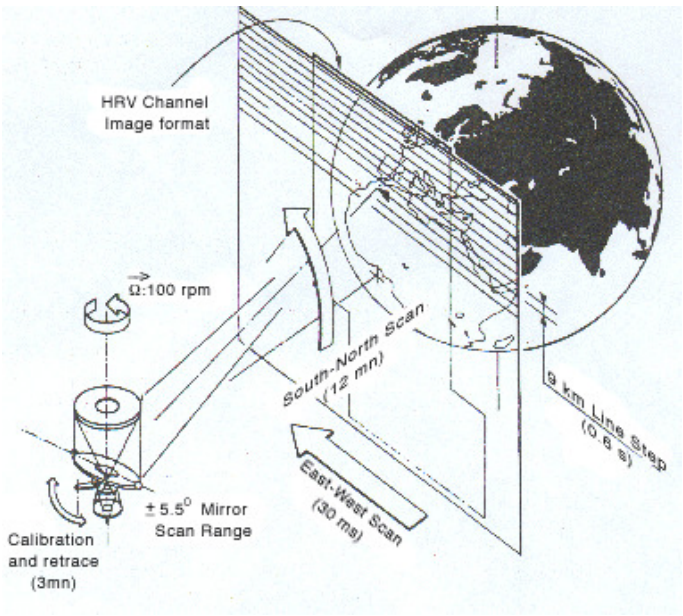


Figure 2.21: Schematic image of SEVIRI acquisition system [100].

Channel number	**Channel name**	**Nominal central wavelenght (μm)**	**Nominal spectral band (μm)**	**Spatial resolution at nadir (km)**
ch01	VIS0.6	0.64	0.56-0.71	3
ch02	VIS0.8	0.81	0.74-0.88	3
ch03	NIR1.6	1.6	1.50-1.78	3
ch04	NIR3.9	3.92	3.48-4.36	3
ch05	WV6.2	6.2	5.35-7.15	3
ch06	WV7.3	7.3	6.85-7.85	3
ch07	IR8.7	8.7	8.30-9.10	3
ch08	IR9.7	9.7	9.38-9.94	3
ch09	IR10.8	10.8	9.80-11.80	3
ch10	IR12.00	12.0	11.00-13.00	3
ch11	IR13.4	13.4	12.40-13.40	3
ch12	HRV0.75	0.75	0.6-0.9	1

Table 2.1: SEVIRI spectral channels and corresponding resolutions.

Field Of View) shape. IFOV areas increase getting away from nadir and became elliptic, with maximum axes dimensions of 51.6×25.3 km^2 (in fig.2.22 are plotted all AMSU IFOV sizes). The antenna provides a cross-track scan, scanning ±48.95° from nadir with a total of 90 Earth IFOV per scan line. This instrument completes one scan every 2.67 seconds. Both sensor spectral bands are reported in tab.2.2.

AMSU-A Channels	
Channel	**Central frequency (GHz)**
ch1	23.8
ch2	31.4
ch3	50.3
ch4	52.8
ch5	53.596±0.115
ch6	54.4
ch7	54.94
ch8	55.5
ch9	57.290
ch10	57.290±0.217
ch11	57.290±0.3222±0.048
ch12	57.290±0.3222±0.022
ch13	57.290±0.3222±0.010
ch14	57.290±0.3222±0.0045
ch15	89.0
AMSU-B Channels	
Channel	**Central frequency (GHz)**
ch16	89.0±0.9
ch17	150.0±0.9
ch18	183.31±1.00
ch19	183.31±3.00
ch20	183.31±7.00

Table 2.2: AMSU-A and AMSU-B spectral channels.

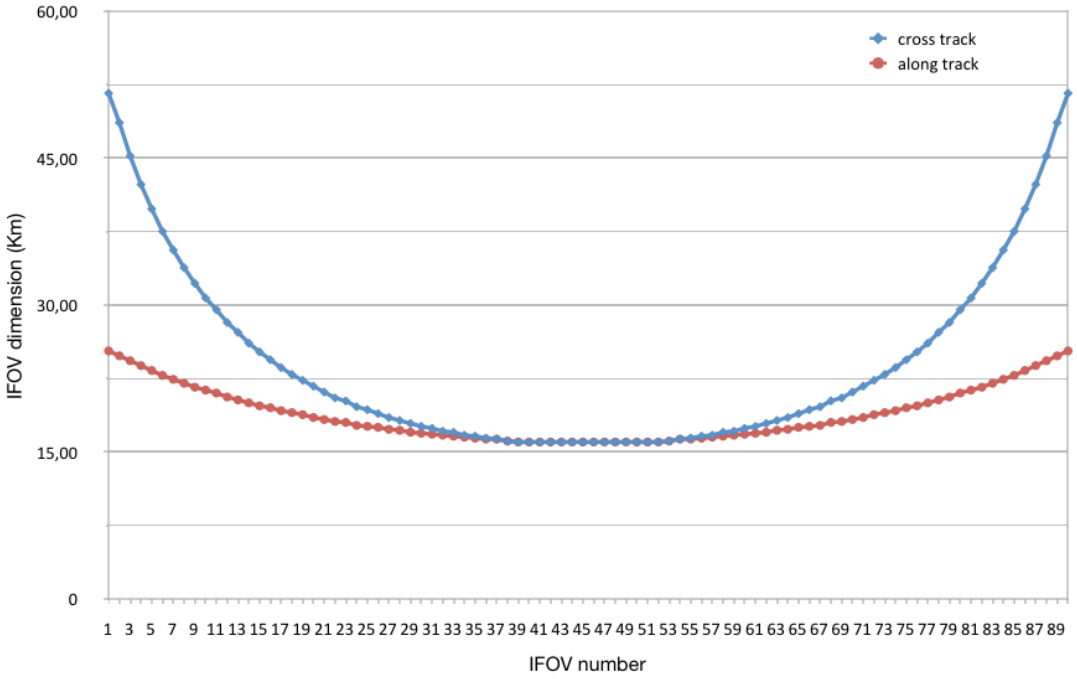


Figure 2.22: AMSU IFOV dimensions along track and cross track.

Chapter 3

Artificial Neural Networks

In this chapter, after an introduction on the use of Artificial Neural Network in satellite rainfall estimation, a basic theory of the ANN will be presented.

3.1 Neural Network in geophysics and remote sensing

The satellite estimate of geophysical parameters is very often the only chance to have information on some process taking place in the Earth system over a wide range of spatial and temporal scales, wiht an accuracy suitable for quantitative analysis. Direct measurements for many parameters of interest are in general not available: even when in situ measurements are available, they are usually sparse and can be at a level of accuracy and reliability difficult to control. The quality of geophysical parameters derived from remote sensing measurements varies significantly depending on the strength and uniqueness of the signal from the geophysical parameters and mathematical methods applied to extract these parameters, i.e. to solve forward and inverse remote sensing problems: neural network techniques are a widely used tool to approach this problem [101].
Precipitation is a clear example of a geophysical parameter highly variable, difficult to measure and nevertheless, crucial for several applications. Moreover, the relationship between remotely sensed electromagnetic signal and precipitation characteristics is unfortunately weak and elusive, as will be shown in the next chapter. We have available today a large number of coincident radiance measurements, at

different wavelengths, from the current satellite systems, each of them carrying in general small information on the precipitation key characteristics: Artificial Neural Networks are often exploited to study these relationships and to establish retrieval techniques [102].
Chen and Staelin (2003) [62] used an ANN with one hidden layer with 5 nodes, using the tan(h) sigmoid function, that accepts in input the radiances of Advanced Microwave Sounding Unit (AMSU) to retrieve precipitation over the U.S., after a training with the ground weather radar data. similar approach was used by Capacci and Porcú (2009) [103] to retrieve precipitation over the British Isles: a Multi Layer Perceptron ANN was implemented with one hidden layer wit ten nodes. As input the radiances of SEVIRI sensor, in geostationary orbit, was used, and the training has been performed on the U.K. radar network.
Geostationary VIS-IR data have been trained by means of TRMM-PR data to retrieve precipitation over tropical ocean by Bellerby et al. (2000) [104]: they used a Multilayer Feed Forward Network with two hidden layers and sigmoidal transfer function. The Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN) is based on a two hidden layers Modified Counter Propagation Network [78]. This system, originally trained with hourly gauges data, applies to geostationary VIS-IR input data and, in its latest version, is able to accept several different inputs and to adapt to different sources of calibration [105].

3.2 ANN introduction

The Artificial Neural Network is a statistical approach based on the human brain mechanism of learning information from experience and apply them to next similar experiences.
The Human Neural Network is composed of neurons, electrically excitable cells that processes and transmits information by electrical and chemical signaling. Chemical signaling occurs via synapses, specialized connections with other cells. A typical neuron possesses a cell body (often called the soma), dendrites and an axon (fig. 3.1). Dendrites are thin structures that arise from the cell body, often extending for hundreds of micrometers and branching multiple times, giving rise to a complex "dendritic tree". An axon is a special cellular extension that arises from the cell body

at a site called the axon hillock and travels for a distance, as far as 1 micrometer in humans or even more in other species. The cell body of a neuron frequently gives rise to multiple dendrites, but never to more than one axon, although the axon may branch hundreds of times before it terminates. At the majority of synapses, signals are sent from the axon of one neuron to a dendrite of another. The axon hillock is the last site in the soma where membrane potentials propagated from synaptic inputs are summated before being transmitted to the axon. If the voltage changes by a large enough amount, an electrochemical pulse, called an action potential, is generated and travels rapidly along the cell's axon. A voltage threshold is in this way beaten and the "neural signal" can be transmitted and activates synaptic connections with other cells.

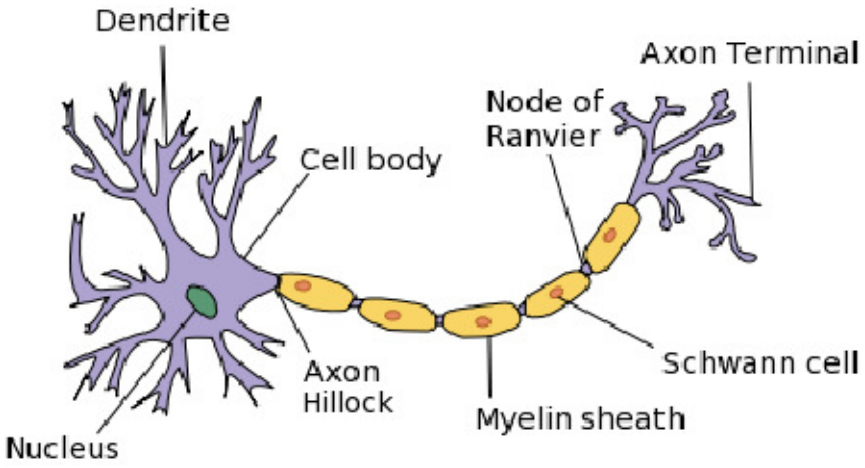


Figure 3.1: Biological neuron schematization.

The Human Neural Network inspired the Artificial Neural Network, where neurons are called perceptrons. Every individual computational element is called node, it receives one or more inputs (representing the one or more dendrites) and sums them to produce an output (representing a biological neuron's axon). Every input is associated to a weight (representing the strength of the synapses) and an activation function (representing the action potential in the hillock axon) is applied to the node in order to obtain the output.

3.3 Single Layer Perceptron

In principle a perceptron can contain an arbitrary number of layers of nodes in addition to the input and output layers. Here we begin with the so-called simple perceptron, the more similar to the biological neuron, where the input layer feeds directly into the output layer through an activation function (see fig. 3.2).

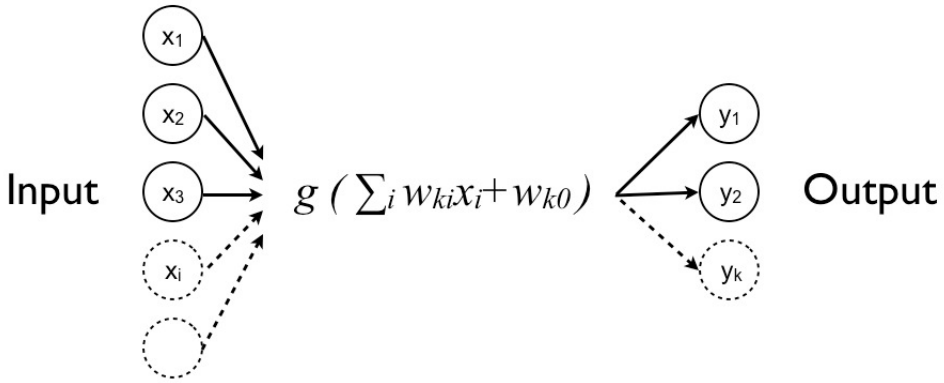


Figure 3.2: Single layer perceptron. g is the activation function with $i = 1, In$ and w_{k0} is the bias term.

Every input x_i $(i = 1, In)$ is multiplied by the weight w_i. The mathematical sum of all those products becomes the argument of the activation function, defined by 3.1.

$$y_k(\mathbf{x}) = g(a_k) = g\Big(\sum_{i=1}^{In} w_{ki}x_i + w_{k0}\Big) \tag{3.1}$$

where $w_{k,0}$ is the bias term (not in a statistical meaning). This bias determines the position of the hyperplane in $\mathbf{x}$-space (fig. 3.3).

It could be convenient to adopt a notation in which input and weights are defined with new $(In + 1)$-dimensional vectors with an extra input variable x_0 (see fig. 3.4) whose value is permanently set at $x_0 = 1$ (eq. 3.2).

$$y_k(x_i) = g(a_k) = g\Big(\sum_{i=0}^{In} w_{ki}x_i\Big) \tag{3.2}$$

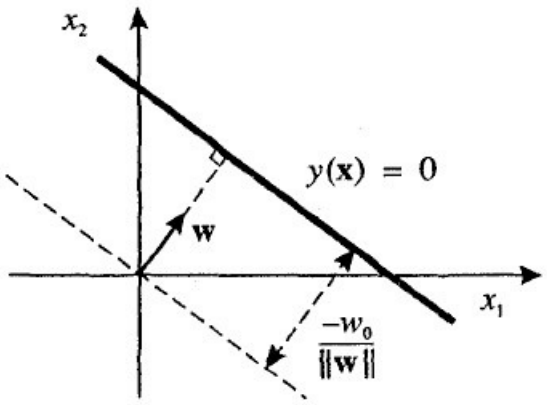


Figure 3.3: A linear decision boundary, corresponding to $y(\mathbf{x}) = 0$, in a two dimensional input space (x_1, x_2). The weight vector $\mathbf{w}$, which can be represented as a vector in $\mathbf{x}$-space, defines the orientation of the decision plane, while the bias w_0 defines the position of the plane in terms of its perpendicular distance from the origin [106].

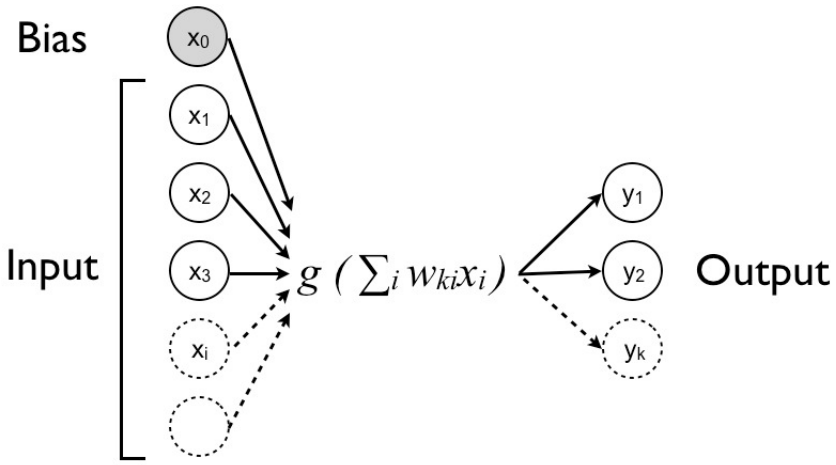


Figure 3.4: Single layer perceptron. g is the activation function with $i = 0, In$.

The activation function is generally chosen to be monotonic, smooth (continuous first derivative) and asymptotic. It can be a step function, a linear function or a sigmoidal function. Usually the logistic sigmoid activation function (eq. 3.3) is the

best choice for ANN because it allows the outputs to be interpreted as posterior probabilities [106].

$$g(a_k) \equiv \frac{1}{1 + \exp(-a_k)} \tag{3.3}$$

Using discriminant functions having a decision boundary which is linear or more generally hyperplanar is however a very restricted choice and we can expect such network to have less than optimal performance for many practical applications.

3.4 Multi-Layer Perceptron

After the starting idea of using fundamentals of the Human Neural Network as base for an Artificial Neural Network, mathematical and statistical exigences force hence to develop the ANN joining several layers of perceptrons.

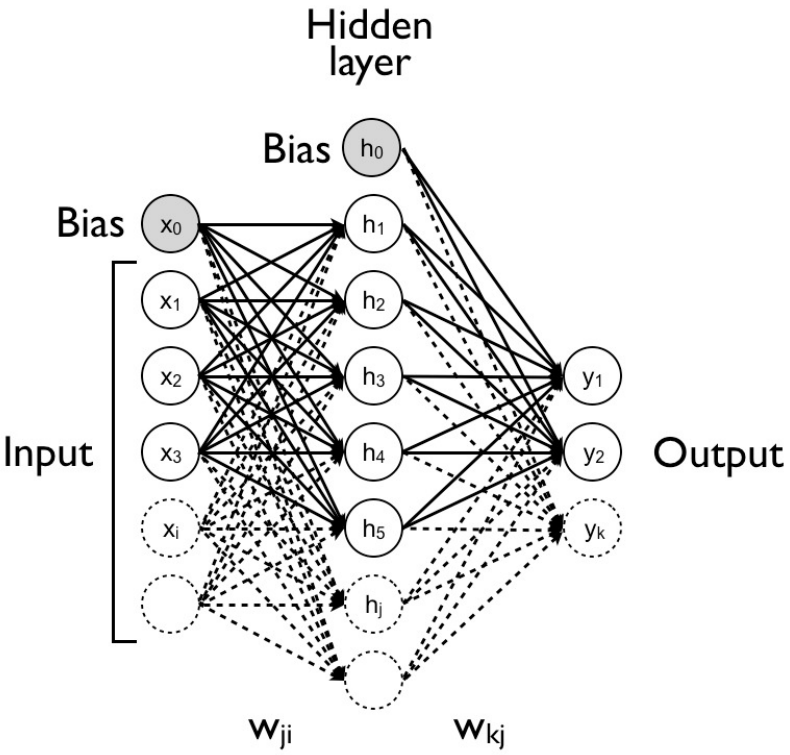


Figure 3.5: Multi layer perceptron.

The way the perceptrons are connected determines the kind of ANN. The multi-layer perceptron (MLP) is a special Network with one or more layers between the input and the output ones (fig. 3.5). These middle layers are called hidden layers and are introduced to solve linearly inseparable problems otherwise not solvable (ex. XOR problem) [106] [107].
The connections allowed in a MLP scheme are only the ones between every node in a layer to every node in the next layer, but not between nodes in the same layer. In addition the diagram must be feed-forward, so that it contains no feedback loops.
The first ANN used in this work is determined by a double application of the activation function. At first the activation function is applied to the linear combination of input and secondly it is applied to the output of the first hidden layer to obtain the output of ANN:

$$a_j = \sum_{i=0}^{In} w_{ji} x_i \tag{3.4}$$

$$\begin{aligned} a_k &= \sum_{j=0}^{H} w_{kj} g(a_j) \\ &= \sum_{j=0}^{H} w_{kj} g\Big(\sum_{i=1}^{In} w_{ji} x_i\Big) \end{aligned} \tag{3.5}$$

$$\begin{aligned} y_k = g(a_k) &= g\Big(\sum_{j=0}^{H} w_{kj} g(a_j)\Big) \\ &= g\Big(\sum_{j=0}^{H} w_{kj} g\Big(\sum_{i=0}^{In} w_{ji} x_i\Big)\Big) \end{aligned} \tag{3.6}$$

The activation function is of the same type described in 3.3. In the contest of classification problems, networks with sigmoidal non-linearities and two layers of weights (input layer and one hidden layer) can approximate any decision boundary to arbitrary accuracy [106].
The number of hidden nodes is determined by the number of inputs: the number of hidden units must grow as the size of the data set grows. If we try to approximate a given function with a network having a finite number of hidden nodes, we will introduce a residual error. This error will decrease increasing the number of hidden units.

3.5 ANN building phases

To make an ANN operative, three building phases are necessary:

1. Data set building phase
2. Training phase
3. Testing phase

3.5.1 Data set building phase

In order to train the Neural Network, it is necessary to provide a supervised data set. An ensemble of input data (features) is prepared and every set of features $(x_1, x_2, \ldots, x_{In})$ is associated to a "true" output value. The value of truth is considered as a "target" ($\mathbf{t}$) and it is necessary in order to "teach" how to interpret input values to the ANN.
The kind of feature $\mathbf{x}$ depends on the problem under examination and the phase where they are chosen and calculated is called features extraction.
Among all the calculated features, there is an ideal group of them which gives the best performance when used as input to the ANN. The departing from this ideal number can lead to a reduction in performances, due to redundant effect of some features or to the "course of dimensionality", which is a poor representativeness of output "true" values if the dimension of the feature space is too high [96]. This ideal group is extracted in the training phase simply trying several combinations of features as ANN inputs and choosing the one giving the best result.

3.5.2 Training phase

The aim of this phase is to define the weights $\mathbf{w}$ in order to obtain an output $\mathbf{y}$ as close as possible to the target $\mathbf{t}$. To obtain the best output, an error function, called *cost function*, is used. Knowing the inputs of the so called *training data set*, the weight are adjusted in order to minimize the cost function which is a differentiable function of the network outputs y_k and the outputs are functions of weights and biases through the activation function g (eq.3.6). If g is differentiable, also y_k are differentiable with respect to the weights and biases and so is the cost function. In

this way, using the differentiability property of the cost function, it is possible to find the set of weights which minimize it.
The function we choose for this work is the sum-of-squares error function:

$$E = \sum_{Out}^{k=1} E_k \tag{3.7}$$

$$\text{in which} \quad E_k = \frac{1}{2} \sum_{n=1}^{N} [y_k(\mathbf{w}, \mathbf{x}^n) - t_k^n]^2 \tag{3.8}$$

Out is the number of output nodes and N is the number of target-features set. The algorithm used to evaluate the derivative of the error function is known as *back-propagation* and consists in a propagation of errors backwards through the neural network [106]. The training algorithm involve an iterative procedure for the minimization of the error function, with adjustments to the weights being made in a sequence of steps (*epochs*). The optimization method used in this work to update weights and find the ones which minimize the cost function is called the *gradient descent technique* and can be summarized in the following equations:

$$w_{kj}^{\tau+1} = w_{kj}^{\tau} - \eta \left. \frac{\partial E}{\partial w_{kj}} \right|_{w^\tau} \tag{3.9}$$

$$w_{ji}^{\tau+1} = w_{ji}^{\tau} - \eta \left. \frac{\partial E}{\partial w_{ji}} \right|_{w^\tau} \tag{3.10}$$

At the step τ the weights are known and they are calculated for the step $\tau + 1$ through equations 3.9 and 3.10. At the beginning ($\tau = 0$) the weights values are randomly assigned. η is a small positive number called the *learning rate parameter*. The choice of the value for η can be critical: if it is too small the reduction in error will be very slow, if it is too large, divergent oscillations can result.

3.5.3 Testing phase

Once defined the best set of weights which minimize the cost function and than which define outputs as close as possible to the targets, a last phase of test is necessary.

The group of features in the input has to be of the same type used in the training phase (with input data associated with targets). With a new dataset (independent from the training dataset) and fixed weights, the output of the ANN is calculated and compared with the expected output. The evaluation is performed using some statistical parameters.

3.6 Interpretation of network outputs

Classification and regression problems are known as supervised learning because they can learn from a sample in which the outputs are known. We talk about regression when we predict quantitative outputs and about classification when we predict qualitative outputs.
It is important to stress that the central goal in network training is not to memorize the training data, but rather to model the underlying generator of the data, so that the best possible predictions for the output vector can be made when the trained network is subsequently presented with a new value for the input vector.

3.6.1 Regression problems

For regression problems the goal is to model the conditional distribution of the output variables, conditioned on the input variables. The most general and complete description of the generator of the data is in terms of the probability density $p(\mathbf{x}, \mathbf{t})$ in the joint input-target space. It is convenient to decompose the joint probability density into the product of the conditional density of the target data, conditioned on the input data, and the unconditional density of input data:

$$p(\mathbf{x}, \mathbf{t}) = p(\mathbf{t}|\mathbf{x})p(\mathbf{x}) \tag{3.11}$$

$$\text{where} \quad p(\mathbf{x}) = \int p(\mathbf{t}|\mathbf{x})d\mathbf{t} \tag{3.12}$$

For the purpose of making predictions of $\mathbf{t}$ for new values of $\mathbf{x}$, we need to model the conditional density $p(\mathbf{t}|\mathbf{x})$.
Considering a network trained by minimizing a sum-of-squares error function, we can derive that the outputs approximate the conditional averages of the target data.

In the limit of N going to infinity, we can replace the sum over patterns in the 3.8 with an integral:

$$E = \lim_{N\to\infty} \frac{1}{2N} \sum_{n=1}^{N} \sum_{k} [y_k(\mathbf{w}, \mathbf{x}^n) - t_k^n]^2 \tag{3.13}$$

$$= \frac{1}{2} \sum_{k} \iint [y_k(\mathbf{w}, \mathbf{x}) - t_k]^2 p(t_k, \mathbf{x})\, dt_k\, d\mathbf{x} \tag{3.14}$$

and remembering the 3.11:

$$E = \frac{1}{2} \sum_{k} \iint [y_k(\mathbf{w}, \mathbf{x}) - t_k]^2 p(t_k|\mathbf{x}) p(\mathbf{x})\, dt_k\, d\mathbf{x} \tag{3.15}$$

Using a conditional averages notation:

$$< t_k|\mathbf{x} > \equiv \int t_k\, p(t_k|\mathbf{x}) dt_k \tag{3.16}$$

$$< t_k^2|\mathbf{x} > \equiv \int t_k^2\, p(t_k|\mathbf{x}) dt_k \tag{3.17}$$

we can rewrite the 3.15:

$$[y_k - t_k]^2 = [y_k - < t_k|\mathbf{x} > + < t_k|\mathbf{x} > - t_k]^2 \tag{3.18}$$

$$= [y_k - < t_k|\mathbf{x} >]^2 + 2[y_k - < t_k|\mathbf{x} >][< t_k|\mathbf{x} > - t_k] + [< t_k|\mathbf{x} > - t_k]^2 \tag{3.19}$$

and substituting the 3.19 in 3.15 we can write the sum-of-squares error in the form:

$$E = \frac{1}{2} \sum_{k} \int [y_k(\mathbf{w}, \mathbf{x}) - < t_k|\mathbf{x} >]^2 p(\mathbf{x}) d\mathbf{x} + \frac{1}{2} \sum_{k} \int [< t_k^2|\mathbf{x} > - < t_k|\mathbf{x} >^2] p(\mathbf{x}) d\mathbf{x} \tag{3.20}$$

The second term on the right-hand side of 3.19 vanishes as a consequence of the integration over t_k. We can also note that the second term in 3.20 is independent

of the network function y_k and is therefore independent of the network weights, so for the purposes of determining the network weights by error minimization, this term can be neglected. Since the integrand in the first term is non-negative, the absolute minimum of the error functions occurs when this first term vanishes, which corresponds to the following result for the network function:

$$y_k(\mathbf{x}, \mathbf{w}^*) =< t_k|\mathbf{x} > \tag{3.21}$$

where $\mathbf{w}^*$ is the weight vector at the minimum of the error function. Equation 3.21 is a key result and says that the network function is given by the conditional average of the target data, in other words by the regression of t_k conditioned on $\mathbf{x}$ [106].

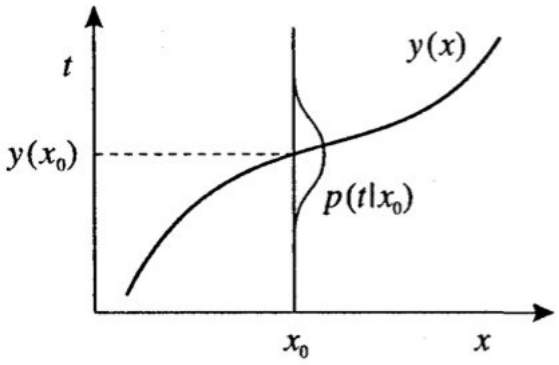


Figure 3.6: A schematic illustration of the property 3.21. Here we consider a mapping from a single input variable x to a single target variable t. At any given value x_0 of the input variable, the network output $y(x_0)$ is given by the average of t with respect to the distribution $p(t|x_0)$ of the target variable, for that value of x [106].

3.6.2 Classification problems

In classification problems the input vector is assigned membership in one of a number of finite groups on the basis of whether they have some property or not. Typically there is one output unit for each possible class and the activation of each output unit represents the corresponding posterior probability $p(C_k|\mathbf{x})$, where C_k is the kth

class and $\mathbf{x}$ is the input vector.
The density function in target space for classification problems can be written as:

$$p(t_k|\mathbf{x}) = \sum_{l=1}^{c} \delta(t_k - \delta_{kl}) P(C_l|\mathbf{x}) \tag{3.22}$$

since $P(C_l|\mathbf{x})$ is the probability that $\mathbf{x}$ belongs to class C_l.
Substituting this last equation in 3.21 and remembering 3.16, we obtain:

$$y_k(\mathbf{x}) = P(C_k|\mathbf{x}) \tag{3.23}$$

which is exactly the Bayesian posterior probability.
If the network outputs represent probabilities, then they should lie in the range (0,1) and should sum to 1. For a network with linear output units, trained by minimizing a sum-of-squared error function, however, there is no guarantee that they will lie in the range (0,1). In fact, the sum-of-squares error function is not the most appropriate for classification problems. It was derived from maximum likelihood on the assumption of Gaussian distributed target data. However, the target values for a 1-of-c coding scheme are binary and hence far from having a Gaussian distribution. However there are advantages in using a sum-of-squares error function, first of all the fact that the determination of the output weights in a network represents a linear optimization problem.
These probabilities can then be used in a subsequent decision-making stage to arrive at a classification. In this work, the decision is taken over a probability rejection threshold value, in others words, if the posterior probability falls below this threshold, then no classification decision is made or, dealing with a binary classification problem, the threshold divides the two classes.
For a two-class problem, as we will deal to in the first part of this thesis, we can consider a single output y and a target coding which sets $t^n \geq threshold$ if $\mathbf{x}^n$ is from class C_1 and $t^n < threshold$ if $\mathbf{x}^n$ is from class C_2 [106].

Chapter 4

Multi-sensor satellite precipitation estimate algorithms: implementation and results

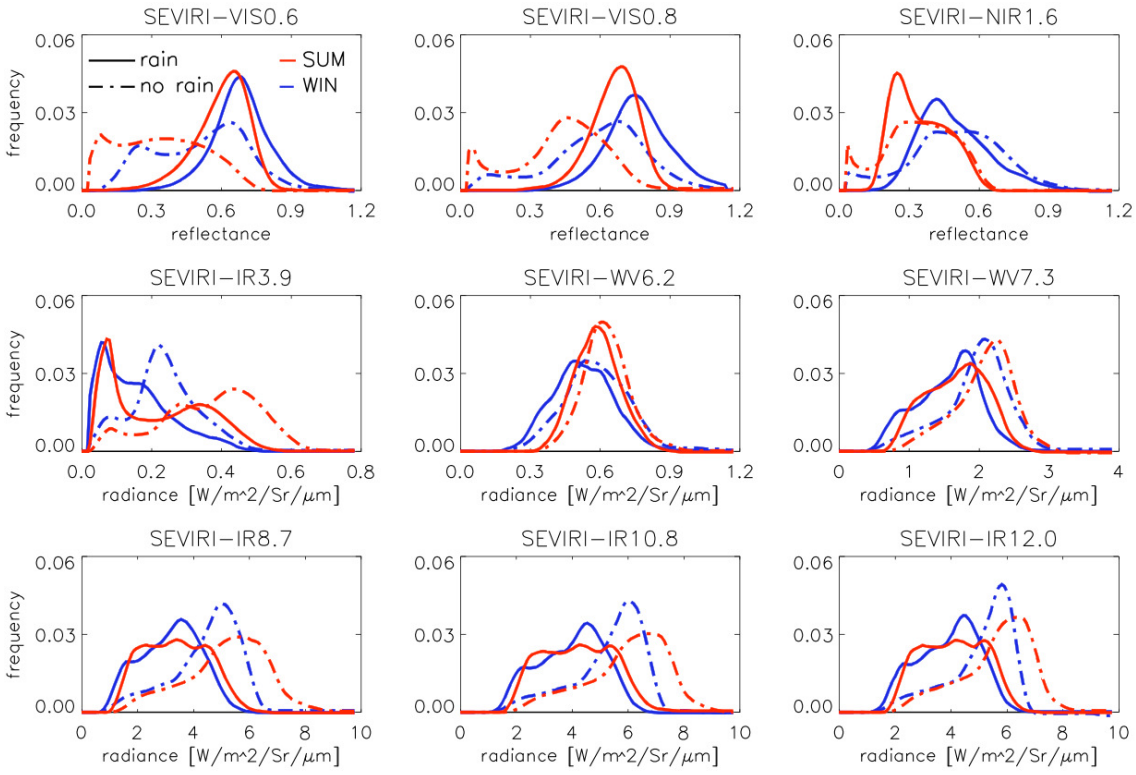
In this chapter is presented a detailed description of the multi-sensor satellite precipitation estimate algorithm developed at the University of Ferrara (Department of Physics - Atmospheric Physics Group) within the PROSA Project, which is the major objective of my PhD Thesis work. The algorithm presented here is an adaptation of the original algorithm developed for United Kingdom area [108] and consists in an Artificial Neural Network designed to find the best correlation between satellite data and precipitation at the ground. The original algorithm and all the further implementations will be described in this chapter.

4.1 The original algorithm

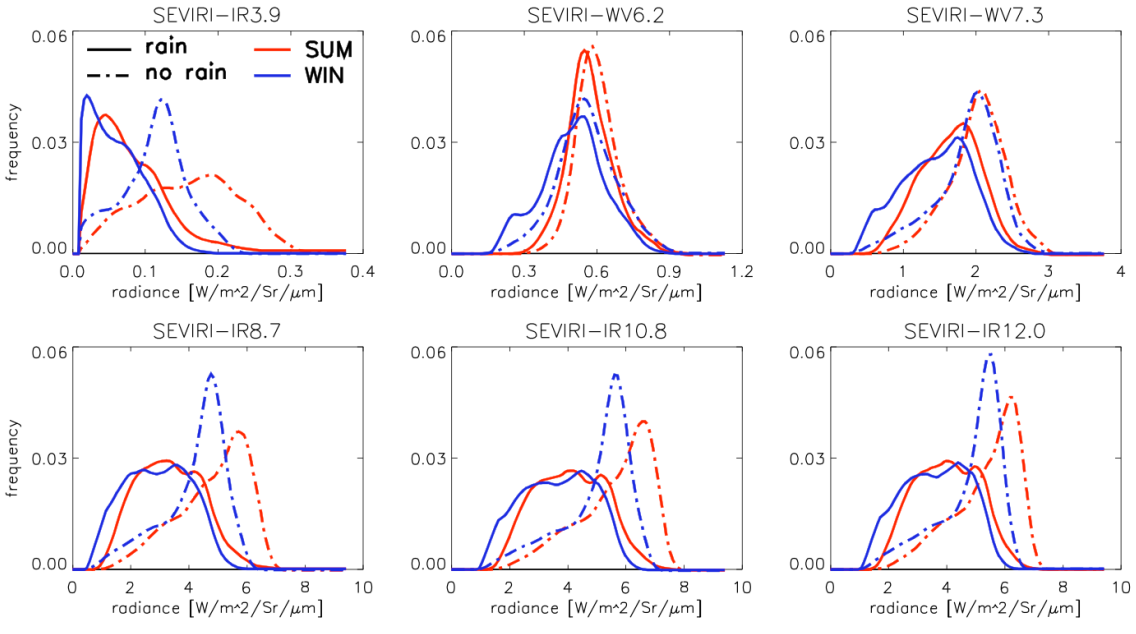
The original idea comes from an algorithm developed for MODIS sensor data at first [108] and for SEVIRI data in a second version [103] and optimized for working over British Isles area.
As described in detail in sec.2.2, the relation between radiances and precipitation at the ground in the VIS-IR bands is very weak and difficult to be physically interpreted and modeled. For this reason a statistical approach is used to find a relation between radiation coming from the top layers of the cloud and measured by satellite

sensors, and the precipitation at the ground as measured by independent instruments.
The algorithm basis consists in an Artificial Neural Network of Multi Layer Perceptron type (described in detail in section 3.4). The ANN is a two layers Network, with an input layer, one hidden layer and an output layer. The hidden layer has 10 hidden nodes and the activation function is sigmoidal (eq. 3.3). The ANN is set up as a classification problem and gives as output a probability of belonging or not to a defined class of precipitation (value 0 for no-rain class and 1 for rain class).
After the assessment of the seasonal dependence of the algorithm [48], it was separately developed for winter and summer seasons.
The original algorithm was set up to work on the UK region and only during daytime, in order to use all satellite sensors available channels (also in the VIS-NIR band, not available during night-time). The training had been carried out with the ground data of the UK network meteorological radar as precipitation *ground truth* [109]. Radar data are particularly appropriate for this type of task because they are areal and instantaneous precipitation estimate, more readily comparable with satellite data also areal and instantaneous radiance measurements.
Three summer months (JJA: June, July and August 2004) and three winter months (DJF: December 2004, January and February 2005) had been used as training dataset and the Equitable Threat Score (ETS) had been used as statistical parameter for training testing and validation. ETS and others statistical parameters used in this chapter will be described in detail in appendix A.
As a preliminary study during the algorithm developing, the sensitivity of SEVIRI input channels to the precipitation at the ground had been carried out, separating day-time performances from night-time ones and for different seasons. For each SEVIRI channel, satellite radiances are compared with nearly simultaneous, co-located radar rainfall maps, and each pixel is classified as rain or no rain. With the term nearly simultaneous we define two observations from different sensors of the same meteorological phenomenon carried on with a time lag small if compared to the typical time scale of the observed phenomenon. The definition of the maximum time lag allowed to compare two observations, in priciple, depends on the spatial resolutions of the two observing systems (the smaller is the resolution, the shorter is the allowed time lag) and on the considered cloud type (shorter time lag for convective clouds).

The frequency of raining and non raining pixels as function of the satellite radi-



(a)



(b)

Figure 4.1: Sensitivity study for the UK-ANN, trained with UK radar network. Daytime (a) and night-time (b) SEVIRI measurements distributions are plotted for every channel considered. Curves are separated for seasons, red lines represent summer distributions and blue ones represent winter distributions. Solid lines represent rain estimates distributions and dash-dot lines represent no-rain estimates.

ance values are plotted in figures 4.1 for reference, since the same sensitivity studies were also been performed for the new technique development. The figure reads as follows: the highest is the separation between corresponding blue and red curves, the highest is the sensitivity of the channel to the precipitation at the ground. It can be noted that the channel sensitivity is highly variable with channel number, seasonal and diurnal cycle.
Generally, training and validating dataset have to be independent, in order to estimate the real capability of the network to generalize its performances after being trained. For this reason the whole dataset had been separated into two sub-sets, one for the training phase and one for the validation one.
In fig.4.2 are shown the ANN performances for the studied periods: it is possible to note the different values of ETS for the two seasons, separately studied, considering that the best value for ETS is 1. Best performances are obtained in summer season with values of ETS reaching 0.6, while during the winter ETS never exceeds 0.5. For both seasons the ETS is always above zero, which represents the no-skill value for this parameter (see appendix A).
In fig.4.3 an example of ANN output is presented: the ANN derived Probability of Precipitation (PoP) image (on the right) is compared to the nearly simultaneous radar observation (center) and the correspondent SEVIRI. Here the rain flag on the PoP image is assigned to those pixels with PoP greater than 0.5, and the resulting ETS is around 0.42: all the major rain areas detected by the radar network are correctly identified by the algorithm, with some overestimation.
After the set up of ANN for the UK region, some tests had been made applying the ANN to other areas. An example is shown in fig.4.4 with the application of the UK-ANN to the Italian Friuli Region, where the Fossalon di Grado weather radar data has been available: for the reported maps the ANN algorithm shows some overestimation of the rain areas and the overall performances of the technique for the considered event over Italy are slightly lower than the performances over the UK, with averaged ETS around 0.35 [110].

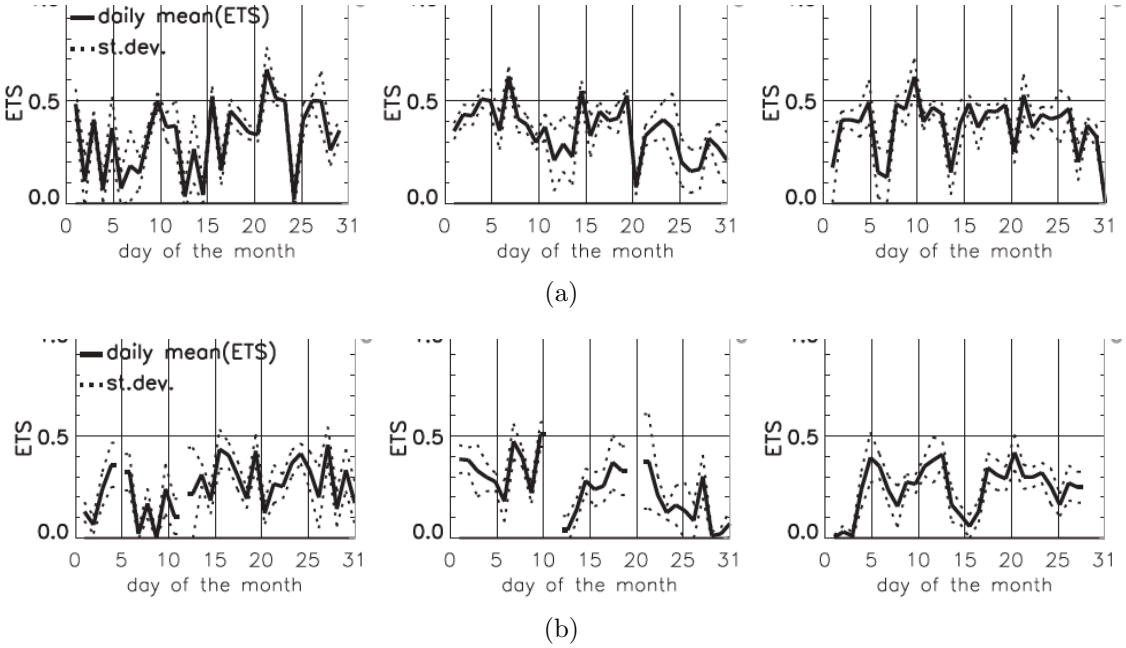


Figure 4.2: ETS trend for summer (JJA 2004) (a) and winter (DJF 2004-2005) (b) seasons. The gaps in winter graphics are due to an absence of precipitation within the radar coverage.

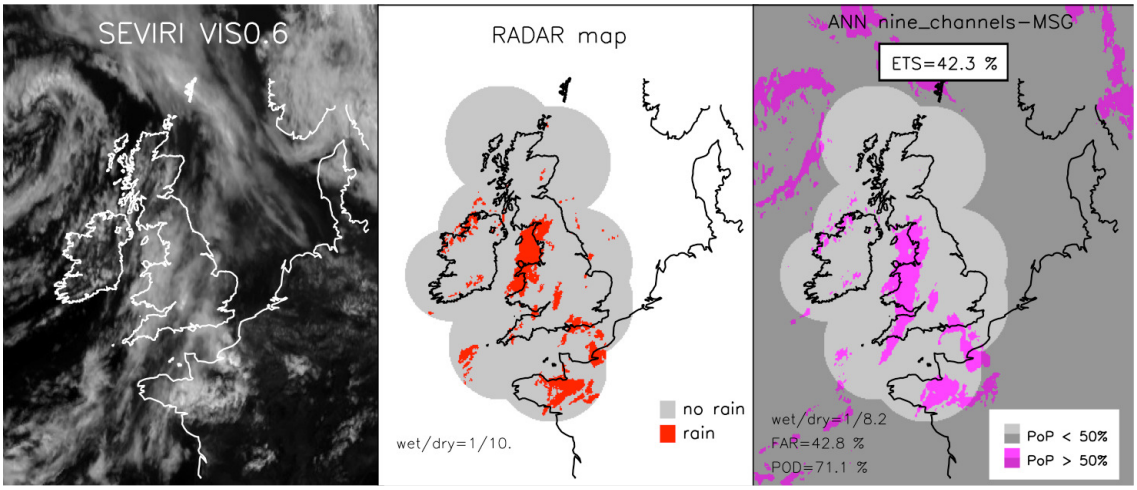


Figure 4.3: 22/07/2004, 14.27 UTC case: SEVIRI image on the left, radar rain-no rain map in the middle and ANN rain-no rain map on the right [111].

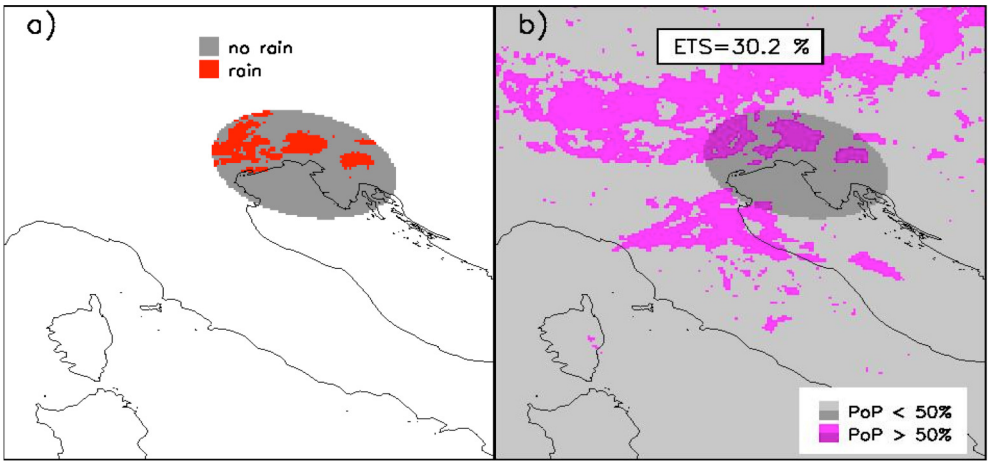


Figure 4.4: 05/05/2004, 16.57 UTC case: Fossalon di Grado (GO, Italy) radar map on the left and ANN classifier on the right.

The logical scheme of the algorithm can be summarized as follows:

- creation of a supervised dataset, with ground *true* precipitation data and corresponding (spatially and temporally) satellite data;

- separation of the supervised dataset in two different sub-datasets (training-testing dataset and validation dataset) using as criteria the division of odd day number from even ones;

- separation of the training-testing dataset in two different sub-dataset (training dataset and testing dataset), randomly picking up a fixed equal number of pixels for each dataset;

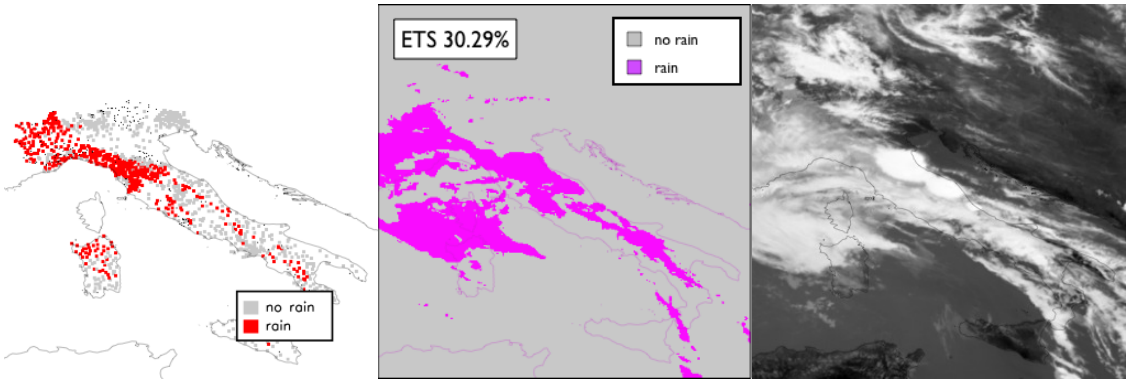
- training of ANN, establishing the ANN weights (w_{ji}, w_{kj}) trough the back propagation algorithm, iterating the process a number n of times (epochs);

- testing the ANN, choosing the weights set which gives the best statistical parameters (see appendix A);

- (optional) validate the ANN using the validation dataset to evaluate the ANN generalizing capability.

4.2 PROSA Product 1

This first algorithm developed for the Italian area, represents the first product of the 'Multi-sensor satellite precipitation estimate' of the PROSA system. Every 15 minutes, a new rain-no rain map is made available to the system few tenths of seconds after the satellite images are available on the PROSA archive. It has to be mentioned that the algorithm output format had to be adapted to fit all the technical requirements imposed by the the PROSA system. In fig.4.5 an example of the algorithm output available from the PROSA website.



(a) (b) (c)

Figure 4.5: 29/05/2008, 12.27 UTC case: precipitation map from rain-gauge network (a), estimation map from ANN (b) and SEVIRI IR12.0 image (c).

This product uses ANNs trained with the UK radar network data and for this reason its performances are not optimized for the Italian area. It presents a couple of problems individuated during a systematic monitoring of the outputs during the demonstration phase of the PROSA Project:

1. First of all, the fact of using two different ANN for day-time and night-time, with different weights and thresholds, leads to an evident discontinuity between the two subsequent precipitation maps produced by the two different ANN. Actually, this anomaly can not be eliminated because of the intrinsic difference between the input available to the two ANN.

2. The second problem is about the wrong classification of snow at ground, often considered as precipitation, mainly during night-time, in which the VIS-NIR

channels can not help for the identification (see an example in fig.4.6). The cold snow temperature on the surface makes the ANN to misunderstand the real non precipitating situation, producing many *false alarm* pixels. This anomaly can be solved with an optimized training of the ANN for the region of interest.

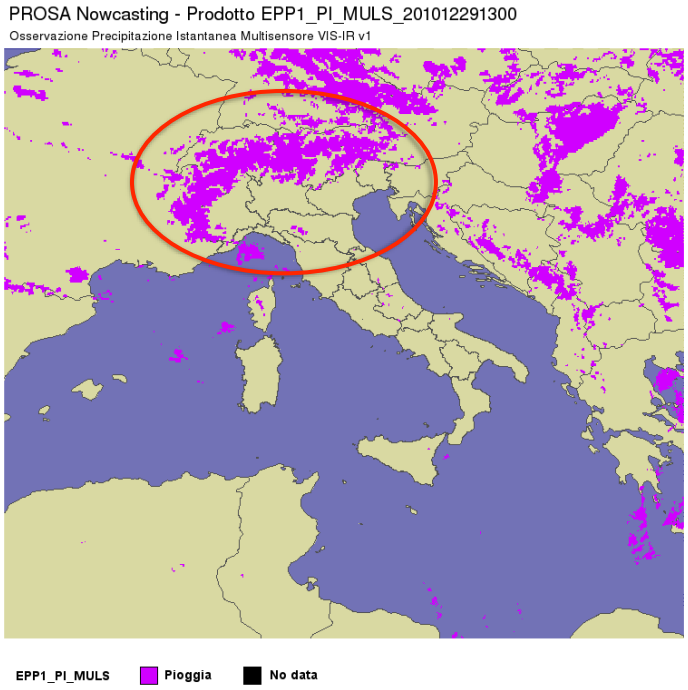


Figure 4.6: 29/12/2010, 13.00 UTC case: example of wrong classification of snow at ground over the Alps region.

4.3 ANN training over Italy

In 2009, I started my PhD program at the University of Ferrara in the Group of Atmospheric Physics and my first task was to adapt the UK-algorithm to the Italian area. After a first test application of the UK-ANN to some Italian regions (i.e. Friuli Region) and then to the whole Italian area with the first release of PROSA, the idea of training the ANN with Italian ground data was considered and a new supervised dataset with Italian rain-gauges rain-rate measures was built.

The use of rain-gauges as *true* values for an ANN training cannot be considered the absolute best choice. As described in sec.2.1.1, rain-gauges are not homogeneously distributed over the surface and they are not representative of large areas because of the small size of the catching area. Furthermore, in the Italian network rain-gauges measure hourly cumulated precipitation values and are difficultly comparable with instantaneous satellite measures (in tab.4.1 a comparison between rain-gauges, radars and satellite measurement characteristics). However, at the time when this work has started, the italian radar network was far to be fully operational and to be reliable enough to constitute a true field for training and validation purposes over the whole Italian Region. For this reason, despite all the described disadvantages, the rain-gauges network was assumed as the best choice for this work as *true* rain-rate values.

	Rain Gauge	**Radar**	**Satellite**
sampling area	30 dm^2	1×1 km^2	3.5×4.5 km^2
instrument sensitivity	0.2 mm	continuous data	continuous data
revisit time	60 min	15 min	15 min
sampling time	60 min	∼1 s	∼1 μs

Table 4.1: Instruments comparison

The pluviometric dataset for the Italian training is made available by the Italian DPC and consists in a variable number of rain-gauges (between 1500 and 1700) and are distribute as shown in fig.2.1. The number of operational stations greatly varies on hourly, daily and monthly basis, because of malfunctions or discontinuity in the data measure, transfer or storage. A relevant characteristics of the spatial distribution of the Italian network is shown in fig.4.7, where the distribution of the minimum distance between closest stations is reported: most of the stations have a minimum distance ranging between 6 and 10 km from the closest station. This distance can be assumed as a measure of the rain gauges network density.

To construct the supervised dataset with spatially and temporally correlated satellite radiances and *truth* values, it was necessary to establish a criterion to spatially and temporally match rain-gauges and satellite observations. Spatial matching was obtained by projecting rain-gauges position on the SEVIRI grid, as computed by routines provided by EUMETSAT; the rain-gauges coordinates were provided by the

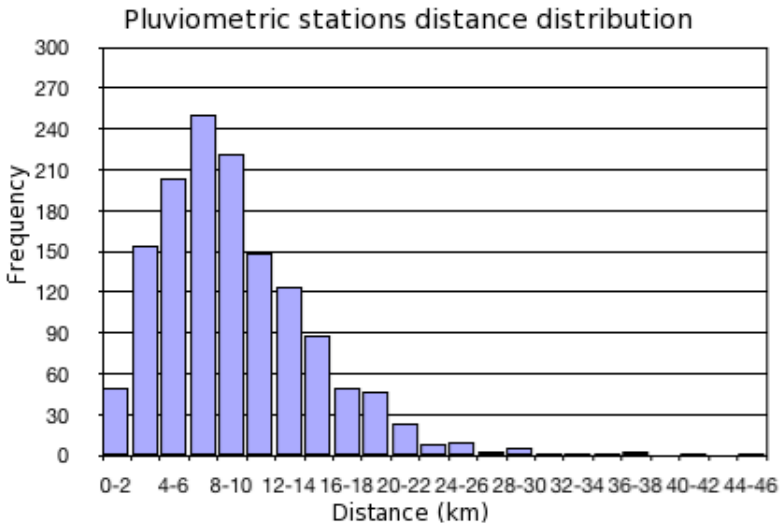


Figure 4.7: Minimum distances distribution between Italian pluviometric stations.

DPC and the World Geodetic System 84 (WGS84) was used for ellipsoid reference values. As for the temporal matching, we know that the hourly cumulated measures of rain-gauges have to be compared with four instantaneous reflectance/radiance every hour. To better understand how to consider this temporal mismatching, two possible solutions had been investigated through a channels sensitivity study:

1. consider only the central satellite slot data (hh:27) and compare them to the hourly cumulated rain-gauge data of the successive hour (hh+1:00);

2. consider all the four slots (of minutes 12, 27, 42, 57), make a distribution of radiance/reflectance for every channel and define a minimum value of reflectance/radiance variation allowed to a pixel to be considered in further processing. A new training dataset is then built, considering only the samples in which the reflectance/radiance variations are under the fixed threshold value.

Two reflectance/radiance distributions for every channel were then made, one for the first solution (fig.4.10) and one for the second (fig.4.8). In fig.4.9 it is shown a particular comparison between the two techniques for one VIS and one IR channel.

It is clear, both for VIS and IR channels, that considering constant reflectance/radiance the raining and no-raining pixels are better resolved. This could lead to

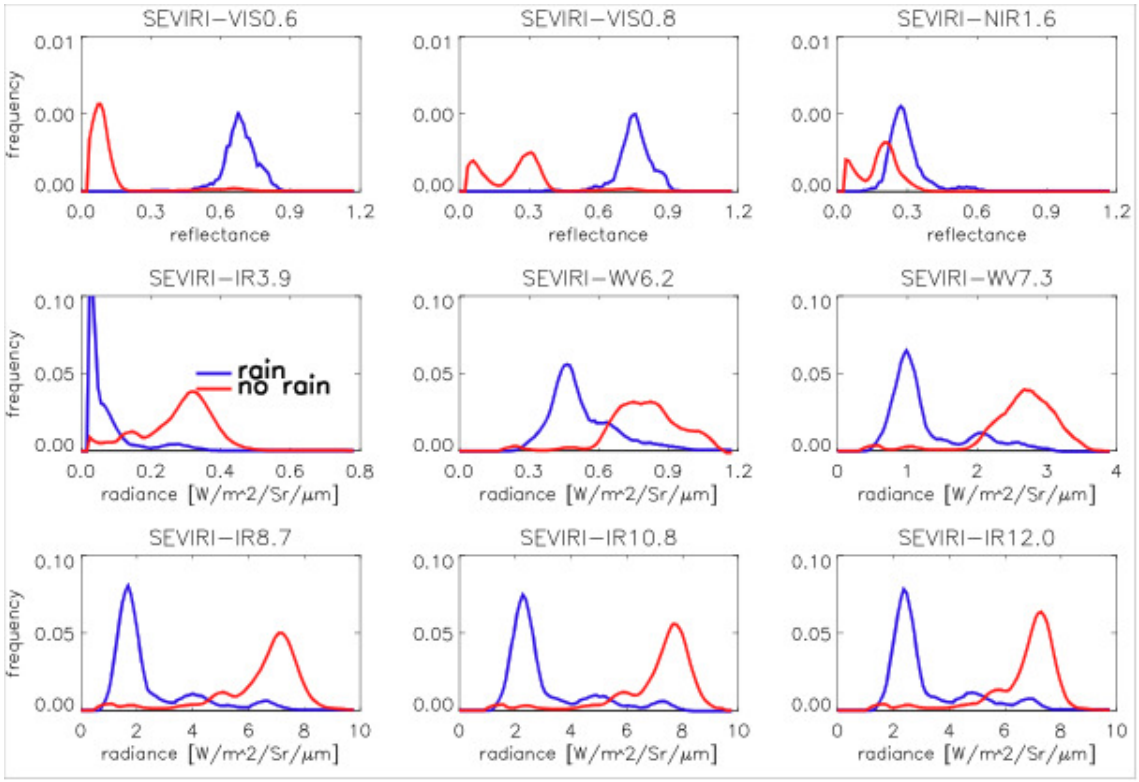


Figure 4.8: Sensitivity study for the Italian-ANN, trained with pluviometer Italian network and reflectances/radiances considered hourly constant. Day-time (a) and night-time (b) SEVIRI measurements distributions are plotted for every channel considered.

choose the constant-reflectance/radiance approach for our aims, but this option has actually more disadvantages than advantages. Considering only the constant radiances/reflectance, the dataset would be drastically reduced, with a consequent reduction in statistic significance. But the main problem is the assumption, using this type of approach, that the cloud structure stays constant during the whole hour, loosing all the rapidly varying systems, as, for examples, thunderstorms.

For these reasons, the Italian ANN uses 10 inputs (9 SEVIRI channels and 1 solar zenith angle) and is trained with hourly rain-gauges data compared with the central slot (hh:27) of SEVIRI measures. The precipitative events considered in the training are representative of a wide typology of meteorological situation, in order to make the network better generalize the further situation it will encounter. The

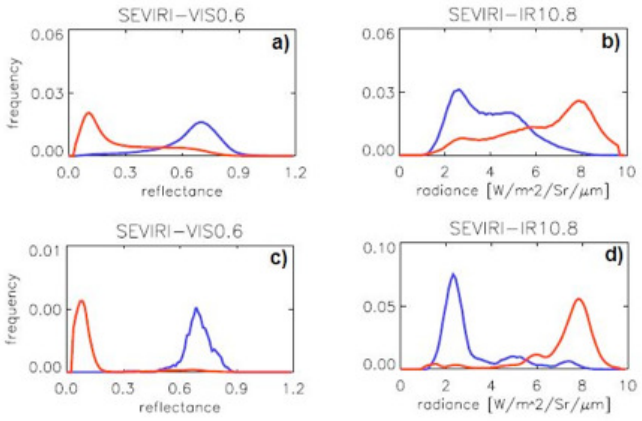


Figure 4.9: Sensitivity comparison between two SEVIRI channels, one in the VIS band and one in the IR. (a) and (b) represent the reflectances/radiances distributions obtained considering for the training dataset only the hh:27 slot. (c) and (d) are the correspondents for the constant reflectance/radiance cases.

selected dataset of about 10^6 pixels is divided in summer and winter events, to train the two seasonal separated ANN, and is here summarized:

Summer cases: 17-19/04/2008, 26-30/05/2008, 11-14/06/2008, 16-18/06/2008, 06-08/07/2008, 12-15/07/2008

Winter cases: 21-24/03/2008, 13-15/01/2009, 23-27/01/2009

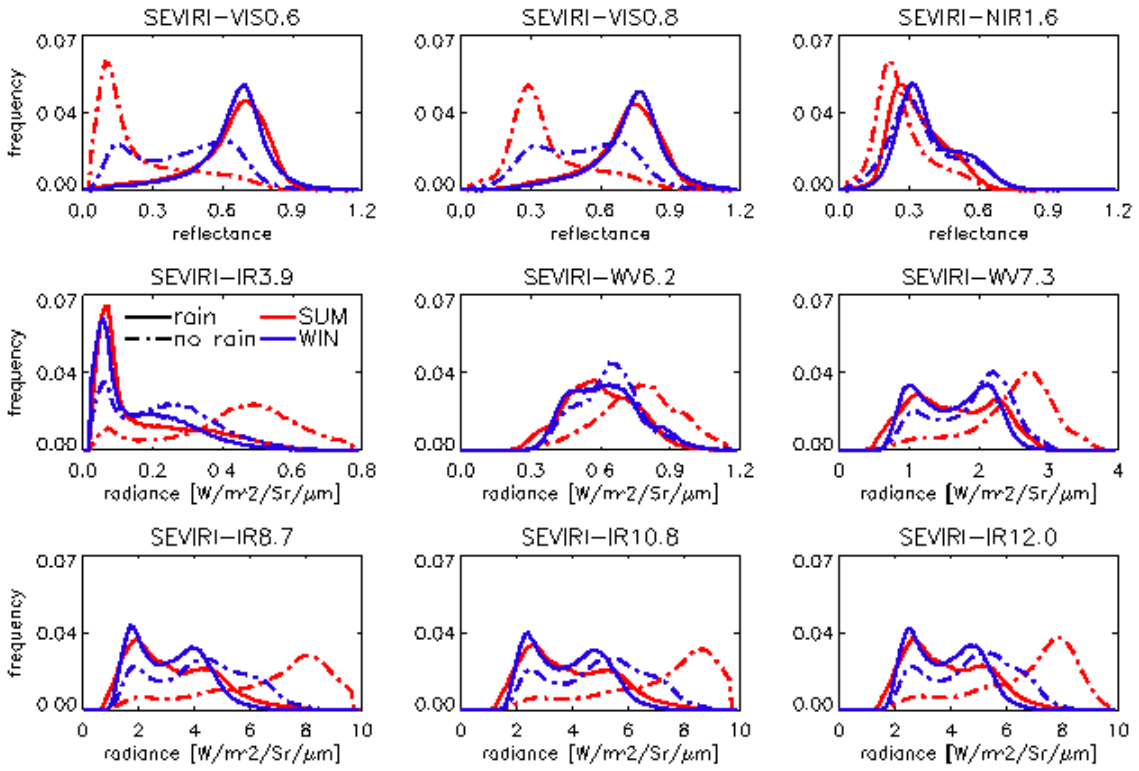
From this total dataset, two smaller datasets are obtained, one used for the training-testing procedure and one for a preliminary validation of the product. The two groups are separated putting the odd days cases in one and the even days cases in the other.
The training-testing phase is done following the procedure described in Capacci and Porcú (2009) [103]. It consists in a random choice of pixels to divide training (50%) and testing (50%) datasets. Then the ANN is trained and tested in order to find the best choice of weights (see sec.3.5.2 and 3.5.3) using 50 epochs. This procedure is repeated many times with different probability thresholds in order to find the one which better classify raining and no-raining pixels (see sec.3.6.2 for the use of thresholds in classification problems).
The validation phase is important to understand the real capability of the network to

generalize what it learnt in the training-testing phase. The ANN is then applied to an independent dataset and statistical parameters are considered (results are shown in tab.4.6).

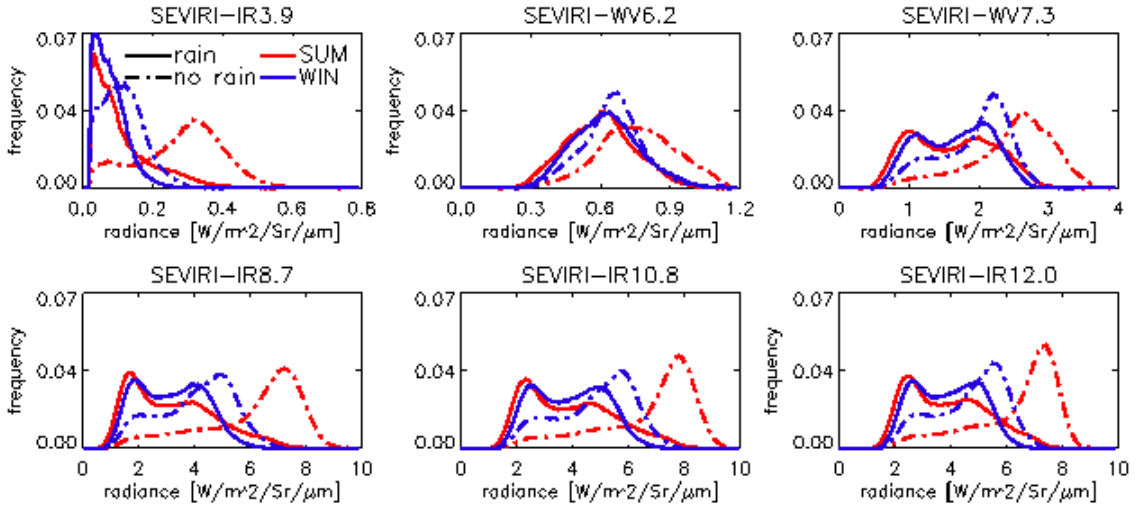
4.4 Night-time ANN

The Italian ANN is set from the beginning to be used 24 hours a day. A night-time ANN is studied and implemented, and the ANNs of the Italian system are thus four (summer-day, summer-night, winter-day, winter-night).
The separation between day-time and night-time networks is automatically performed by a test on the solar zenith angle: the night-time ANN is used when the zenith solar angle is larger than 85° for more than 5% of considered pixels or is not available. Input radiances are in this case 6: VIS-NIR data (0.65, 0.85 and 1.6 μm) and zenith solar angle are not available during night-time. In addition, the 3.9 μm channel, in night-time case, does not contain the reflection contribution and for this reason we have a different behavior from it in the ANN.
VIS-NIR channels give important informations about particles effective radius in the cloud top layers [98] and about cloud optical thickness. These two parameters are linked to the cloud structure and show a good correlation with precipitation areas. Another important factor to be considered is that the reflection channels can help in the understanding of ambiguous situations. Cirrus are very cold clouds, but being thin they are not precipitating. Considering them only in the IR channels, there could be a misunderstanding of the real situation due to their brightness temperature, referring to them as precipitating clouds and overestimating raining areas. On the contrary, low and so relatively warm clouds can be thick enough to be precipitating, but, because of their relatively high brightness temperature, they can be missed using only IR channels. In general during night-time, being the Earth surface colder then in day-time, the clouds-surface thermal contrast is lower and performance of ANN can reduce (see sec.2.2.2, 2.2.4 and 2.2.5 for physical details).
As for the UK-ANN, also for the Italian-ANN, with a rain-gauges training, a channel sensitivity study is made. The resulting distributions are shown in fig.4.10.

Comparing day and night-time distributions, the IR channels show almost the same trend, whereas the big difference can be observed in the 3.9 μm channel. In the night-time distribution, because of the scattered radiation lack, radiances values



(a)



(b)

Figure 4.10: Sensitivity study for the Italian-ANN, trained with Italian pluviometers network and considering the hh:27 slot of SEVIRI data. Day-time (a) and night-time (b) SEVIRI measurements distributions are plotted for every channel considered. Curves are separated for seasons, red lines represent summer distributions and blue ones represent winter distributions. Solid lines represent rain estimates distributions and dash-dot lines represent no-rain estimates.

are smaller and, more important, the information on the cloud top particles r_{eff} is missing. Furthermore, there is a better separation between rain and no-rain for the summer season, because of the good capability of this channel to emphasize cloud-surface temperature differences (greater in summer season).
Generally comparing this last sensitivity study with the one made for UK-ANN (fig.4.1), we cannot appreciate marked differences. Both distributions are normalized, so the different number of samples used in the two trainings does not constitute an element of difference, also if the lower samples number represents a lower statistic. On the contrary, focusing on single channels, some differences can be noted. In the two visible channels (VIS0.6 and VIS0.8), the Italian-ANN can better resolve the difference between raining and no-raining pixels and this difference is greater in summer season (red lines). This behavior strongly depends on the solar zenith angle, smaller at Italian latitudes (37-47°N) than at UK ones (around 55°N) and in summer season than in winter one. At Italian latitudes we expect a greater occurrence of convective systems (better resolved at VIS-IR wavelengths) and, as said before, a smaller solar zenith angle, so with a greater dynamics in VIS channels which strongly helps in the rain-no rain discrimination. Again some differences are noted for the 3.9 μm channel. Better performances are observed for the Italian ANN in the summer season, mainly in the day-time distribution where the scattering contribution can profit of the smaller zenith angle at Italian latitude, with consequent better performances. On the contrary, in winter season, the UK-ANN presents better performance. One of the possible reasons is the use of rain-gauges as training instruments for the Italian ANN: as an example, in case of snow events, pluviometers usually underestimate the real precipitation rate, giving false rates when the snow melts.
In tab. 4.6 are shown the validation results divided for the four different networks considered.

Network	**ETS (%)**	$\sigma_{\mathbf{ETS}}(\%)$
Summer day	35	14
Summer night	19	16
Winter day	30	9
Winter night	10	7

Table 4.2: ETS values and variances for the first Italian ANN.

As we expected, the night-time ANN gives lower values of ETS with respect to the day-time one, due to the lack of VIS-NIR data. The dispersion of the ETS data relatively increases for nigth-time, indicating higher variability of the performances, and also the possibility to have ETS values lower than zero for some slot. The quality of the estimate decreases for both day and night moving from summer to winter due to different precipitation types and higher solar zenith angle.

4.5 Classes of precipitation intensity

A further development of the Italian algorithm consists in the introduction of precipitation classes as ANN output instead of the simple precipitation probability. The basic ANN is the same as for the PoP extraction, but in this case it is used a cascade scheme of several ANNs in order to separate pixels belonging to different precipitation classes. As described in Capacci and Porcú (2009) [103], a first ANN separates dry pixels (no rain) from wet ones (rain) in the same way of the precedent algorithm version. After that, wet pixels satellite values are used as input of a second ANN, which separates slight rain pixels from others. This scheme is repeated 4 times, in order to obtain 5 classes of rain-rate. In fig.4.11 it is shown a schematic description of the cascade system.
Each ANN is trained separately, with the same procedure: the weights of the selected ANN are those that maximize the ETS, that in this case measures the capability of the ANN in correctly separate pixels between precipitation classes. The choice of rain-rate classes limits is made on the Probability Density Function (PDF). The continuous probability parameter has to be divided in such a way to guarantee almost the same population in all classes, in order to have enough cases statistic to train the networks. Starting from a lower limit of 0.2 $\mathrm{mm\,h^{-1}}$ (the sensitivity of most tipping bucket rain-gauges in the Italian network), classes are chosen following a power law (see fig.4.12) and are summarized as follows:

Class 0 (no rain): $0.0 \leq \mathrm{rr} \leq 0.2\ \mathrm{mm\,h^{-1}}$
Class 1 (slight rain): $0.2 < \mathrm{rr} \leq 0.6\ \mathrm{mm\,h^{-1}}$
Class 2 (slig/mod): $0.6 < \mathrm{rr} \leq 2.4\ \mathrm{mm\,h^{-1}}$
Class 3 (moderate rain): $2.4 < \mathrm{rr} \leq 9.6\ \mathrm{mm\,h^{-1}}$
Class 4 (heavy rain): $\mathrm{rr} > 9.6\ \mathrm{mm\,h^{-1}}$

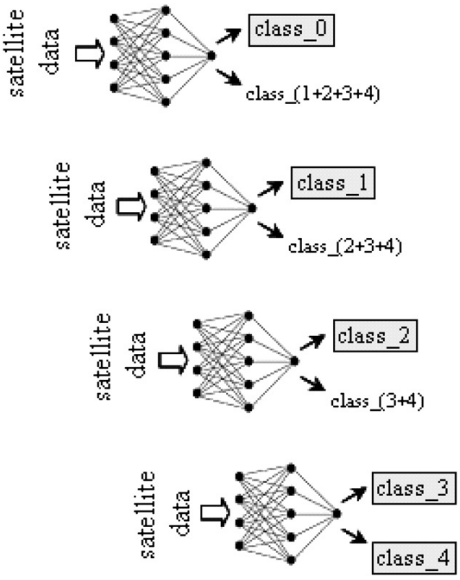


Figure 4.11: Schematic description of the ANN cascade system which divides the precipitation pixels in 5 different classes.

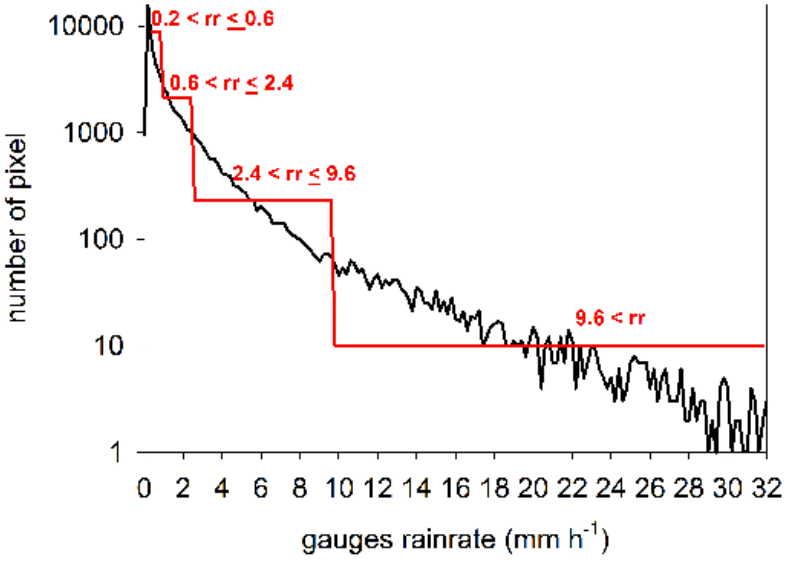


Figure 4.12: Probability Density Function (PDF) of precipitation retrieved from rain-gauges (black line) and precipitation integration in the five classes chosen (red line). The PDF refers to April, May, June and July 2008 Italian pluviometers network data.

The dataset results then divided in the five classes as resumed in tab.4.3 for the training-testing dataset and in tab.4.4 for the validation dataset:

Network		Class 0-(1,2,3,4)	Class 1-(2,3,4)	Class 2-(3,4)	Class 3-4
Summer day	tot. 137290				
	wet	8645	6130	2396	403
	dry	128645	2515	3734	1993
Summer night	tot. 185763				
	wet	11512	8398	3393	403
	dry	174251	3114	5005	2990
Winter day	tot. 44702				
	wet	10440	7259	1796	41
	dry	34258	3185	5463	1755
Winter night	tot. 104250				
	wet	17849	11999	3218	123
	dry	86401	5850	8781	3095

Table 4.3: Training-testing dataset: for every season and daytime, it is indicated the number of pixels considered for every ANN. The first ANN divides class 0 pixels from others (first column), the second divides class 1 from 2-3-4, the third separates class 2 from 3-4 and the last one class 3 from 4. The total number of pixels considered for every season/daytime is also reported.

Network	tot.	Class 0	Class 1	Class 2	Class 3	Class 4
Summer day	87805	78980	2321	3699	2309	496
Summer night	117775	107852	2660	4066	2659	538
Winter day	60983	51128	2669	4950	2055	181
Winter night	143763	123947	6772	9095	3653	296

Table 4.4: Validation dataset: for every season and daytime, it is indicated the number of pixels considered for every class. The total number of pixels considered for every season/daytime is also reported.

In general, also with this type of algorithm, mainly during night-time, the precipitation amount is overestimated because of the misunderstanding of non-precipitating cold clouds as precipitating. In this case the estimation error, made for the first class is propagated to the others because of the cascade architecture of the algorithm, with consequent mistakes in the rain-rate classification.
An other problem of this algorithm version, is the low representativeness of the dataset. Although the efforts made for the rain-gauges dataset division, the high rate events (with $rr > 9.6\,mm\,h^{-1}$) were not enough for the statistic. Looking tab.4.3 it is clear that the ANN separating the last two classes (last column) has a number of samples very low with respect to the others and with respect to the total number of samples. The ANN, for this reason, was not able to complete the training in most of the cases. The summer-day ANN was the only one completed, but probably the ANN weights are not considered reliable for the work. Others ANN did not complete the training phase and are for this reason without the highest class (only $rr > 2.4\,mm\,h^{-1}$ are considerd). Statistical parameters about the training-testing phase are resumed in tab.4.5 (see appendix A for detailed descriptions about statistical parameters).

A preliminary observation about the capability of satellite data to discriminate among rain-rate classes can be made considering the SEVIRI channels sensitivity with respect to the class division (fig.4.13).

For all channels excepts WV6.2 and NIR1.6, the no-rain class is clearly separated from the others, while the rain classes are more similarly distributed, especially in VIS channels. Channel NIR3.9 seems to be the most sensitive to rain-rate: increasing the rain-rate, the low radiance peak increases significantly. Moreover, the high rain-rate curve is better separated from other rain curves, especially for IR channels. A preliminary check on the performance of the algorithm can be done by looking at the ETS (and other parameters) values obtained by each ANN used in cascade. In tab.4.5 are reported for the four ANN cascades the statistical parameter values for each of the four ANN operating to distribute pixels among the precipitation classes, as computed during the testing phase. The best performances are obtained by the ANN that discriminate wet and dry pixels, as expected also by looking at fig.4.13, and the other ANN have markedly lower ETS, but always greater than zero. The decrease of the ETS is caused by both decrease in the probability of detection (POD) and an increase of the false alarm ratio (FAR). The multiplicative BIAS is very high

Network		Class 0-(1,2,3,4)	Class 1-(2,3,4)	Class 2-(3,4)	Class 3-4
Summer day	ETS (%)	33.06	8.16	6.21	4.21
	FAR (%)	55.22	21.37	50.30	76.76
	POD (%)	74.78	69.74	50.34	31.85
	BIAS	1.67	0.89	1.01	1.37
Summer night	ETS (%)	18.93	5.17	5.89	-
	FAR (%)	73.70	23.85	50.87	-
	POD (%)	90.25	71.36	61.28	-
	BIAS	3.43	0.94	1.25	-
Winter day	ETS (%)	23.06	10.96	9.79	-
	FAR (%)	62.04	16.78	59.65	-
	POD (%)	77.99	55.44	61.40	-
	BIAS	2.05	0.67	1.52	-
Winter night	ETS (%)	10.16	5.37	6.44	-
	FAR (%)	76.99	29.43	64.20	-
	POD (%)	83.32	56.52	68.98	-
	BIAS	3.62	0.80	1.93	-

Table 4.5: Statistical parameters of training-testing phase.

also for the first class (except for summer day ANN) indicating large overestimation, which is reduced for the higher rain-rates ANN. In general, the diurnal cycle has higher impact than the seasonal cycle in decreasing the capability of the ANN to resolve precipitation classes.
To have a more useful information of he overall performance of the technique, the categorical Heidke Skill Score (HSS) statistical parameter is computed the four ANN cascades for the validation set and the results are reported in tab.4.6. The HSS can be considered as the equivalent for matching data separated in classes of the Pearson correlation coefficient for continuous variables and the maximum value for highest correlation is 1 (see appendix A for definition). Results show values comparable to those found in the literature for daytime performances, while during the night-time, especially for winter, HSS becomes markedly lower.

In fig.4.14 are shown HSS distributions separated by season and day-time. As mentioned above, best results are obtained for the summer-day ANN in which most

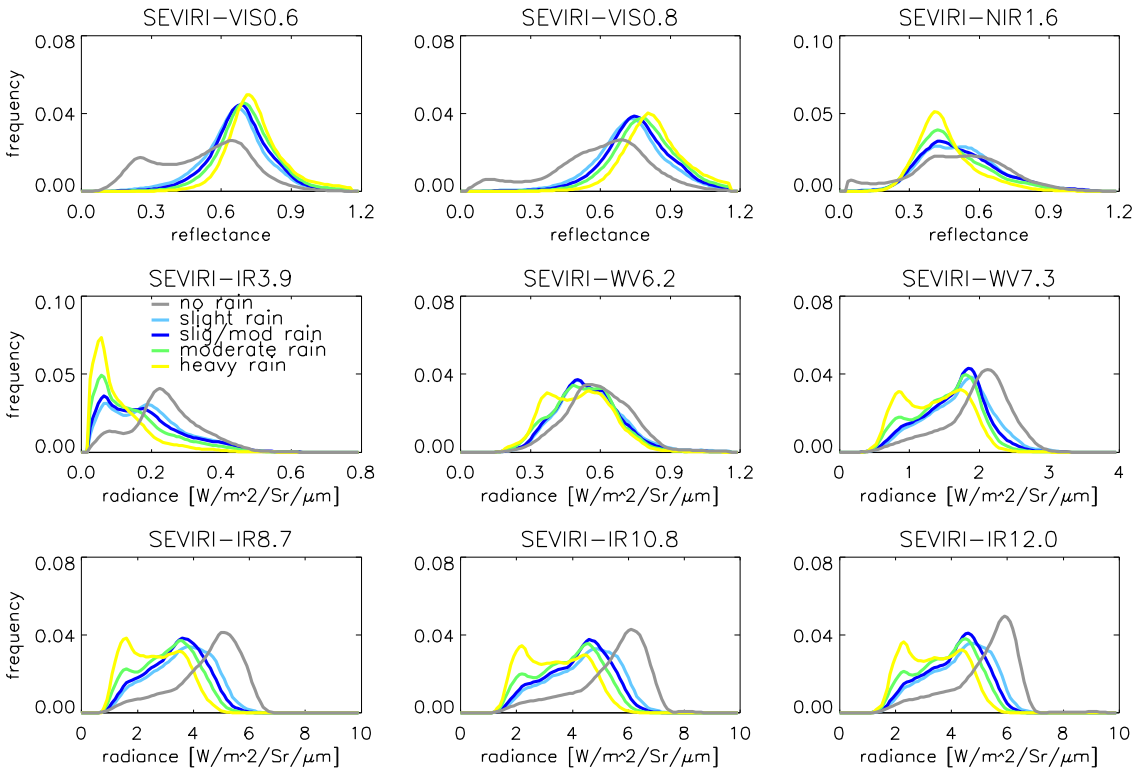


Figure 4.13: Winter-day radiance and reflectance distributions for five precipitation classes.

Network	**HSS** (%)	$\sigma_{\mathbf{HSS}}$(%)
Summer day	27.87	10.58
Summer night	16.40	10.89
Winter day	20.38	7.44
Winter night	9.80	4.63

Table 4.6: HSS mean values and variances for the class-Italian ANN.

of the considered cases (26) have an HSS value between 25 and 35% (fig.4.14a). On the contrary, the worst results are obtained for the winter-night ANN in which most cases have an HSS between 5 and 15% (fig.4.14d). This is a consequence of the already described situation of VIS-NIR channels lack during night-time (sec.4.4). Winter performances are in general worse than summer ones because of the clouds typology that leads to precipitation: stratiform clouds, typical of winter, have rela-

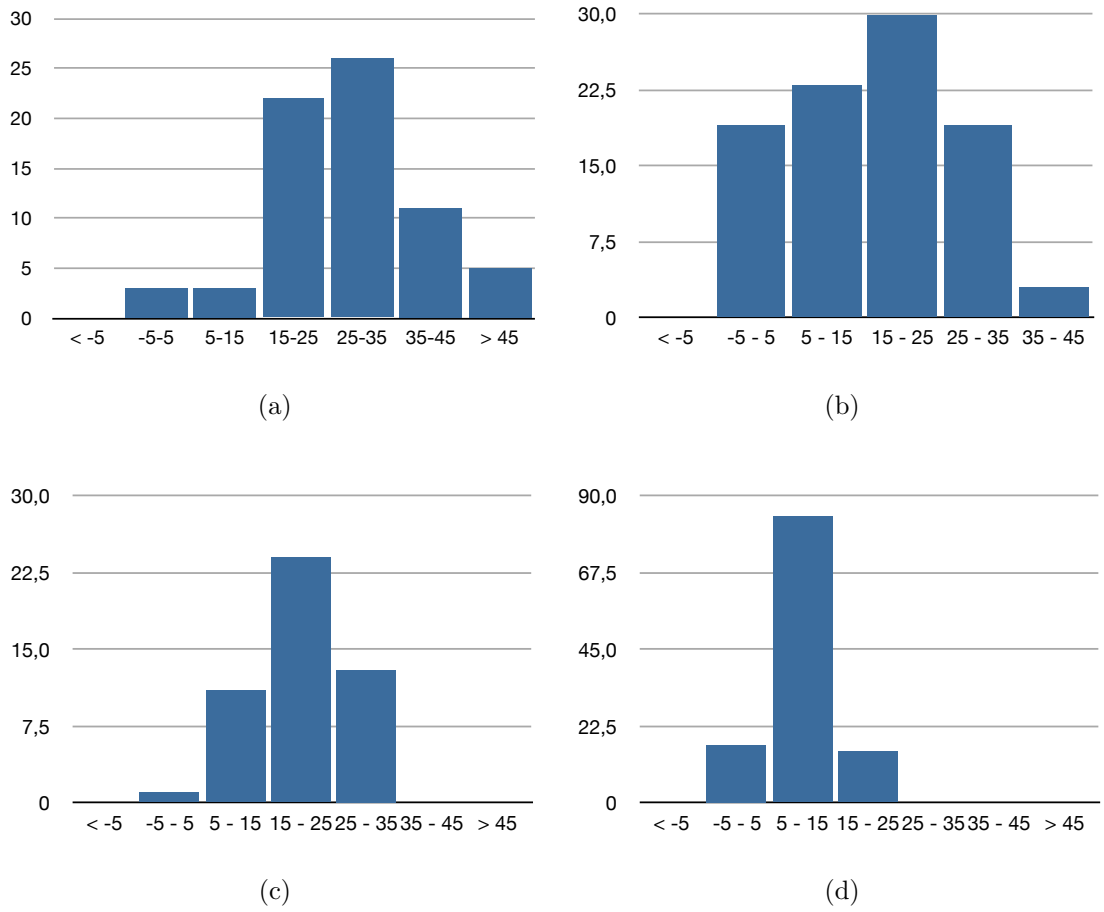


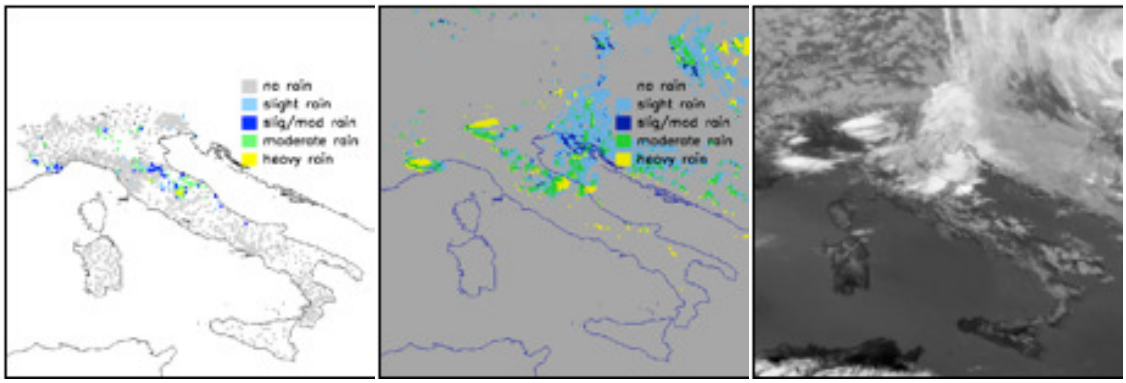
Figure 4.14: HSS distributions separated by season and day-time: day summer (a), night summer (b), day winter (c) and night winter (d).

tively low top and because of the cold ground, the difference in radiance can be very low. This can lead to a general underestimation of precipitation and in a consequent lower performances of the ANN.

4.6 PROSA Product 2

This second algorithm developed for the Italian area, represents the second product of the "Multi-sensor satellite precipitation estimate" of the PROSA system. It separates the estimate rain-rate in five precipitation classes as described in the previous

section and is trained with the Italian rain-gauges network as *true* values for the supervised dataset. In fig.4.15 an example of the algorithm output.



(a) (b) (c)

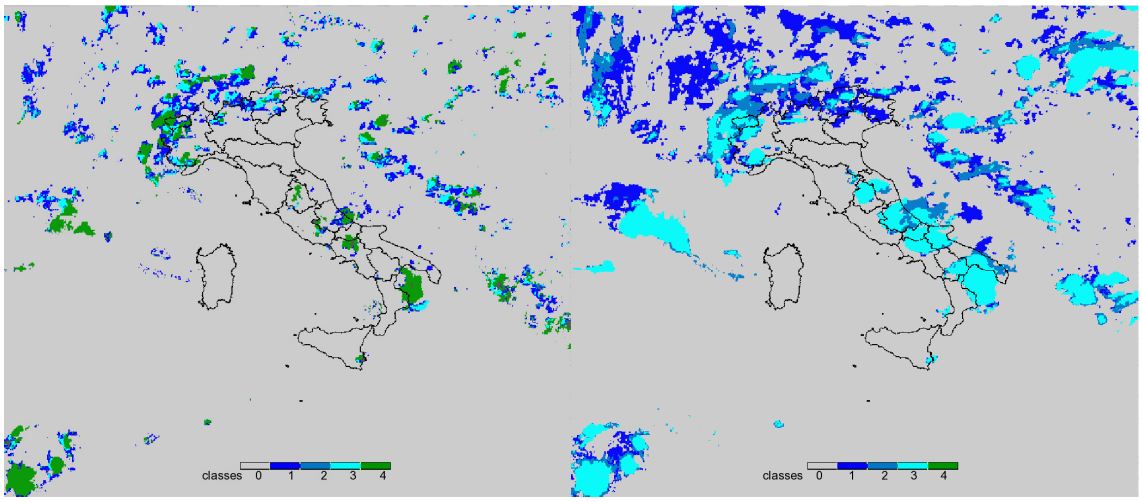
Figure 4.15: 13/06/2008, 15.27 UTC case: precipitation map from rain-gauge network (a), estimation map from ANN (precipitating areas are divided in 5 precipitation classes) (b) and SEVIRI IR12.0 image (c).

This release of the algorithm from one side is able to solve problems of the precedent product version, but new different problems are noted:

1. Again is present the not yet solved problem of day/night-time transition which leads to discontinuities in the precipitation maps sequence (in fig.4.16 an example of two subsequent estimations, where it is clear the discontinuity between the two ANN).

2. A second problem is about the last class of precipitation training. Because of the low statistic, as explained above, the class with rr > 9.6 mm h^{-1} could not be trained for the summer-night, winter-day and winter-night ANNs and also the summer-day one, although trained, can not be considered reliable. A good solution is to increase the samples number in the supervised dataset (in fig.4.16 it is possible to note the absence of the last class of rain-rate (green) in the night-ANN image).

3. The problem of snow at the ground highlighted in the first release product, is greatly reduced in this version. This confirms the hypothesis of using an optimized dataset to train the ANN and correct the anomaly.



(a) (b)

Figure 4.16: 05/07/2009, 17.00 (a) and 17.15 (b) UTC cases: example of switch between day (a) and night (b) ANN and absence of the last class of rain-rate in the night-ANN (b).

4.7 Precipitation rate as a continuous variable

The last version of the algorithm aims to overcome the classes problems establishing a mathematical relation between output probability of the ANN and rain-rate values (in $\mathrm{mm\,h^{-1}}$). This idea is developed in order not to radically change the basic algorithm, and so to maintain the classification structure of the ANN.
After systematic testing work on several summer and winter events, we noted a general relationship between the PoP value and the corresponding precipitation rate as measured by the rain-gauge network. The probability is greater in correspondence of higher precipitation rates (see fig.4.17). This can be explained with the fact that more the cloud is precipitating, more correctly the PoP ANN should classify it and give to this cloud an high PoP value.

The law connecting probability to rain-rate is defined studying the four probability distributions of the four separate ANN (fig.4.18). A set of piecewise-defined functions can be introduced as fit of the four distributions: a constant, nominal value ($0.2\ \mathrm{mm\,h^{-1}}$) is assigned for PoP between the threshold value (as computed during the ANN training phase) and the maximum of the distribution, while a linear

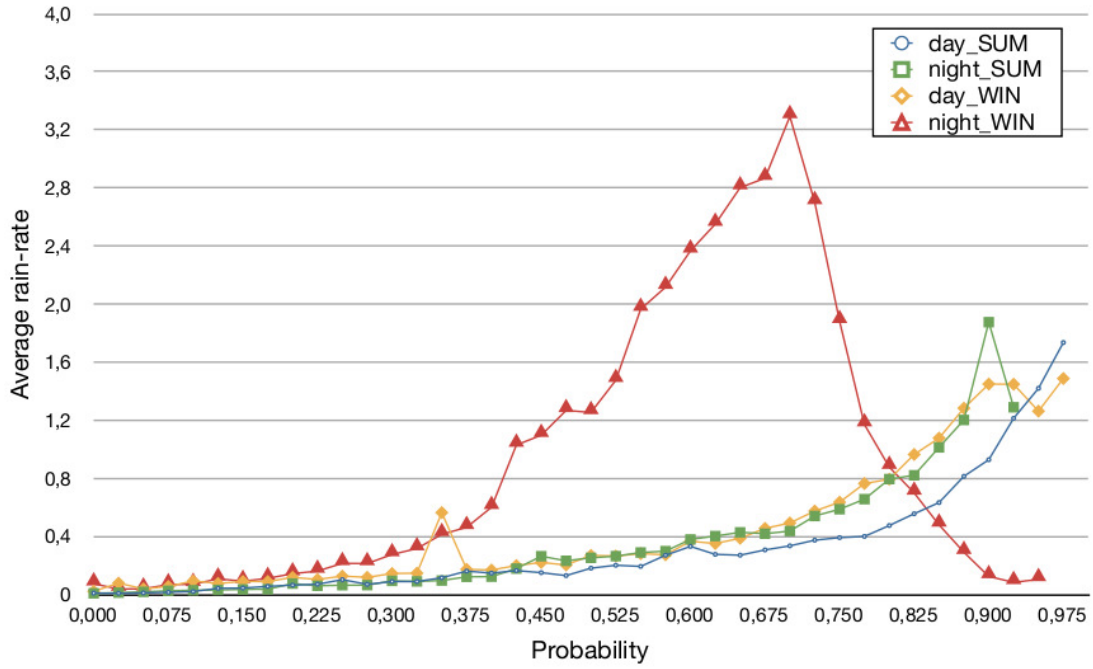


Figure 4.17: Output probability of ANN vs. average rain-rate measured from rain-gauges.

relationship is imposed for PoP values ranging from the maximum to the 95% of rain-rate distribution.

PoP$\leq$ threshold: rr = 0.0 mm h^{-1}
threshold < P < maxPoP: rr = 0.2 mm h^{-1}
maxPoP $\leq$ PoP $\leq$ 1: rr=m PoP+q

PoP is the output probability of the ANN, threshold is the fixed threshold of the ANN, maxPoP is the probability with the greatest number of pixels, rr is the rain-rate value, m and q are the coefficients of the linear dependance.
An example of this particular fit is shown on the summer-day ANN probability distribution in fig.4.19. The rain-rate value corresponding to the 95% of rain-rate distribution is extrapolated from the rain-rate distribution shown in fig.4.20.

The considered parameters of the four ANN are resumed in tab. 4.7.

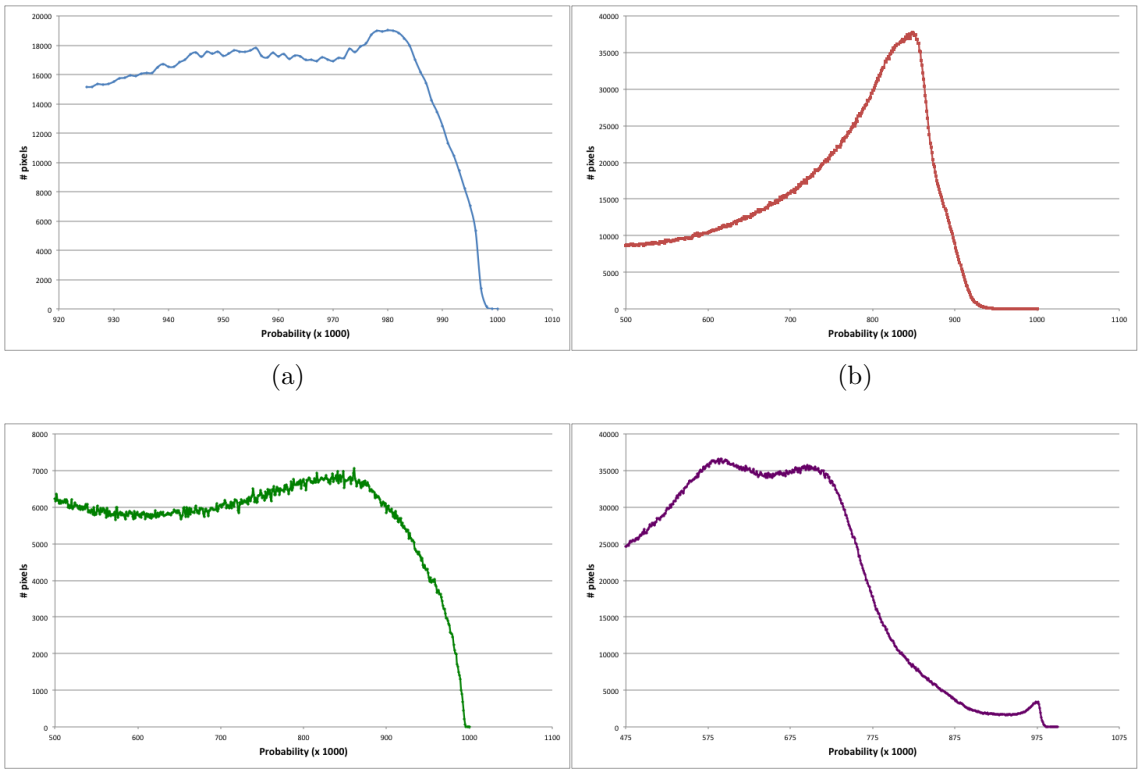


Figure 4.18: Probability distributions used to find the relation law between probability ANN output and rain-rate, separated by season and day-time: day summer (a), night summer (b), day winter (c) and night winter (d).

	Probability threshold	**maxP**	**rr at 95% of distribution**	**linear relation between rain-rate and probability**
Summer-day	51%	95%	39	rr=780.2P-741.2
Summer-night	50%	85%	42	rr=276.8P-234.8
Winter-day	50%	86%	21	rr=149.6P-128.6
Winter-night	47.5%	70%	21	rr=69.3P-48.3

Table 4.7: Considered parameters for the four ANN.

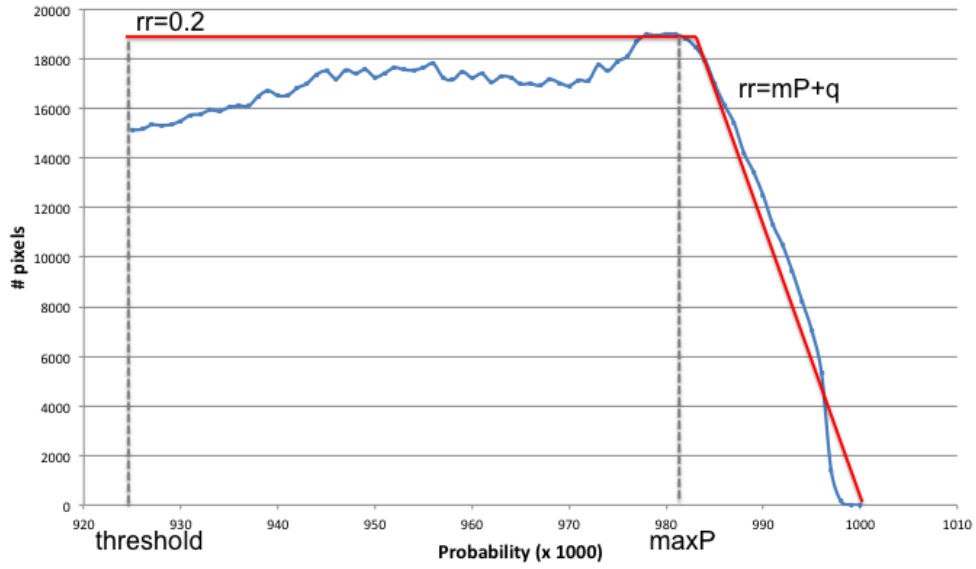


Figure 4.19: Example of fit used for the relation between ANN output probability and rain-rate.

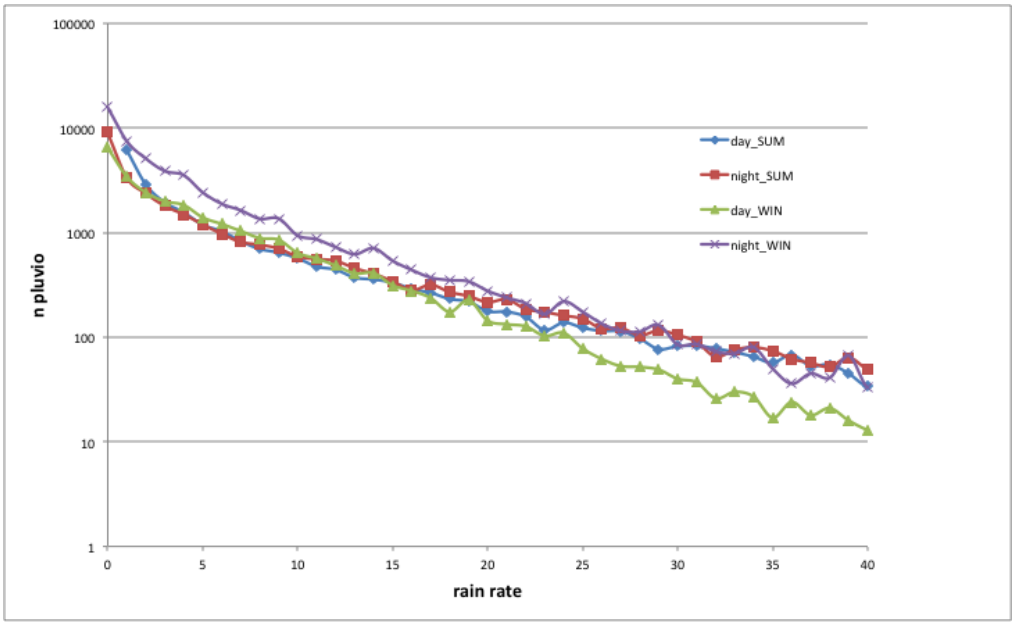


Figure 4.20: Rain rate distributions used to find the rain-rate values corresponding to the 95% of the distribution for every season and day-time.

4.8 Microwave data for ANN online calibration

The last algorithm improvement is the implementation of a new calibration strategy to be flanked to the rain-gauges one used for the previous releases of the algorithm. It is based on precipitation estimate products made with Passive Microwave data and is based on a so called *online* re-calibration of the ANN. As described in sec. 2.2.6, the microwave radiation carries informations about clouds precipitating layers. This characteristic allows to develop algorithms based on more physical approaches than shorter waves ones.
A preliminary study was made over UK area and an algorithm had been developed for the use of AMSR-E data for precipitation retrieval [103]. The estimate technique was similar to the one developed for VIS-IR bands and is based on ANN. Radiance temperatures in the 6.9, 10.7, 18.7, 23.8, 36.5 and 89.0 GHz channels of AMSR-E sensor were used as ANN inputs and, for the training phase, a supervised dataset with UK radar network as *true values* is used.
In fig.4.21 the outputs for daytime slots are compared with the outputs of the first SEVIRI ANN (see sec.4.1) and it is possible to note a light better performance of AMSR-E algorithm with respect to the SEVIRI one mainly in the summer season when convective systems are dominant and better recognized by MW algorithms.

It has to be noted that during the night the PMW performance (very little sensitive to diurnal cycle) are expected to be much bettere than the nighttime IR ANN.

This preliminary study demonstrate that the use of MW data for a precipitation estimate product can be considered a good instrument for nowcasting aims, but the problem of low spatial and temporal resolution of MW data, force to find an other solution. A good compromise can be a blending technique which uses both VIS-IR and MW data.
The third algorithm developed for this thesis work, and discussed above about the continuous values of estimation (sec.4.7), uses MW products, when available, as *true* values to build a temporary supervised dataset for an online rapid update of the ANN. The technique is divided in two main parts, an off-line and an on-line phase.

Off-line phase:

1. a calibration window is defined, in order to specify how many minutes before

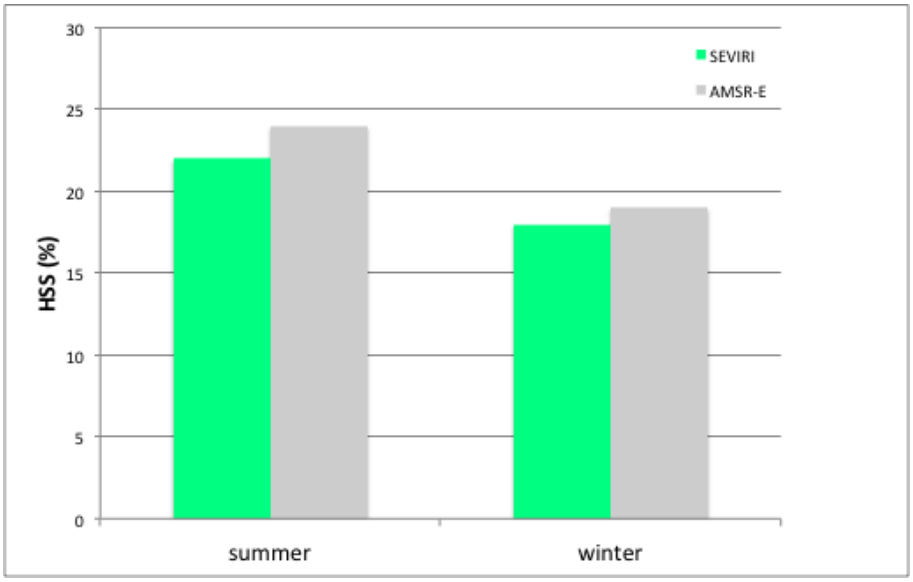


Figure 4.21: HSS daytime values for summer and winter seasons for the SEVIRI (green) and AMSR-E (grey) estimates divided in four classes.

the SEVIRI data availability it is correct to consider the MW data as nearly simultaneous;

2. the four ANNs are trained with a *climatological* dataset (same samples used to train the ANN of the first and second algorithms, tab.4.3 and 4.4), in order to obtain weights and thresholds for every ANN.

On-line phase:

1. the system is activated at every SEVIRI data availability and produces a precipitation probability using weights and thresholds present at the moment in the system;

2. the system looks for MW products in the calibration window (defined off-line);

3. if there are available MW products within the calibration window, the system builds a new supervised dataset with the present slot SEVIRI data and the MW products as *true* values;

4. the ANN is re-trained with the new supervised dataset and climatological weights and thresholds are substituted with new ones;

5. at the following SEVIRI data availability, the system produces a precipitation probability using updated weights and thresholds.

In fig. 4.22 the flow chart of the technique.

The lineup of this procedure results from a system analysis and sensitivity work. Every step can be explained in terms of physical characteristics and system optimization.

The choice of calibration window finds a reason in the improvement of ANN sensitivity. To better resolve precipitant systems it is useful to calibrate the ANN with systems structurally similar to the ones the ANN would find some instants after the training. In this way it is not necessary a great generalization capability of the network, but simply it is required to recognize the same system it is trained with. Actually, a different calibration window should be chosen for different seasons and day-time, but this would involve a greater computational and analytical effort. We thus evaluated this calibration window to be 150 minutes, assuming this as a compromise between the need to have a statistically significant training set and the hypothesis of limited variability of cloud structures during this time. This assumtion is certainly questionable at single cloud scale, especially in the case of convective systems that very often in 150 minutes can complete thier intere lifecycle, but it is more acceptable at "scene" scale. Considering a PMW swath wide cloudy scene, we expect that the cloud structures in the scene, mainly detemined by slow-varying environmental conditions, can be assumed as stationary for the selected time interval. For the nature of MW data, a consideration is made also about orbits. Being MW sensors on board of polar satellites, could happen that some orbits are partially coincident. In order to save computational time, a preliminary filtering is made and in case of partially coincident orbits, the one with the greater number of IFOV is used. A second filter is applied to MW products with less than 20 raining IFOV which are not considered because a small number of data leads to a poor statistic for the ANN training.

After the preliminary products filtering, the attention is focused on the spatio-temporal matching of MW product pixels with VIS-IR ones. SEVIRI data are available every 15 minutes, so an eventual MW product can have a difference in time with the VIS-IR one of maximum 7.5 minutes to be considered in near-real-time (see tab.4.8 for the detailed temporal matching). This time difference is acceptable considering the low spatial resolution of MW sensors: we can assume that at this

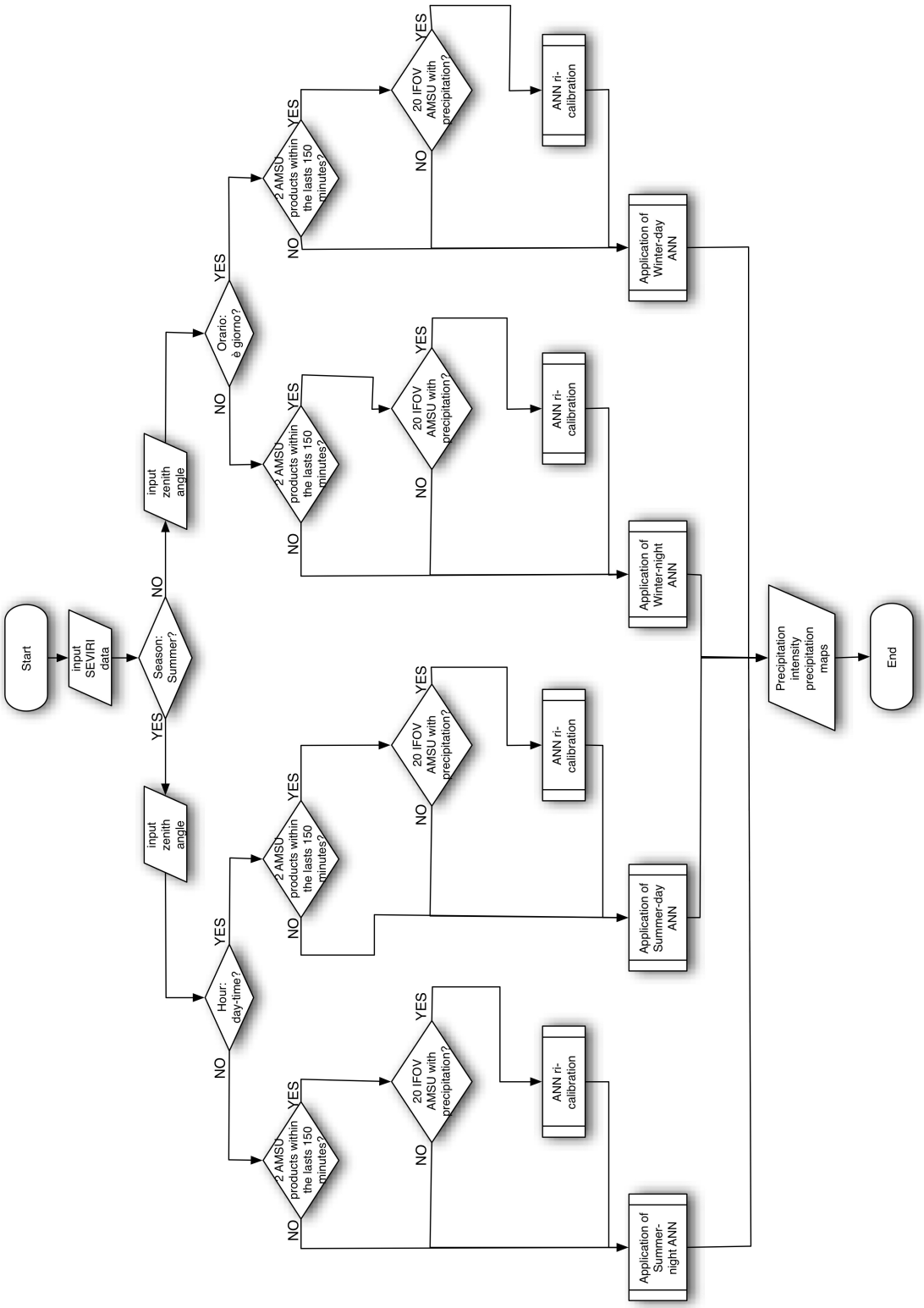


Figure 4.22: Flow-chart of the technique used for the last algorithm.

spatial scale and at this time order, a cloud structure and its precipitating properties do not vary.

MW-sensor time (minute)	**SEVIRI time (minute)**
$03 \leq \text{MWtime} < 18$	12
$18 \leq \text{MWtime} < 33$	27
$33 \leq \text{MWtime} < 48$	42
$48 \leq \text{MWtime} < 03$	57

Table 4.8: Temporal matching criteria between SEVIRI and MW-sensors data.

An other big problem is represented by the acquisition geometry of the sensor. In the case of conical acquisition, there is no IFOV deformation along the swath, the constant acquisition angle of the sensor with Earth's surface, guarantees a constant IFOV dimension. On the contrary, in the case of cross-track acquisition, the IFOV dimensions depend on the nadir-IFOV angle. At nadir the IFOV is circular and its dimensions could vary from about 6×4 km^2 to about 74×43 km^2 depending on the sensor and on the frequency considered. The IFOV shape changes getting away from the nadir and becomes an ellipse, increasing its axes measures till the swat border.
Furthermore, the antenna pattern of this type of sensors acquisition is gaussian and the IFOV diameter refers to the full width at half maximum of the curve. We expect that the sensor can collect also radiation from outside the IFOV nominal dimensions. An other problem are the geolocation parallax errors which increase getting away from the nadir. Contrarily, limb-broadening are corrected in the MW estimate algorithm. In order to reduce this type of problems, since AMSU is a cross-track scanning sensor, we chose to consider only the central 50 IFOVs of the swath, reaching maximum IFOV dimensions of about 21×18 km^2.
The spatial resolution difference between SEVIRI and MW-sensors data poses a pixels/IFOV matching problem that requires a study in terms of errors and performances. To solve this problem, two options had been investigated:

1. Calculate the mean radiance of SEVIRI pixels contained in the MW-sensor IFOV and assign this average as input value in the ANN, with the rain-rate estimate of that particular MW-sensor IFOV as *true* value.

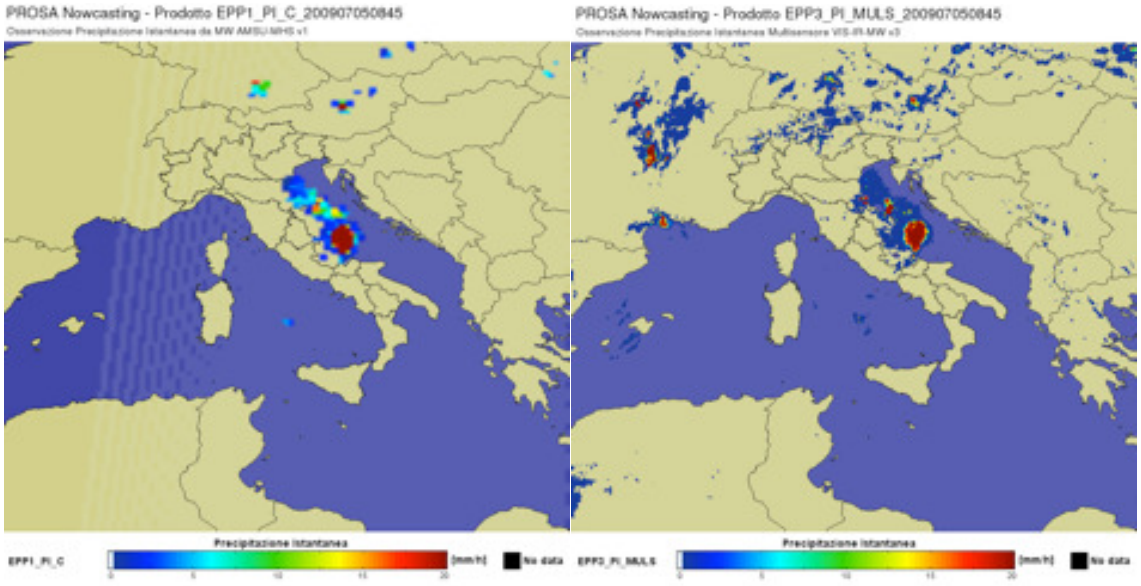
2. Assign the same rain-rate estimate, calculated for the MW-sensor IFOV considered, to every SEVIRI pixel contained in it.

Both solutions present physical problems, in the first case we would make a linear average of SEVIRI pixels radiances, although they are not in a linear relationship with the precipitation estimate, and in the second case we would associate the same precipitation rate to 9 different SEVIRI pixels with different radiances values, ignoring the beam filling problem. We chose the second option, with the only explanation of a greater number of data for the supervised dataset.

4.9 PROSA Product 3

The third release of PROSA product consists in an ANN with continuous rain-rate values as output and is implemented with an online training procedure with MW-sensors products as *true* precipitation values. Also if the algorithm can work with any type of MW product, in this case are considered only AMSU sensors data, processed by the CNR-Roma group to obtain an instantaneous precipitation product which represents the input *true* precipitation value for the online calibration of our product. In origin, the idea was to use many MW-sensors (such as AMSR-E, SSM/I, SSMI/S, AMSU-A, AMSU-B and TMI) for the online updating, but analyzing the time availability of data, many of these result to have a delay too big to respect the nowcasting request of near-real-time data and make short range forecasting.
In fig.4.23 it is shown a comparison between AMSU and SEVIRI products, the latter trained online with the MW product. It is possible to note a substantial difference in the precipitation pattern resolution, the MW product can define a precise rain-rate differentiation also in the same precipitating nucleus, while the VIS-IR one presents an almost constant rain-rate. On the contrary, the AMSU product can not differentiate spatially localized precipitation systems because of its low spatial resolution.

Also for this product release, it is generally noted a better performance of the summer-day ANN with respect to the others, for the same reason explained in detail for the other versions of the product.
The official validtion inside the PROSA project has been carried out by a project partner not involved in algorithm development (the University of Camerino): for the



(a) (b)

Figure 4.23: 05/07/2009, 08.45 UTC case: AMSU precipitation map (a) and ANN precipitation estimate produced after the ri-calibration of the ANN with the AMSU precipitation map as input *true* value (b).

last release of the products, the Hanssen and Kuipers skill score has been computed for two case studies, and the results are reported in table 4.9.

Product	data used	5 July 2009	7-9 November 2009
EPP3_PI_A	PMW	0.168	0.116
EPP3_PI_B	PMW	0.072	0.102
EPP1_PI_C	PMW	0.466	0.109
EPP3_PI_BLE	IR+PMW	0.602	0.267
EPP3_PI_IBR	IR+PMW	0.644	0.497
EPP3_PI_MULS	IR+PMW	0.327	0.324

Table 4.9: HK values for precipitation products of the third release of PROSA.

Results show rather poor performance of PMW products, without clear seasonal signal, except for EPP1_PI_C, the product used in all the IR+PMW techniques, that shows high performance for the summer case. Among the IR+PMW technique, the

MULS (our algorithm) has comparably worst performance for the summer case and average performance for the winter case. MULS is less dependent on the season, probably given the input of PMW. The case studies include also night-time slots, so the reported values are an "averaged" value between night and day ANN.
Other statistical indicators can be applied to validate quantitative satellite estimates, often referred as continuous verification scores: Fractional Standard Error (FSE) and Pearson Correlation Coefficient (r), and normalized bias (NBIAS) (see appendix for definition). We computed these parameters for the two case studies used for the PROSA validation, but, unfortunately, we cannot compare our values with other algorithms. Results are reported in table 4.10, with also the Binary Bias (BB), a dichotomic indicator, for reference. The FSE values are well larger than 100% and r very close to zero, especially for the summer case, a quantitative overestimate, much larger for the November case, is also evident, while the rain area is slightly overestimated in July and strongly underestimated in November.

	FSE	r	**NBIAS**	**BB**
5 July	2.25	0.02	-1.01	0.76
7-9 November	3.82	0.14	-4.55	1.30

Table 4.10: Statistical parameters of product 3 validation.

The values of these parameters, rarely considered for instantaneous precipitation estimates validation given their sensitivity to small misplacement of precipitation peaks and to rainrate oultiers, indicate that an acceptable quality of the quatitative estimate is not reached in these case studies. The large overestimate of the rainrate for both cases for the November case correspond to an overestimate of rainareas, while for July the raining area is underestimated. Such quantitative rainrate overestimation is responsible for high FSE values, especially for the November case.
After the demonstration activity carried on under PROSA, some monitoring of the products, and the validation activity, some problems are detected and are summarized here:

1. the known problem of transition between day and night-time is not solved, despite the MW update support;

2. the fact that AMSU data are not continuously available involves a not systematic online calibration of the ANN with a consequent continuous switching between climatological and updated ANN parameters. This continuous change of ANN parameters leads to a discontinuous pattern sequence in the estimate maps;

3. The overall performances of the technique is comparable to other similar techniques in PROSA and in H-SAF, as far as dichotomic or categorical indicators are used;

4. the quantitative estimate, as shown by continuous statistics indicators, are far from being satisfactory: a part of the error comes from the PMW estimate, that have to be more frequent and more accurate, as it will be expected after the full operational exploitation of the Global Precipitation Measurement mission (GPM);

5. the task of any ANN algorithm is to replicate at best the input received during the training phase, and the trainig and validation carried on using hourly raingauge data it is not the first choice, given the different nature of the two datasets;

6. no parallax correction is applied to co-locate SEVIRI pixel with raingauges and with PMW IFOV neither to build the training set neither to validate products: this could be responsible for a significant part of the error, since most skill indicators are very sensitive to small misplacement between product and reference maps.

Conclusions

This work aims to describe the development of three different algorithms for multi-sensor satellite precipitation estimate. All the three algorithms are based on Artificial Neural Networks, having as main input satellite radiances in the VIS-IR part of the spectrum, set up as classification problem and using rain-gauges data as *true* values of precipitation at the ground for network training.

The goal of the first algorithm is the classification of raining and non raining areas over the Italian region. A preliminary study about satellite channels sensitivity to precipitation is done to understand the different performances of the ANN at day-time and night-time and to evaluate which channels carry more information on precipitation. Better results are observed in the day-time case, when VIS channels, providing information on cloud optical thickness and hydrometeors effective radius, are available.
The algorithm is also tuned according to seasonal cycle, and best performances are observed for summer season when solar zenith angle is lower and precipitating events are mainly convective, better recognized in VIS-IR bands.

The second algorithm is focused on the precipitation classification in rain-rate classes. After a study about the distribution of rain-rate, five classes are defined. The classifier uses a cascade scheme of four ANNs in order to separate pixels belonging to different precipitation classes. The first ANN separates dry pixels (no rain) from wet ones (rain) in the same way of the first algorithm version. Then, wet pixels are used as input of the second ANN, which separates slight rain pixels from others. This scheme is repeated 4 times in order to obtain 5 rain-rate classes. In general, best performances are observed again for the summer day-time ANN. Some problems of the first version algorithm, such as the misinterpretation of snow at the

ground, are corrected.

The last algorithm gives as output precipitation maps expressed in $\mathrm{mm\,h^{-1}}$ values. In order to define a relation between probability of detection (PoP), output of the ANN, and rain-rate values, a study about probability distributions of the ANN is done. A set of seasonal and diurnal cycle dependent relationships between PoP and rain-rate are introduced. A further relevant improvement of the algorithm is performed by using MW precipitation data for the online calibration of the ANN. The online calibration aims to provide the ANN with a set of weights, trained on the newest PMW estimate, optimized for the current cloud scene. In order to do that, a specific work on SEVIRI and MW data matching is done, because of the different spatial and temporal resolutions of the two type of sensors. The PMW estimates are available from LEO AMSU sensor in near real time.

In general, better performances are always observed for day-time algorithms because a greater number of inputs is available for this ANN (VIS+IR+zenith) and more informations are taken into account. Considering the seasonal cycle, summer is usually the best one because of the greater occurrence of convective systems, better resolved at VIS-IR wavelengths.
It has to be highlighted that algorithms performances strongly depend on *true* rain-rate values available for the ANN calibration, since ANN tends to replicate the reference data processed during the training phase. In order to improve performances it is necessary to improve the *true* precipitation data quality and the matching with satellite data. MW sensors data could be considered a good solution because of their physical informations more linked to the precipitation structure, especially in the view of the forthcoming Global Precipitation Measuring mission era, that will ensure a more dense PMW satellite constellation and an unprecedented precipitation radar covering mid-latitudes.
Finally, we remark the wide flexibility of the proposed approach in terms of input data: in principle it is possible to feed the proposed on-line ANN training strategy with any near real time reliable precipitation data, such as well calibrated ground weather radar, to extend the estimate outside the radar coverage with small decrease of the estimate quality.

Appendix A

Satellite estimates validation

A significant part of the work done for the development and tuning of a satellite precipitation estimation technique is often labeled as "validation". Validation means basically to compare the estimated precipitation values to a corresponding observation of what actually occurred, or some good estimate of the true outcome. In the latest years new sensors and algorithms have produced a number of precipitation products that aims to provide quantitative values reliable enough to compete with other precipitation measuring instruments (such as rain-gauges and weather radars) for operational or research purposes. As a consequence, the demand for a definition of the error structure of the estimated precipitation field is also growing, to allow the user to decide how to make use of satellite products in his activity.
Most of the initiatives mentioned in Chapter 1, aiming to improve satellite techniques, include validation activities, but two international efforts should be mentioned for their relevance.
In 2003 the International Precipitation Working Group (IPWG) began a project to validate and intercompare operational and semi-operational satellite rainfall estimates over Australia and the US in near real time. A European verification was added in 2004, and other regions may be included in the future. This study focuses on the large-scale validation of daily rainfall estimates, for two reasons. First, the large number of rainfall observations from rain-gauges at the 24-hour time scale provides good quality verification data on a large scale. Second, daily rainfall estimates are required as input to a large number of climate and other applications. For comparison, 1-day forecasts from a limited number of numerical weather prediction

models, namely the ECMWF, the US (NCEP), and US Navy global models, and the Australian regional model, are also verified [112].
Within the H-SAF initiative, the group of Atmospheric Physics of Ferrara University has the responsibility to validate precipitation products over Italy, and also is charing the Working Group on rain-gauges. The challenging task of the Validation activity is to validate high resolution satellite products to assess the performances of instantaneous satellite estimates at the finest available spatial resolution (ranging from 5 km to 20 km depending on sensor used). During the H-SAF validation activity several studies have been carried out to improve the quality and the significance of the matching between satellite estimates and ground references. A number of statistical indicators are tested to better describe the products error, and some of them have also been used in the development of the algorithms presented in this work. They are defined here highlighting some of their properties. Most, if not all, have a long historical background but they are still used very commonly. It has also to be mentioned that in no case it is sufficient to apply only just one single verification measure.

A.1 Dichotomic parameters

A class of parameters is designed to assess the product capability to discriminate between wet and dry pixels. To this end, a contingency table (see tab.A.1) is defined after the careful spatial co-location of product and reference nearly simultaneous rainfall maps.

After the number of observations to be evaluated is fixed, all the pixels (N) with valid data are used to fill the table. The values in the table are then used to compute the following dichotomic indicators.

Probability of Detection (POD):

$$POD = \frac{hits}{hits + misses} \tag{A.1}$$

POD ranges from zero (worst) to one (best). POD is sensitive to hits but takes no account of false alarms. It can be artificially improved by producing excessive *yes*

Event estimated	Event observed		
	Yes	No	Marginal total
Yes	Hit	False Alarm	Es Yes
No	Miss	Correct negative	Es No
Marginal total	Obs Yes	Obs No	Sum total

Table A.1: 2×2 contingency table showing the frequency of *yes* and *no* estimations and corresponding observations.

estimations to increase the number of hits (with a consequence of numerous false alarms). While maximizing the number of hits and minimizing the number of false alarms is desirable, it is required that POD be examined together with False Alarm Ratio (FAR):

$$FAR = \frac{false\,alarm}{false\,alarm + hits} \tag{A.2}$$

Range of FAR is one to zero, a perfect score FAR = 0, i.e. FAR has a negative orientation. FAR is also very sensitive to the climatological frequency of the event. Contrarily to POD, FAR is sensitive to false alarms but takes no account of misses. Likewise POD, it can be artificially improved, but now by producing excessive *no* estimations, i.e. to reduce the number of false alarms. Because the increase of POD is achieved by increasing FAR and decrease of FAR by decreasing POD, POD and FAR must be examined together.

An overall skill score, that consider both ability in detecting rain and in not detecting no rain is the Critical Success Index (CSI), defined as:

$$CSI = \frac{hits}{hits + false\,alarm + misses} \tag{A.3}$$

Range of CSI is zero to one, a perfect score has CSI = 1. CSI is sensitive to hits and takes into account both false alarms and misses and can be seen as a measure for the event being estimated after removing correct (simple) *no* estimations from

consideration. CSI is sensitive to the climatological frequency of events (producing poorer scores for rarer events), since some hits can occur due to random chance. A random assigned precipitation classification could produce CSI of about 0.2 or 0.3, depending on precipitation pattern. To overcome this effect the Equitable Threat Score (ETS) adjusts for the number of hits associated with random chance, and is defined as:

$$ETS = \frac{hits - hits_{random}}{hits + false\,alarm + misses + hits_{random}} \tag{A.4}$$

Range of ETS is -1/3 to one, for a perfect score ETS = 1, no skill estimate ETS = 0. The ETS has been chosen as main skill indicator in the whole training testing validation procedure for the set up of the various ANN algorithms presented in this work.

An other popular skill score is the Hanssen-Kuipers Skill Score defined as:

$$HK = \frac{hits}{hits + misses} - \frac{false\,alarm}{false\,alarm + correct\,negative} \tag{A.5}$$

Range of HK is minus one to one, a perfect score = 1, no skill estimation = 0. For rare events, the frequency of correct negative cell is typically very high in the contingency table compared to the other cells, leading to a very low False Alarm Rate and, consequently, HK is close to POD.
The Binary Bias (BB) of binary estimates compares the frequency of positive estimates to the frequency of actual occurrences and is computed by the ratio:

$$BB = \frac{hits + false\,alarm}{hits + misses} \tag{A.6}$$

Range of BB is zero to infinity, an unbiased score = 1. With B>1 (<1), the estimate exhibits overestimation (underestimation) of the rain area.

A.2 Multicategorical parameters

Categorical events are naturally not limited to binary estimates of two categories and the associated 2×2 contingency tables. The general distributions approach in rainfall estimation verification studies the relationship among the elements in multi-category contingency tables. One can consider rain-rate in several mutually

exhaustive classes, at it is done for the release 2 of the ANN product described in this work.

It is advisable to initiate verification again by constructing a contingency table where the frequencies of estimates and observations are collected in relevant cells as illustrated in tab.A.2 for a 3×3 category case. A perfect estimation system would (again) have all the entries along the diagonal (r, v, z, in the example), all other values being zero. The verification measures introduced in the previous section are valid only with the binary yes/no estimate situation. To be able to apply these measures, one must convert the k>2 contingency table into a series of 2×2 tables. Each of these is constructed by considering the *estimated event* distinct from the complementary *non-estimated event*, which is composed as the union of the remaining k-1 events. The off-diagonal cells provide information about the nature of the estimation errors. For example, Binary Biases (BB) reveal if some categories are under- or over-estimated, while PODs quantify the success of detecting the distinct categorical events. The categorical Heidke Skill Score (HSS) gives in a single parameter an overall measure of the estimate accuracy, and can be interpreted as the discrete version of the Pearson Correlation coefficient. HSS for multi-category cases can be written as:

$$HSS = \frac{\sum p(e_i, o_i) - \sum p(e_i)p(o_i)}{1 - \sum p(e_i)p(o_i)} \tag{A.7}$$

where the subscript i denotes the dimension of the table, $p(f_i, o_i)$ represents the joint distribution of estimations and observations (i.e. the diagonal sum count divided by the total sample size), and $p(f_i)$ and $p(o_i)$ are the marginal probability distributions of the estimations and observations (i.e. row and column sums divided by the sum total), respectively.

Also HK can be generalize for a multi-categorical case:

$$HK = \frac{\sum p(e_i, o_i) - \sum p(e_i)p(o_i)}{1 - \sum (p(e_i))^2}, \tag{A.8}$$

A.3 Continuous variables

The verification of continuous variables typically provides statistics on how much the estimated values differ from the observations and, thereafter, computation of relative

Estimated	Observed			
	o_1	o_2	o_3	Es $\sum$
e_1	r	s	t	$\sum e_1$
e_2	u	v	w	$\sum e_2$
e_3	x	y	z	$\sum e_3$
Obs $\sum$	$\sum o_1$	$\sum o_2$	$\sum o_3$	$\sum$

Table A.2: Multi categorical contingency table showing the frequency of *yes* and *no* estimations and corresponding observations.

measures against some reference estimating systems. Among the continuous statistic parameters is the Pearson Correlation Coefficient, defined as:

$$r = \frac{\sum(e_i - \overline{e_i})(o_i - \overline{o_i})}{\sqrt{\sum(e_i - \overline{e_i})^2}\sqrt{\sum(o_i - \overline{o_i})^2}} \tag{A.9}$$

Range of r is minus one to one, a perfectly correlated score has $r = 1$.

Two further error measurements are then presented: Fractional Standard Error (FSE) and Normalized Bias (NBIAS). FSE is a measure of the relative error made in the estimate, NBIAS quantifies the relative departure between estimations and observations indicating over or underestimation in therm of quantitative values. These errors measures are given by:

$$FSE = \frac{\sqrt{\frac{1}{N}\sum(o_i - e_i)^2}}{\frac{1}{N}\sum o_i} \tag{A.10}$$

$$NBIAS = \frac{\frac{1}{N}\sum(o_i - e_i)}{\frac{1}{N}\sum o_i} \tag{A.11}$$

Acknowledgments

I want first of all to express my gratitude to ASI (Italian Space Agency) which financially supported my PhD within PROSA and to Prof. Franco Prodi, scientific director of the Project.
I wish to thank DPC (Italian Civil Protection Departement) for the data they made available for my work, Dr. Stefano Dietrich and Dr. Francesco di Paola for the support in the use of their PMW precipitation estimate product and Marco Vimercati for his precious help in the practical informatics set up of my algorithm within the PROSA System. I would like to thank also Prof.Marco Borga and Dr.Paolo Antonelli for their review of my work.

I want then to express my sincere and deep gratitude to my tutor, Dr. Federico Porcú. He introduced me to Atmospheric Physics and made me appreciate it. His scientific and technical support, his constant presence, his patience and in many moments his willingness to share thoughts provided me with a friendly and relaxed work environment.

A great thanks to all the people passed in room 427 during my PhD years, for their friendship and their "atmospheric thoughts". In particular I wish to thank Claudia and Martina because our friendship, started and grown at the "fourth floor" is something that now goes out from there. Thanks to my strange but dear colleague Uarda, who shared with me the office. Our diversity has been a richness in our PhD works and our lives.

Outside work environment, I wish to express my heartfelt gratitude to my parents, to which this thesis is dedicated. With my PhD started my "alone life", in my own house, with many satisfactions and many efforts, but without their silent precious presence and their support in every circumstances I would never be able to

do what I done.

And finally, last but not least, a huge thanks to Federico, for his patience and his constant presence. We met thanks to this PhD experience. With all its pleasures and its difficulties we can consider it a good training for our future life together.

Bibliography

[1] R.A Scofield and C.E Weiss. A report on the Cheasapeake Bay region nowcasting experiment. In *NOAA technical memorandum NESS*, page 52, Washington D.C., 1977. NOAA.

[2] F.Porcù and A.Carrassi. Toward an estimation of the relationship between cyclonic structures and damages at the ground in Europe. *Nat. Hazards Earth Syst. Sci.*, 9:823–829, 2009.

[3] B.J.Conway, L.Gerard, J.Labrousse, E.Liljas, S.Senesi, J.Sunde, and V.Zwatz-Meise, editors. *Meteorology - Nowcasting, a survey of current knowledge, techniques and practice*, volume Phase1 Report of *EUR 16861*, Brussels, 1996. COST 78.

[4] K.Lagouvardos, E.Liljas, B.Conway, and J.Sunde, editors. *Improvement of nowcasting techniques*, volume Final Report of *EUR 19544*, Brussels, 2001. COST Action 78.

[5] N.Hilker, A.Badoux, and C.Hegg. The Swiss flood and landslide damage database 1972-2007. *Nat. Hazards Earth Syst. Sci.*, 9:913–925, 2009.

[6] F.Guzzetti and G.Tonelli. Information system on hydrological and geomorphological catastrophes in Italy (SICI): a tool for managing landslide and flood hazards. *Nat. Hazards Earth Syst. Sci.*, 4(2):213–232, 2004.

[7] M.C.Llasat, M.Llasat-Botija, and L.López. A press database on natural risks and its application in the study of floods in Northeastern Spain. *Nat. Hazards Earth Syst. Sci.*, 9:2049–2061, 2009.

[8] M.C.Llasat, L.López, M.Barnolas, and M.Llasat-Botija. Flash-floods in Catalonia: the social perception in a context of changing vulnerability. *Adv. Geosci.*, 17:63–70, 2008.

[9] F.Porcú, C.Caracciolo, and F.Prodi. Cloud systems leading to flood events in Europe: an overview and classification. *Meteorol. Appl.*, 10(3):217–227, 2003.

[10] TOPICS GEO, Natural catastrophes 2010, Analyses, assessments, positions. Münchener Rückversicherungs-Gesellschaft, Königinstrasse 107, 80802 München, Germany, 2010. Munich RE.

[11] A.Buzzi and S.Tibaldi. Cyclogenesis in the lee of the Alps: A case study. *Q. J. R. Meteorol. Soc.*, 104(440):271–287, 1978.

[12] R.Rotunno and R.Ferretti. Orographic effects on rainfall in MAP cases IOP 2b and IOP 8. *Q. J. R. Meteorol. Soc.*, 129(588):373–390, 2003.

[13] D.Mastrangelo, K.Horvath, A.Riccio, and M.M.Miglietta. Mechanisms for convection development in a long-lasting heavy precipitation event over southeastern Italy. *Atmos. Res.*, 100(4):586–602, June 2011.

[14] S.Federico, C.Bellecci, and M.Colacino. Quantitative precipitation forecast of the Soverato flood: the role of orography and surface fluxes. *Nuovo Cimento C*, 26:6–22, 2003.

[15] F.Grazzini. Predictability of a large-scale flow conducive to extreme precipitation over the western Alps. *Meteorol. Atmos. Phys.*, 95:123–138, 2007.

[16] P.Reichenbach, F.Guzzetti, and M.Cardinali. Carta delle aree colpite da movimenti franosi e da inondazioni. Progetto AVI, 2^a edizione. Map of sites affected by floods and landslides, the AVI Project, 2nd edition. Scala 1:1.200.000 (in Italian and English) pub.n.1786, GNDCI, 1998.

[17] P.Salvati abd C.Bianchi, M.Rossi, and F.Guzzetti. Societal landslide and flood risk in Italy. *Nat. Hazards Earth Syst. Sci.*, 10:465–483, 2010.

[18] 13th International Conference on Cloud and Precipitation (ICCP). *MEFFE satellite and combined satellite-radar techniques in meteorological forecasting for flood event; research activities and results*, Reno, 2000.

[19] M.A.Shapiro and A.J.Thorpe. THORPEX: a global atmospheric research programme for the beginning of the 21st century. *WMO Bulletin*, 53:222–226, 2004.

[20] C.G.Collier. Weather radar precipitation data and their use in hydrological modelling. In M.B.Abbott and J.C.Refsgaard, editors, *Distributed Hydrological Modelling*, volume ch 8, pages 143–163, Dordrecht, The Netherlands, 1996. Kluwer Academic Publishers.

[21] A.Rossa, M.Bruen, D.Frühwald, B.Macpherson, I.Holleman, D.Michelson, and S.Michaelides, editors. *Use of radar observations in hydrological and NWP models*, volume Final Report. COST 717, 2005.

[22] J.Joss and U.Germann. Solutions and problems when applying qualitative and quantitative information from weather radar. *Phys. Chem. Earth Part B*, 25(10-12):837–841, 2000.

[23] R.Minciardi, R.Sacile, and F.Siccardi. Optimal Planning of Weather Radar Network. *J. Atmos. Oceanic Technol.*, 20(9):1251–1263, 2003.

[24] U.Germann, M.Berenguer, D.Sempere-Torres, and M.Zappa. REAL - ensemble radar precipitation estimation for hydrology in a mountainous region. *Q. J. R. Meteorol. Soc.*, 135(639):445–456, 2009.

[25] K.Schröter, X.Llort, C.Velasco-Forero, M.Ostrowski, and D.Sempere-Torres. Implications of radar rainfall estimate uncertainty on distributed hydrological model predictions. *Atmos. Res.*, 100(2-3):237–245, 2011.

[26] K.Köck, T.Leltne, W.L.Randeu, M.Divjak, and K.-J. Schrelber. OPERA: Operational programme for the exchange of weather radar information. First results and outlook for the future. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceand and Atmosphere*, 25(10):1147–1151, 2000.

[27] A.Apituley, H.W.J.Russchenberg, and W.A.A.Monna, editors. *News on the European Weather Radar Network (OPERA)*, I.Holleman, Delft, The Netherlands, 2009. 8th International Symposium on Tropospheric Profiling.

[28] C.G.Collier, C.A.Fair, and D.H.Newsome. International Weather-Radar Networking in Western Europe. *Bull. Amer. Meteor. Soc.*, 69(1):16–21, 1988.

[29] A.Rossa, K.Liechti, M.Zappa, M.Bruen, U.Germann, G.Haase, C.Keil, and P.Krahe. The COST 731 Action: A review on uncertainty propagation in advanced hydro-meteorological forecast systems. *Atmos. Res.*, 100:150–167, 2011.

[30] C.Klepp and S.Bakan. Satellite Derived Energy and Water Cycle Components in North Atlantic Cyclones. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceand and Atmosphere*, 25(2):65–68, 2000.

[31] F.Porcú. Classification of frontal systems by SSM/I water vapor measurements. In K.Lagouvardos, E.Liljas, B.Conway, and J.Sunde, editors, *Improvement of nowcasting techniques*, pages 243–250, EUR 19544, Brussels, 2001.

[32] F.Porcú, F.Prodi, S.Dietrich, A.Mugnai, and R.Bechini. Multisensor estimation of severe rainfall events. In *Eumetsat Meteorological Satellite Data Users' Conference*, pages 371–378. EUM P 29, Eumetsat, Darmstadt, 2000.

[33] G.Panegrossi, S.Dietrich, F.S.Marzano, A.Mugnai, E.A.Smith, X.Xiang, G.J.Tripoli, P.K.Wang, and J.P.V.Poiares Baptista. Use of Cloud Model Microphysics for Passive Microwave-Based Precipitation Retrieval: Significance of Consistency Between Model and Measurement Manifolds. *J. Atmos. Sci.*, 55:1644–1673, 1998.

[34] F.Porcú, M.Borga, and F.Prodi. Rainfall estimation by combining radar and infrared satellite data for nowcasting purposes. *Meteorol. Appl.*, 6:289–300, 1999.

[35] R.Bechini, E.D'Acunzo, S.Dietrich, R.Fabbo, A.Mugnai, S.Natali, F.Porcú, F.Prodi, and A.Tassa. Multisensor analysis of Friuli flood event (5-7 October 1998). In P.Pampaloni and S.Paloscia, editors, *Microw. Radiomet. Remote Sens. Earth's Surf. Atmosphere*, pages 379–386, Zeist, 2000. VSP.

[36] S.Dietrich, R.Bechini, C.Adamo, A.Mugnai, and F.Prodi. Radar calibration of physical profile-based precipitation retrieval from passive microwave sen-

sors. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceand and Atmosphere*, 25(10-12):877–882, 2000.

[37] E.Toth, A.Brath, A.Montanari, F.Porcú, and F.Prodi. Use of satellite derived rainfall maps as input to distributed rainfall-runoff models. pages 353–358. 3rd EGS Plinius Conference, 2001.

[38] V.Levizzani, P.Bauer, A.Buzzi, S.Davolio, D.E.Hinsman, C.Kidd, F.S.Marzano, F.Meneguzzo, A.Mugnai, J J.P.V.Poiares Baptista, F.Porcú, F. Prodi, J.F.W.Purdom, D.Rosenfeld, J.Schmetz, E.A.Smith, F.Tampieri, F.J.Turk, and G.A.Vicente. EURAINSAT - Looking into the future of satellite rainfall estimations. In *EUMETSAT Meteorological Satellite Data Users' Conf.*, pages 375–384, EUM P 33EUMETSAT, Darmstadt, 2001.

[39] V.Levizzani, P.Bauer, and J.F.Turk, editors. *Measuring precipitation from space: EURAINSAT & the future (Advances in global change research)*, volume 28. Springer, 2007.

[40] D.Rosenfeld, R.Lahav, A.Khain, and M.Pinsky. The Role of Sea Spray in Cleansing Air Pollution over Ocean via Cloud Processes. *Science*, 297:1667–1670, 2002.

[41] A.Khain, D.Rosenfeld, and A.Pokrovsky. Aerosol Impact on Precipitation from Convective Clouds. In V.Levizzani, P.Bauer, and J.Turk, editors, *Measuring precipitation from space: EURAINSAT and the Future*, volume 28, pages 421–434. Springer, 2007.

[42] S.Dietrich, R.Solomon, C.Adamo, and A.Mugnai. Rainfall monitoring at geostationary scale: potential of lightning data in a rapid update approach. In *Proceedings 2001 EUMETSAT Meteorological Data Users' Conference*, pages 393–398, Antalya, 2001.

[43] C.Kidd, F.J.Tapiador, V.Sanderson, and D.Kniverton. The University of Birmingham global rainfall algorithms. In V.Levizzani, P.Bauer, and J.Turk, editors, *Measuring precipitation from space: EURAINSAT and the Future*, volume 28, pages 255–267. Springer, 2007.

[44] F.S.Marzano, N.Cimini, and F.J.Turk. Multivariate Probability Matching for Microwave Infrared Combined Rainfall Algorithm (MICRA). In V.Levizzani, P.Bauer, and J.Turk, editors, *Measuring precipitation from space: EURAINSAT and the Future*, volume 28, pages 269–280. Springer, 2007.

[45] F.Torricella, V.Levizzani, and F.J.Turk. Application of a Blended MW-IR Rainfall Algorithm to the Mediterranean. In V.Levizzani, P.Bauer, and J.Turk, editors, *Measuring precipitation from space: EURAINSAT and the Future*, volume 28, pages 497–507. Springer, 2006.

[46] M.Celano, F.Porcú, P.P.Alberoni, and F.Prodi. On the combined use of satellite multispectral and radar polarimetric measurements to infer cloud microphysics. *Meteorol. Atmos. Phys.*, 101:111–126, 2008.

[47] A.Jurczyk, K.Osrodka, and J.Szturc. Research studies on improvement in real-time estimation of radar-based precipitation in Poland. *Meteorol. Atmos. Phys.*, 101:159–173, 2008.

[48] F.Porcú and D.Capacci. Seasonal sensitivity af a VIS-NIR-IR rain-no rain classifier. *Meteorol. Atmos. Phys.*, 101:147–157, 2008.

[49] A.Mugnai, E.A.Smith, G.J.Tripoli, S.Dietrich, V.Kotroni, K.Lagouvardos, and C.M.Medaglia. Explaining discrepancies in passive microwave cloud-radiation databases in microphysical context from two different cloud-resolving models. *Meteorol. Atmos. Phys.*, 101:127–145, 2008.

[50] M.Dorninger, S.Schneider, and R.Steinacker. On the interpolation of precipitation data over complex terrain. *Meteorol. Atmos. Phys.*, 101:175–189, 2008.

[51] M.Verdecchia, E.Coppola, C.Faccani, R.Ferretti, A.Memmo, M.Montopoli, G.Rivolta, T.Paolucci, E.Picciotti, A.Santacasa, B.Tomassetti, G.Visconti, and F.S.Marzano. Flood forecast in complex orography coupling distributed hydrometeorological models and in-situ and remote sensing data . *Meteorol. Atmos. Phys.*, 101:267–285, 2008.

[52] T.Diomede, S..Davolio, C.Marsigli, M.M.Miglietta, A.Moscatello, P.Papetti, T.Paccagnella, A.Buzzi, and P.Malguzzi. Discharge prediction based on multi-model precipitation forecasts. *Meteorol. Atmos. Phys.*, 101:245–265, 2008.

[53] A.Tafferner, C.Forster, M.Hagen, C.Keil, T.Zinner, and H.Volkert. Development and propagation of severe thunderstorms in the Upper Danube catchment area: Towards an integrated nowcasting and forecasting system using real-time data and high-resolution simulations. *Meteorol. Atmos. Phys.*, 101:211–227, 2008.

[54] T.Zinner, H.Mannstein, and A.Tafferner. Cb-TRAM: Tracking and monitoring severe convection from onset over rapid development to mature phase using multi-channel Meteosat-8 SEVIRI data. *Meteorol. Atmos. Phys.*, 101:191–210, 2008.

[55] A.Bartzokas, J.Azzopardi, L.Bertotti, A.Buzzi, L.Cavalieri, D.Conte, S.Davolio, S.Dietrich, A.Drago, O.Drofa, A.Gkikas, V.Kotroni, K.Lagouvardos, C.J.Lolis, S.Michaelides, M.Miglietta, A.Mugnai, S.Music, K.Nicolaides, F.Porcú, K.Savvidou, and M.I.Tsirogianni. The RISKMED project: philosophy, methods and products. *Nat. Hazards Earth Syst. Sci.*, 10:1393–1401, 2010.

[56] A.Buzzi, M.C.Llasat, C.Obled, and R.Romero. The EU Interreg IIIB project HYDROPTIMET. *Nat. Hazards Earth Syst. Sci.*, 5:447–448, 2005.

[57] V.Levizzani, F.Porcú, F.S.Marzano, A.Mugnai, E.A.Smith, and F.Prodi. Investigating a SSM/I Microwave algorithm to calibrate Meteosat infrared instantaneous rainrate. *Meteorol. Appl.*, 3:5–17, 1996.

[58] F.Prodi, D.Capacci, D.Casella, E.Cattani, M.De Rosa, S.Dietrich, F.Di Paola, M.Formenton, S.Laviola, V.Levizzani, G.P.Marra, F.S.Marzano, L.Milani, A.Mugnai, F.Porcú, P.Sanó, and M.L.Tampellini. Precipitation Retrieval and NOwcasting from Satellite within Project "Nowcasting-PROSA" of the Italian Space Agency. Coimbra, Portugal, 23-25 June 2010. International Precipitation Conference (IPC10).

[59] D.Casella. *Analysis and Verification of the Cloud Dynamics and Radiation Database (CDRD) Methodology for Bayesian Retrieval of Precipitation with Passive Microwave Radiometer Instruments Onboard Satellites.* PhD thesis, Department of Science, Technical Information and Communication, University of Rome "La Sapienza", Rome, ITALY, 2010.

[60] P.Sanó. *The Cloud Dynamics and Radiation Database (CDRD) Approach for Precipitation Retrieval by Means of Satellite Based Microwave Radiometry.* PhD thesis, Department of Computer, Systems and Production Engineering, Tor Vergata University of Rome "Tor Vergata", Rome, ITALY,, 2010.

[61] P.Sanó, D.Casella, A.Mugnai, G.Schiavon, E.A.Smith, and G.J.Tripoli. Bayesian Estimation of Precipitation from Satellite Passive Microwave Measurements using CDRD Technique. Part 1: Algorithm Description and Applications over Italy. *IEEE Trans. Geo. Rem. Sens.*, submitted, 2012.

[62] F.Chen and D.H.Staelin. AIRS/AMSU/HSB precipitation estimates. *IEEE Trans. Geosci. Remote Sens.*, 41:410–417, 2003.

[63] F.S.Marzano, G.Rivolta, E.Coppola, B.Tomassetti, and M.Verdecchia. Rainfall Nowcast from Multi-Satellite Passive Sensor Images using Recurrent Neural Network. *IEEE Trans. Geosci. Remote Sens.*, 45:3800–3812, 2007.

[64] R.J.Joyce, J.E.Janowiak, P.A.Arkin, and P.Xie. CMORPH: A method that produces global precipitation estimated from passive microwave and infrared data at high spatial and temporal resolution. *J. Appl. Meteor.*, 5:487–503, 2004.

[65] A.Buzzi, M.Fantini, P.Malguzzi, and F.Nerozzi. Validation of a limited area model in cases of Mediterranean cyclogenesis: Surface fields and precipitation scores. *Meteorol. Atmos. Phys.*, 53:137–153, 1994.

[66] S.Davolio and A.Buzzi. A nudging scheme for the assimilation of precipitation data into a mesoscale model. *Wea. Forecasting*, 19:855–871, 2004.

[67] R.Tettamanti, P.Malguzzi, and D.Zardi. Numerical simulation of katabatic winds with a non-hydrostatic meteorological model. *Polar Atmospheres*, 1:1–95 ISSN 1591–3902, 2002.

[68] O.V.Drofa and P.Malguzzi. Parameterization of microphysical processes in a non hydrostatic prediction model. In *14th Intern. Conf. on Clouds and Precipitation (ICCP)*, pages 1297–3000, Bologna, 19-23 July 2004.

[69] W.F.Krajewski. Ground network: are we doing the right thing? In V.Levizzani, P.Bauer, and J.Turk, editors, *Measuring precipitation from space: EURAINSAT and the Future*, pages 403–417. Springer, 2007.

[70] E.E.Ebert. Methods for verifying satellite precipitation estimates. In V.Levizzani, P.Bauer, and J.Turk, editors, *Measuring precipitation from space: EURAINSAT and the Future*, pages 345–356. Springer, 2007.

[71] F.J.Turk, P.Arkin, M.R.P.Sapiano, and E.E.Ebert. Evaluating High-Resolution Precipitation Products. *Bull. Amer. Meteor. Soc.*, 89:1911–1916, 2008.

[72] B.W.Golding. Nimrod: a system for generating automated very short range forecasts. *Meteorol. Appl.*, 5(1):1–16, March 1998.

[73] W.H.Hand. An object-oriented technique for nowcasting heavy showers and thunderstorms. *Meteorol. Appl.*, 3(1):31–41, March 1996.

[74] C.E.Pierce, P.J.Hardaker, C.G.Collier, and C.M.Haggett. GANDOLF: a system for generating automated nowcasts of convective precipitation. *Meteorol. Appl.*, 7(4):341–360, December 2000.

[75] N.E.Bowler, C.E.Pierce, and A.W.Seed. STEPS: a probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Q. J. R. Meteorol. Soc.*, 132:2127–2155, 2006.

[76] D.Grimes, E.Pardo-Iguzquiza, and R.Bonifacio. Optimal areal rainfall estimation using raingauges and satellite data. *J. Hydrol.*, 222(1-4):93–108, 1999.

[77] R.S.Chadwick, D.I.F.Grimes, R.W.Saunders, P.N.Francis, and T.A.Blackmore. The TAMORA algorithm: satellite rainfall estimates over West Africa using multi-spectral SEVIRI data. *Adv. Geosci.*, 25:3–9, 2010.

[78] K.L.Hsu, X.Gao, S.Sorooshian, and H.V.Gupta. Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks. *J. Appl. Meteor.*, 36(9):1176–1190, September 1997.

[79] S.Sorooshian, K.L.Hsu, H.V.Gupta, B.Imam, and D.Braithwaite. Evaluation of PERSIANN System Satellite-Based Estimates of Tropical Rainfall. *Bull. Amer. Meteor. Soc.*, 81(9):2035–2046, 2000.

[80] PERSIANN webpage: http://chrs.web.uci.edu/research/satellite_precipitation/activities00.html.

[81] Y.Hong, K.L.Hsu, S.Sorooshian, and X.Gao. Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks Cloud Classification System. *J. Appl. Meteor.*, 43(12):1834–1853, December 2004.

[82] PERSIANN-CCS webpage: http://chrs.web.uci.edu/research/satellite_precipitation/activities01.html.

[83] J.M.Fernandez, H.Legleau, V.Zwatz-Meise, and A.I.Dybbroe. The Satellite Application Facility (SAF) of EUMETSAT to support nowcasting - An introduction. In *9th Conference on Satellite Meteorology and Oceanography*, pages 319–322, Paris, France, 25-29 May 1998.

[84] A.Tokay, D.B.Wolff, K.R.Wolff, and P.Bashor. Rain Gauge and Disdrometer Measurements during the Keys Area Microphysics Project (KAMP). *J. Atmos. Oceanic Technol.*, 20(11):1460–1477, 2003.

[85] C.E.Duchon and C.J.Biddle. Undercatch of tipping-bucket gauges in high rain rate events. *Adv. Geosci.*, 25:11–15, 2010.

[86] C.E.Duchon and G.R.Essenberg. Comparative rainfall observations from pit and aboveground gauges with and without wind shields. *Water Resour.*, 37(12):3253–3263, 2001.

[87] G.Leitinger, N.Obojes, and U.Tappeiner. Accuracy of winter precipitation measurements in alpine areas: snow pillow versus heated tipping bucket rain gauge versus accumulative rain gauge. In *EGU General Assembly Conference Abstracts*, volume 12 of *EGU General Assembly Conference Abstracts*, page 5076, Vienna, Austria, May 2010.

[88] K.Sugiura, D.Yang, and T.Ohata. Systematic error aspects of gauge-measured solid precipitation in the Arctic, Barrow, Alaska. *Geophys. Res. Lett.*, 30(4):1192, 2003.

[89] L.G.Lanza, E.Vuerich, and I.Gnecco. Analysis of highly accurate rain intensity measurements from a field test site. *Adv. Geosci.*, 25:37–44, 2010.

[90] A.M.Rossa, F.L.Del Guerra, M.Borga, F.Zanon, T.Settin, and D.Leuenberger. Radar-driven high-resolution hydro-meteorological forecasts of the 26 September 2007 Venice flash flood. *J. Hydrol.*, 394:230–244, 2010.

[91] C.J.Smith. Reduction of Errors Caused by Bright Bands in Quantitative Rainfall Measurements Made Using Radar. *J. Atmos. Oceanic Technol.*, 3:129–141, March 1986.

[92] M.Celano, P.P.Alberoni, V.Levizzani, and A.R.Holt. Analysis of severe convective events from two operational dual polarisation doppler radars. *Nat. Hazards Earth Syst. Sci.*, 6:397–405, 2006.

[93] V.N.Bringi and V.Chandrasekar. *Polarimetric Doppler Weather Radar*. Cambridge University Press, 2001.

[94] A.V.Ryzhkov, S.E.Giangrande, and T.J.Schuur. Rainfall Estimation with a Polarimetric Prototype of WSR-88D. *J. Appl. Meteor.*, 44(502-515), 2005.

[95] J.J.Gourley, S.E.Giangrande, Y.Hong, Z.L.Flamig, T.Schuur, and J.A.Vrugt. Impacts of Polarimetric Radar Observations on Hydrologic Simulation. *J. Hydrometeor.*, 11(3):781–796, June 2010.

[96] D.Capacci. *Estimation of Probability of Precipitation using VIS/NIR/IR satellite measurement*. PhD thesis, Università degli studi di Ferrara, 2003-2004.

[97] F.L.Chang and Z.Li. Estimating the vertical variation of cloud droplet effective radius using multispectral near-infrared satellite measurements. *J. Geophys. Res.*, 107(D15):4257, 2002.

[98] I.M.Lensky and D.Rosenfeld. Estimation of Precipitation Area and Rain Intensity Based on the Microphysical Properties Retrieved from NOAA AVHRR Data. *J. Appl. Meteor.*, 36(3):234–242, 1997.

[99] D.H.Staelin and F.W.Chen. Precipitation Observations Near 54 and 183 GHz Using the NOAA-15 Satellite. *IEEE Trans. Geosci. Remote Sens.*, 38(5):2322–2332, September 2000.

[100] SEVIRI image on ESA webpage:
http://www.esa.int/msg/sev_zeiss_bad.html.

[101] V.M.Krasnopolsky and H.Schiller. Some neural network applications in environmental sciences. Part I: forward and inverse problems in geophysical remote measurements. *Neural Networks*, 16(Special Issue):321–334, 2003.

[102] F.J.Tapiador, C.Kidd, K.-L.Hsu, and F.Marzano. Neural Networks in satellite rainfall estimation. *Meteorol. Appl.*, 11:83–91, 2004.

[103] D.Capacci and F.Porcú. Evaluation of a Satellite Multispectral VIS-IR Daytime Statistical Rain-rate Classifier and Comparison with Passive Microwave Rainfall Estimates. *J. Appl. Meteor. Climat.*, 48:284–300, 2009.

[104] T.Bellerby, M.Todd, D.Kniveton, and C.Kidd. Rainfall Estimation from a Combination of TRMM Precipitation Radar and GOES Multispectral Satellite Imagery through the use of an Artificial Neural Network. *J. Appl. Meteor.*, 39:2115–2128, 2000.

[105] A.Behrangi, K.-L.Hsu, B.Imam, S.Sorooshian, G.J.Huffman, and R.J.Kuligowski. PERSIANN-MSA: a Precipitation Estimation Method from Satellite-Based Multispectral Analysis. *J. Hydrometeor.*, 10(6):1414–1429, December 2009.

[106] C.M.Bishop. *Neural Networks for Pattern Recognition.* Clarendon Press - Oxford, 1995.

[107] B.Müller, J.Reinhardt, and M.T.Strickand. *Neural Networks: an introduction.* Springer, 1995.

[108] D.Capacci and J.Conway. Delineation of precipitation areas from modis visible and infrared imagery with artificial neural networks. *Meteorol. Appl.*, 12:291–305, 2005.

[109] C.A.Fair, P.K.James, and P.Larke. The United Kingdom weather radar network. In C.G.Collier and M.Chapuis, editors, *Weather Radar Networking, Seminar on COST Project 73*, pages 147–154. Kluwer Academic, 1990.

[110] F.Porcú, D.Capacci, F. Prodi, S.Dietrich, and E.Santorelli. La precipitazione da satellite: un prodotto alternativo per le applicazioni agrometeorologiche di monitoraggio ambientale. *Rivista Italiana di Agrometeorologia*, 3:51–55, 2005.

[111] D.Capacci, F.Porcú, P.N.Francis, B.Conway, and F.Prodi. Probability of precipitation estimation using seviri data and artificial neural networks. In *Proceedings of the 2005 EUMETSAT Meteorological Satellite Conference*, 2005.

[112] *WWRP/WGNE Joint Working Group on Verification, 2005: Forecast Verification*, volume Methods and FAQ. 2005.