Universidad de Concepción
Dirección de Postgrado
Facultad de Ingeniería Agrícola
Programa de Doctorado en Ingeniería Agrícola con
mención en Recursos Hídricos en la Agricultura

Agricultural drought in Chile: From the assessment toward prediction using satellite data

Sequía Agrícola en Chile: De la evaluación hacía la predicción usando datos satelitales

Tesis para optar al grado de Doctor en Ingeniería Agrícola con mención en Recursos Hídricos en la Agricultura

FRANCISCO JAVIER ZAMBRANO BIGIARINI
CHILLÁN-CHILE
2017

Profesor Guía: Mario Lillo Saavedra
Dpto. de Mecanización y Energía, Facultad de Ingeniería Agrícola
Universidad de Concepción

Agricultural drought in Chile: From the assessment toward prediction using satellite data.

Thesis approved by

Mario Lillo Saavedra Ing. Civil Eléctrico, Doctor	Profesor Guia
Octavio Lagos Ing. Civil Agrícola, Ph.D.	Evaluador Interno
Aldo Montecinos Oceanógrafo, Doctor	Evaluador Interno
Francisco Meza Ing. Agrónomo, Ph.D.	Evaluador Externo
Gabriel Merino Licenciado en Física, Ph.D.	Director (S) de Programa

Contents

List of Tables

List of Figures

Preface

> I think it only makes sense to seek out and identify structures of authority, hierarchy, and domination in every aspect of life, and to challenge them; unless a justification for them can be given, they are illegitimate, and should be dismantled, to increase the scope of human freedom.
> – Noam Chomsky

I started doctoral studies because I realized after having worked for multiple government institutions through Chile, that what I really like the most and I am passionate about it, is to research and in fact was the job I was really doing. I have found some people that seems to be more worried about the Doctoral diploma, which would like to think that to get one is like to get a noble title and they want his throne and crown and that the people call them *Doctor*, rather than just a research degree made to improve some research field. By the way, you do not need to be a genius to get a Doctoral diploma, just a myth. Nevertheless, the truth be told, you can have a doctorate and not to be a good researcher or not have one and be a tremendous researcher. Regrettably, I have found several cases from the first and huge egos during my life as researcher. Thankfully, some of the best researchers and Doctors I have found in Chile and abroad, are the most humble people. For me, the passion for doing what I like the most has been the engine that has kept me running forward, no matter what.

The doctorate has been a long journey with a litlle of bitter and a lot of sweet taste. I have saved for me the sweet and I have learned from the bitter. In some stage of the path, I felt like a humble soccer team, which has to play against a powerful team (in terms of have power but not regard quality). The adversary team has doping his players, paid to the referee, play a foul game, and even in the middle of the game has turned off the stadium lights and makes a goal. Then, like in a movie, the humble team have to stand up, shake off the dirt and play masterfully to finally win the match. In this sense, I share a part of a speech from the movie *Rocky* when Rocky is talking to his son about life, that reflect this feeling :

> Let me tell you something you already know. The world ain't all sunshine and rainbows. It's a very mean and nasty place and I don't care how tough you are it will beat you to your knees and keep you there permanently if you let it. You, me, or nobody is gonna hit as hard as life. But it ain't about how hard ya hit. It's about how hard you can get hit and keep moving forward. How much you can take and keep moving forward. That's how winning is done!

Besides, I wanted to become Doctor because I thought that was a career in which meritocracy can work, and mostly it does despite a few outliers. In that regard, as a Doctor I compromise to always judge a research work for its merit, be always learning and hold a high ethical standard as researcher, no matter what.

Dedication

To my beloved sons Matteo and Luca

Acknowledgement

> Nobody can say anything about you. Whatsoever people say is about themselves. But you become very shaky, because you are still clinging to a false center. That false center depends on others, so you are always looking to what people are saying about you. And you are always following other people, you are always trying to satisfy them. You are always trying to be respectable, you are always trying to decorate your ego. This is suicidal. Rather than being disturbed by what others say, you should start looking inside yourself...
> – Osho

I started to research about agricultural drought thanks to Mr. Isaac Maldonado, who brought me to the *Agricultural Research Institute (INIA)* of Chile to improve the work about vegetation indices for monitoring agricultural drought over Chile. Thanks to this work the Professors of the University of Concepción Dr. Mario Lillo and Dr. Diego Rivera offered me to start doctoral studies which I did, then Dr. Mario Lillo become in my advisor. I appreciate the freedom that he gave me to develop my research. The Thesis path was as a long ladder with several steps. One of the first step that help me to improve it, was thanks to the collaboration of Dr. Koen Verbist from the International Hydrological Programme (IHP) at UNESCO, who was advisor for the Ministry of Agriculture of Chile regarding climate emergencies. His vision and knowledge helped me for the second chapter of this Thesis in the sense of gave it more scientific relevance thanks to which later it could be published. Also, I would like to thanks to Dr. Octavio Lagos who put my in contact with the University of Nebraska, Lincoln, in the U.S.; which made possible my stay there as Visiting Scholar, which allowed me to gave the second big step in improving my research and publish the second article. But, none of this would have been possible without funds, thanks to the National Commission for Scientific and Technological Research (CONICYT) for the funding and support in Chile and abroad along the doctorate. Also, thanks to the additional funding that was provided through the project *Agua y Energa: Una integración interdisciplinaria para el fortalecimiento del Programa de Doctorado en Ingeniera Agrícola (UCO-1407)* by the University of Concepción. Thanks to the postgraduate direction of the University of Concepción for his support and recognition. Lastly, but not less, I am much grateful of Dr. Francisco Meza from the Pontifical Catholic University in Chile for accepting to be part of this Thesis at last moment.

Thanks to my broad family. My parents who have given me housing and support in between changes through my life. To my Mom Giulia, a wonderful woman, I think the best person that I know, who tough me principles with her example. To my elder brother Mauricio, for have been such a good example of life as well as a researcher. We are different, but I recognize you as a kind and nice person as well as a top researcher, who has known how to be ethic, honest, hard working and perfectionist. Thanks to your many pieces of advice. To your wife Carolina and your sons Francesco and Dario for remain close despite the distance and time. To my sister Loreto and her beloved daughters Antonella and Giulia, for her joy and spontaneous way of being. To my brother Felipe, her wife Rosita and the little fighter Luciano, for given me housing in their home when I had to travel to Santiago for a conference, reunion, embassy issue, or any matter; during the time I have been doing the doctorate. To my elder nephew Matias, for being part of the family. To my beloved son **Matteo**, for being the flame on my heart, and the kind of person I would like to become. To my second son **Luca** with whom I could not be so close as I would have like, but you are part of me and I love you.

During the time I was doing my doctorate I had the luck to be abroad in the United States and in The Netherlands as visiting researcher. Thanks, to the people from the *Water For food Institute*, *Center for*

Advanced Land Mangagement Information Technologies (CALMIT), and *National Drought Mitigation Center (NDMC);* all from the University of Nebraska, Lincoln, U.S. where I was working during the first semester of 2016. Thanks to Christopher Neale and his family for helping me in my arriving to Lincoln and all the help during my stay there. Thanks to Dr. Brian Wardlow and Dr. Tsegaye Tadesse, for gave me his feedback and help to write the manuscript that we co-authored during my time there. This work gave me the confidence that I was doing was right and I was following the correct path in my research. Thanks to the people of Lincoln Literacy where I periodically assist for improving my English (still is not good but always improving), where I found people all around the world, Russia, Japan, Spain, Eritrea, Turkey, Vietnam and many other countries more.

Also, I was as a visiting researcher for four months in the Netherlands, what a wonderful society!. What I love the most was the bicycle culture and its egalitarian and respectful society. This was just a brief time but highly productive. I was invited by Dr. Andy Nelson from the Natural Resources Department (NRS) of the Faculty of Geo-Information Science and Earth Observation (ITC) of the University of Twente. There, I had the luck to know and be working with Dr. Anton Vrieling. To understand how important in my research this was, I need to extend the story. The first article I read which become in the cornerstone of my research those days, by which I was able to derive the first agricultural drought index I used (now used for the Ministry of Agriculture of Chile) was one published in the journal Remote Sensing of Environment and co-authored by Oscar Rojas, Anton Vrieling, and Felix Rembold. Years later, when I sent my first article to be published in the Remote Sensing (MDPI) journal, Anton was Academic Editor and who deeply review the manuscript allowing to making several improvements. Well, then at the ITC we worked together with Anton and Andy in a research about agricultural drought prediction in Chile. This work only was possible thanks to the knowledge of Anton about agriculture monitoring using remote sensing and his advice in how to improve the research. Thanks to pushing my always two step beyond. Maybe was a brief time in The Netherlands but I learn a lot, and I could highly improve my research work. I was a hard working person during my time there because I feel the needed to make something significant, which I think I did. The ITC is a such a great environment to work, all the people very nice and helpful, sorry that for me was such a brief time. Finally, thank to my Chilean roommate and friend who met in Enschede, Coti. She shows me some thing about living in The Netherlands that help me to make the life easier. Thanks for the time we share, the Wednesday night dinners, and all the talk.

Abstract

Climate change is occurring and there is a scientific consensus that human being is playing a key role by pouring greenhouses gases to the atmosphere. Temperature has been increasing globally and the precipitation patterns are changing. Regionally, since the year 2010 Chile has been experiencing which has been called a mega drought, however, it has been seen mostly in meteorological terms by analyzing precipitation deficits. Further, the future projection for Chile indicates that the precipitation will decrease in Central-South Chile, this addded to the increase on temperature likely could increase drought frequency and intensity. Also, in this regard crop yield of corn and wheat decreases are forecasted by 2050 for Chile.

The study on how climate variability and human activity impact agriculture has been known as agricultural drought. One of the main factors that trigger this drought conditions is precipitation deficit, thus is crucial to understand how this depletion relates to agriculture development. Although, since 2010 Chile has been facing water shortage mostly as results of the analysis of annual precipitation, but still there is a lack of knowledge about how this mega drought is affecting agriculture over Chile. Moreover, during the growing season 2007-2008 a large part of the country experienced decreases in crop yield for which these areas were declared under drought emergency by the government. However, by analyzing the total amount of annual precipitation these years are not seen as relevant drought years. This happens in part because for vegetation is more important the timing of the rainfall deficit rather than the cumulative over a year. Thus, the study and understanding of agricultural drought and methods that could help to anticipate it are challenging.

The study of agricultural drought at regional and global scale brings the problem of having enough data that allow to analyze it spatially and temporally. Nonetheless, since the 70's the use of remote sensing data obtained from satellite to monitor the environment at global and regional scale has been highly improved, and nowadays are a key data source to support climatic and environmental studies. In that regard, there is an important amount of satellite-derived data publicly available. One of this dataset that provided useful data for the monitoring of vegetation is provided for the National Aeronautics and Space Administration (NASA) and its sensor the *Moderate-Resolution Imaging Spectroradiometer* (MODIS) which is coupled to the TERRA and AQUA satellites. Further, multiple microwave and infrared satellites have allowed the development of precipitation estimates products at different temporal and spatial resolutions Between them, highlight the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) having data since 1998, and also has been derived long-term precipitation products such as the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Climate Data Record (PERSIANN-CDR) with data since 1983, and the Climate Hazards Group InfraRed Precipitation with Station data version 2 (CHIRPS v2) providing estimates since 1981. These vegetation and precipitation satellite products are valuable data sources for agricultural drought studies, allowing to evaluate the interaction vegetation-precipitation at regional and global scale.

Accordingly, in this thesis was studied the usefulness of satellite data for the assessment and prediction of agricultural drought over Chile. The main research question is: How well the satellite data of vegetation and precipitation together with climatic oscillation indices can be used to predict agricultural drought before the end of the growing season? To achieve this, the work was developed in three stages: 1) assessment of vegetation response to water shortage in the BioBío Region of Chile, 2) the evaluation of long-term satellite precipitation data over Chile for use in drought studies, and 3) prediction of agricultural drought in Chile from one to four month before the end of the growing season for 2000-2016. For the first stage, was used

the Vegetation Condition Index (VCI) which was derived from the Normalized Difference Vegetation Index (NDVI) provided by the MODIS Vegetation Indices (VI) product MCD12Q1.005. The land cover product MCD12Q2.051 from MODIS was used to derive a cropland mask for the BioBio region. Besides, twenty six weather stations were used to derive the Standardized Precipitation Index (SPI) at time-scales from one to six months. To understand how the vegetation response to water shortage, the VCI over the growing season was correlated against the SPIs. For this the data was aggregated for the whole region and per administrative unit, and also was analyzed the point-to-point correlation. For the second stage, were used 758 weather station over Chile having monthly precipitation data since 1981 to 2015. Then, were selected the long-term satellite data products, CHIRPS v2 for 1981-2015, and the PERSIANN-CDR for 1983-2015. The accuracy when measuring monthly precipitation was evaluated for these two long-term precipitation products together with the TMPA product. These three products were compared with the point precipitation data obtained at weather stations. Also, the long-term products were evaluated for drought monitoring using the SPI for one, three, and six months. Finally, for the third stage, the study area was defined for 29-41°S through Chile. Over this region, the product MCD12Q2.051 from MODIS was used to derive a single cropland mask over Chile. From the 2221 census unit used for the Ministry of Agriculture of Chile for the 10-year agricultural census over the study area, 758 were selected by filtering using the cropland mask. Per census unit was extracted the growing season start (SOS) and end (EOS) from the product MCD12Q2.005 from MODIS. Over each census unit was calculated a proxy of biomass production, the anomaly of cumulative NDVI (zcNDVI) for 2000-2016. As predictors per census unit were calculated from one to four months before EOS the zcNDVI and SPI at one, three, six, twelve, and twenty-four months. Also, were used as a predictor for one to four months before EOS the three-month average Pacifical Decadal Oscillation (PDO) and Multivariate ENSO index (MEI), which were also lagged at 0, 3, and 6 months; from the corresponding prediction timing months before EOS. Two methods were assessed for the prediction, an Optimal Linear Regression (OLR) per census unit which selects the predictor that produce less error with a cross-validated linear regression. The second method corresponds to a Multi Layer Feed-Forward Neural Network also called Deep Learning (DL), for which were created cross-validated models considering all units and using as additionals predictor latitude and longitude.

Results from the first stage showed that the 3-month SPI (SPI-3), calculated for the modified growing season (Nov-Apr) instead of the regular growing season (Sept-Apr), has the best Pearson correlation with VCI values with an overall correlation of 0.63 and between 0.40 and 0.78 for the administrative units. These results show a very short-term vegetation response to rainfall deficit in September, which is reflected in the vegetation in November, and also explains to a large degree the variation in vegetation stress. It is shown that for the last 16 years in the BioBío Region we could identify the 2007-2008, 2008-2009, and 2014-2015 seasons as the three most important drought events; this is reflected in both the overall regional and administrative unit analyses. These results concur with drought emergencies declared by the regional government. Next, from the second stage, results showed that the monthly analysis for all satellite products highly overestimated rainfall in the arid North zone. However, there were no major differences between all three products from North to South-Central zones. Further, in the South zone, PERSIANN-CDR shows the lowest fit with high underestimation, while CHIRPS 2.0 and TMPA 3B43 v7 had better agreement with in-situ measurements. The accuracy of satellite products were highly dependent on the amount of monthly rainfall with the best results found during winter seasons and in zones (Central to South) with higher amounts of precipitation. PERSIANN-CDR and CHIRPS 2.0 were used to derive SPI at time-scales of 1, 3 and 6 months, both satellite products presented similar results when were compared in-situ against satellite SPIs. Because of its higher spatial resolution that allows better characterizing of spatial variation in precipitation pattern, the CHIRPS 2.0 was used to mapping the SPI-3 over Chile. Finally, from the third stage, results from the two prediction models evaluated (OLR and DL) showed similar and good prediction accuracy, with mean R^2cv values for OLR of 0.94, 0.79, 0.63 and 0.51, and for DL of 0.93, 0.79, 0.63 and 0.51, for one, two, three and four months before EOS respectively. Also, was discussed potential model improvements and how the method could contribute to an early warning system for agricultural drought in Chile.

Chapter 1

Introduction

According to the last report of the Intergovernmental Panel on Climate Change (IPCC) (IPCC, 2013), indicate that there is no doubt about that climate change was occurring. Besides, is extremely likely that human influence has been the dominant cause of the observed warming since the mid-20th century which was caused by the anthropogenic increase in greenhouse gas concentrations and other anthropogenic forcings (IPCC, 2013). Moreover, has been quantified the scientific consensus, studies indicate that ninety-seven percent of climate scientists agree that climate-warming trends over the past century are due to human activities (Cook et al., 2013, 2016) (Fig. 1.1a). Worldwide, climate change cause variations in climatic patterns (Van der Wiel et al., 2017; Dore, 2005), which makes vary the normal precipitation and temperature conditions. There is evidence that global temperature has been increasing since 1880 (Fig. 1.1b) (Hansen et al., 2010; Mann et al., 2007) and the last years has been the warmest. Regarding precipitation, it will increase and decrease both in intensity and frequency for different zone all around the globe, which has begun to appreciate in different parts of the world (Van der Wiel et al., 2017) including Chile (Garreaud et al., 2017). These changes in precipitation and temperature will impact the occurrence of natural hazards such as drought (Dai, 2012).

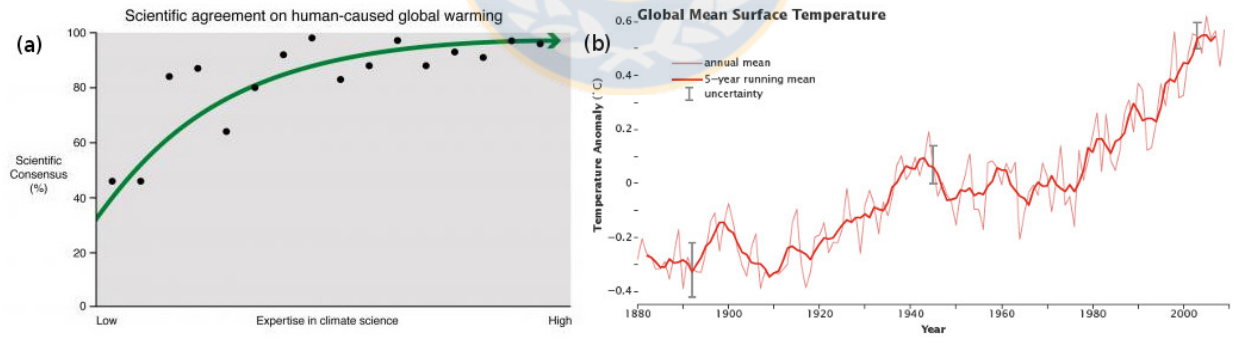

Figure 1.1: (a) Scientific agreement vs expertise in climate based in the research of Cook et al. (2013), (b) Global mean surface temperature since 1880 (Hansen et al., 2010)

One of the most complex natural hazards is drought, due that there is not a recognized start neither an end, these timings are hard to identify because drought has a slow onset, for which has often been called a *creeping* phenomenon (Gillette, 1950). Also, its complexity it is due to the different hydrometeorological parameters with which it is related. For this, there is no universal definition of drought (Mishra and Singh, 2010). However, an accepted simple definition is an exceptional lack of water respect to normal condition. Climate variability has been seen as the main factor that triggers drought conditions and the lack of water at regional and global scale has been referred mainly to rainfall. Thus, depending on the time-scale at which the shortage of rainfall occur drought has been classified mainly in three groups (Wilhite and Glantz, 1985). At

short-term scale (< 3 months) is considered a meteorological drought, when the rainfall deficit remains and start to affect soil moisture and consequently vegetation development (< 9 months) has been denominated agricultural drought, finally the persistence of the deficit (> 12 months) will affect river, reservoirs and groundwater levels, and this has been called hydrological drought. Nonetheless, the mentioned classification of drought has been seen as a natural phenomenon only, recently some authors (Loon et al., 2016) suggest that research about drought need to recognize that human activities modify the hydrological processes underlying drought propagation (Fig. 1.2), and this can be recognized explicitly in the definition of drought by distinguishing between climate-induced drought, human-induced drought and human-modified drought (Fig. 1.3).

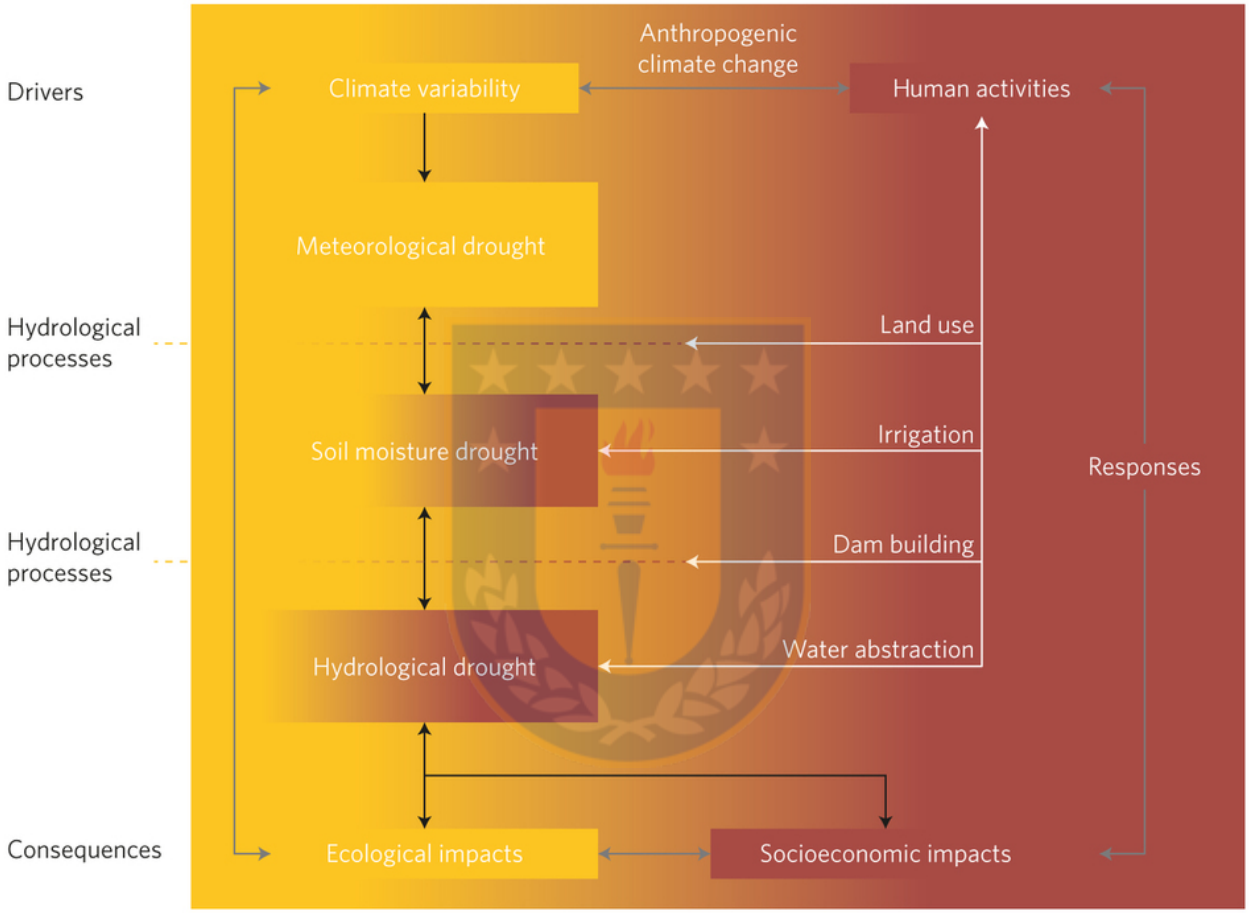

Figure 1.2: Drought propagation in the Anthropocene. The propagation from meteorological drought to soil moisture and hydrological drought (black arrows) is initiated by climatic (left; yellow) and human (right; red-brown) drivers. Drought is modified by hydrological catchment processes (dotted lines) that are altered by human activities (white arrows). The resulting ecological and socioeconomic impacts initiate responses, which in turn result in changes to the human influence on drought and the climate variability (grey arrows). Source: Loon et al. (2016)

Regionally, for the Center-South zone of Chile is forecasted a decreasing of precipitation (IPCC, 2013). Furthermore, since the year 2010 an unprecedented precipitation deficit has taken place in Central Chile (30-38°S) (Garreaud et al., 2017) which has been called a *mega drought*. Findings from the study of Boisier et al. (2016) estimate that a quarter of this rainfall deficit is from an anthropogenic origin. Also, Garreaud et al. (2017) indicate that the *mega drought* has reached even farther south (>38°S). This is even more relevant if is considered that the Center and South of Chile has a humid climate and concentrated an important extension of crops. Then, the drought impact in this zone potentially would be more difficult to overcome

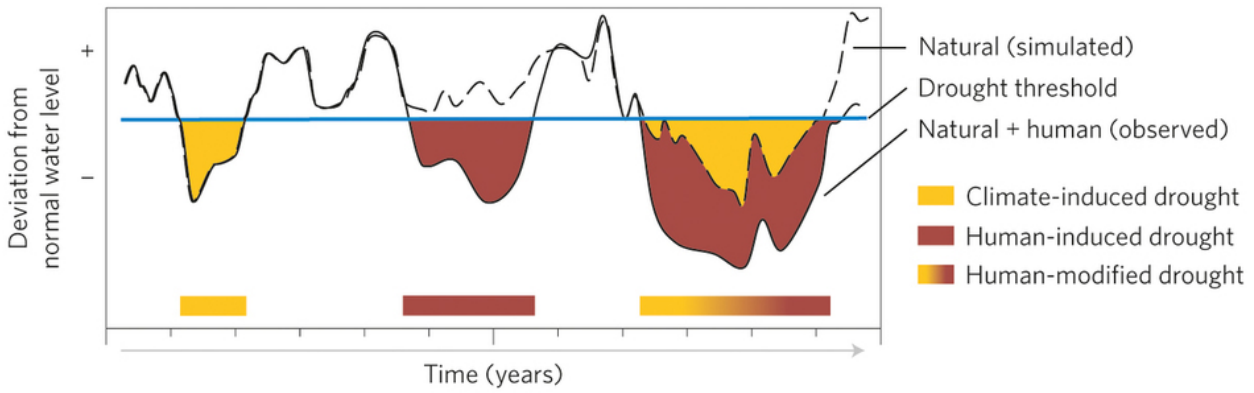

Figure 1.3: Hypothetical schematic of the distinction between climate-induced drought, human-induced drought and human-modified drought. Observed water levels (solid black line) that are influenced by both natural and anthropogenic factors are compared to simulated water levels from virtual model experiments (dashed line) that consider only natural drivers. Note that human-modified drought can be aggravated or alleviated with respect to the natural situation. Source: Loon et al. (2016)

than in other zones with higher preparedness to cope drought. Moreover, studies forecast a decreasing in wheat and corn yield for Chile for 2050 from about 15 to 20% (Meza and Silva, 2009). Even though, others studies (Cohn et al., 2016) consider that impacts on agricultural production could be larger when taking the effects on cropping frequency and area into account. The study of agricultural drought seeks to understand the relationship for multiple environmental, climatic and human-induced factor and its effect on agriculture conditions.

Precipitation deficit is the main driver that trigger agricultural drought at regional and global scale. Nowadays, several indices exist which allows understanding how different time-scales of rainfall deficit are related to agriculture. One of the most used indices which have been recommended for the WMO (World Meteorological Organization) as an index to characterize droughts (Hayes et al., 2011) is the SPI (Standardized Precipitation Index; McKee et al. (1993)). This index uses historical rainfall data ($>$ 30 years) which is then transformed to a normal distribution providing values in term of standard deviations from an average value equal to zero. The main feature of this index is its multi-scale calculation capability, for which could be used considering different ranges of accumulated precipitation. This characteristic allows it to be related to different classes of drought. Short-term SPI ($<$ 9 months) are showed to have more strong relation with soil moisture and vegetation response. In the last decade, a new index the SPEI (Standardized Precipitation Evapotranspiration Index; Vicente-Serrano et al. (2010)) was developed to improves SPI. This index adds the effect of temperature to the SPI, as a simplified water balance (precipitation less evapotranspiration). As well as the SPI, SPEI is a multi-scale drought index. Perhaps, one of the main restriction to use these indices it is the availability of enough spatial data for its calculation. For Chile, there is a lack in the amount of weather station as well as historical records of precipitation and temperature. However, recently has been developed satellite-derived products with a long-term record of precipitation (Funk et al., 2015; Ashouri et al., 2015) and spatially distributed at differents spatial resolution which could be used to overcome this issue. Then, to understand how much of the vegetation response could be explained by the SPI/SPEI we need a measure or proxy of vegetation biomass to support the analysis.

Data about crops such as type, yield, and phenology is very difficult to collect and spatialize to be used at regional scale. Nonetheless, since the space age, the increase on research findings has been large for a broad number of matters, including those regarding climate and how it affects vegetation. In the early 70's with the launch of the Landsat-1 mission, results in the derivation and first uses of the NDVI (Normalized Difference Vegetation Index; Rouse et al. (1974)), which by using the absorbed and reflected radiation from vegetation gives a measure of vegetation development and quality. This index has been widely used as a proxy for biomass productivity of vegetation, (Jung et al., 2008; Rigge et al., 2013), thus allowing to fill the gap for the spatial analysis of vegetation. Further, for agricultural drought monitoring and analysis

using NDVI has been derived other indices. For example, the z-score NDVI (zNDVI; Peters et al. (2002)) and the VCI (Vegetation Condition Index; Kogan (1995a)). Nevertheless, the satellite products of NDVI give a spatial measure of a broad extension mixing vegetated and non-vegetated areas. But, for agricultural drought analysis, the results will be better as we are capable of considering the surface which is effectively used for crops and during the time in which there are growing. Thus, there are satellite products about land cover type, which allow identifying at regional scale the extension of croplands (Friedl et al., 2010). Also the growing season for crops can be obtained whether from the analysis of time-series of NDVI (Vrieling et al., 2017; Meroni et al., 2014b; Vrieling et al., 2011) or from satellite-derived product for phenology dynamics (Ganguly et al., 2010).

Climate change arises big challenges worldwide under the actual forecasted climate, and particularly for Chile, which is facing *mega drought* condition since 2010 (Garreaud et al., 2017) which has been unprecedented to the South. Satellite-derived agricultural drought indices have been helping to monitor drought condition allowing to stakeholder the taking of decision regarding the spatial extension, intensity, severity, and duration of drought. The undergoing research about drought has allowed advances in the understanding of the relation between the different components involved: precipitation, vegetation, soil moisture, climatic oscillation indices; and thus identify the more relevant parameters that allow explaining the variability regarding the vegetation development and condition. However, there is a scope in this research field, for the study of using satellite data to anticipate agricultural drought conditions.

1.1 Hypothesis

Regarding the current and forecasted climate condition, the hypothesis of the thesis state that vegetation satellite data can be used to derive a interannual proxy for biomass production spatially and timely relevant as an indicator of agricultural drought over Chile, and this proxy could be predicted before the end of the growing season using precipitation long-term and vegetation satellite-derived data together with climatic oscillation indices.

1.2 Research objectives

1.2.1 General objective

The aim of this research work is to assess multiple time-series of satellite data publicly-available for vegetation and precipitation retrievals together with climatic oscillation indices toward analyzing its predictive power for agricultural drought in Chile.

1.2.2 Specific objectives

1. To assess VCI derived from MODIS data at 250m spatial resolution as an effective indicator for monitoring agricultural drought and evaluate its observed impact in the BioBío Region over the period 2000-2015.
2. To evaluate the performance and fit of monthly, long-term satellite-based precipitation products over Chile for mapping and quantifying historical rainfall and drought patterns.
3. Derive a proxy of biomass production during the growing season as indicator of agricultural drought and assess its prediction using multiple spatiotemporal satellite data and climatic oscillation indices.

1.3 Thesis outline

This thesis consists of five chapters. The three core chapters (2-4) focus on the aforementioned three research objectives respectively. Chapters 2 and 3 have been published, and chapter 4 is nearly to be submitted for publication, all of them as peer-reviewed papers in one of the Web of Science journals.

Chapter 1 first describes the rational behind the selection of the research topic. The concepts used, hypothesis and objectives are then presented.

Chapter 2 describe the assessment of agricultural drought for 2000-2015 over the cropland area of the BioBío región in Chile using the VCI and analyzing its relation with different time-scales of SPI derived from weather stations.

Chapter 3 present the evaluation of two long-term satellite-derived products from which monthly rainfall was compared against data extracted from weather stations, and also was assess its application for drought monitoring.

Chapter 4 was derived a proxy of biomass production during the growing season, the cumulative anomaly of NDVI (zcNDVI), to be used as an index for agricultural drought in Chile. Further, were analyzed two methods for the prediction of agricultural drought from satellite-derived precipitation indices, climatic oscillation indices together with the zcNDVI at different prediction timing before the end of the season.

Chapter 5 present the overall and specific conclusion of the thesis. Finally, in the Appendix A was included a brief review of other useful satellite data to be considered for future works.

Chapter 2

Agricultural drought in the BioBío Region of Chile

Zambrano, F.; Lillo-Saavedra, M.; Verbist, K. & Lagos, O. Sixteen years of Agricultural Drought Assessment of the BioBío Region in Chile using a 250m resolution Vegetation Condition Index (VCI) Remote Sensing, 2016, 8, 530.

Abstract

Drought is one of the most complex natural hazards because of its slow onset and long-term impact; it has the potential to negatively affect many people. There are several advantages to using remote sensing to monitor drought, especially in developing countries with limited historical meteorological records and a low weather station density. In the present chapter, we assessed agricultural drought in the croplands of the BioBío Region in Chile. The vegetation condition index (VCI) allows identifying the temporal and spatial variations of vegetation conditions associated with stress because of rainfall deficit. The VCI was derived at a 250m spatial resolution for the 2000-2015 period with the Moderate Resolution Imaging Spectroradiometer (MODIS) MOD13Q1 product. We evaluated VCI for cropland areas using the land cover MCD12Q1 version 5.1 product and compared it to the in situ Standardized Precipitation Index (SPI) for six-time scales (1-6 months) from 26 weather stations. Results showed that the 3-month SPI (SPI-3), calculated for the modified growing season (Nov-Apr) instead of the regular growing season (Sept-Apr), has the best Pearson correlation with VCI values with an overall correlation of 0.63 and between 0.40 and 0.78 for the administrative units. These results show a very short-term vegetation response to rainfall deficit in September, which is reflected in the vegetation in November, and also explains to a large degree the variation in vegetation stress. It is shown that for the period 2000-2015 in the BioBío Region we could identify the 2007/2008, 2008/2009, and 2014/2015 seasons as the three most important drought events; this is reflected in both the overall regional and administrative unit analyses. These results concur with drought emergencies declared by the regional government. Future studies are needed to associate the remote sensing values observed at high resolution (250m) with the measured crop yield to identify more detailed individual crop responses.

2.1 Introduction

Drought is considered one of the most complex natural hazards because of its slow onset and long-term impact; it has the potential to negatively affect many people. Drought is caused by various environmental factors that mainly changes the pattern and amount of rainfall. This situation is expected to intensify over time because of climate change (Dore, 2005). The fifth report from the Intergovernmental Panel on

Climate Change (IPCC) (IPCC, 2013) projects an increase in global temperature and indicates that rainfall in south-central Chile will decrease. This will likely increase both drought frequency and intensity.

According to Wilhite and Glantz (1985), drought can be classified into four categories: (1) Meteorological drought, (2) Hydrological drought, (3) Agricultural drought, and (4) Socio-economic drought. To monitor drought, many different indices were developed and applied to research areas, such as meteorology, hydrology, agriculture, and water resource management. Currently, there are more than 100 drought indices (Niemeyer, 2008; Amin et al., 2011). Among the most popular drought indices is the multiscalar Standardized Precipitation Index (SPI) (McKee et al., 1993), which is used to characterize meteorological drought. The SPI is estimated by frequency analysis of rainfall records; this requires a long-term record of precipitation data, preferably more than 30 years, to select an appropriate probability distribution (Mishra and Singh, 2010). The Standardized Precipitation Evapotranspiration Index (SPEI) is a more recent drought index developed by Vicente-Serrano et al. (2010), which not only includes precipitation but also the effect of temperature. Another important meteorological drought index was developed by Palmer (1965) with temperature and precipitation data to estimate moisture supply and demand in a two-layer soil model known as the Palmer Drought Severity Index (PDSI). However, as noted by Alley (1984), this index has some limitations, such as the use of arbitrary rules to quantify it and limited methodology to standardize it. Using temperature and evapotranspiration (ET) data, Palmer (1968) developed a crop moisture index (CMI), which was one of the first agricultural drought index. There are also hydrological drought indices, such as the surface water supply index (SWSI) (Shafer and Dezman, 1982) and standardized streamflow index (SSI) (Vicente-Serrano et al., 2012b). However, most of the aforementioned indices depend on the availability of temporal and spatial field data, thus complicating their implementation in data-scarce developing countries where historical record availability is limited and meteorological station density is insufficient. This situation does not allow adequate spatial mapping of the index that needs to be generated (Caccamo et al., 2011). Furthermore, using discrete, point-based meteorological measurements collected at weather station locations has resulted in a restricted level of spatial precision for monitoring drought patterns (Wu et al., 2013). Remote sensing, therefore, offers significant advantages for monitoring agricultural drought at the regional and local levels by allowing both spatial and temporal evaluations.

Remote sensing vegetation indices (VI) have been widely used to assess vegetation and drought conditions (Rojas et al., 2011; Rhee et al., 2010; Logan et al., 2010; Kogan, 1995a; Tonini et al., 2012; Skakun et al., 2016; Rembold et al., 2013, 2015). The vegetation health index (VHI) (Kogan, 1995a, 1997), temperature condition index (TCI) (Kogan, 1995a), and vegetation condition index (VCI) (Kogan, 1995b) are among the main drought indices based on remote sensing; they have been successfully applied in numerous case studies under many different environmental conditions around the globe (Zhang and Jia, 2013; Rojas et al., 2011; Gebrehiwot et al., 2011; Singh et al., 2003; Seiler et al., 1998; Unganai and Kogan, 1998). There are drought indices based on the spatial feature of land surface temperature (T_s) and the normalized difference vegetation index (NDVI), as well as the vegetation temperature dryness index (TVDI) suggested by Sandholt et al. (2002) and the vegetation temperature condition index (VTCI) developed by Wang et al. (2001), which are time-dependent and usually region specific. All of these indices use NDVI as input, which is the most widely used index to monitor vegetation quantity, quality, and development. The indices based on NDVI are more useful during the plant growing seasons (Wan et al., 2004; Vicente-Serrano, 2007). Given the physical complexity of drought, there is an ongoing development and improvement of remote sensing drought indices. Zhang et al. (2013) used VCI to construct the time-integrated vegetation condition index (TIVCI), which considers the time lag effect on NDVI from climate factors; however, the time-lag effect of NDVI on meteorological data when monitoring drought requires more attention (Zhang et al., 2013). Du et al. (2013) integrated multi-source remote sensing data with a moderate resolution imaging spectroradiometer (MODIS), and the tropical rainfall measuring mission (TRMM) explained the synthesized drought index (SDI), which is defined as a principal component of VCI, TCI and precipitation condition index (PCI). One limitation of this index occurs at a temporal scale that is shorter than 1 month (Du et al., 2013). Meanwhile, Mu et al. (2013) introduced the drought severity index (DSI) to monitor and detect drought on a global scale with a 1km spatial resolution and 8-day, monthly, and yearly frequencies; this new index integrates satellite ET and NDVI. Recently, Enenkel et al. (2016) develop the Enhanced Combined Drought Index (ECDI) which link rainfall, soil moisture, land surface temperature and vegetation status.

The VCI (Kogan, 1990) was one of the first remote sensing drought indices widely used to monitor agricultural drought (Du et al., 2013; Rojas et al., 2011; Rhee et al., 2010; Logan et al., 2010); it is derived from NDVI, is easily calculated, and accessible for different spatial and temporal resolutions. The VCI concept was originally designed to extract the weather component from NDVI values (Kogan, 1990), considering vegetative variations by climate factors rather than seasonality. Kogan (1997) developed the TCI and then combined both VCI and TCI in the VHI drought index (Kogan, 1995a, 1997) to increase the accuracy of drought monitoring and explain the contribution of temperature in drought analysis, which also provided useful information for monitoring vegetation stress caused by soil saturation. In the present chapter, VCI was selected for its multiple advantages and mainly because it not only reflects spatial and temporal vegetation variability but also allows identifying the impact of weather on vegetation (Kogan, 1995b; Unganai and Kogan, 1998). However, care is needed with unusual extreme events. For example, if most of the NDVI values in a particular location are close to the minimum and there is an unusual event with high NDVI values, most of the calculated VCI will be very low. Also, NDVI are hindered by noise arising from varying atmospheric conditions and sun-sensor-surface viewing geometries (Hird and McDermid, 2009; Klisch and Atzberger, 2016; Julien and Sobrino, 2010; Atkinson et al., 2012), to minimize the possible impacts of undetected clouds and poor atmospheric conditions (Klisch and Atzberger, 2016) a smoothing technique should be applied (Hird and McDermid, 2009; Klisch and Atzberger, 2016; Julien and Sobrino, 2010; Atkinson et al., 2012). To overcome this issue, Klisch and Atzberger (2016) estimates uncertain for NDVI values, which is used to downweight uncertain observation while calculating VCI. Moreover, the VCI in the present chapter was calculated at a 250m spatial resolution to monitor and evaluate agricultural drought, which would be useful because of the geographic and agricultural conditions of the BioBío in Chile (Fig. 2.1).

The SPI is a multiscale drought index that could be used to measure short-term rainfall deficit related to agricultural drought; this is not always true because agricultural drought is mainly affected by soil moisture stress (Mishra et al., 2015) which depends on rainfall as well as many other factors. The VCI is a more direct measurement of agricultural drought because it reflects vegetation health scaled according to long-term NDVI variability for the period under study. Some authors (Zhang et al., 2013; Zhang and Jia, 2013; Wu et al., 2010; Quiring and Ganesh, 2010) compared vegetation indices and SPI to evaluate the correlation between agricultural and meteorological drought. For example, Quiring and Ganesh (2010) evaluated the usefulness of VCI to monitor meteorological drought in Texas, analyzed the relationship between VCI and SPI and its result shows that VCI is most strongly correlated with the 6-month and 9-month SPI, and showed that VCI is strongly influenced by spatially varying environmental factors. Gebrehiwot et al. (2011) also used VCI and SPI to evaluate the spatial and temporal characteristics of vegetative and meteorological drought; they found a time lag between the peak VCI period and precipitation values obtained from the meteorological stations. Ji and Peters (2003) analyzed the relationship between NDVI (derived from AVHRR) and SPI in croplands in the northern Great Plains, which have similar weather and agricultural conditions as those found in the BioBío Region, and showed that the relationship between NDVI and SPI is significant in grasslands and croplands if the seasonal effect is taken into account.

In collaboration with the Ministry of Agriculture of Chile, the Food and Agriculture Organization of the United Nations (FAO), the International Research Institute for Climate and Society (IRI), and some other government institutions and research centers, the United Nations Educational, Scientific and Cultural Organization (UNESCO) implemented the Chilean Agroclimatic Observatory, which collects different meteorological, hydrological, and agricultural information as well as various indices to monitor drought and support decision makers when Chile faces drought conditions.

The present chapter evaluated the VCI drought index proposed by Kogan (1990), calculated it from a time series of MODIS data, and applied it in the BioBío Region. The VCI is compared to SPI at different time scales to identify how long would be the monthly rainfall deficit that has the major impact on agricultural drought in the region. Additionally, the usefulness of monitoring agricultural drought is assessed by comparing VCI with the agricultural drought emergency declared by the Chilean government in the study area. The aim of this chapter was to assess VCI derived from MODIS data at 250m spatial resolution as an effective indicator for monitoring agricultural drought and evaluate its observed impact in the region over the period 2000-2015. Based on this analysis, recommendations are made to include the index in the Chilean Agroclimatic Observatory to support climate-informed decision making.

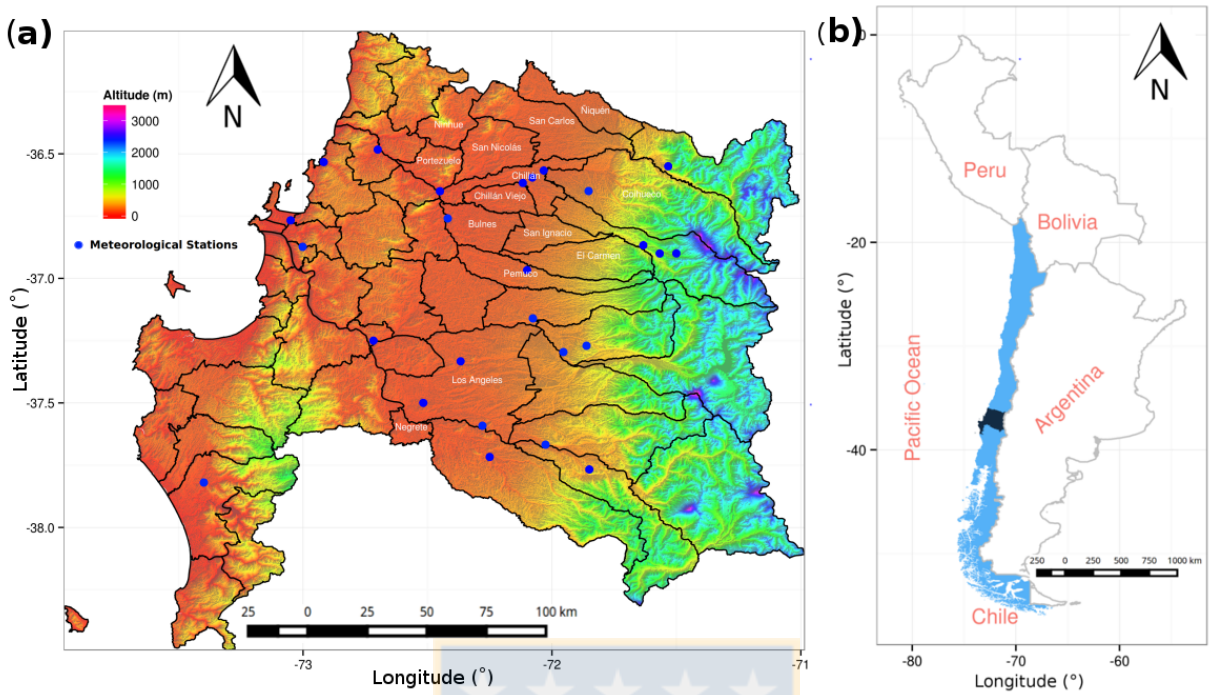

Figure 2.1: (a) BioBio Region administrative units in a digital terrain model with 26 weather stations. (b) Location of the BioBio Region, Chile.

2.2 Study Area

The BioBío Region in Chile is located between 36° 00' and 38° 30' south latitude and between 71° 00' and 74° 00' west longitude along the South Pacific Ocean (Fig. 2.1b) with a total area of 37.068,7 km^2. This area is characterized by the transition from a warm Mediterranean climate to a humid and temperate climate. In terms of agricultural, this region produces a significant amount of annual crops, including wheat, oats, barley, sugar beet and corn, which have a growing season between September and April. The region has 54 administrative units (Fig. 2.1a).

The bioclimatic variables estimated by Hijmans et al. (2005) were used to describe the spatial climatic characteristics of the study area. These variables were derived from monthly temperature and rainfall values to generate more biologically meaningful variables. Precipitation in the driest month is generally $<$ 35mm and a significant portion of the region has $<$ 20mm rainfall in that month (Fig. 2.2). The wettest month has precipitation between 150mm and 350mm while total annual precipitation is between 750mm and 2000mm. The temperature in the warmest month is over 25°C mainly in the central part of the region, as shown in Fig 2.3. The temperature in the coldest month is between 0°C and 7°C from the center to the west and below 0°C to the east and the annual mean temperature is usually above 10°C.

Based on land cover (MCD12Q1.51), changes in the cropland area for the main administrative units are shown in Table 2.1. In general, the cropland area decreased from 2001 to 2013, and the administrative units of San Nicolas and Chillán Viejo showed the largest changes. The land cover map (2013) for the region, illustrated in Fig. 2.4, indicates that most of the region is covered by forest and this is followed by cropland and grassland.

Table 2.1: Percentage of cropland area (%) of 15 administrative units of the BioBío Region, Chile, with cropland area $\geq$ 10% from 2001 to 2013 and the 13-year mean.

Comuna	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
San Ignacio	79	79	80	78	76	73	73	72	75	75	73	72	76
Bulnes	82	82	81	79	78	76	77	77	77	76	74	73	73
San Carlos	79	74	71	66	66	66	68	66	76	74	71	62	68
Chillán	75	74	74	70	69	67	67	65	69	68	66	62	63
Ñiquén	76	67	63	54	54	55	57	59	72	72	64	54	59
Negrete	58	59	50	47	41	48	42	49	53	44	36	34	47
El Carmen	52	54	52	51	50	49	48	45	46	45	43	44	46
Chillán Viejo	65	61	60	54	51	46	48	44	57	54	52	39	42
San Nicolás	76	62	56	45	44	45	49	45	63	58	54	36	40
Los Angeles	31	35	32	33	27	28	26	28	31	25	22	26	32
Pemuco	36	37	36	32	30	24	25	24	27	25	23	18	21
Coihueco	25	26	25	23	22	20	19	19	21	21	20	19	20
Yungay	24	26	24	22	20	20	20	18	20	17	16	14	17
Quillón	32	24	21	14	16	16	15	16	21	18	17	13	15
Pinto	18	16	14	12	13	12	12	11	12	13	11	10	11

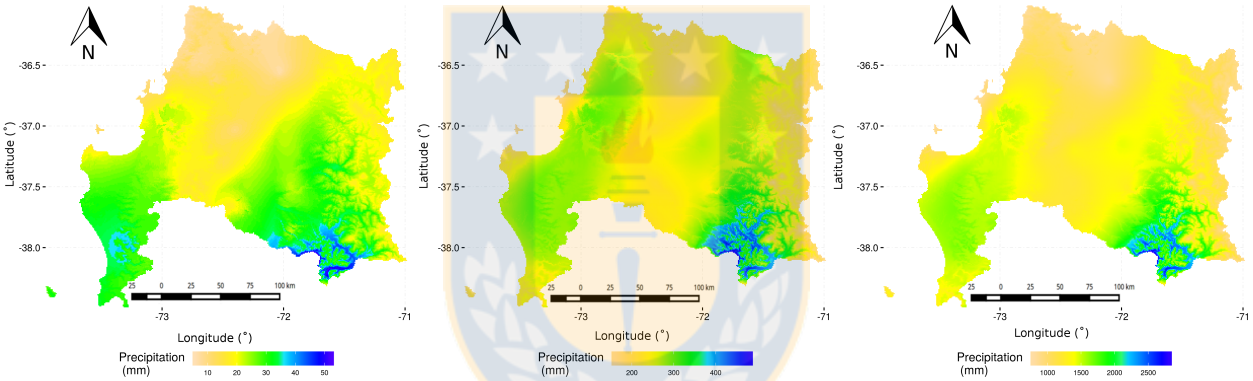

Figure 2.2: Bioclimatic precipitation variables of the BioBío Region, Chile

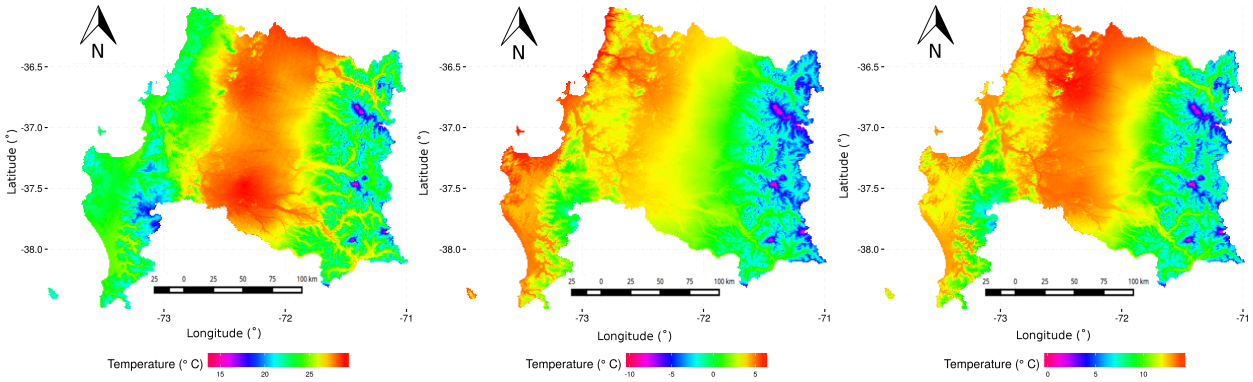

Figure 2.3: Bioclimatic temperature variables of the BioBío Region, Chile

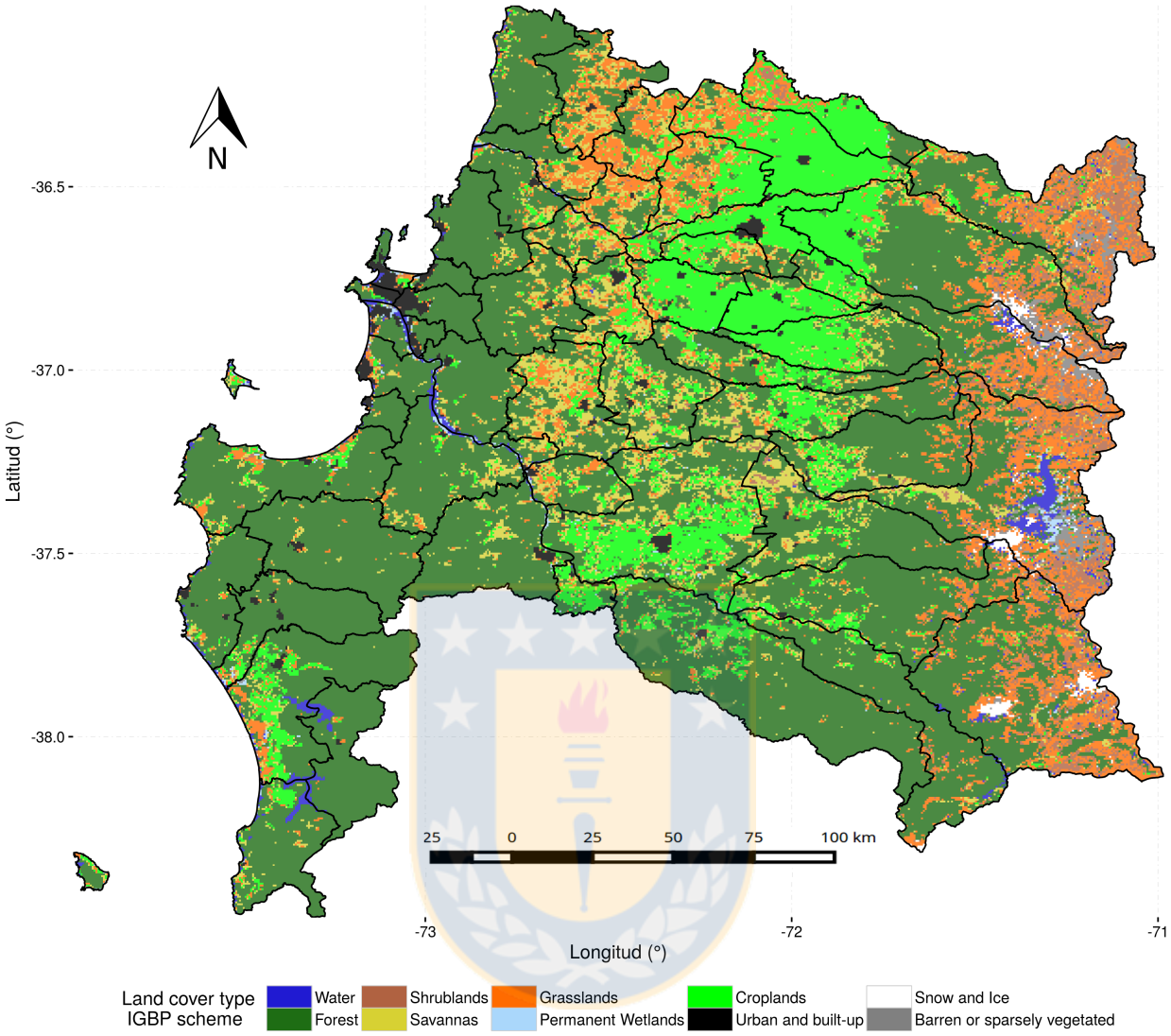

Figure 2.4: Land cover classes in the BioBío Region, Chile, based on the IGBP land cover scheme for the MODIS MCD12Q1 version 5.1 product.

2.3 Data

The MODIS has been a key environment remote sensing tool for more than 16 years; it has been used in countless studies of different disciplines all over the world. The MODIS instrument was developed to improve heritage sensors in terms of its spectral, spatial, and temporal resolutions, as well as more stringent calibration requirements. This instrument takes observations in 36 spectral bands covering wavelengths from 0.41 to 14.4 μm and at three nadir spatial resolutions: 250m, 500m, and 1km (Xiong et al., 2009).

The usefulness of NDVI for evaluating vegetation response is well known (Huete et al., 2002). In the present case, the vegetation indices (VI) were obtained from the MODIS *Vegetation Indices 16-Day L3 Global 250m* short name *MOD13Q1* product (Didan, 2015). Huete et al. (2002) present the NDVI analysis and the Enhanced Vegetation Index (EVI); their results demonstrate the scientific usefulness of MODIS VI. Moreover, Miura et al. (2008) compared MODIS VI with the high-resolution Advanced Thermal Emission and Reflection Radiometer (ASTER) (15 m) and showed that they coincided well on a global scale.

Several land cover products are available; the most frequently used are GLC-2000 (Fritz et al., 2003), Globcover (Bontemps et al., 2011), and MODIS Collection 5 land cover (Friedl et al., 2010). Comparative studies have shown large spatial discrepancies among these three products. Two of the main advantages of the MODIS Collection 5 are its 500m spatial resolution and, according to Friedl et al. (2010), the product overall accuracy is approximately 75%. Therefore, the MODIS cropland cover *Land Cover Type Yearly L3 Global 500 m SIN Grid* short name *MCD12Q1* version 5.1 was used. Data for the present chapter were obtained through the online Data Pool at the NASA Land Processes Distributed Active Archive Center (LP DAAC) and USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota.

Meteorological data were collected from 53 meteorological stations in the BioBío Region from the General Water Authority (DGA) and the Chilean Meteorological Directorate (DMC). A total of 26 stations were selected from this dataset; stations had more than 30 years of records and few missing data (Fig. 2.1).

2.4 Methods

2.4.1 Procedure for calculating VCI in cropland areas.

A time series of vegetation index products (MOD13Q1 version 5) was used to derive the VCI index, and the land cover product (MCD12Q1 version 5.1) (Friedl et al., 2010) was used to determine the spatial extension of croplands; both products are from the Moderate-Resolution Imaging Spectroradiometer (MODIS) sensor. Meteorological stations with a long record (more than 30 years) were used to calculate the SPI index.

All processing and calculations with the raster data were performed with the **R** software (R Core Team, 2016) and the *raster* package (Hijmans, 2015). Once the MOD13Q1 and MCD12Q1 satellite data were obtained, they were reprojected to the WGS84 datum and geographic projections with the Modis Reprojection Tool (MRT) (Dwyer and Schmidt, 2006) using nearest neighbor resampling. A smoothing process was required to reduce noise in the NDVI time series. Multiple techniques are available in the literature to do this (Hird and McDermid, 2009; Klisch and Atzberger, 2016; Julien and Sobrino, 2010; Atkinson et al., 2012). In the present case, a locally-weighted polynomial regression (Lowess) (Cleveland, 1981) was used. Fig. 2.5, shows time-series of NDVI compared with those smoothed by Lowess. In future studies the smoothing could be improved using for example an adapted lowess (Moreno et al., 2014) or also with a modified Whittaker smoother as proposed by (Klisch and Atzberger, 2016).

Then, with the smoothed NDVI,NDVI_{max}, and NDVI_{min} from 2000-2015, the VCI values were calculated using Eq. (2.1) for each pixel in the BioBío Region every 16-day. Finally, the VCI time series (2000-2015) were masking out, using the cropland mask.

2.4.2 Cropland mask

The *cropland* type in the present study is the IGBP classification scheme of the MCD12Q1 collection 5.1, and it was used because cropland class reliability is $> 92\%$ according to Friedl et al. (2010). In addition, the cropland mask derived from IGBP scheme concurs well with the cropland data from the 2007 national agriculture and livestock census (INE, 2007). From the MCD12Q1 product, the land cover class that corresponds to croplands (class 12 and 14 of the IGBP scheme) was used to create an agricultural mask with a 500m spatial resolution for the 2001-2013 period. These data had to be resampled from a 500m to 250m spatial resolution using the *raster* package (Hijmans, 2015) to coincide with the VCI resolution. Thirteen yearly cropland masks were created; the mask for 2001 was used with VCI data for 2000 and 2001, and the mask for 2013 was used for 2013, 2014, and 2015. A sub-selection of 15 units with $> 10\%$ cropland was established from the total of 54 administrative units.

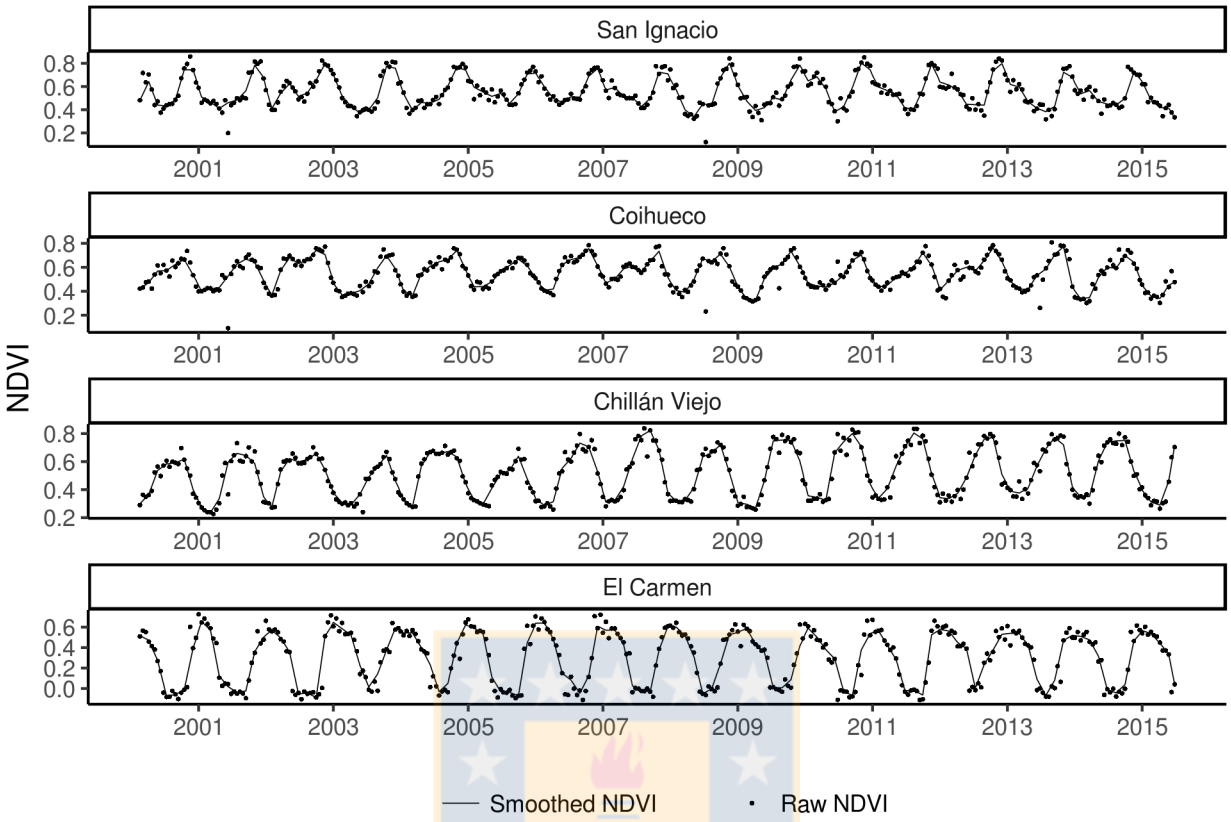

Figure 2.5: Time-series comparison between raw NDVI (points) and smoothed NDVI by Lowess (lines) for five points on five administrative units in cropland areas of the BioBío Region.

2.4.3 Vegetation condition index (VCI)

The VCI (Kogan, 1990, 1995b) is used to monitor agricultural drought and is derived from NDVI. It scales NDVI between its maximum and minimum values for a given period and can be expressed as:

$$VCI_{(i,p,j)} = \frac{NDVI_{(i,p,j)} - NDVI_{min_{(i,p)}}}{NDVI_{max_{(i,p)}} - NDVI_{min_{(i,p)}}} \tag{2.1}$$

where $NDVI_{(i,p,j)}$ is the smoothed NDVI for pixel i, period p and year j; in the present chapter, the period is 16-day (from 1 to 23 for each year) from 2000-2015. $NDVI_{max(i,p)}$ and $NDVI_{min(i,p)}$ are the multi-annual maximum and minimum, respectively, calculated for each pixel i and 16-day period p from 2000-2015. To compare VCI values extracted in the 26 weather stations with SPI monthly data, a weighted mean was applied to convert VCI to monthly values.

According to Kogan et al. (2003), NDVI represents two environmental signals, the ecosystem, which explains long-term changes in vegetation (driven by climate, soils, vegetation type, topography, etc.), and the weather (short-term), which explains intra- and inter-annual variations in each ecosystem in response to weather fluctuations. Given that the weather component is much smaller than the ecosystem component, the algorithm was developed to enhance the weather component.

2.4.4 Standardized Precipitation Index (SPI)

Since VCI incorporates both climatic and ecological components (Kogan, 1990), an analysis was required to understand the effect of the precipitation variability on VCI for croplands in the study area. The SPI drought index was therefore used to analyze the correlation of VCI with rainfall departure. The SPEI (Vicente-Serrano et al., 2010) is a more significant measure because it incorporates the temperature effect. However, in the present chapter, SPEI could not be used because of a scarcity of temperature measurements; this situation did not allow calculating reference ET in each one of the weather stations.

The SPI (McKee et al., 1993) is a meteorological drought index that is estimated from long-term precipitation records. These long-term records are fitted to a probability distribution (usually Gamma or Pearson III) which is then transformed into a normal distribution so that mean SPI for the location and desired period is zero. Positive SPI values indicate that precipitation is higher than the median while negative values indicate precipitation is lower than the median. The SPI can be computed for different time scales where shorter scales (1-6 months) are related to short-term deficit, such as soil water content (vegetation response) and longer scales (12-36 months) with a long-term deficit that is generally associated with groundwater and reservoirs. The *spi* function from the *SPEI* **R** package (Beguería and Vicente-Serrano, 2013) was used to calculate SPI for time scales between 1 and 6 months for the 26 meteorological stations. To fit a *Gamma* distribution on the data, the *spi* function was set by the method of unbiased probability weighted moments.

The classification scheme used for VCI and SPI was similar to the classification scheme used by Bhuiyan et al. (2006) and proposed by Du et al. (2013), as shown in Table 2.2.

Table 2.2: Drought classification scheme for SPI and VCI (Bhuiyan et al., 2006; Du et al., 2013)

Drought classes	SPI	VCI
Extreme	SPI < -2.0	$0 \leq$ VCI < 10
Severe	-2.0 $\leq$ SPI < -1.5	10 $\leq$ VCI < 20
Moderate	-1.5 $\leq$ SPI < -1.0	20 $\leq$ VCI $\leq$ 30
Mild	-1.0 $\leq$ SPI < 0.0	30 $\leq$ VCI $\leq$ 40
No drought	0.0 < SPI	40 < VCI $\leq$ 100

2.4.5 Correlation between VCI and SPI

The standardized VCI anomalies were used for the Pearson correlation test:

$$STD_{ijk} = \frac{X_{ijk} - \overline{X}_{ij}}{\sigma_{ij}} \tag{2.2}$$

where X_{ijk} is the VCI value in pixel i, period j and year k, $\overline{X_{ij}}$ is the mean value of VCI in pixel i and period j; and σ_{ij} is the standardized deviation of pixel i and period j.

To identify the lag-time period that is more sensitive to rainfall deficit, the correlation between standardized VCI and SPI for the cropland area of the BioBío Region for three different seasons was tested: 1) January to December, 2) September to April (growing season), and 3) November to April (modified growing season). A modified growing season was tested because when is considered the normal growing season (Sept-Apr), and SPI-3 is calculated in September (start of growing season), it refers to the accumulated effect of rainfall deficit from July to September. Instead, if is considered the start of the period in November (start of modified growing season), then the SPI-3 reflect the rainfall deficit from September to November. In this way, is possible evaluate the impact of rainfall deficit which occurs during the months of the growing season.

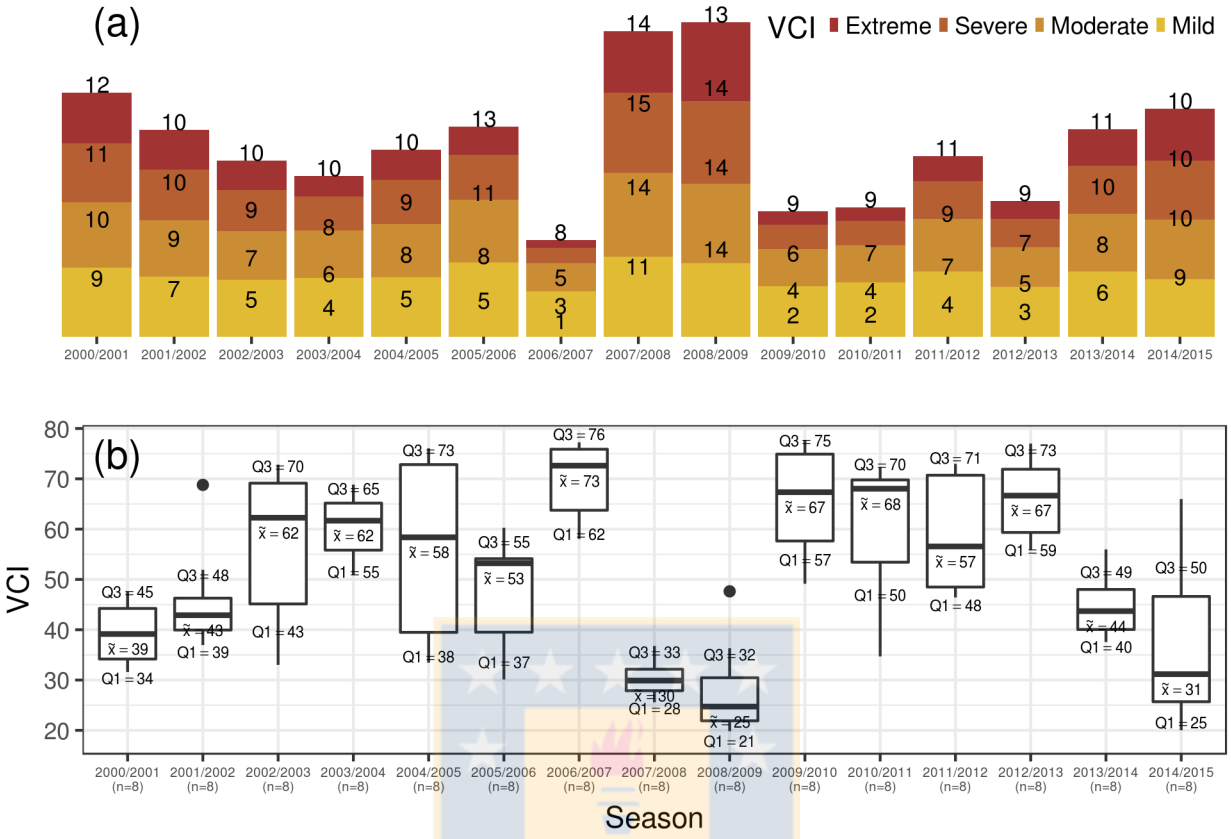

Figure 2.6: Variation of the (a) global VCI percentage (%) of cropland surface with different VCI classes and (b) boxplot of global VCI intensity (%) for the growing seasons between 2000/2001 and 2014/2015 in the Biobio Region, Chile.

2.5 Results and discussion

2.5.1 Spatio-temporal variation of VCI and comparison with drought declaration

At the regional level, the variation of drought intensity percentages in the study area for the cropland area between the 2000/2001 and 2014/2015 growing seasons is illustrated in Fig. 2.6a. The BioBío Region reached the lowest percentage of the cropland area under drought in 2006/2007 (17%) and the highest in 2008/2009 (55%). The 2007/2008 and 2008/2009 periods exhibited almost the same percentage of cropland surface under drought; however, a difference was observed in the extreme intensity class covering more cropland surface in 2008/2009 (14%), compared to the 2007/2008 period (11%). Fig. 2.6b depicts the box-plot of mean VCI values in the BioBío Region considering the cropland growing season. The box-plot shows the upper and lower quartiles (Q3 and Q1) and the median value. The 2007/2008 and 2008/2009 periods showed the lowest VCI values, followed by 2014/2015 (Fig. 2.6a). Moreover, 2008/2009 had 75% of its values with VCI $< 32\%$ and 50% of the data had VCI values between 21% and 32%, causing it to be the most severe period, which was noted by a considerable increase of the surface under extreme drought.

On the administrative unit level, Fig. 2.7 illustrates a heatmap of the time series VCI intensity values on the cropland surface Fig. 2.7a and Fig. 2.7c and the percentage of cropland surface under drought conditions ($VCI \leq 40$) (Fig. 2.7b and Fig. 2.7d) for 15 administrative units between 2007-2009 and 2014-2015. The dashed white line corresponds to the times in the past 16 years when the Chilean government has

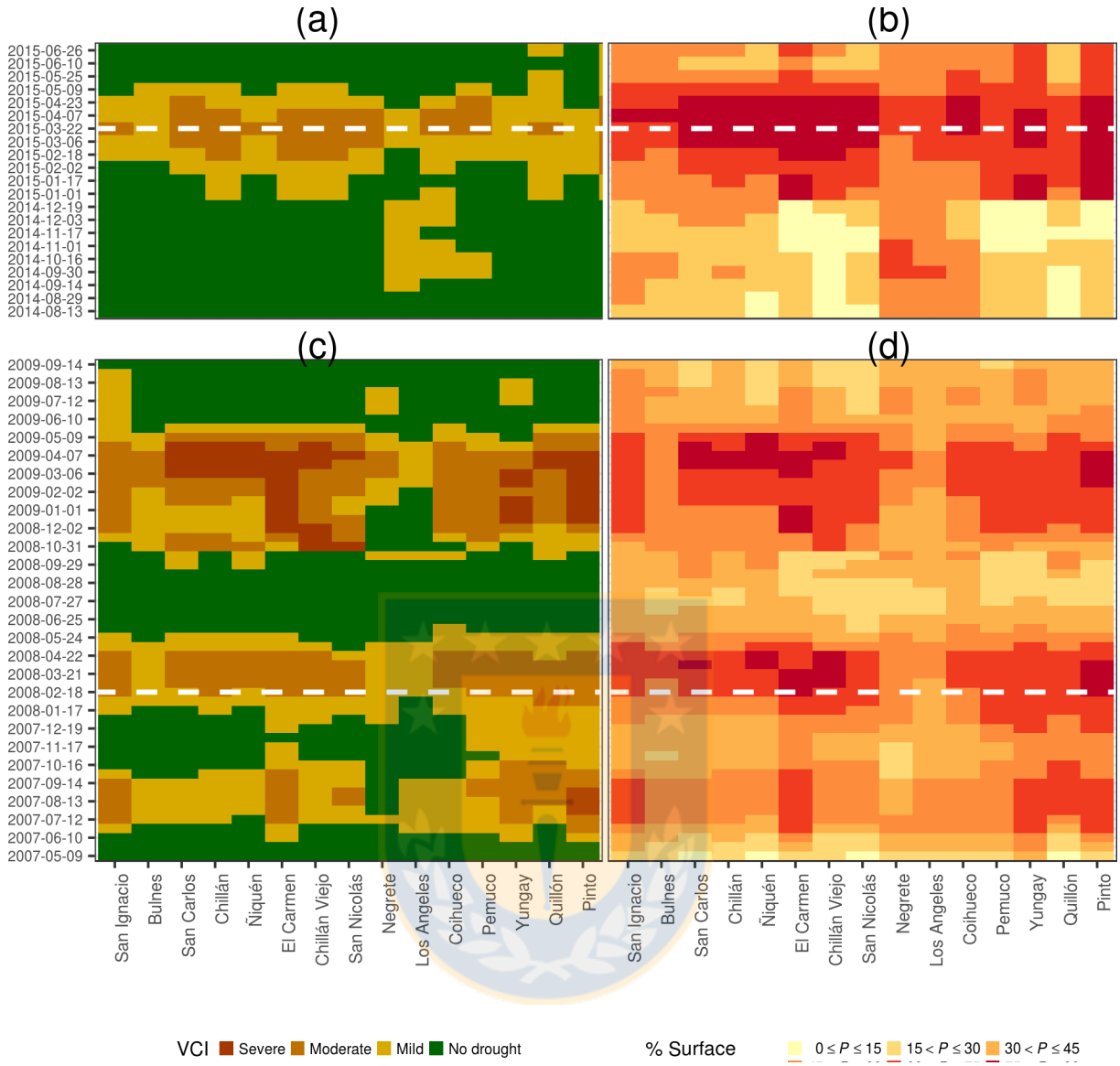

Figure 2.7: Heatmap of VCI conditions (a and c) and percentage of cropland area < VCI = 40 (b and d) for the principal drought periods in the main cropland administrative units in the study area for the 2007-2009 (c and d) and 2014-2015 (a and b) seasons. The dashed white line corresponds to the date when the agricultural drought emergency was declared by decision makers.

declared an agricultural drought emergency in the region. According to the VCI values, the administrative unit was under moderate and mild drought conditions between 2007 and 2008 while in the late growing season from January to May, VCI intensity was mostly moderate drought, and the surface percentage of drought was > 60% for each administrative unit. At the beginning of February 2008 (2007/2008 season) the government declared an agricultural drought emergency. However, during the next season, 2008/2009, drought conditions were similar but with a longer duration (Sept-May) and more severe intensity, but the government did not declare drought emergency. The last drought emergency was declared in March 2015 and it seems that is was declared late according to VCI intensity and cropland surface affected. As displayed in Fig. 2.7b, the emergency was declared in the middle of the drought period (Jan-May) when the percentage of surface affected by drought was between 60% and 90%. The drought emergency declared by the Ministry of Agriculture of Chile do not consider the intensity levels but rather only the conditions with or without

drought.

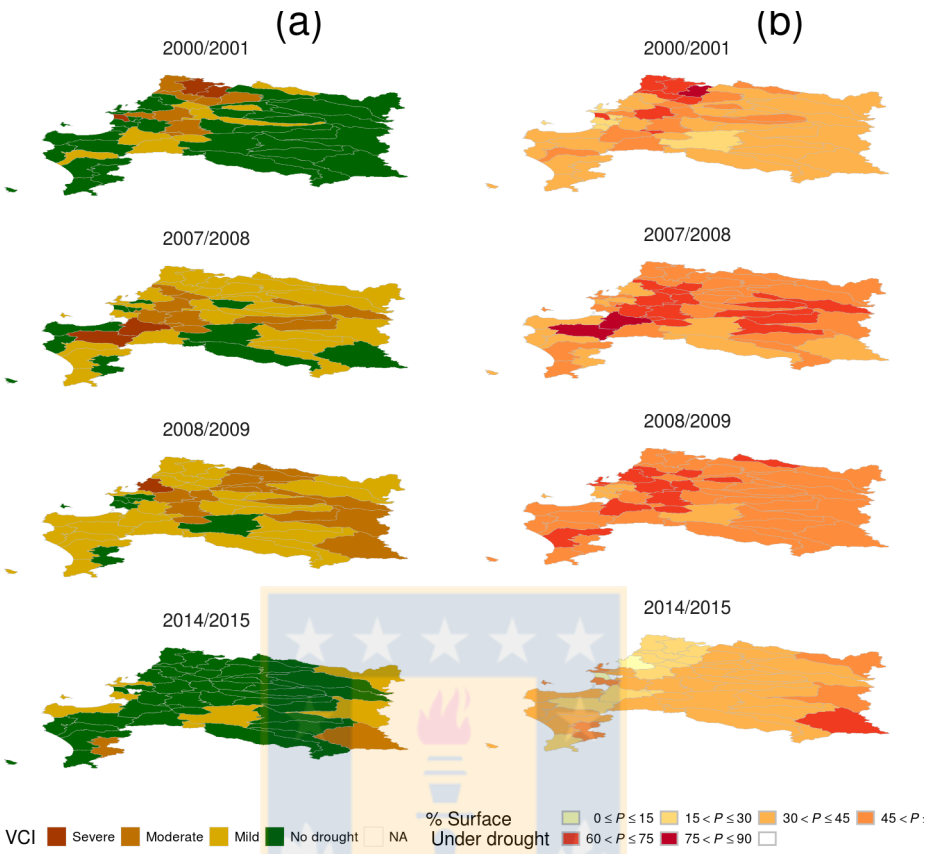

Figure 2.8: (a) Mean VCI conditions and (b) percentage of cropland surface with VCI $\leq$ 40% in the administrative units of the Biobio Region, Chile, for the 2000/2001, 2007/2008, 2008/2009, and 2014/2015 growing seasons (Sept-Apr).

The periods with the lowest VCI means for each unit during the growing seasons in the last 16 years are mapped in Fig. 2.8. Fig 2.8a shows that the 2000/2001 season had partial drought conditions, which mostly affected the western part of the region, with three, eight, and nine units with severe, moderate, and mild drought conditions, respectively. It is possible to note that the 2007/2008 and 2008/2009 seasons had drought conditions in almost every unit in the region with mild and moderate drought intensity as the main condition. These periods also had a unit surface percentage under drought > 45% (Fig. 2.8b). The 2014/2015 season had mostly a surface percentage between 30% and 45% under drought conditions in each unit, also had two and nine units with moderate and mild drought, respectively.

Table 2.3: Pearson correlation value (r) between time scale 1 to 6 and standardized VCI for 15 administrative units in the BioBío Region, Chile, with cropland area > 10% and considering the mean values between November and April.

No.	Adm. unit	Station name	SPI-1	SPI-2	SPI-3	SPI-4	SPI-5	SPI-6
1	Los Angeles	DGA Las Achiras	0.46	0.69	0.78	0.73	0.67	0.64
2	Chillán	DMC Chillán	0.38	0.56	0.70	0.66	0.59	0.53
3	Bulnes	DGA Chillancito	0.37	0.59	0.66	0.59	0.47	0.34
4	Negrete	DGA Los Angeles	0.47	0.69	0.74	0.69	0.62	0.55
5	Chillán Viejo	DGA Chillán Viejo	0.41	0.59	0.67	0.64	0.55	0.45

No.	Adm. unit	Station name	SPI-1	SPI-2	SPI-3	SPI-4	SPI-5	SPI-6
6	El Carmen	DGA Diguillin	0.29	0.48	0.58	0.55	0.46	0.36
7	San Ignacio	DGA Pemuco	0.31	0.48	0.56	0.51	0.43	0.36
8	San Nicolas	DMC Chillán	0.31	0.47	0.56	0.53	0.47	0.39
9	San Carlos	DMC Chillán	0.34	0.49	0.59	0.56	0.50	0.45
10	Pinto	DGA Las Trancas	0.24	0.40	0.49	0.45	0.35	0.29
11	Coihueco	DGA Coihueco	0.33	0.49	0.58	0.52	0.43	0.38
12	Yungay	DGA Cholguan	0.19	0.37	0.43	0.44	0.40	0.33
13	Pemuco	DGA Pemuco	0.21	0.34	0.40	0.37	0.29	0.18
14	Quillon	DGA Chillancito	0.38	0.55	0.62	0.59	0.51	0.37
15	Ñiquen	DGA San Fabián	0.38	0.55	0.57	0.48	0.38	0.28

The VCI temporal mean at pixel level for the 2007/2008, 2008/2009, and 2014/2015 growing seasons (Sept-Apr) in the cropland area of the BioBío Region is shown in Fig. 2.9.

2.5.2 Correlation VCI vs SPI

Globally, in the BioBío Region we compared three averaged periods (Sept-Apr, Nov-Apr, and Jan-Dec) and the correlation between SPI and VCI at different time scales. We found the highest regional correlation between SPI-3 and VCI during the growing season (Sept-Apr); with a Pearson correlation value of 0.54. The period between November and April shows a higher Pearson correlation value of 0.63. This is comparable to the results presented by Vicente-Serrano (2007), who indicated that the vegetative drought index is useful for monitoring drought during the growing season.

The administrative units with a cropland area $> 10\%$ were correlated with the nearest meteorological station. Pearson correlation values for SPI-1 to SPI-6 are listed in Table 2.3. As previously shown, the higher correlation values were at SPI-3 (0.40 to 0.78) (Table 2.3). Vegetation had a short-term response to rainfall, and 3-month departures explained between 16% and 61% of the variance in cropland health. Mean VCI anomaly and SPI-3 values were compared in Fig. 2.10 for the period between November and April (modified growing season). The SPI-3 and VCI values were negative and similar for all 15 administrative units in the 2007/2008, 2008/2009, and 2014/2015 seasons, which corresponded to the three periods in the last 16 years in which the BioBío Region was under the most severe drought conditions (Fig. 2.6). However, there was an opposite correlation with negative VCI and positive SPI-3 in the 2000/2001 season. Rainfall is the main variable among others affecting vegetation response. Management, irrigation, and plant disease also affect agricultural drought, and they must be analyzed in greater detail.

Monthly correlation values calculated at 26 weather stations during the growing season are displayed in Fig. 2.11a; September and October (beginning of the season) showed the lowest correlation between SPI-3 and VCI whereas from November to April (middle and end of season) the Pearson correlation value was always $>$ 0.6; the lowest value during this period was in February. The correlation of VCI with SPI-1 was the highest in October, November, and February; and with SPI-3 the correlation was around ≈ 0.6 from November to March. This indicated that rainfall deficit in September and October had a higher impact on vegetation with an accumulated effect beginning in November during the growing season. On the other hand, rainfall during July, August, and September had a lower impact on the agricultural drought conditions in September and October. The period from November to April was therefore identified as being more sensitive to water scarcity because of the accumulated effect of three months of rainfall on the cropland growing season. This result concurs with observations by Ji and Peters (2003), who found that the correlation between vegetation and SPI-3 was stronger in the middle of the season and weaker at the beginning and end of the growing season. Cropland vegetation in this region mainly had a short-term response (3 months) to rainfall deficit. In terms of the monthly rainfall deficit (SPI-1), October, November, December, and February were more significant with Pearson correlation values of 0.49, 0.49, 0.43, and 0.48, respectively. The 6-month accumulated effect of rainfall deficit on vegetation (SPI-6), which began increasing in December and peaked in April, confirms the accumulated effect from November to April (r=0.68). In addition, if we want to monitor the croplands

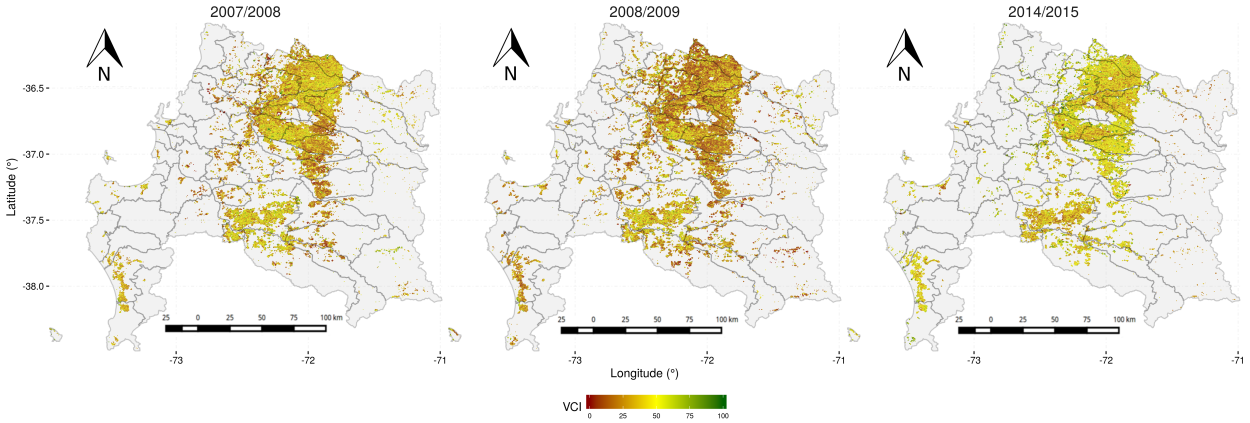

Figure 2.9: VCI mean values for croplands during growing season (Sept-Apr) in the Biobío Region, Chile, for 2007/2008, 2008/2009, and 2014/2015 seasons.

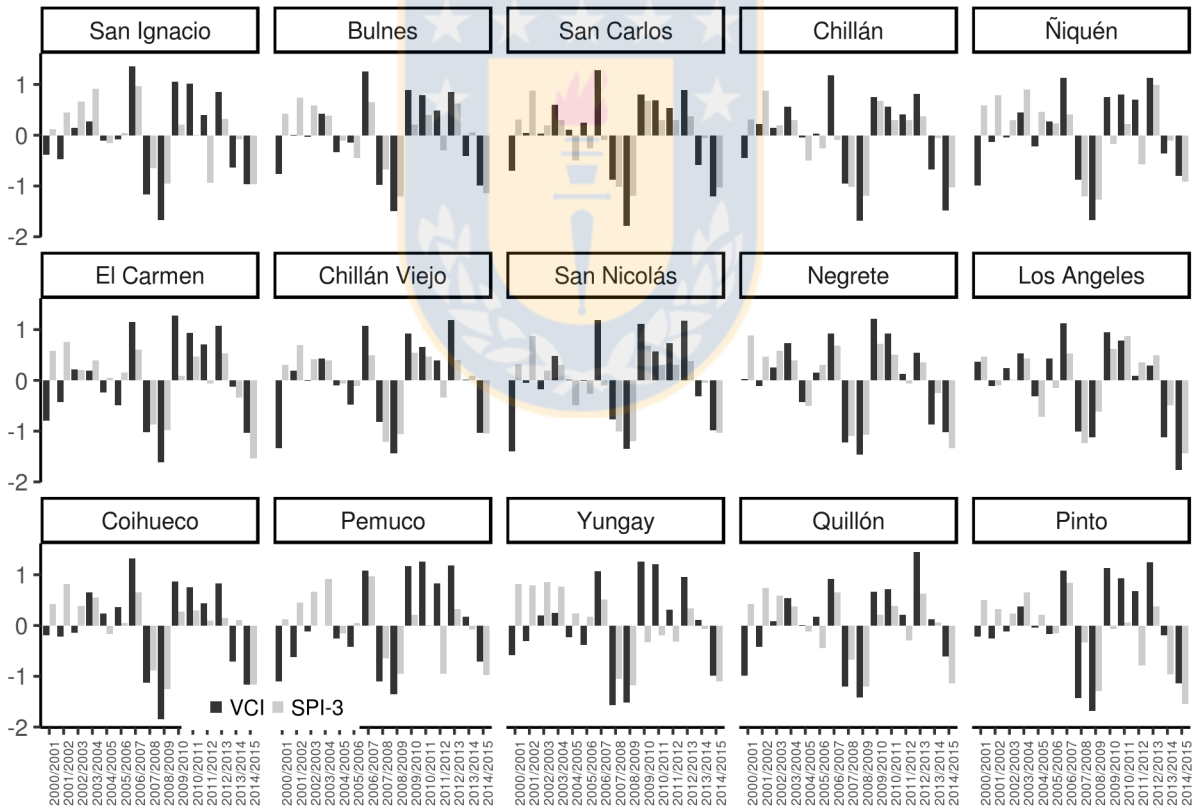

Figure 2.10: Comparison of SPI-3 and VCI anomaly for 15 administrative units with percentage cropland > 10% from 2000/2001 to 2014/2015 modified growing seasons (Nov-Apr).

in more detail, we could consider SPI-1 in October, SPI-3 from November to February, and SPI-6 in March

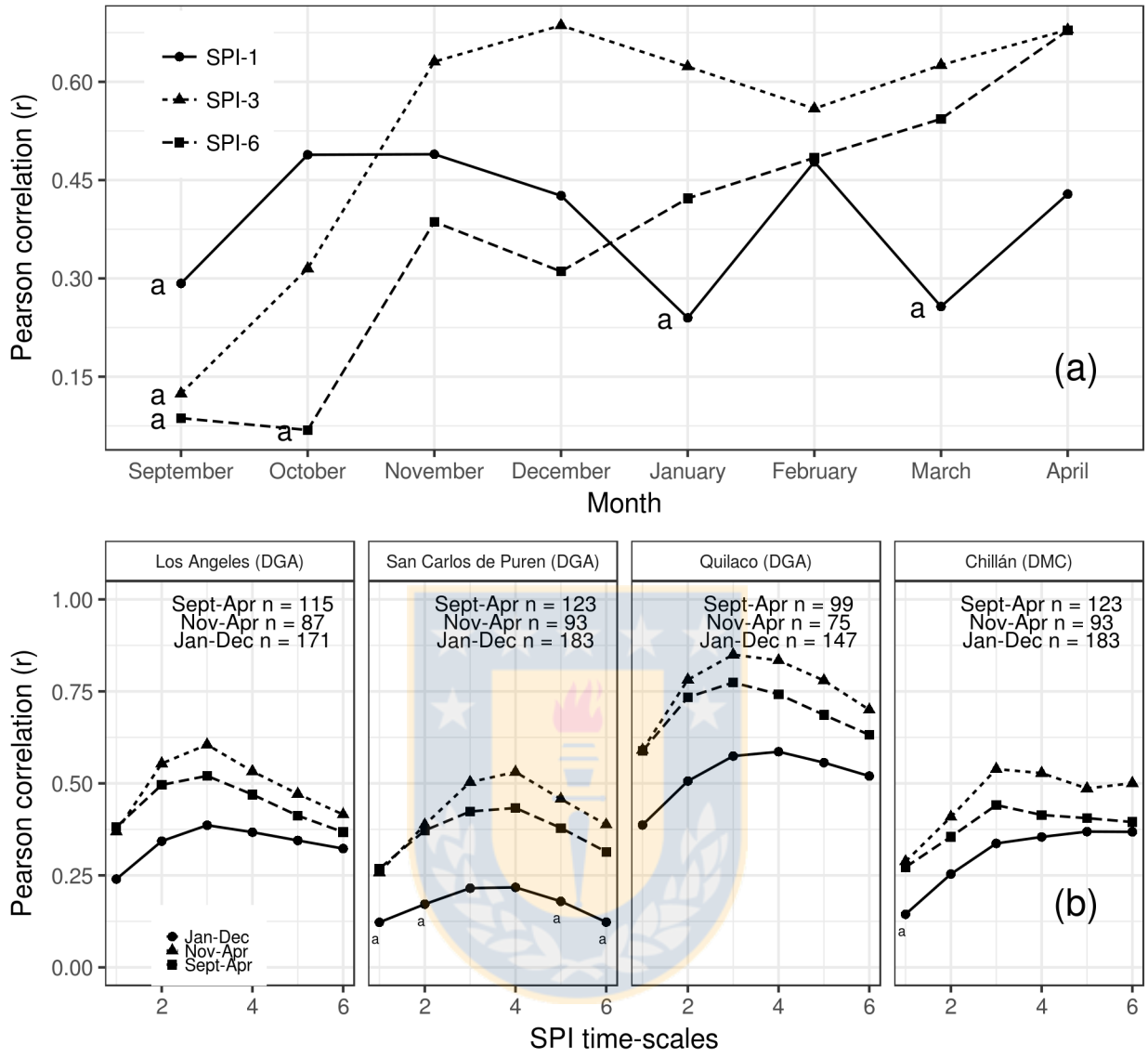

Figure 2.11: (a) Correlation between mean VCI (croplands) and mean SPI at time scales between 1 to 6 for meteorological station Number 4 in the Biobío Region, Chile, for three different periods; (b) Monthly correlation between SPI-1, SPI-3 and SPI-6 with VCI in the growing season. Letter (a) in the plot means significance at $p > 0.01$.

and April.

Comparative results of correlations between VCI and SPI time scales of 1 to 6 months for three periods and four meteorological stations are shown in Fig. 2.11b. The pattern is similar for all the periods and November to April had the highest Pearson correlation between SPI and VCI. This correlation peaked at the 3- or 4-month time scales and then decreased. However, for the station with lower correlations (*San Carlos de Puren*), the SPI-4 showed clearly the highest correlation. The SPI-3 in the *Quilaco* station was prominent with $r = 0.77$.

Several studies have compared the vegetation drought index with meteorological conditions (Gebrehiwot et al., 2011; Bajgiran et al., 2008; Ji and Peters, 2003; Quiring and Ganesh, 2010; Wu et al., 2010). Our observations are consistent with those of Gebrehiwot et al. (2011), who found a strong correlation between

VCI and precipitation deficit of the last 3 months at the station and regional levels. Similarly, Bajgiran et al. (2008) established the highest correlation with 3-month precipitation in stations where land use is predominantly cropland (r=0.8) and grassland (0.81). Ji and Peters (2003) indicated a Pearson correlation value of 0.47 between VCI and SPI-3 for the croplands of the northern Great Plains, and Wu et al. (2010) found a very good correlation between VCI and SPI-3 for grassland. Quiring and Ganesh (2010) identified the correlation between SPI and VCI in Texas as having high spatial variability; however, they found that VCI is most strongly correlated with the 6-month and 9-month SPI. They also established permeability, irrigation, landcover type and water table depth as the most important independent variables besides rainfall, which explain the variation in vegetation health (Quiring and Ganesh, 2010).

It will be important to consider the effect of temperature on agricultural drought in future studies by using, for example, SPEI (Vicente-Serrano et al., 2010); this is a multiscalar drought index which takes into account potential ET as a measurement of water demand. This was not possible in the present chapter because historical temperature data were scarce. One option to overcome the lack of temperature data could be the use of remote sensing data, such as the MOD16 ET product proposed by Mu et al. (2007).

2.6 Conclusions

In a country where the impact of agricultural drought is increasing, detailed monitoring and early warning tools are required to trigger responses that allow mitigating the drought effect. Therefore, a thorough understanding of the drivers of agricultural drought are needed even when detailed field observations are lacking and meteorological stations are scarce, as is the case in Chile. This result makes remote sensing datasets for vegetation monitoring a particularly powerful tool under these circumstances.

The present chapter assessed the agricultural drought dynamics using the vegetation condition index (VCI) at 250m spatial resolution and evaluated the cropland area of the BioBío Region in Chile from 2000 to 2015.

A VCI analysis for croplands could identify the spatial distribution of stressed vegetation associated with drought conditions. Comparing cropland VCI for all the administrative units during the growing stage indicated that, according to the selected drought intensity classification, three drought episodes have occurred in the last 16 years that coincide with the years in which agricultural emergency funding was provided to the farmers in the region. The VCI indicator shows the potential to further tailor the drought emergency response and identifies more objectively the stakeholders who are the most affected even when detailed local observations are lacking.

The correlation between rainfall deficit (SPI) on short- and long-term scales and VCI values shows that SPI-3 exhibited the highest correlation values for the BioBío Region between November and April, defined as a modified growing season, instead of between September and April, which is the normal growing season. This result indicates that vegetation responds rapidly to rainfall deficit beginning in September and this is evidenced in vegetation in November.

Based on these findings, we can conclude that VCI is useful for monitoring agricultural drought in the BioBío Region and is closely correlated with SPI-3 during the modified growing season (Nov. to Apr.), which indicate that rainfall deficit beginning on September it is when has a larger impact on vegetation health, this would be related with crops types in the region, what it should be evaluated in future studies. This makes it a relevant indicator for agricultural drought monitoring and response plans. Further research is needed to associate the remote sensing values observed at high resolution (250m) with the measured crop yield (Seiler et al., 2007) and individually identify more detailed crop responses. This identification will gradually construct an effective drought management tool for the agricultural sector in Chile.

Chapter 3

Satellite monthly precipitation products for use in Chile

Zambrano, F.; Wardlow, B.; Tadesse, T.; Lillo-Saavedra, M. & Lagos, O. Evaluating satellite-derived long-term historical precipitation dataset for drought monitoring in Chile Atmospheric Research, 2017, 186, 26-42

Abstract

Precipitation is a key parameter for the study of climate change and variability and the detection and monitoring of natural disaster such as drought. Precipitation datasets that accurately capture the amount and spatial variability of rainfall is critical for drought monitoring and a wide range of other climate applications. This is challenging in many parts of the world, which often have a limited number of weather stations and/or historical data records. Satellite-derived precipitation products offer a viable alternative with several remotely sensed precipitation datasets now available with long historical data records (+30 years), which include the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR) datasets. This chapter presents a comparative analysis of three historical satellite-based precipitation datasets that include Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) 3B43 version 7 (1998-2015), PERSIANN-CDR (1983-2015) and CHIRPS 2.0 (1981-2015) over Chile to assess their performance across the country and for the case of the two long-term products the applicability for agricultural drought were evaluated when used in the calculation of commonly used drought indicator as the Standardized Precipitation Index (SPI). In this analysis, 278 weather stations of in-situ rainfall measurements across Chile were initially compared to the satellite data. The study area (Chile) was divided into five latitudinal zones: North, North-Central, Central, South-Central and South to determine if there were a regional difference among these satellite products, and nine statistics were used to evaluate their performance to estimate the amount and spatial distribution of historical rainfall across Chile. Hierarchical cluster analysis, k-means and singular value decomposition were used to analyze these datasets to better understand their similarities and differences in characterizing rainfall patterns across Chile. Monthly analysis showed that all satellite products highly overestimated rainfall in the arid North zone. However, there were no major difference between all three products from North to South-Central zones. Though, in the South zone, PERSIANN-CDR shows the lowest fit with high underestimation, while CHIRPS 2.0 and TMPA 3B43 v7 had better agreement with in-situ measurements. The accuracy of satellite products were highly dependent on the amount of monthly rainfall with the best results found during winter seasons and in zones (Central to South) with higher amounts of precipitation. PERSIANN-CDR and CHIRPS 2.0 were used to derive SPI at time-scale of 1, 3 and 6 months, both satellite products presented similar results when it was compared in-situ against satellite SPI's. Because of its higher spatial resolution that allows better

characterizing of spatial variation in precipitation pattern, the CHIRPS 2.0 was used to mapping the SPI-3 over Chile. The results of this chapter show that in order to use the CHIRPS 2.0 and PERSIANN-CDR data sets in Chile to monitor spatial patterns in the rainfall and drought intensity conditions, these products should be calibrated to adjust for the overestimation/underestimation of rainfall geographically specially in the North zone and seasonally during the summer and spring months in the other zones.

3.1 Introduction

Precipitation is one of the key parameters for climate monitoring, particularly to detect climatically-extreme events such as drought, which impacts most regions of the world. A simple definition of drought is an extended period of abnormal dryness that has negative impacts on agricultural and water resources (WMO, 1986). Conceptually, there are several different sectoral definitions of drought that are defined by the duration of the precipitation deficit including short-term meteorological drought (spanning days to weeks), agricultural drought (month to several months), and hydrological drought (months to years) (Wilhite and Glantz, 1985). According to IPCC (2013), changes in the patterns of precipitation is expected globally over the next few decades. The change in precipitation patterns coupled with the sustained increase in global temperature since 1880 and the anthropogenic factors due to human activities (e.g. increased emission of the Green House Gasses and land use change such as cutting down forests to create farmland) is likely to increase the frequency and intensity of natural disasters like drought throughout the world (IPCC, 2013; Loon et al., 2016). Knowledge of the amount and spatial variability of precipitation historically is important to map and monitor drought condition globally. In situ-based rainfall measurements at weather station locations have traditionally been used for this application, but the number, geographic distribution, and length of record of these measurements are often lacking in many countries including Chile. The creation of such data sets is challenging because it is costly to maintain a dense network of weather stations over a long period of time as a result, there are often spatial gaps and a lack of local resolution in the rainfall data and the drought patterns mapped from these point-based weather station data using spatial interpolation techniques. Many weather stations have a relatively short or incomplete historical record of observations, which is problematic for determining the magnitude of specific precipitation deficit period and the severity of the corresponding drought.

Accurate historical precipitation data and effective drought monitoring tools are of considerable interest for Chile. The IPCC (2013) indicate that precipitation is expected to decrease in the near-future in the central part of Chile, which is a primarily agricultural area within the country. Studies made during the last years from North to Central Chile found important results about drought frequency. A trend in the increase of drought frequency in the Coquimbo region of northern Chile, particularly in Limarí Valley was identified (Meza, 2013). Also, was found that a rainfall deficit of 40% had a return interval of approximately once every 4 years in the northern, semi-arid Coquimbo region of Chile and a 22-year return interval in the more humid O'Higgins regions of central Chile (Núñez et al., 2011). Lately, Zambrano et al. (2016) evaluated agricultural drought using satellite-based vegetation index data and found that in the Bío-Bío region (South-Central Chile) over the last sixteen years had experienced three drought event during the 2007-2008, 2008-2009 and 2014-2015 growing seasons. All these studies were carried out using precipitation data obtained from a limited number of weather stations across Chile. These results could be extended both on the spatial and temporal scale to improve the findings using accurate satellite long-term precipitation datasets.

Satellite datasets are becoming increasingly important to fill in the spatial and temporal data gaps for climate-based applications such as drought monitoring. Several global remotely sensed datasets now have historical records spanning 18 years or more, and lately the long-term products having more than 30 years which are appropriate for climate studies and represent a viable information source in many parts of the world. One widely used remotely sensed precipitation dataset has been acquired by the Tropical Rainfall Measuring Mission (TRMM), which is jointly supported by National Aeronautics and Space Administration (NASA) and Japan Aerospace Exploration Agency (JAXA). The TRMM precipitation datasets (Huffman et al., 2007) spans since November 1997 until present, although the mission comes to its end on April 2015, but thanks to its successor the Global Precipitation Measurement (GPM) which is the continuity of TRMM the

dataset have continued. The Global Precipitation Measurement (GPM) is an international satellite mission to provide next-generation observations of rain and snow worldwide starting from temporal resolution of three hours and at spatial resolution as high as 0.1°-30 minute. Other precipitation datasets have been produced using a combination of infrared (IR) and passive microwave (PMW) observations from multiple satellite sensors using different precipitation estimation methods. These include Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) (Hsu et al., 1997) and Climate Prediction Center Morphing (CMORPH) technique (Joyce et al., 2004). Most products have short-term data and have been evaluated in different part of the world such as: South America (Salio et al., 2015), Colombia (Dinku et al., 2009), Saudi Arabia (Almazroui, 2011), Greece (Nastos et al., 2016), Ethiopia (Duan and Bastiaanssen, 2013), China (Guo et al., 2016b), India (Shah and Mishra, 2015), Iran (Moazami et al., 2016) and Himalayas (Bharti and Singh, 2015) and many other more (Pipunic et al., 2015; Kenawy et al., 2015; Dinku et al., 2007; Tan et al., 2015).

The study of climate change and climate variability requires a long-record data to permit the evaluation of climate and associated natural disasters like drought. The National Research Council (NRC) defined the Climate Data Records (CDR) as time-series measurements of sufficient length, consistency, and continuity to determine variability and climate change (National Research Council, 2004). Thus, two new satellite products for long-record precipitation studies were considered in this chapter They are the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR) (Ashouri et al., 2015) and CHIRPS 2.0 (Funk et al., 2014) datasets which both have more than 30-year records of data. Both products represent potentially valuable data sources for monitoring drought in data-limited countries such as Chile because they have an adequate length of record to detect and quantify drought conditions within a longer historical context. These products are relatively new and still there are only a limited number of studies evaluating their performance of estimating the amount and spatial distribution of precipitation. Miao et al. (2015) evaluated PERSIANN-CDR over China and found that the agreement between the dataset with in-situ measurements in dry regions is not strong. Ashouri et al. (2015) tested PERSIANN-CDR during the Hurricane Katrina (2005) and the flooding on Sydney, Australia (1986); and found in both that PERSIANN-CDR is performing reasonably well when compared to radar and ground-based observations. However, Ashouri et al. (2015) also examined the frequency distribution of precipitation from PERSIANN-CDR as compared to those of CPC gauge observations and TMPA v7 concluding that generally PERSIANN-CDR tends to underestimate the frequency distribution. Lately, Guo et al. (2016a) analyzed PERSIANN-CDR for the assessment of meteorological drought over China using ground-based gridded China monthly Precipitation Analysis Product (CPAP) from 1983 to 2014, their results shows that 6-month SPI has the best agreement with CPAP in identifying drought months; however, large differences between PERSIANN-CDR and CPAP in depicting drought patterns and identifying specific drought events were found over northwestern China. Katsanos et al. (2016) evaluated CHIRPS over Cyprus for a 30-year period and their results showed good correlation with in situ measurement with an overestimation noted during the decade 2001-2010 possibly due to the incorporation in the latter of TRMM estimates, which tend to overestimate rainfall (Katsanos et al., 2016).

The aim of this chapter is to evaluate the performance and fit of monthly, long-term satellite-based precipitation products over Chile for mapping and quantifying historical rainfall and drought patterns. In this chapter, the CHIRPS 2.0 and PERSIANN-CDR datasets are compared with a well-studied satellite-based product, the TMPA 3B43 v7 and with in-situ measurements obtained from weather stations across Chile. The Chilean territory was divided in five zones primarily based on climate characteristics to evaluate the performance of these datasets regionally across the country. The goal was to evaluate the accuracy and applicability of these products to characterize rainfall patterns across Chile and transform the data into a precipitation-based drought index, the Standardized Precipitation Index (SPI, McKee et al. (1993)) technique, for agricultural drought monitoring in Chile.

3.2 Study area

According to the Köppen climate classification system (Kottek et al., 2006), Chile has five primary climate regimes that include the cold desert climate (Bwk) in the North, temperate Mediterranean climate (Csb) in the Central part, temperate oceanic climate (Cfb) in the South-Central, cool oceanic climate (Cfc) in the South and tundra climate in the austral part of the extreme South. To regionally evaluate the satellite-derived precipitation data with in-situ measurements, the Chilean territory (for areas north of 50° South Latitude) were divided into five zones that capture geographic variations in climatic conditions. The five zones were geographically defined as: (1) North (17.6° S to 28° S latitude), (2) North-Central (28° S to 32° of S latitude), (3) Central (32° S to 36° S latitude), (4) South-Central (36° S to 40° S latitude) and (5) South (40° S to 50° S latitude). Fig 3.1a shows the five zones with the 278 weather stations locations used in this chapter. Annual rainfall is mostly below 80 mm in the North zone, below 100 mm in the North-Central zone, around 800 mm in the Central zone, from 1000 mm to 1500 mm in the South-Central zone and reaching up to 2000 mm in the South zone. To describe vegetative features of Chile, the land cover MCD12Q1 product with the scheme IGBP was used (Friedl et al., 2010). The land cover of the study area is shown in Fig. 3.1b. The North zone is mostly barren and in the North-Central region is dominated by shrubland with some isolated cropland along the river valleys. Agriculture in Chile is primarily concentrated from the 32° S to 40° S in the Central and Central-South zones where forest, cropland and grassland are the dominant land cover types. Between 40° S and 45° S there are mainly forest and some areas covered with cropland which are close to the 40° S. Finally, to the South of 45° S there are a mixed vegetation of forest and grassland as the principal land cover type. Fig. 3.1c shows the altitude difference through Chile, highest altitudes are in the North and North-Central zones, range mainly from 700m to more than 3500m. In the Central and South-Central zones is possible to note the valleys in the middle part with altitudes between 100m and 700m. South zone close to 40° S latitude has topography around 10m to 200m. Also, toward eastern from 40° S to the North the highest altitudes corresponds to the Andes Mountains.

3.3 Methods

3.3.1 Data

Monthly time-series of precipitation (mm month^{-1}) from the three satellite products and the weather station-based in-situ rainfall measurements were compared and analyzed. The first satellite product evaluated was the 0.25-degree, spatial resolution TMPA products with near-global (between 50°S - 50°N) coverage. The standard monthly TMPA product 3B43 version 7 was already temporally aggregated to a monthly time-step and the data required no additional temporal modifications prior to analysis. The TMPA data record spanned a period from January 1998 to present (Huffman et al., 2007, 2010). The second satellite product was the PERSIANN Precipitation Climate Data Record (PERSIANN-CDR), which is a daily quasi-global precipitation product covering the period starting from January 1983. The PERSIANN-CDR data covers from 60°S to 60°N latitude and 0° to 360° longitude at 0.25° spatial resolution (Ashouri et al., 2015). The third satellite products evaluated was the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) dataset, which has a 30+ year, quasi-global rainfall dataset, available at the monthly, dekad, pentad and daily time steps. CHIRPS spatial data coverage spans from 50° S to 50° N (over all longitudes) with the data record extending from 1981 to present. CHIRPS data have a 0.05° spatial resolution with the satelitte data calibrated with in-situ station data to create gridded rainfall time-series appropriate for trend analysis and seasonal drought monitoring (Funk et al., 2014).

The General Water Directorate of Chile (DGA) has the densest weather stations network through Chile and with the longest historical records too. The network consists of stations with observers, which, recorded on a daily basis each data. Also, it has datalogger equipment to store data and transmit to a central database. About the quality control, the first verification is performed regionally for experienced staff which checks coarse errors and data gaps. Then, the quality control is made by spatial consistency comparing rainfall measurement with weather stations of neighboring sub-basin of the same time interval. Distance and altitude

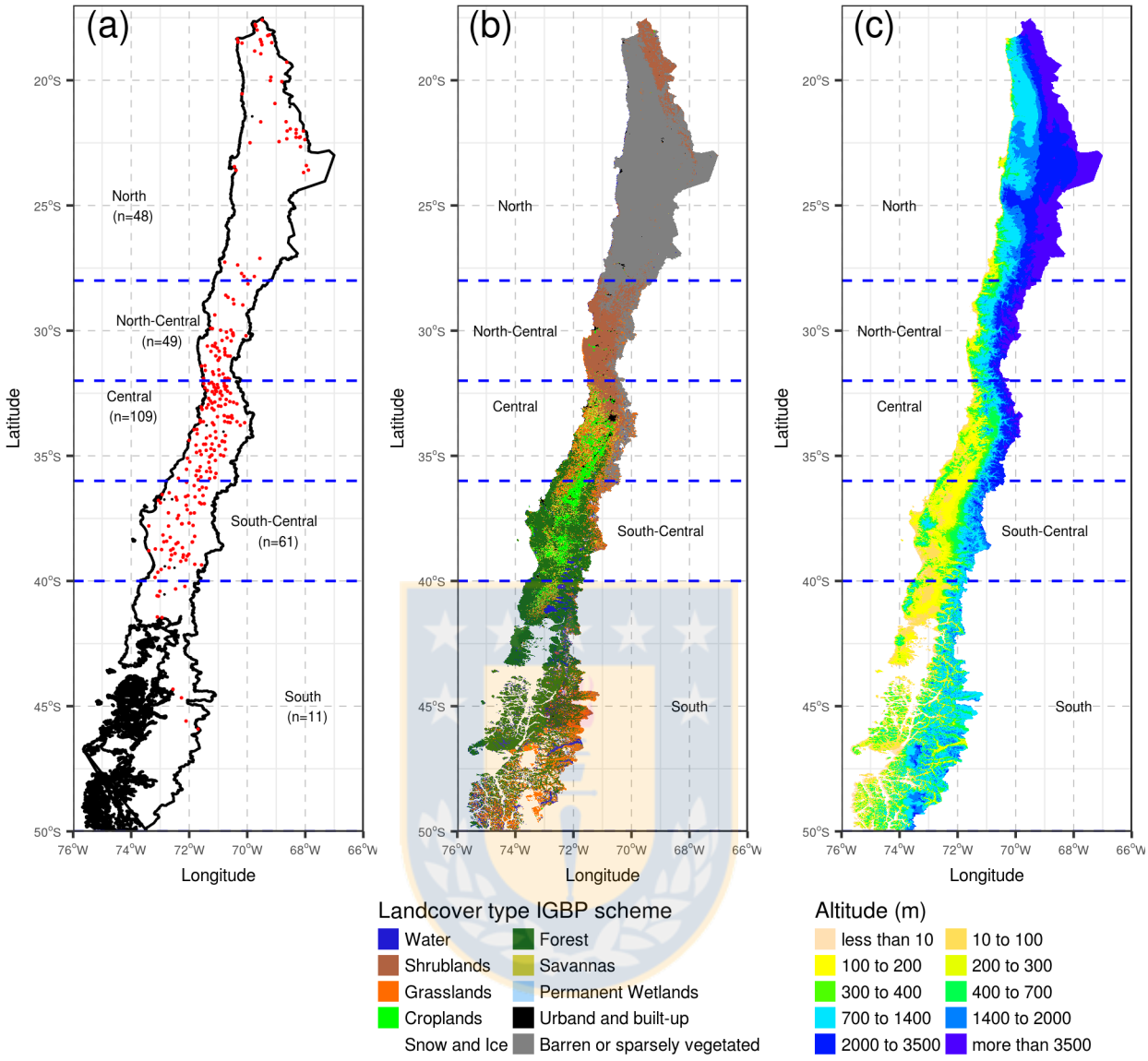

Figure 3.1: Study area map with (a) 278 weather stations, (b) landcover type and (c) terrain elevations.

considerations are included in the selection of neighboring stations. The effect of an air mass over the area and the impact on the weather stations values are analyzed. For measurements validation, the analyst experience allows discovering the errors of direct readings from an a observer or datalogger malfunction.

For the evaluation of satellite precipitation products, ground observation of rainfall data were obtained from the database of the Center for Climate and Resilience Research (CR2) at the University of Chile. The database consisted of 780 stations with monthly rainfall data collected from 1940 to 2015. They are comprised of data records consolidated from the DGA and the Meteorological Directorate of Chile (DMC) for use in research. From the total of 780 available meteorological stations with in-situ data, only 278 stations were selected for analysis in this chapter based on the criteria that a station must have at least a 25 year rainfall data record (1981 to 2015), less than 10% missing data, and be geographically located between 17° S and 50° S in continental Chile. The dramatic reduction in the number of stations that could be used in this analysis illustrates the relative lack of long-term, temporally-complete rainfall records from in-situ gauges that are available in Chile.

3.3.2 Preparation and data analysis

With the daily precipitation data from the product PERSIANN-CDR from January 1983 to June 2015, monthly cumulative precipitation was calculated obtaining a total of 402 precipitation grids. In the case of CHIRPS 2.0 and TMPA 3B43 v7, these datasets are in monthly accumulated precipitation. For TMPA 3B43 v7, the data spanned from 1998 to 2015 and 210 data grid were used between January 1998 and June 2015. For CHIRPS 2.0, a total of 414 data grids were used from January 1981 to June 2015. Monthly precipitation values across the historical record were extracted from the grid cell locations that corresponded to the 278 weather stations locations from the TMPA 3B43 v7, CHIRPS 2.0 and the PERSIANN-CDR precipitation datasets, respectively. The extracted precipitation data from each remote sensing and in-situ data set were then spatially averaged with the five regional zones defined for the regional analysis part of this chapter. The aggregation was done by averaging the values of precipitation (in-situ and satellite) for each monthly period for all weather station locations within each zone. The aggregated data were then statistically analyzed using box-plot and heatmap graphs, as well as summary statistics that included mean, median, first quartile ($Q_{25\%}$), third quartile ($Q_{75\%}$), maximum value and number of observations (n).

The monthly precipitation time-series data were compared for each zone and between the satellite products and in-situ measurements. Using the spatially-aggregated data, the monthly difference between satellite products and in-situ data for each year was calculated to measure the level of underestimation and overestimation among datasets. To complement this analysis, the data were also temporally aggregated calculating the mean cumulative season for winter, summer, spring and autumn time periods to determine if there were differences in the inter-relationships among the datasets at different times of the year.

To further evaluate and compare the time-series precipitation datasets nine statistics were calculated for each of the 278 weather station locations that included: lineal coefficient of correlation (CC), the magnitude of underestimation (ME) (Eq. (3.1)), mean absolute error (MAE) (Eq. (3.2)), multiplicative bias (bias) (Eq. (3.3)), efficiency (E_{ff}) (Eq. (3.4)), the Root Mean Square Error (RMSE) (Eq. (3.5)), frequency bias (FBS) (Eq. (3.6)), the probability of detection (POD) (Eq. (3.7)) and false-alarm ratio (FAR) (Eq. (3.8)). These station-based statistics were then aggregated for each zone generating 135 statistics in total among the five zones. These statistics were analyzed using hierarchical cluster analysis, singular value decomposition (SVD) and k-mean to better understand their similarities and differences in characterizing spatial and temporal rainfall patterns. To evaluate the behavior during the year, the statistics were monthly calculated, obtaining 1620 statistics (i.e., 12 months $\times$ 9 statistics $\times$ 3 products $\times$ 5 zones), which were also analyzed using hierarchical cluster, k-means and singular value decomposition. Also, we used the Root Mean Squared Error (RMSE) statistic to have a measure of the error between the in situ time-series of rainfall and SPI values against those derived from satellite products.

Finally, to analyze the application of these satellite products and evaluate the spatial variability of agricultural drought patterns, the CHIRPS 2.0 precipitation data were used to calculate the SPI a time-scales (i.e., one, three and six months SPI) for the period of 1981 to 2015. For SPI calculation, a Gamma distribution was adjusted using a method for parameter fitting based on unbiased Probability Weighted Moments (Vicente-Serrano et al., 2010). To compute the SPI the *SPEI* (Beguería and Vicente-Serrano, 2013) package within the **R** environment (R Core Team, 2016) was used.

3.4 Results

3.4.1 Satellite and rain gauge precipitation

The summary of monthly rainfall and distribution comparison for 278 selected weather stations with extracted data from the grids of satellite precipitation products estimates are presented in Table 3.1, Fig. 3.2a and Fig. 3.2b. In the North, the monthly precipitation is mainly under 200 mm and to the South is possible to observe the winter pattern where the amount of rainfall increase to more than 200 mm in most locations and as high as more than 400 mm in the South. There are some extreme precipitation values (i.e., more than 1200 mm) from the Central zone and southward. This extreme was 1232 mm on May 1981, 1241 mm

Table 3.1: Summary statistics for precipitation data from weather stations (in-situ) and PERSIANN-CDR, TMPA 3B43 v7, CHIRPS 2.0 satellite products, in five zones of Chile. Total number of observations (n), weather stations by zone (Stations), missing observations, mean ($\bar{X}$), first quartile ($X_{25\%}$), median ($X_{50\%}$), third quartile ($X_{75\%}$) and maximum (X_{max}). The time-period used was 1981-2015, 1983-2015 and 1998-2015 for in-situ, CHIRPS 2.0, PERSIANN-CDR and TMPA 3B43 v7, respectively.

		Summary statistics							
Zone	Product	n	Stations	Missing	$\bar{X}$ [mm mo^{-1}]	$X_{25\%}$ [mm mo^{-1}]	$X_{50\%}$ [mm mo^{-1}]	$X_{75\%}$ [mm mo^{-1}]	X_{max} [mm mo^{-1}]
North	in-situ	19872	48	473	6.95	0.00	0.00	0.50	413.40
	PERSIANN-CDR	18720	48	0	11.90	1.24	4.50	12.44	213.64
	TMPA 3B43 v7	10080	48	0	10.19	0.84	3.31	10.23	284.99
	CHIRPS 2.0	19872	48	0	10.47	1.43	3.90	10.93	221.72
North-Central	in-situ	20286	49	226	12.40	0.00	0.00	7.50	597.50
	PERSIANN-CDR	19110	49	0	18.85	2.02	8.25	24.10	390.18
	TMPA 3B43 v7	10290	49	0	16.26	0.52	3.23	16.03	568.67
	CHIRPS 2.0	20286	49	0	10.72	1.23	3.41	11.55	220.66
Central	in-situ	45126	109	853	46.06	0.00	9.20	54.20	1394.10
	PERSIANN-CDR	42510	109	0	45.99	5.62	22.12	64.29	504.89
	TMPA 3B43 v7	22890	109	0	45.33	1.80	14.91	61.71	517.03
	CHIRPS 2.0	45126	109	0	44.34	5.22	14.74	52.53	787.23
South-Central	in-situ	25254	61	500	124.93	29.00	81.00	177.60	1258.60
	PERSIANN-CDR	23790	61	0	85.49	19.13	55.87	127.58	554.02
	TMPA 3B43 v7	12810	61	0	105.30	28.26	77.21	159.97	688.65
	CHIRPS 2.0	25254	61	0	121.81	34.80	82.31	170.37	1238.03
South	in-situ	4554	11	130	145.56	57.18	114.90	201.12	1139.00
	PERSIANN-CDR	4290	11	0	107.73	50.44	86.80	147.28	441.86
	TMPA 3B43 v7	2310	11	0	142.40	69.64	125.33	197.01	594.76
	CHIRPS 2.0	4554	11	0	128.34	56.29	102.70	170.85	827.94

on July 1987, 1236 mm on July 1987, 1373 mm on June 2000, 1394 mm on June 2000 and 1258 on July 2001 for the weather stations of *Rio Malleco y Vergara*, *Rio Maule Medio*, *Rio Loncomilla*, *Rio Maule Medio*, *Rio Loncomilla* and *Rio Bío-Bío Alto*, respectively. The seasonal variations are also evident for the weather station located in the central and northern zones as shown by the alternating seasonally higher (red color) and lower (green color) for each year in the multi-year historical record presented. Annual drier periods also were more pronounced across the stations moving northward from central to northern parts of Chile.

The boxplots of Fig. 3.2b compare the statistical distribution of rainfall data between in-situ and satellite products for the five zones. In the North zone, all three products overestimate the distribution of precipitation as shown by Fig. 3.2b where the box of three satellite products are higher than the box for in-situ measurements. In-situ rainfall totals for 75% of the stations in the North were below 0.5 mm, whereas for the three satellite rainfall products the totals ranged between 0.84 mm and 12.44 mm as shown in Table 3.1 and Fig. 3.2b. In the North-Central zone, the satellite products overestimate the lower precipitation values with $X_{25\%}$ equal to 0.52 mm in the case of TMPA 3B43 v7 compared to the in-situ data that equaled 0 mm, but the three satellite products capture most of the remaining rainfall distribution across the intermediate and higher values. In the Central zone, the three satellite products had a $X_{75\%}$ between 53 mm (CHIRPS 2.0) to 64 mm (PERSIANN-CDR) close to in-situ values which has 54 mm. However, like in the North-Central zone, the lower rainfall values ($X_{25\%}$) for the PERSIANN-CDR and CHIRPS 2.0 datasets were approximately 5 mm and TMPA 3B43 v7 was 1.8 mm compared to 0 mm for in-situ data (see Fig. 3.2a and Table 3.1). The South and South-Central zone shows rainfall distributions very similar between satellite products and in-situ measures, in the range for interquartile range (IQR) of 29 mm to 178 mm for South-Central and between 50 mm to 200 mm in the South zone (see Fig. 3.2b and Table 3.1).

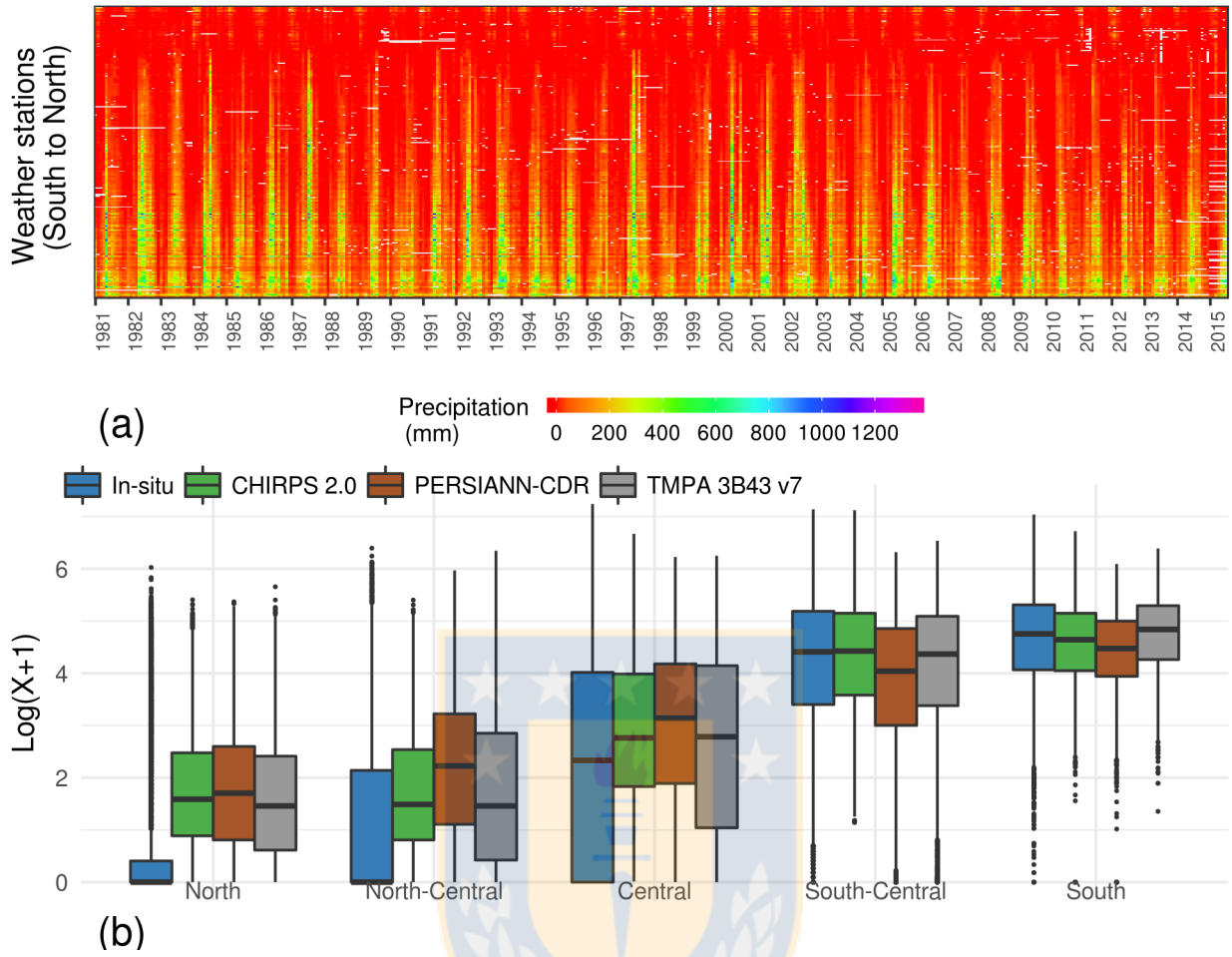

Figure 3.2: (a) Monthly rainfall data from 1981 to 2015 for 278 weather stations, and (b) Comparison of distribution for precipitation data with logarithmic transformation, from in-situ and satellite products TMPA 3B43 v7, PERSIANN-CDR and CHIRPS 2.0; for North, North-Central, Central, South-Central and South zones in Chile. The y-axis presents the 278 weather stations from North (top) to South (bottom), and the x-axis presents the period from 1981 to 2015. The white spaces represent missing data.

3.4.2 Time series, annual difference and seasonal variation of rainfall

Fig. 3.3 shows the time series of the spatially-averaged precipitation datasets for the five zones. Like the previous results, the values of in-situ with satellite products had the close agreement in the Central zone as compared to the North zone, where the largest differences were found, particularly when there is low precipitation near 0 mm. The South-Central and South zones exhibited good correspondence between in-situ and satellite data across the range of precipitation totals. Values of *RMSE* for the North zone was around 7 mm for all three products, in the North-Central TMPA 3B43v7 shows the lowest values with 12.2 mm and CHIRPS 2.0 the highest with 20.8 mm. Similar values had PERSIANN-CDR and CHIRPS 2.0 in the Central zones reaching around 20 mm. In the South-Central and South zones, TMPA 3B43 v7 has the lowest *RMSE* value with 37.6 mm and 24.2 mm, respectively; and PERSIANN-CDR reaches the highest values with 54.8 mm and 48 mm, respectively (Fig. 3.3).

Fig. 3.4a shows the yearly averaged values of monthly precipitation difference between the three satellite products (S) and the in-situ rain gauge (G) data for the five regional zones. In the North zone, the rainfall totals were consistently overestimated by as much as 10 mm in all years except 2000 and 2001. This pattern was also observed for all satellite products over the North zone. The North-Central and Central zones had

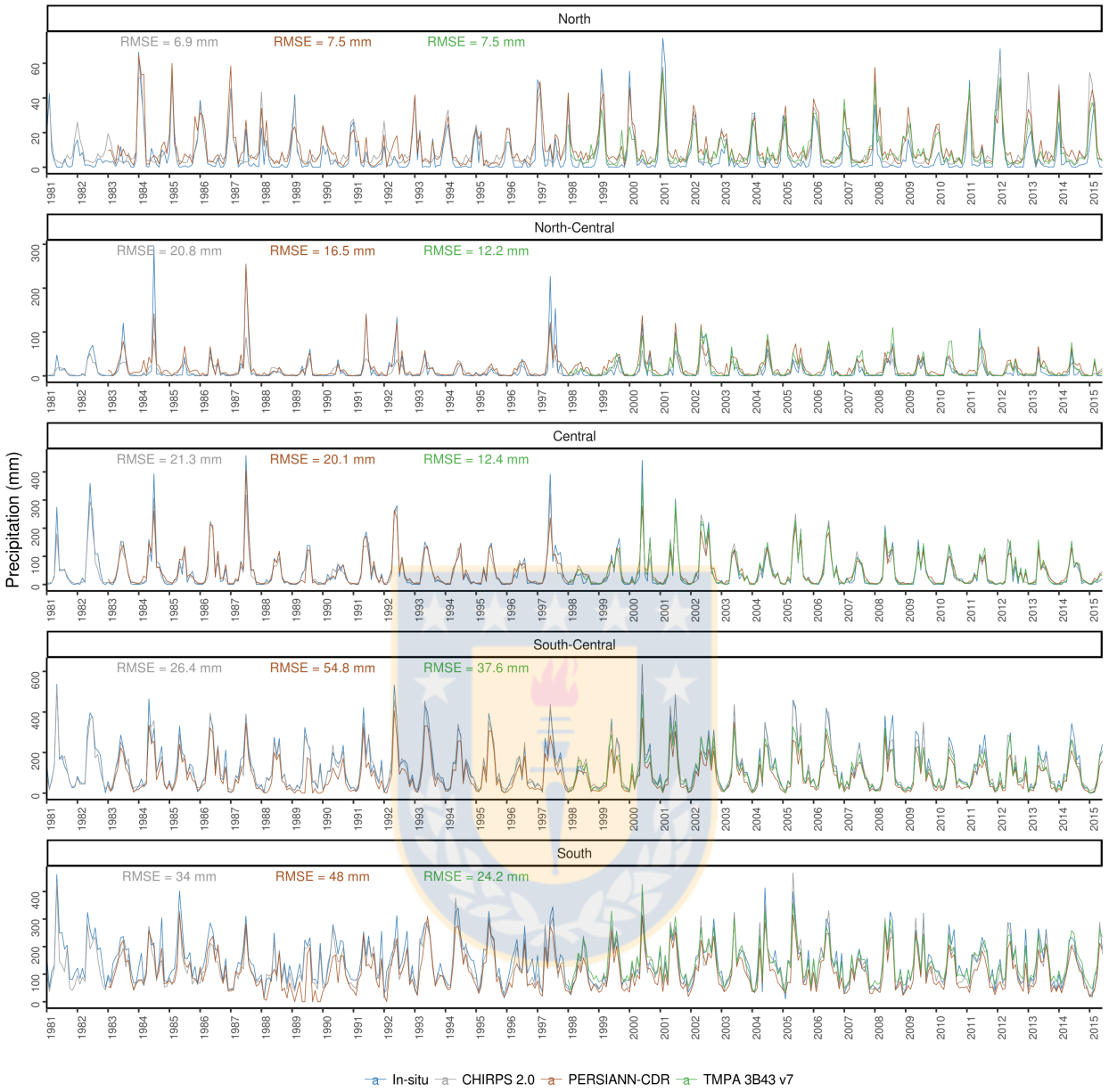

Figure 3.3: Spatial-averaged monthly time-series of in-situ rainfall data and extracted from satellite products TMPA 3B43 v7, PERSIANN-CDR, CHIRPS 2.0 in five zones of Chile: North, North-Central, Central, South-Central and South.

a similar rainfall pattern, in most of the years with satellite products overestimating rainfall. The CHIRPS 2.0 data had the lowest difference values and the best fit among the remotely sensed products, but there were still some anomaly years (1987 and 1997) in which CHIRPS 2.0 data had differences by as much as 20 mm in North-Central and Central zones where the mean annual precipitation are around 80 mm (~5%) and 500 mm (~4%), respectively. In the South-Central and South zones, the satellite products consistently underestimated monthly precipitation with differences as high as 69 mm and 83 mm, respectively; in the case of PERSIANN-CDR which correspond as much as 5% for both. In contrast, CHIRPS 2.0 has the best correspondence in these zones with in-situ measurements difference in South-Central and South zones of 20 mm (1%) and 47 mm (2%), respectively (Fig. 3.4a).

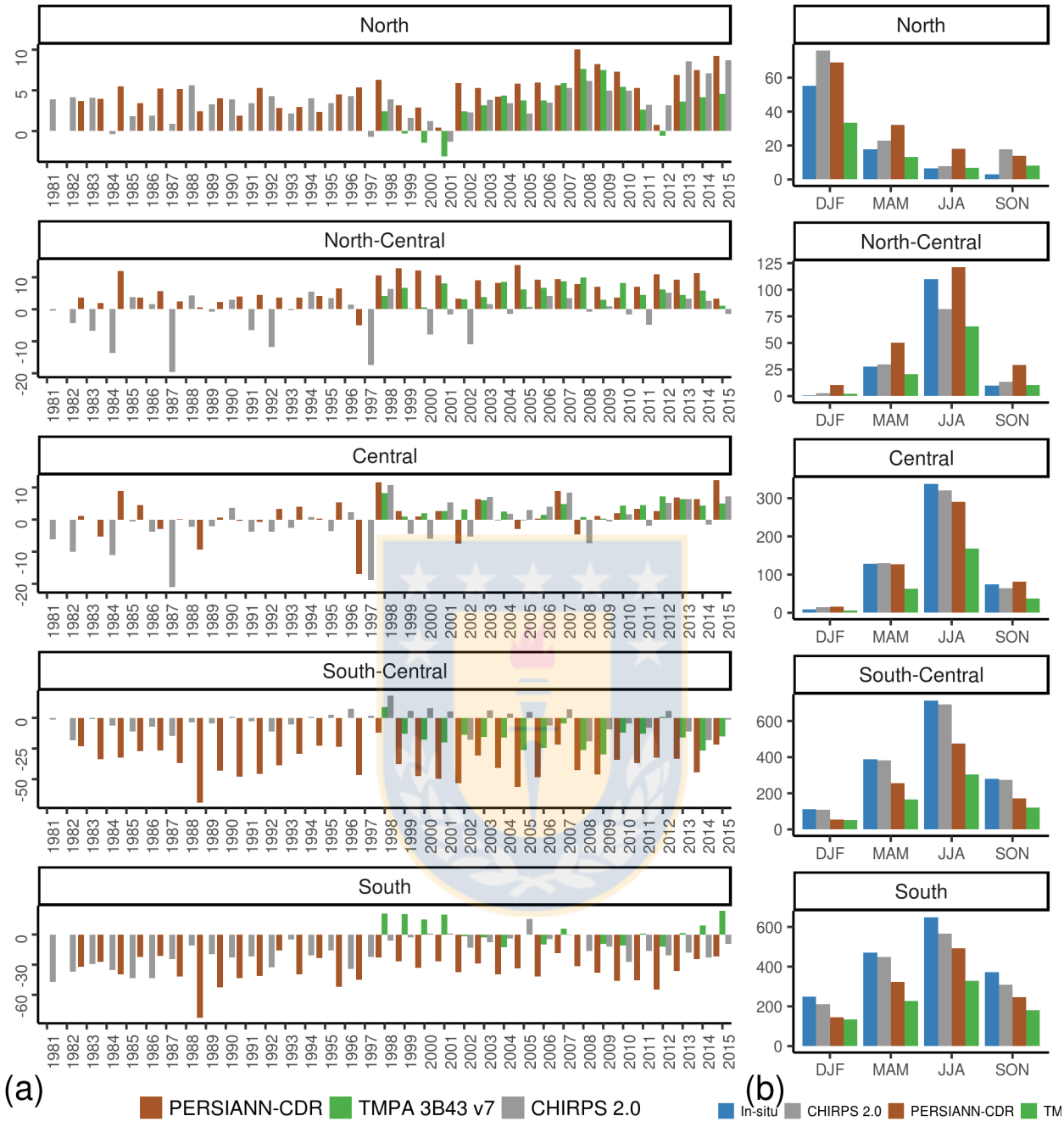

Figure 3.4: (a) Annual averaged difference of monthly precipitation for in-situ data and from satellite products TMPA 3B43 v7, PERSIANN-CDR and CHIRPS 2.0. (b) Seasonal comparison of precipitation for five zones of Chile and between in-situ, CHIRPS 2.0, PERSIANN-CDR and TMPA 3B43 v7. The seasonal periods were: March, April and May (MAM) for Autumn; June, July and August (JJA) for Winter; September, October and November (SON) for Spring; and December, January and February (DJF) for Summer.

In addition, a comparison of the seasonal variation of precipitation between products for each zone is presented in Fig. 3.4b. The greatest difference with in-situ data was during the Spring season in the North zone, where each product had accumulated seasonal precipitation values higher than the in-situ measurements.

Also, PERSIANN-CDR and CHIRPS 2.0 overestimates, while TMPA 3B43 v7 underestimates seasonal rainfall. Over the North zone, the precipitation amount is higher during the Summer season, which corresponds to the commonly known as *Bolivian winter* (Romero et al., 2013). In the North-Central zone, there was a significant difference with in-situ measurements, CHIRPS 2.0 underestimates by 30mm and PERSIANN-CDR overestimates by 10mm during the Winter season and in the other zones, the behavior is similar with those found in the North zone. However, in the other three seasons (i.e., Spring to Autumn), CHIRPS 2.0 had the best agreement with in-situ data while PERSIANN-CDR highly overestimates seasonal rainfall. For the Central, South-Central and South zones, CHIRPS 2.0 had the best agreement for all seasons with the lowest precipitation in Summer and the highest in Winter (see Fig. 3.4b), followed by PERSIANN-CDR and ending TMPA 3B43 v7 with the lowest seasonal agreement.

3.4.3 Statistics of comparison between satellite-derived and in-situ measurements

The statistics used to compare the different precipitation datasets were calculated for each 278 meteorological stations and the average values of the satellite products over the five zones are presented in Table 3.2. Hierarchical cluster analysis, k-means and singular value decomposition were used to analyze these datasets to better understand their similarities and differences in characterizing rainfall patterns across Chile. As an important source of variation monthly precipitation in rain gauges (G) was included to the 9 statistics for the analysis. The results revealed no significant difference between the products from North to South-Central zones and differences exists between them in the South zone. The singular value decomposition indicated that the CC, E_{ff}, POD, G and $RMSE$ are the major contributors to the variation among these datasets. The statistics showed that differences were greater in areas that receive lower precipitation than areas with high rates of precipitation. For example, the South and South-Central zones, which normally receive between 1500 mm and 1700 mm of precipitation annually instead of 80 mm and 150 mm in the North and North-Central zones. The South and South-Central zones had the best results reflected by the FAR and high POD values among the zones, as well as the highest CC and good E_{ff} values. PERSIANN-CDR in the South and South-Central zones presents higher underestimations reaches -36 mm of ME in both zones. Cluster analysis identifies two big groups, one corresponding to North zone, and the second from North-Central to South zones. This reflects that North zone (group 1) presents the lower agreement with in-situ measurements and this was similar for the three satellite products. The second group has two main sub-groups, one for North-Central and Central, showing better results than in North zone, and the second sub-group with the higher performance for the three satellite products showed in South-Central and South zones.

In order to further analyze the goodness of fit of the satellite products during the year, several monthly statistics were also studied. Fig 3.5 presents the heatmap diagrams for the hierarchical cluster analysis performed by month for five zones and 9 statistics along with in-situ precipitation (G) that was added as an additional measure of accuracy variation. For Fig. 3.5, the top horizontal axis represents the dendrogram cluster by statistics and the left vertical axis corresponds to dendrogram cluster by specific month. Also the vertical color palette between the left vertical dendrogram and the first column of the heatmap shows the specific remote sensing precipitation product that corresponds to each row. The colors are grey, brown and green, corresponding to CHIRPS 2.0, PERSIANN-CDR, and TMPA 3B43 v7, respectively. As example, in Fig. 3.5a, in first row, the left dendrogram indicates that was included in one of the main clusters, next the green color in the vertical palette indicate that correspond to TMPA 3B43 v7 product, then the nine cells shows the nine scaled statistics values, beside that on the right the month is showed, in the case of the first row was 'November'.

The major contributors statistics in the North zone were ME, MAE, $bias$, FBS and FAR as showed in the horizontal dendrogram in Fig. 3.5a. TMPA 3B43 v7 has the highest values among these statistics from September to November, and CHIRPS 2.0 in September; which collectively formed the first group with the lowest fit. In vertical dendrogram in Fig. 3.5a, the second and third group are also defined. The second group had the best seasonal results from December to March, which corresponds to the 'Bolivian winter', when higher precipitation are received during this Summer period over the North zone. The third group that corresponded to the April to November time period had the poorest fit among all products.

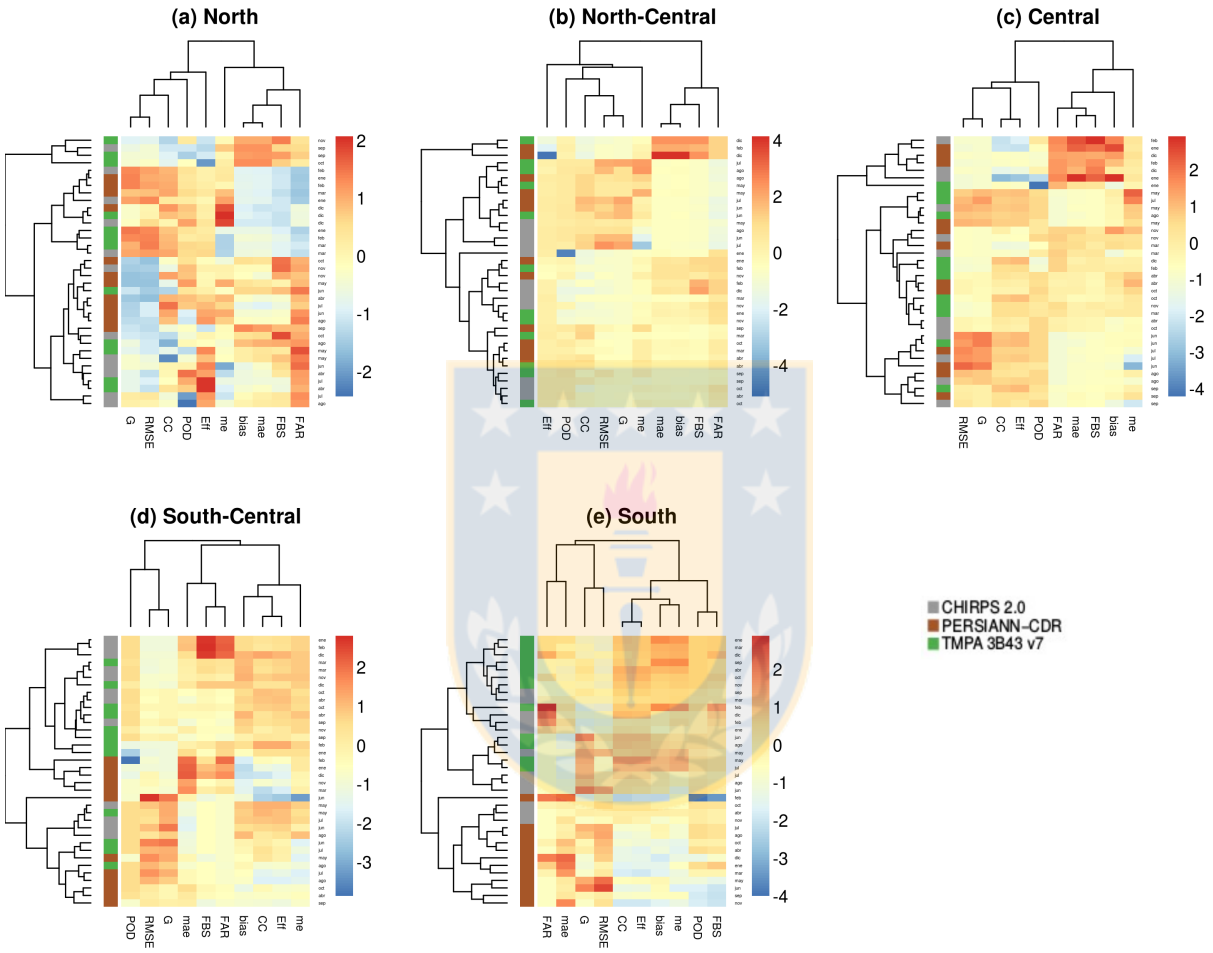

Figure 3.5: Heatmap for hierarchical cluster analysis of monthly statistics of RMSE, G, CC, E_{ff}, POD, FAR, MAE, FBS, bias, and ME; for North, North-Central, Central, South-Central and South zones. In the left vertical axis, the dendrogram shows groups made by month. The top axis present the dendrogram by statistics. On the left of the column showing POD, there is a palette legend showing the satellite product at which each row correspond.

Table 3.2: Summary of statistics aggregate for zones North, North-Central, Central, South-Central and South; for products TMPA 3B43 v7, PERSIANN-CDR and CHIRPS 2.0. CC, ME, MAE, bias, E_{ff}, FBS, POD, FAR and HSS

		Statistics of comparison								
Zone	Product	CC	ME [mm mo^{-1}]	MAE	bias	E_{ff}	RMSE [mm mo^{-1}]	FBS	POD [%]	FAR [%]
North	CHIRPS 2.0	0.48	3.34	10.69	10.89	-44.44	15.32	7.31	0.87	0.73
	PERSIANN-CDR	0.59	4.64	8.86	9.33	-30.90	13.95	6.66	0.95	0.72
	TMPA 3B43 v7	0.51	2.97	9.20	9.43	-36.57	14.78	7.86	0.91	0.72
North-Central	CHIRPS 2.0	0.72	-1.71	0.86	0.82	0.43	24.20	2.14	0.95	0.54
	PERSIANN-CDR	0.80	6.45	1.07	1.63	0.57	20.99	2.38	0.99	0.56
	TMPA 3B43 v7	0.75	5.59	1.04	1.55	0.05	21.29	1.94	0.97	0.48
Central	CHIRPS 2.0	0.88	-1.50	0.51	1.01	0.72	38.25	1.57	0.99	0.35
	PERSIANN-CDR	0.91	1.16	0.55	1.19	0.73	37.53	1.48	0.98	0.32
	TMPA 3B43 v7	0.93	3.47	0.46	1.19	0.76	31.23	1.32	0.99	0.24
South-Central	CHIRPS 2.0	0.90	-3.00	0.31	1.01	0.76	58.49	1.05	1.00	0.05
	PERSIANN-CDR	0.91	-36.94	0.39	0.78	0.60	77.47	1.02	0.97	0.04
	TMPA 3B43 v7	0.91	-15.08	0.32	0.93	0.74	63.20	1.02	0.99	0.03
South	CHIRPS 2.0	0.81	-17.52	0.30	0.93	0.56	64.76	1.00	0.99	0.00
	PERSIANN-CDR	0.79	-35.45	0.42	0.89	0.23	82.60	0.98	0.97	0.00
	TMPA 3B43 v7	0.88	2.56	0.28	1.08	0.69	50.86	1.00	1.00	0.00

In the *North-Central* and *Central* zones, the statistics that made major contribution to the variation were *CC*, E_{ff}, *RMSE*, *G* and *POD* as showed in Fig. 3.5b and Fig. 3.5c, respectively. For the North-Central zone (Fig. 3.5b), the TMPA 3B43 v7 and PERSIANN-CDR had the poorest fit and with the PERSIANN-CDR having the lowest fit in February. Collectively for these months, the precipitation datasets poor fit is shown by the high values of *MAE*, *bias*, *FBS* and *FAR*. The best fit was found during the period from May to August (mainly Winter season) for the second group across all the precipitation products. The third group had lower results particularly from September to April (Spring and Summer seasons). The Central zone (Fig. 3.5c) exhibited similar results than North-Central zone, with a lower fit in December and February for CHIRPS 2.0 and PERSIANN-CDR, and during January for all three products. Also the Central zone, shows that from June to September had the best fit mainly for PERSIANN-CDR and CHIRPS 2.0; and the highest values of *ME*, *bias*, *FBS*, *MAE* and *FAR* showing lower agreement from October to May, similar for the three products.

The indicators that were most relevant for the South-Central zone were *RMSE*, *G* and *POD*; and for the South zone were *MAE*, *RMSE*, *G* and *FAR* as presented in Fig. 3.5d and Fig. 3.5e. PERSIANN-CDR was grouped from November to January with high values of *MAE*, *FBS* and *FAR*; and low values of *bias* (indicating underestimation), *CC*, and E_{ff} showing the lowest agreement with in-situ measurements in South-Central zone (Fig. 3.5d). CHIRPS 2.0 and TMPA 3B43 v7 had better results from September to March with high values of *CC*, E_{ff}, *bias* and *ME*. The best results for all three precipitation data products ocurred from May to September. Finally, in the South zone (Fig. 3.5e) the major contributors were *MAE*, *RMSE*, *G* and *FAR*, and the cluster shows that PERSIANN-CDR had the lowest fit during all the year, similar was CHIRPS 2.0 in April, October and November. In this zone better result were found for CHIRPS 2.0 and TMPA 3B43 v7 with high values of *CC*, E_{ff}, *bias* (showing overestimation), *POD* and *FBS*, during the year.

Fig. 3.6 shows the analysis of the variation of the main statistics with respect to the amount of monthly rainfall. In the *North* and *North-Central* zones, low monthly precipitation below 5 mm had a clear impact on the linear correlation coefficient (CC), bias, *FAR* and *FBS*, compared to the other zones. In the zones where the days with rain are very limited, the ability of the satellite products to detect rain is reduced and often detected precipitation when there was no rain, which is reflected in high values of *FAR* that produce high values of *bias* and *FBS* too. Precipitation estimates were improved as shown in the statistical results as the rainfall increased, which was consistent across all zones. The *South-Central* and *South* statistics shows good agreement when compared to observed precipitation primarily because of the higher monthly rainfall

amount received over these areas. Further, there are considerably more days with rain than in the other zones, that has direct impact on the detection of rain, showing very high *POD* value of close to 1 and a *FAR* value close 0.

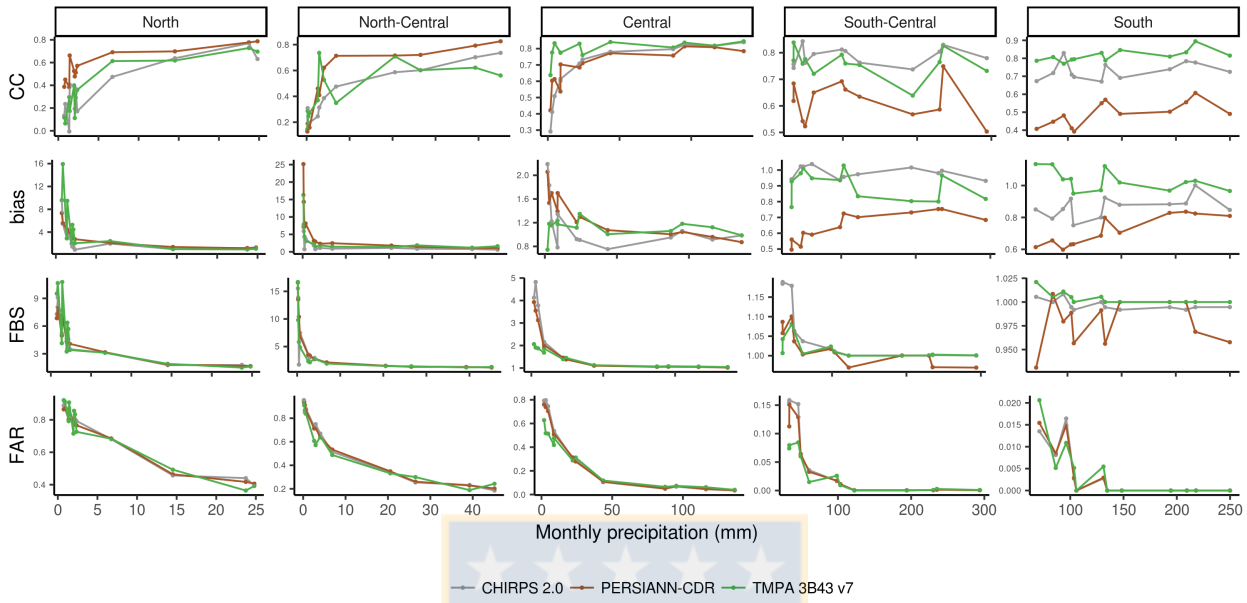

Figure 3.6: Variation of statistics linear coefficient of correlation (CC), bias, false-alarm ratio (FAR) and frequency bias (FBS), with the amount of monthly rainfall, for North, North-Central, Central, South-Central and South zones of Chile and products TMPA 3B43 v7, PERSIANN-CDR and CHIRPS 2.0

3.4.4 Spatial variation and comparison of products with long data-record

Spatial variations of the CC and E_{ff} for the 278 stations and the three satellite precipitation products across Chile are shown in Fig. 3.7 and compared with the spatial variation of in-situ rainfall (G). The values of CC are very high in central Chile, mainly over 0.78 (Fig. 3.7. The TMPA 3B43 v7 data had the highest correlation with values around 0.92 over this area with correlations decreasing in a northward direction for all products as seen in Fig. 3.7a. The E_{ff} was below 0.45 from 30° latitude and northward with values increasing from this latitude northward reaching peak value of 0.7 in central Chile (Fig. 3.7b) with similar values for all three satellite products. When these statistical results (CC and E_{ff}) are compared with the monthly average of precipitation (G) in weather stations, the spatial variation patterns in monthly rainfall are very similar, indicating that the fit of the precipitation products was strongly related to the amount of rainfall mainly for CHIRPS 2.0, followed by PERSIANN-CDR and TMPA 3B43 v7 with the greatest spatial pattern differences.

The climatology of monthly averaged precipitation maps from 1983 to 2015 were compared between the high spatial resolution, CHIRPS 2.0 and the coarse spatial resolution, PERSIANN-CDR in Fig. 3.8. Both products were compared to evaluate the difference on how well these products measure the spatial patterns variation of precipitation. TMPA 3B43 v7 was not included because its shorter historical data, which has limited utility for climatological analysis. As would be expected, the high resolution CHIRPS 2.0 data captured more spatial variability in precipitation patterns than the PERSIANN-CDR. Months with the greatest spatial pattern differences between these datasets occurred in February, March, September, October, November and December; particularly in the southern part of Chile. Most of the spatial differences in the precipitation patterns captured in these two products occurred mainly in the Central to South part of Chile. One of the notable discrepancies occurred in eastern Chile, where the Andes Mountains is located and few, meteorological stations are located at the higher altitudes. During May to August in eastern Chile near

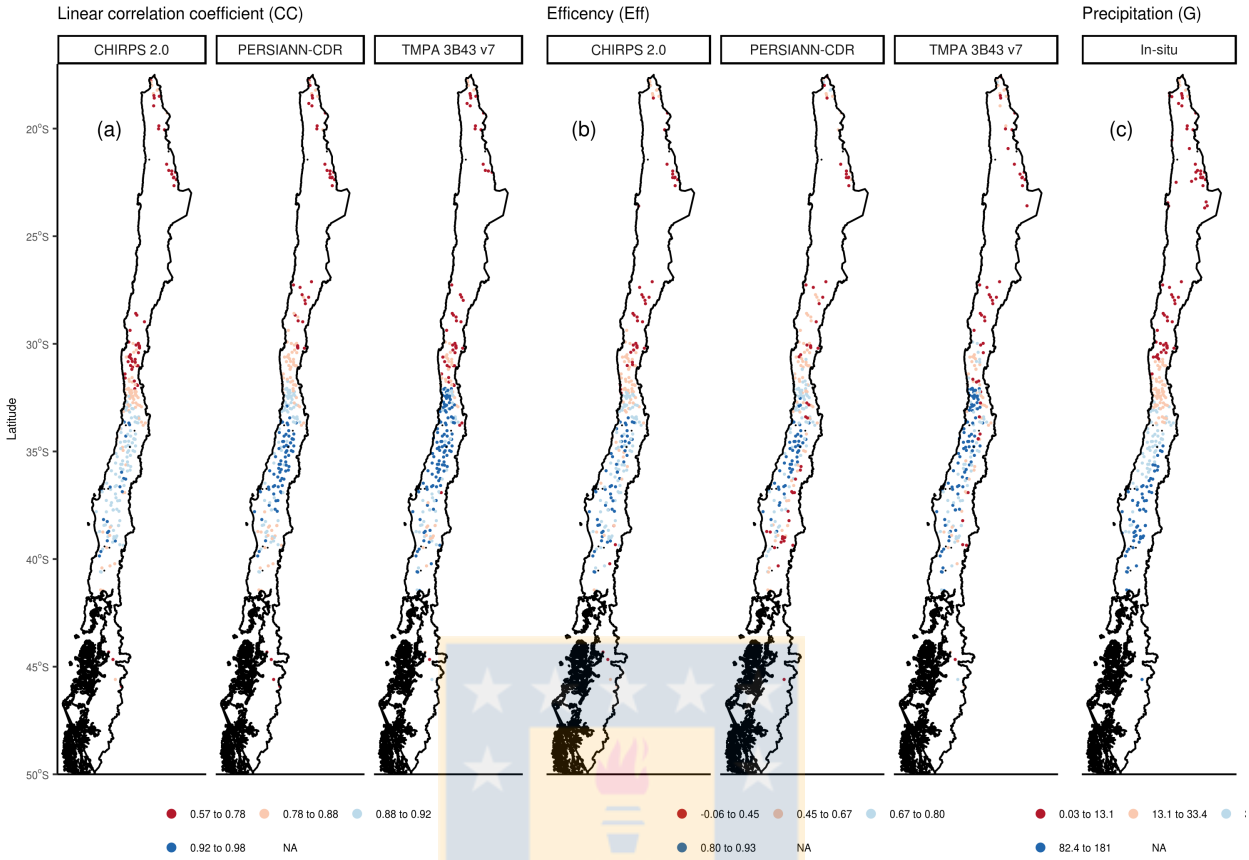

Figure 3.7: Spatial variation of statistics of (a) linear correlation coefficent (CC), (b) efficiency (Eff) for products CHIRPS 2.0, PERSIANN-CDR and TMPA 3B43 v7 and (c) In-situ precipitation in 278 rain gauge (G) over Chile

the 40° S latitude, precipitation can exceed 400 mm as estimated for CHIRPS 2.0, compared to 200 mm estimated by PERSIANN-CDR. In the South between 45° S and 50° S in far eastern Chile, CHIRPS 2.0 had lower precipitation values of less than 50 mm and PERSIANN-CDR estimated approximately 100 mm. It is well known that precipitation raises as altitude increases, further, Garreaud (2009) indicates that orographic air uplift produces 2-3 times more annual precipitation up in the Andes Mountains relative to the coastal values at the same latitudes. CHIRPS 2.0, captures the increasing of monthly precipitation climatology as showed in Fig. 3.8b during Winter months from May to July and toward the Andes Mountains between 35° and 45° South latitude. Fig. 3.8b shows that monthly rainfall raises up to 340 mm toward Andes Mountains and picking more than 550 mm during Winter. Those patterns were not captured by PERSIANN-CDR product as showed in Fig. 3.8a.

When the spatial variation of annual precipitation is considered for the maps presented in Fig. 3.9, together with the wettest month and driest month during the 30+ year historical period estimated from the satellite-based CHIRPS 2.0 and PERSIANN-CDR products, they were similar to the monthly averaged results in Fig. 3.8. The majority of differences occurred mainly south of 35° S and over the Andes Mountain range in far eastern Chile. The annual precipitation pattern showed in Fig. 3.9a is similar with those of Fig. 3.9b that shows the maximum monthly rainfall. This pattern reflects mainly the monthly rainfall from May to August, also shows that from 35° S and northern there is the driest pattern and in the other direction from 35° S to the South, there is a wettest pattern. Moreover, CHIRPS 2.0 capture precipitation increasing with elevation toward the east, which is not reflected by PERSIANN-CDR. Fig. 3.9c, minimum monthly rainfall reflects the driest pattern of Spring and Summer seasons as shown in Fig. 3.8 (September to December and

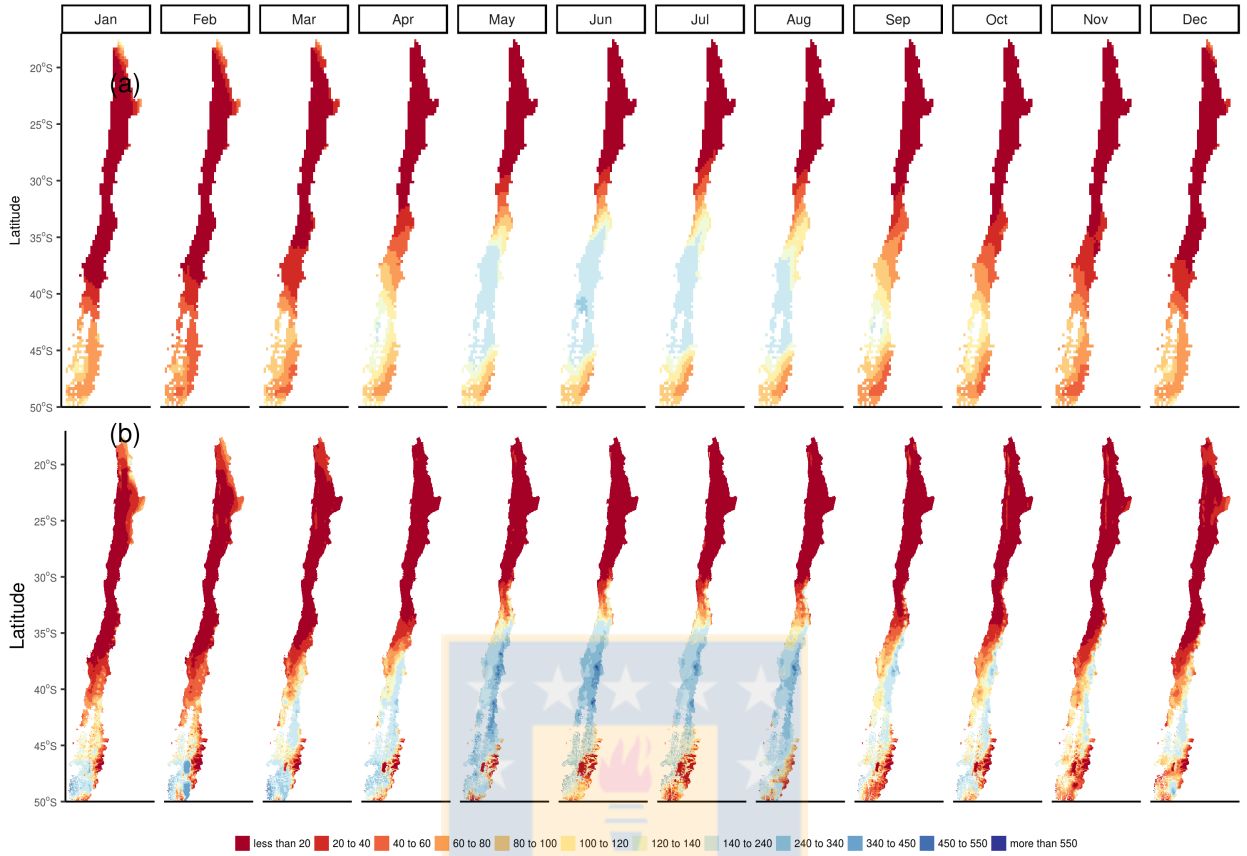

Figure 3.8: Monthly satellite-derived precipitation in Chile during 1983 to 2015 for products (a) PERSIANN-CDR and (b) CHIRPS 2.0.

January to March, respectively), and in this case, the driest pattern moves it to the 40° S latitude. The areas southern to 35° S had higher precipitation variability and the amount of monthly rain is also high as depict maps of annual precipitation (Fig. 3.9a), wettest month (Fig. 3.9b) and driest month (Fig. 3.9c). CHIRPS 2.0 was found to capture the spatial variation of rainfall better than PERSIANN-CDR, as reflected in Fig. 3.9a, Fig. 3.8b and 3.9c, where the spatial pattern of rainfall are highly variable specially in the East part of Chile as showed by CHIRPS 2.0. In Central to South zones in Chile, the PERSIANN-CDR showed more homogeneous spatial variability compared to CHIRPS 2.0, where the gradient of precipitation are more coarse due to the low spatial resolution of PERSIANN-CDR, this is readily apparent in Fig. 3.9a and Fig. 3.9b.

3.4.5 Application for agricultural drought analysis

In this section, historical time-series of SPI data (McKee et al., 1993) derived from both the PERSIANN-CDR and CHIRPS 2.0 precipitation datasets were evaluated. This index was selected because is recommended for the World Meteorological Organization (WMO) as an index to characterize droughts (Hayes et al., 2011). TMPA 3B43 v7 was not included because of its short historical record. The SPI at 3 time-scales (i.e., 1-, 3-, 6-month SPI) commonly associated to agricultural drought were produced and the correlation with SPI derived from weather station observations was evaluated. However, the analysis was mainly focused on the time-scales of three months or less because agricultural drought has generally short-term response. Further, in the South-Central zone of Chile, Zambrano et al. (2016) found the 3-month SPI (SPI-3) had the best correlation over cropland areas.

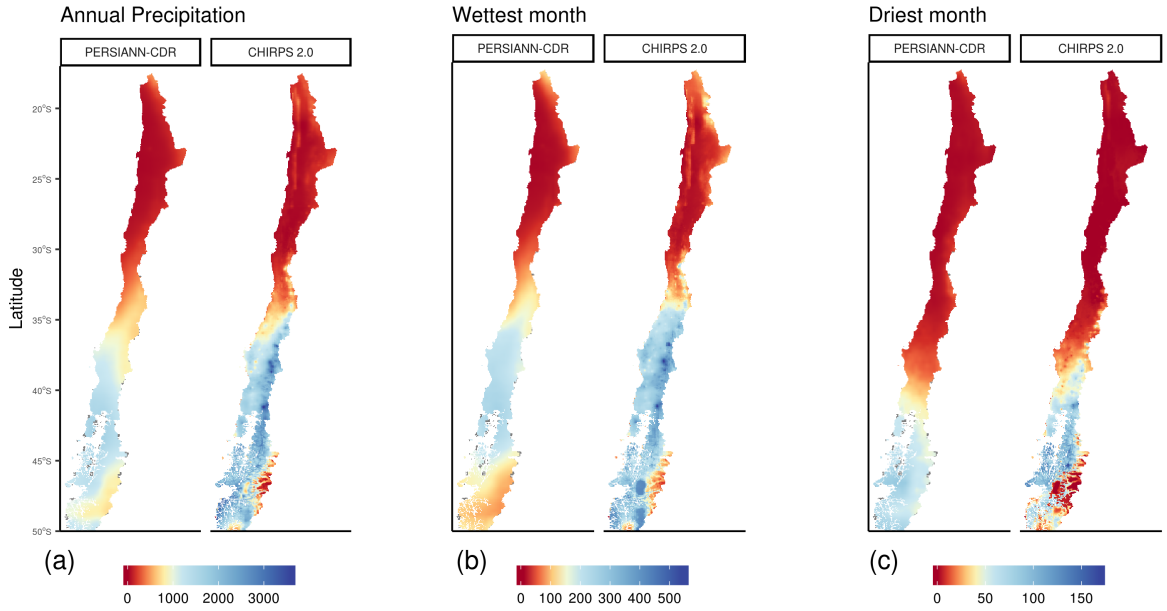

Figure 3.9: Rainfall climatology for period 1983 to 2015 of (a) Annual precipitation, (b) Wettest month and (c) Driest month, for satellite products CHIRPS 2.0 and PERSIANN-CDR

The SPI-3 times-series data is presented for the in-situ, PERSIANN-CDR and CHIRPS 2.0 precipitation data in Fig. 3.10. The time-series line graphs represent the geographically aggregated SPI values over each of the five zones. Also, the *RMSE* indicator was calculated to evaluate the error for derived SPI between satellite and in-situ values as showed in Fig. 3.11. From the results in the previous section, found the precipitation products had a better fit in the Central to South zones. Results in Fig. 3.11 supports these findings, showing that the lowest error was found in South-Central and Central zones for SPI-1, SPI-3, and SPI-6; and they were very similar for both satellite products. However, the lowest value of *RMSE* is 0.6 which is also high and could induce to find a place under drought when it is not.

The variation of SPI-3 during the growing season (September-April) from 1983 to 2015 were compared in Fig. 3.12, for the index data derived from the in-situ, PERSIANN-CDR and CHIRPS 2.0 data. In the North and North-Central zones, major difference are apparent between the satellite and in-situ derived SPI-3 data. Between 1988-1989 and 1996-1997 seasons, PERSIANN-CDR present 8 seasons with greatest differences, showing lower values than those from in-situ SPI-3. On the other hand, from 2001-2002 to 2010-2011 CHIRPS 2.0 has highest differences with in-situ SPI-3 values in 11 seasons in the North and North-Central zones. The SPI-3 bar plot calculated from the PERSIANN-CDR and CHIRPS 2.0 data for Central and southern zone had better fits with the in-situ SPI-3, which is consistent with the previous results presented. However, large discrepancies with in-situ SPI-3 were found for the PERSIANN-CDR-based SPI-3 data during 1989-1990 growing season. In 1989 the PERSIANN-CDR was found to largely underestimate precipitation as shown earlier (see Fig. 3.4), which resulted in the larger discrepancies of SPI-3 that year. In the Central, South-Central and South zones, the 1998-1999, 2007-2008, 2008-2009 and 2014-2015 growing seasons exhibited the lowest SPI-3 values that were calculated in each of the three SPI-derived from the in-situ, PERSIANN-CDR and CHIRPS 2.0, respectively. The SPI-3 results for these years are consistent with Zambrano et al. (2016), who found in the South-Central zone of Chile, severe drought occurrence during the 2007-2008, 2008-2009 and 2014-2015, note that the Zambrano et al. (2016) study did not evaluate the period 1981-1999 that low SPI values were calculated from the various precipitation datasets; however, the three later events were well represented in the SPI-3 time series over this area.

Because agricultural drought is associated with abnormal dryness over shorter-term time scales (< 6 months), and considering the previous results that found high monthly variability of adjustment between satellite products and in-situ data. Fig. 3.13 presents the monthly CC that compared the two satellite-based sets of

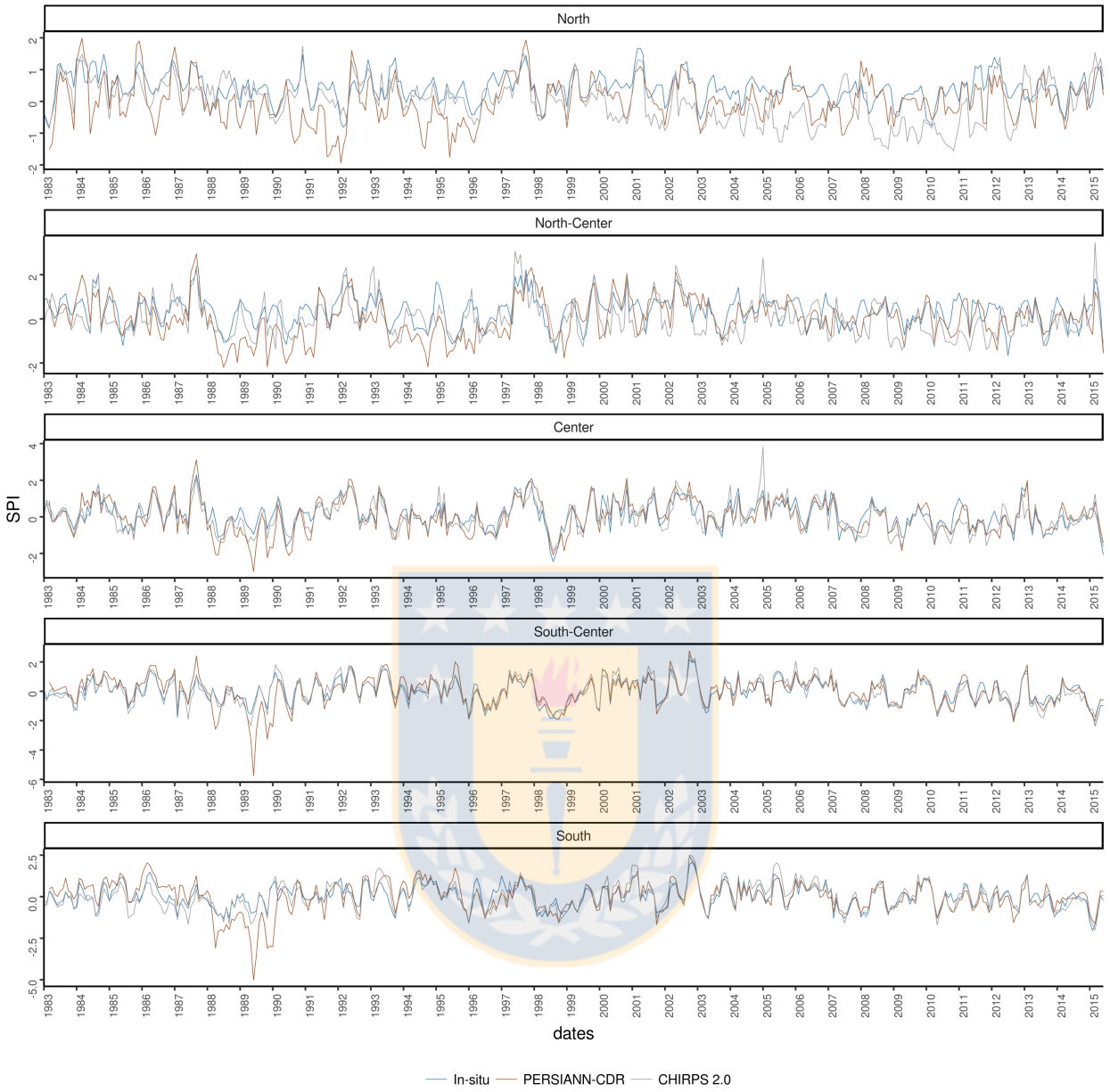

Figure 3.10: Time-series of SPI-3 for in-situ precipitation data and satellite products PERSIANN-CDR and CHIRPS 2.0 data with spatial aggregation for five zones of Chile: North, North-Central, Central, South-Central and South.

SPI datasets with the in-situ SPI data at time-scales of one, three and six months. For SPI-1, the results were very similar between the in-situ and satellite-based index results. Higher correlations among these datasets might be expected for the SPI-1 because calculation is a standardization of monthly precipitation, which was the time set of all three input datasets. In the North-Central and Central zones, the correlation improved as the time-scales of SPI lengthened from one month (SPI-1) to six months (SPI-6), with the greatest improvement in the last months in the year. That could be explained because as seen in section 3.4.3, the higher correlation of rainfall with CHIRPS 2.0 and PERSIANN-CDR were in winter months on the middle of the year (May to August), and this has an accumulated effect in the following months, which allow to improve the correlation on SPI as the time-scales increase until six months. This did not occur in

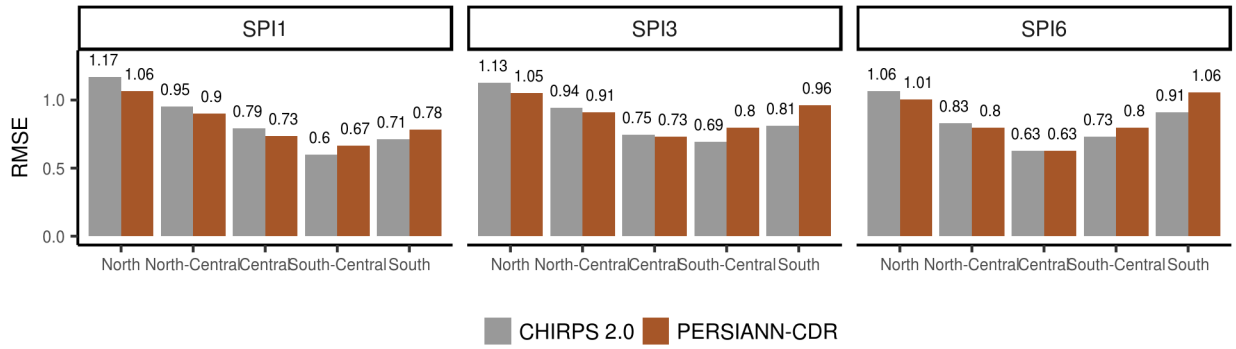

Figure 3.11: Comparison of RMSE of in-situ SPI with satellite derived SPI, for time-scales of 1, 3 and 6 months and zones North, North-Central, Central, South-Central and South

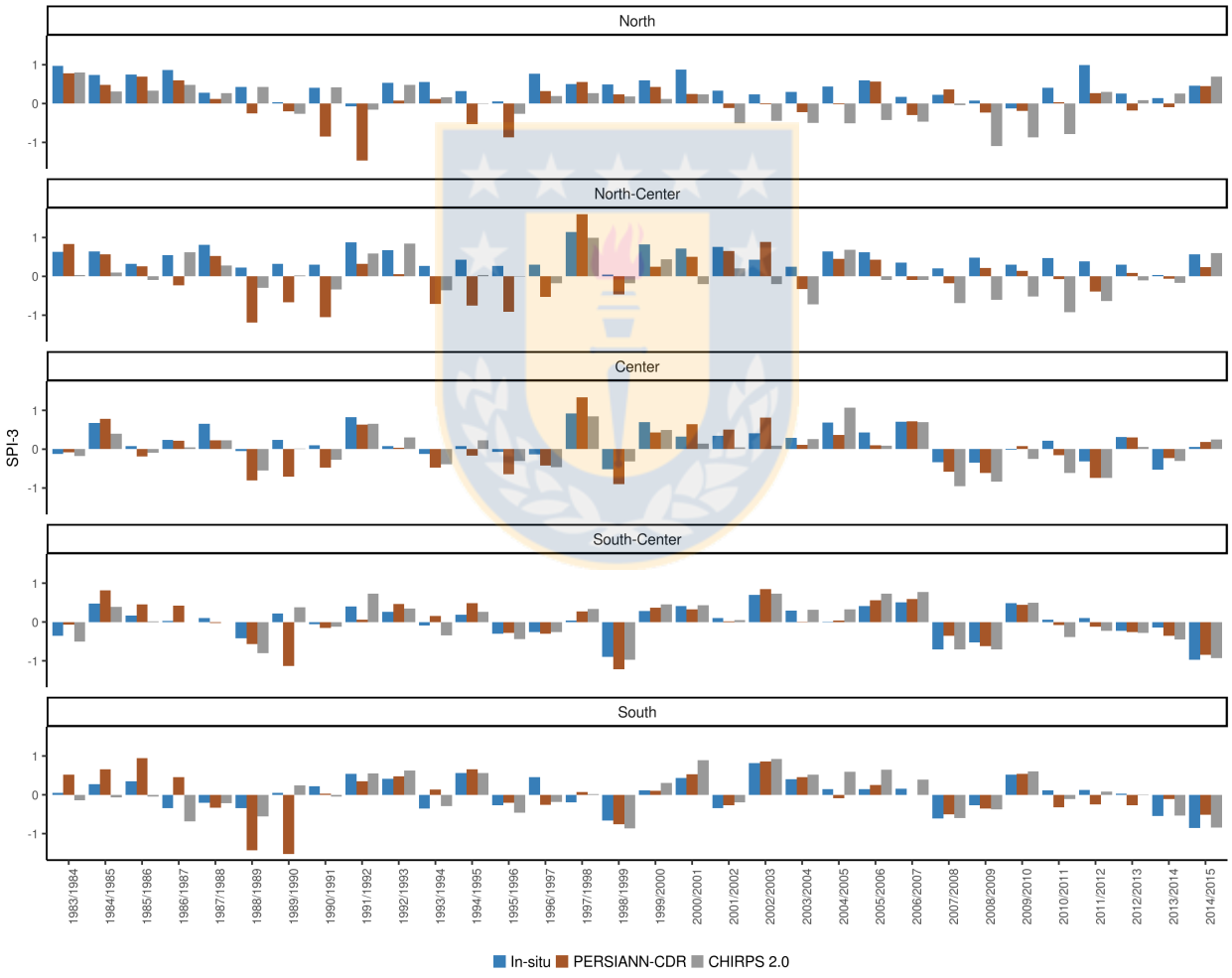

Figure 3.12: Averaged SPI-3 during growing season (September to April) spatially aggregated for zones North, North-Central, Central, Central-South and South; compared for in-situ, PERSIANN-CDR and CHIRPS 2.0 data for period 1983 to 2015.

the South and South-Central zones where the correlation does not have as high variability between the years, this may have been because in these zones the seasonality difference of rainfall between Winter and Summer is lower than in Central and North-Central zones.

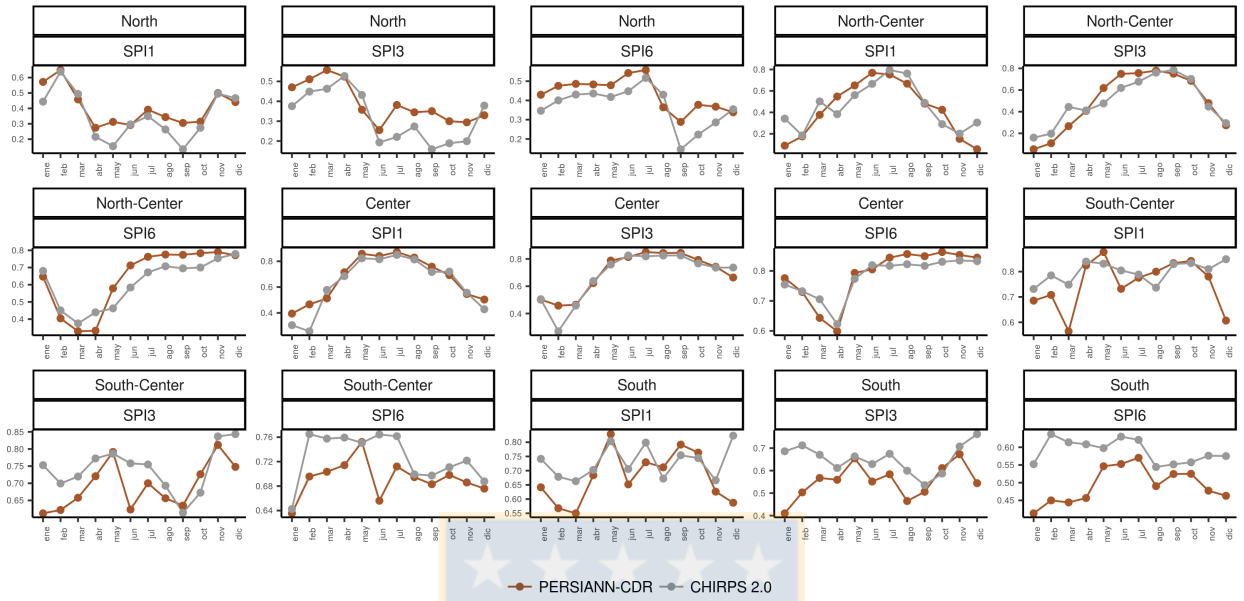

Figure 3.13: Monthly correlation of SPI-1, SPI-3 and SPI-6 between satellite products PERSIANN-CDR and CHIRPS 2.0 with SPI from in-situ data.

In Fig. 3.14, the mapping of SPI-3 derived from CHIRPS 2.0 is presented. CHIRPS 2.0 map results are presented and discussed in more detail here because this remote sensing dataset has a higher spatial resolution applicable for regional applications and had a better fit in the precipitation estimates with in-situ observations than PERSIANN-CDR. From the results presented in Fig. 3.12, the 1998-1999, 2007-2008, 2008-2009 and 2014-2015 growing season were selected as the four most severe drought events because of significant rainfall deficit over a three-month period (or longer) during the last thirty years in Chile. In Fig. 3.14, the large spatial extent of drought conditions is clearly identified during the 2007-2008 growing season compared to the other drought years. However, Fig. 3.12 shows that the CHIRPS 2.0 derived SPI products overestimated the severity of drought during the 2007-2008 and 2008-2009 growing seasons in North-Central and Central zones which is supported in Fig. 3.4 with the monthly precipitation difference with in-situ measurements. North-Central zone in 2007-2008 season has a in-situ SPI-3 of 0.2 against CHIRPS 2.0-derived SPI-3 of -0.68 and in 2008-2009 the in-situ SPI-3 has 0.49 against CHIRPS 2.0-derived SPI-3 of -0.6. Then, Central zone in 2007-2008 season has -0.33 and -0.95 for in-situ and CHIRPS 2.0 derived, SPI's, respectively; and 2008-2009 season has -0.35 and -0.84 for in-situ and CHIRPS 2.0-derived SPI's, respectively. The drought condition during the 2014-2015 growing season over the South-Central and South zones of Chile, are evident with SPI-3 values corresponding to *mild dry* to *moderate dry* conditions. The result for North and North-Central zone are unreliable due the higher discrepancies found in the precipitation estimated from CHIRPS 2.0 as reported earlier. For example, in Fig. 3.12, the North zone in 2007-2008 and 2008-2009 growing season had a high underestimation of SPI-3, similar results has the North-Central zone.

3.5 Conclusions

In this chapter, three satellite-based precipitation products with varying spatial resolutions (i.e., 0.05° and 0.25°), and extended historical data records (ranging from 18 to 30+ years) were evaluated for their accuracy of estimating the amount and spatial patterns of precipitation across Chile and their applicability for

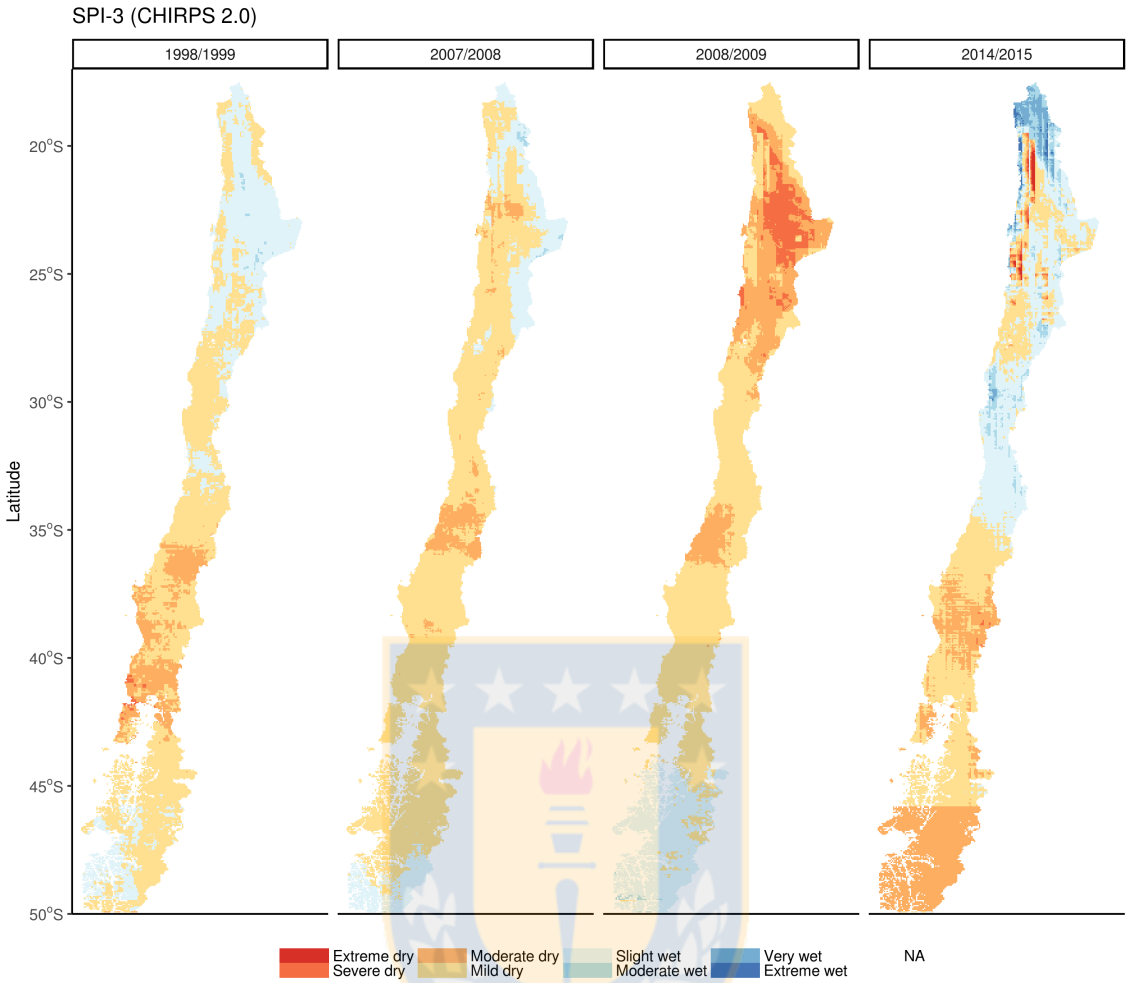

Figure 3.14: Maps of SPI-3 for the growing seasons (September-April) with the most severe values from 1983 to 2015, using CHIRPS 2.0 dataset

calculating the SPI to monitor agricultural drought. The precipitation and SPI estimates were compared to in-situ precipitation measurements and nine statistics indices were calculated for 278 selected weather station locations across Chile. In order to assess potential regional variations of precipitation values across Chile, the country was divided into five latitudinal regional zones based on climate.

Satellite precipitation products with long, 30+ year historical records such as PERSIANN-CDR and CHIRPS 2.0 when compared to the TMPA 3B43 v7 data with a shorter 17-year historical data record and station-based precipitation, for the five latitudinal zones with nine statistics, were found to have similar results from North to South-Central zones. However, differences exist in the South zone of Chile in the case of PERSIANN-CDR which highly underestimate rainfall.

The monthly statistics analysis, showed that for the North zone, the precipitation estimates were more accurate from December to March during the so called period of the 'Bolivian winter', when higher rainfall amounts occur. For the North-Central the precipitation results from satellite were more reliable specially during the Winter months (May to August) than in either the Summer (December to March) or Spring (September to December). The Central zone exhibited similar results than North zone, with the best agreement with in-situ measurement from June to September in the case of PERSIANN-CDR and CHIRPS 2.0; and in May, July, August for TMPA 3B43 v7 . The South-Central zone had the best results from May to September for the three products. In the South zone PERSIANN-CDR had the lowest fit and the highest was

achived by CHIRPS 2.0 and TMPA 3B43 v7, during the year. This was reflected in the results of hierarchical clustering, singular value decomposition and k-means, carry over the 9 elements of statistical comparison (i.e., CC, MAE, ME, bias, E_{ff}, FBD, POD, *FAR* and RMSE) for the three satellite products over the five zones and for each month of the year. Also, results showed that the clusters were made according with seasonality, where the months with higher monthly precipitation (Winter/Autumn) were grouped together and the satellite products had better fit with in-situ measurement.

The spatial variation pattern for CC and E_{ff} showed that the fit of the precipitation satellite products with ground measures was strongly related to the amount of monthly rainfall for the three satellite products and CHIRPS 2.0 showing the highest agreement. Also, results showed that in the North zone where the average monthly rainfall is less than 1 mm, the results were highly inaccurate. Further, during the so called 'Bolivian Winter' which is when the amount of rainfall is higher (wet season) the accuracy of satellite products improves. For the North-Central to southern zones, the accuracy of satellite products were more accurate from May to August (Autumn-Winter) when higher amounts of rainfall is received and the estimates were less accurate from September to April (Spring-Summer), which is a seasonally drier period. More detailed spatial rainfall patterns were captured by the higher spatial resolution CHIRPS 2.0 data than the PERSIANN-CDR data for most areas of Chile. Results in the higher altitudes locations of the Andes Mountains were difficult to evaluate because of the lack of in-situ measurements to validate the satelitte data.

The averaged time-series analysis by zone of the satellite-derived SPI-1, SPI-3 and SPI-6, showed that PERSIANN-CDR and CHIRPS 2.0 have similar results for each zone, and the lowest error was for the South-Central zone for SPI-1 and both products; South-Central for CHIRPS 2.0 and Central for PERSIANN-CDR, with SPI-3; and Central zone for SPI-6. Then, the averaged SPI-3 for zones and growing season, shows high discrepancies in the North and North-Central zones and better results in South-Central zone. Also, identified three most severe drought events between 1983-1984 and 2014-2015 growing seasons, these were 1998-1999, 2007-2008, 2008-2009 and 2014-2015.

Overall, the CHIRPS 2.0 precipitation dataset with its high spatial resolution ($0.05° \times 0.05°$) and long (+30 years) historical record was found to be a very useful dataset for characterizing precipitation patterns across Chile. It also provided a valuable data source to calculate a precipitation-based drought index like the SPI, which is commonly used to monitor drought. In this chapter, two satellite-derived datasets were tested for monitoring agricultural drought by transforming the precipitation data into the SPI over multiple time intervals (1, 3, and 6 months). The results of SPI analysis, particularly derived from the CHIRPS 2.0 data, were promising as the SPI-3 results identified drought events during the growing season that have occurred in Chile over the past thirty years. Most of the lowest SPI-3 values, which represent severe to extreme drought conditions, were in close agreement with the SPI-3 values calculated from ground-based rainfall measurements at most weather station locations across the country. The best performance of the SPI-3 calculations from CHIRPS 2.0 geographically occurred in the Central and South zones of Chile. SPI-3 results for North-Central and North zones were highly inaccurate, showing years in which the condition measure in rain gauges were wet and satellite-derived SPI-3 shows dry condition. The SPI-3 results derived from the CHIRPS 2.0 identified the key drought years of 1998-1999, 2007-2008, 2008-2009 and 2014-2015, particularly in the Central and South zones.

The long-record (more than 30 years) precipitation satellite datasets evaluated in this chapter, the PERSIANN-CDR and the CHIRPS 2.0, were found to be viable options for precipitation information for countries such as Chile, which have limited in-situ precipitation measurements both in number and historical length. These datasets provide alternative data options for the scientific community to improve and extend the hydro-meteorological models and analysis in data poor countries and regions of the world. For drought object, future research must be done to evaluate the use of these datasets to derive drought indices spatially distributed with a higher resolution than traditional based drought maps that are produced from the spatial interpolation of in situ measurements. Like the SPI used here, these satellite-based precipitation data sets could also be used to derive the Precipitation Condition Index (Du et al., 2013) and the Standardized Precipitation Evapotranspiration Index (Vicente-Serrano et al., 2010). Collectively, the increased use of these remotely sensed precipitation data sets has the potential to improve the drought monitoring and early warning tools in Chile and other parts of the world.

However, in order to use the PERSIANN-CDR and CHIRPS 2.0 data for monitoring drought in Chile, these product should be calibrated specifically for all months in the North zone and for the summer and spring months in the North-Central, Central, South-Central and South zones of Chile. To calibrate, and considering the higher correlations found between in-situ and satellite products different techniques could be used such as, regression models, bias correction or spatial interpolations like regression-kriging Hengl (2009). In order to use long-term precipitation satellite products CHIRPS 2.0 and PERSIANN-CDR in other countries and regions of the world similar studies need to be done to firstly assess its accuracy.

Appendix

3.5.1 Data processing and data analysis

Table 3.3: Packages from **R** environment used for the processing of the remote sensing data and the data analysis.

Procedure	Package
Remote sensing processing	*raster* (Hijmans, 2015)
Data analysis	*data.table, dplyr* (Dowle et al., 2015; Wickham and Francois, 2015)
Data visualization	*ggplot2* (Wickham, 2007)
Data transformation	*reshape2* (Wickham, 2007)

3.5.2 Statistics

To compare between in-situ data from measured in situ using the rain gauge (G) and the estimates from the satellite products (S), nine statistics were used, following Eq. (3.1) for magnitude of underestimation (ME), Eq. (3.2) for mean absolute error (MAE), Eq. (3.3) for multiplicative bias (bias), Eq. (3.4) for efficiency (E_{ff}), and Eq. (3.5) for the Root Mean Square Error (RMSE). These statistics evaluate the performance of the satellite products in estimating the amount of the rainfall (Dinku et al., 2009).

$$ME = \frac{1}{N}\sum(S-G) \tag{3.1}$$

$$MAE = \frac{\frac{1}{N}\sum(|S-G|)}{\bar{G}} \tag{3.2}$$

$$bias = \frac{\sum S}{\sum G} \tag{3.3}$$

$$E_{ff} = 1 - \frac{\sum (S-G)^2}{\sum (G-\bar{G})^2} \tag{3.4}$$

$$RMSE = \sqrt{\frac{1}{N}\sum(S-G)^2} \tag{3.5}$$

To evaluate the rainfall detection capabilities of the satellite products, several statistics were calculated including the frequency bias (FBS; Eq. (3.6)), probability of detection (POD; Eq. (3.7)) and false-alarm ratio (FAR; Eq. (3.8)). Table 3.4 shows a contingency table, where A, B, C and D represent hits, false alarms, misses, and correct negatives, respectively (Dinku et al., 2009).

Table 3.4: Contingency table for comparing rain gauge measurements and satellite rainfall estimates. The threshold correspond to the value above which rainfall is considered detected. In this case a value of 1mm was used.

	Gauge $\geq$ threshold	Gauge $<$ threshold
Satellite $\geq$ threshold	A	B
Satellite $<$ threshold	C	D

$$FBS = \frac{A+B}{A+C} \tag{3.6}$$

$$POD = \frac{A}{A+C} \tag{3.7}$$

$$FAR = \frac{B}{A+B} \tag{3.8}$$

Chapter 4

Agricultural drought prediction for Chile

Zambrano, F.; Vrieling, A.; Nelson, A.; Meroni, M. & Tadesse, T. Prediction of agricultural drought in Chile from multiple spatio-temporal data sources. Manuscript in preparation, 2017.

Abstract

Agricultural production in Chile is affected by drought. Climate projections predict that drought frequency and intensity may increase in many parts of the country. Early season forecasts about drought occurrence and severity could help to better mitigate its negative consequences. The objective of this chapter was to assess if agricultural drought in Chile can be accurately predicted from freely-available near real-time data sources. As the response variable, was used the z-score of cumulative NDVI (zcNDVI) based on 2000-2016 data from MODIS as a proxy for net primary productivity over the growing season. For each of the 758 census units considered, the response and predictor variables were averaged for agricultural areas resulting in a 16-year time series per unit for each variable. The prediction timing used was one, two, three, and four months before the unit-specific end of season (EOS). Predictor variables included zcNDVI (before EOS); standardized precipitation indices derived from satellite rainfall estimates for time-scales of one, three, six, twelve and twenty-four months; two climate indices being the Pacific Decadal Oscillation (PDO) and the Multivariate ENSO index (MEI); and latitude and longitude. Two prediction approaches were used: optimal linear regression (OLR) whereby for each census unit the single predictor was selected that provided the best relationship with the interannual zcNDVI variability, and a multi-layer feedforward neural network architecture often called deep learning (DL) where all predictors were combined for all units in a single spatio-temporal model. Both approaches were evaluated with a leave-one-year-out cross-validation procedure. Results for both methods showed similar and good prediction accuracy, with mean R^2cv values for OLR of 0.94, 0.79, 0.63 and 0.51, and for DL of 0.93, 0.79, 0.63 and 0.51, for one, two, three and four months before EOS respectively. Was discussed potential model improvements and how the method could contribute to an early warning system for agricultural drought in Chile and elsewhere

4.1 Introduction

Droughts result in major agricultural production losses worldwide (Campbell et al., 2016). The amplification of the hydrological cycle due to global warming is expected to lead to longer drought periods, even in regions for which overall precipitation increases are expected (IPCC, 2013). For Chile, a precipitation decrease is predicted for the Central-South part (IPCC, 2013) where agriculture activities are concentrated. Between

2010 and 2015, an unprecedentedly long period of relatively dry conditions persisted over Central Chile (30-38°) and has been termed a *mega drought* (Garreaud et al., 2017). Reports on economic impacts of drought are scarce in Chile, but Aldunce and González (2009) indicated that during the 1998 drought 30,000ha of wheat was completely lost, resulting in economic losses of about 31 million USD. Future climate scenarios project that wheat and maize yields will decrease by 5% to 20% by 2050 (IPCC, 2014; Meza and Silva, 2009), but impacts on agricultural production could be larger when taking the effects on cropping frequency and area into account (Cohn et al., 2016). Planning for effective adaptation strategies is crucial to mitigate future impacts (Roco et al., 2014). In addition, the ability to anticipate the impact of drought early in the season and take in-season mitigation measures could help to diminish crop losses (Pulwarty and Sivakumar, 2014; Wilhite et al., 2014, 2000).

Satellite image time series have been widely used for monitoring agricultural drought (AghaKouchak et al., 2015). Commonly-derived parameters from such time series include vegetation indices and rainfall estimates (RFEs), which can be translated into anomalies by comparing the parameters from the current year with the historic distribution (Ashouri et al., 2015; Funk et al., 2015; Huffman et al., 2007). The most commonly-used vegetation index for this purpose is the NDVI (Normalized Difference Vegetation Index; Rouse et al. (1974)) from which multiple anomaly measures have been derived (Kogan, 1990; Peters et al., 2002; Sandholt et al., 2002) and applied for monitoring agricultural drought (Cunha et al., 2015; Rojas et al., 2011; Zambrano et al., 2016; Zhang and Jia, 2013). An often-used approach to translate rainfall information into a drought measure is through the calculation of the multi-scale Standardized Precipitation Index (SPI; McKee et al. (1993)), an anomaly measure that when applied for short time scales (<9 months) is closely related to soil moisture availability (Quiring and Ganesh, 2010). While the SPI can be calculated from weather station data, in countries with a low station density and short historical records (like Chile), RFEs can be a good alternative source for SPI calculation (Tapiador et al., 2012; Zambrano et al., 2017). Other satellite-derived products that have relevance for drought monitoring include those that estimate soil moisture and evapotranspiration (Hao and AghaKouchak, 2013; Sheffield et al., 2004; Mu et al., 2013; Tsakiris et al., 2007). Drought indices are also constructed by combining multiple parameters. For example, the SPEI (Standardized Precipitation Evapotranspiration Index) considers both precipitation and evapotranspiration to account for the effects of temperature variability on drought assessment (Vicente-Serrano et al., 2010) and has been used in various studies for monitoring agricultural drought (Vicente-Serrano et al., 2012a; Moorhead et al., 2015; Potopová et al., 2015). While can be accurately assess and monitor agricultural drought as it occurs with a variety of indices, early prediction of drought is more complex.

The prediction of vegetation conditions in the near future is challenging for three reasons: 1) the underlying uncertainties in weather and climate prediction (Morssm et al., 2008); 2) changes in precipitation patterns (Dore, 2005), and; 3) the effect of both of these on vegetation growth (Sykes, 2001; Knapp et al., 2008) and land management decisions. Several studies have used a single predictor to explain interannual variability in seasonal vegetation productivity. Early prediction is thus achieved by using lagged relationships whereby the predictor is available before the end of the season. Meroni et al. (2014a) evaluated the probability of experiencing a seasonal biomass production deficit by examining the similarity between the current-year and historical fAPAR temporal profiles. Similarly, for East Africa, Vrieling et al. (2016) evaluated if the interannual variability of seasonal productivity, based on cumulative NDVI over the season as a proxy, can be accurately predicted before the end of season by cumulating NDVI over shorter periods. Alternatively, rainfall has been used as a predictor of seasonal vegetation productivity. For example, Meroni et al. (2017) analyzed when and to what extent the SPI derived from gridded RFE could explain anomalies of seasonal vegetation productivity in the Sahel, and found that on average about 40% of the variability in productivity could be explained by selecting, per-pixel, the optimal time-scale and timing of SPI. In addition, climatic oscillation indices, (e.g., Pacific Decadal Oscillation (PDO) and Multivariate ENSO Index (MEI)), have been shown to affect weather across the globe and as such can explain variability in agricultural productivity (Hansen et al., 1998; Reilly et al., 2003; Marj and Meijerink, 2011; Montecinos and Aceituno, 2003; Garreaud and Battisti, 1999; Boisier et al., 2016). Brown et al. (2010) demonstrated that the growing season and cumulative NDVI depended significantly on the PDO and the MEI across multiple locations in Africa. Van Leeuwen et al. (2013) showed how the MEI and Antarctic Oscillation (AAO) index explained interannual variability in annual productivity and phenology for South America, including Chile. While these studies assess the explanatory power of climatic oscillation indices on vegetation variability, they do not specifically address

the prediction of vegetation productivity shortfalls before they occur. Although studies focusing on single parameters offer interesting directions for prediction of vegetation productivity, combining multiple predictors could increase the predictive power.

The use of multiple predictors to estimate vegetation response has been evaluated by applying different techniques such as multiple linear regression models and regression trees. Using historical NDVI data, root-zone soil moisture and the ENSO index, Asoka and Mishra (2015) developed multiple linear regression models that could explain about 60% of the weekly NDVI anomalies one month before their occurrence, but the models were less accurate for two and three-month lags. They found NDVI to be the main predictor in their models, explaining most of the variability, but significant prediction power was added when incorporating satellite-derived soil moisture data and the Niño 3.4 climatic index. Tadesse et al. (2014),Tadesse et al. (2010) used a regression tree approach to predict standardized monthly NDVI anomalies from lagged NDVI data, several biophysical and climatic indices, as well as the SPI at the 3-month time scale. Their model explained between 50% (three months before) and 90% (one month before) of the NDVI variability in Ethiopia (Tadesse et al., 2014). More recently, machine learning methods have been widely used for predicting monthly and daily rainfall (Deo and Şahin, 2015; Abbot and Marohasy, 2014; Nastos et al., 2014; Abbot and Marohasy, 2012), mostly because they can accommodate a large number of input variables and non-linear relationships. Despite these apparent advantages for vegetation related studies, they have been most frequently applied for crop classification and estimation of crop biomass (Ali et al., 2015) and yield (Jia et al., 2013; Johnson et al., 2016; Kouadio et al., 2014; Panda et al., 2010). Given the mentioned advantages of machine-learning methods and the lack of current drought prediction tools in Chile, there is scope to evaluate if machine learning methods could provide more accurate and early prediction of drought than linear regression models.

The goal of this chapter is to assess if interannual variability in crop biomass productivity in Chile can be accurately predicted using freely-available near real-time data sources as input. These sources included NDVI time series, anomalies of cumulative rainfall at different monthly time-step derived from satellite RFEs, and climatic indices. As a proxy for seasonal crop biomass productivity (the explanatory variable) was took the cumulative NDVI over the season, aggregated for agricultural census units. First, predictions were made by an optimal linear regression (OLR) model which selected the best predictor for each unit. Second, was trained a feed-forward multi-layer neural network so called deep learning (DL) that combined all predictors to create a single model for all units. The two model predications were then compared using a cross-validation approach to assess the optimal approach for early prediction of agricultural drought in Chile.

4.2 Study area

Was focused on the main agricultural area of Chile which comprises about 90% of the cultivated land in the country (INE, 2007) between 29° and 41°S. The dominant climate is temperate Mediterranean (Csb) according to the Köppen climate classification system (Kottek et al., 2006; Peel et al., 2007). Despite of some crops in the coast, the main proportion of agriculture in Chile is located in the Central Valley, which corresponds to the depression between the Chilean Coastal Range and the Andes Mountains, with altitude ranging from 200m to 400m (Fig. 4.1a). Mean annual precipitation varies from less than 300mm in the North to 1800mm in the South (Fig. 4.1b). Cultivated land north from 32°40'S is used for fruit production (e.g, grapes, avocado, clementine), vineyards (e.g., mainly for 'pisco' production and a little for wine), and horticulture (e.g., lettuce, green bean, artichoke, corn, carrot) in the 'transversal' valleys that run east-west from the Andes to the Pacific. Fruits (e.g., walnut, grapes, avocado, plum, cherry, apple, hazelnut, blueberry, raspberry), vineyards, industrial crops (e.g., corn, rice, wheat, oats), and horticulture (e.g., corn, lettuce, onion, pumpkin, melon, carrot, watermelon, asparagus) dominate between 32°40' and 37°42'S. Further South, the land is mainly used for raising cattle for beef and dairy production, and includes croplands with cereals (e.g., wheat, oats, barley), and to a lesser extent fruits (e.g., hazelnut, apples, blueberry, cherry) (ODEPA, 2015).

As the spatial units for analysis, was chose the current census units that are used for the agricultural census (INE, 2007) carried out every 10 years by the Chilean government. This choice was motivated by several

reasons: 1) the units were large enough so that interannual variability in spatially-aggregated vegetation indices were not greatly affected by crop rotation signals as would be the case when assessing individual pixels; 2) the units were small enough to capture spatial variability in the main cropping systems; and 3) these units are smaller than the units currently used by the government for agricultural drought declaration. The study area contains a total of 2212 census units. The spatial distribution of their cropland percentage is shown in Fig. 4.1c, which indicates that the most intensively cultivated area is located between 34° and 37°S.

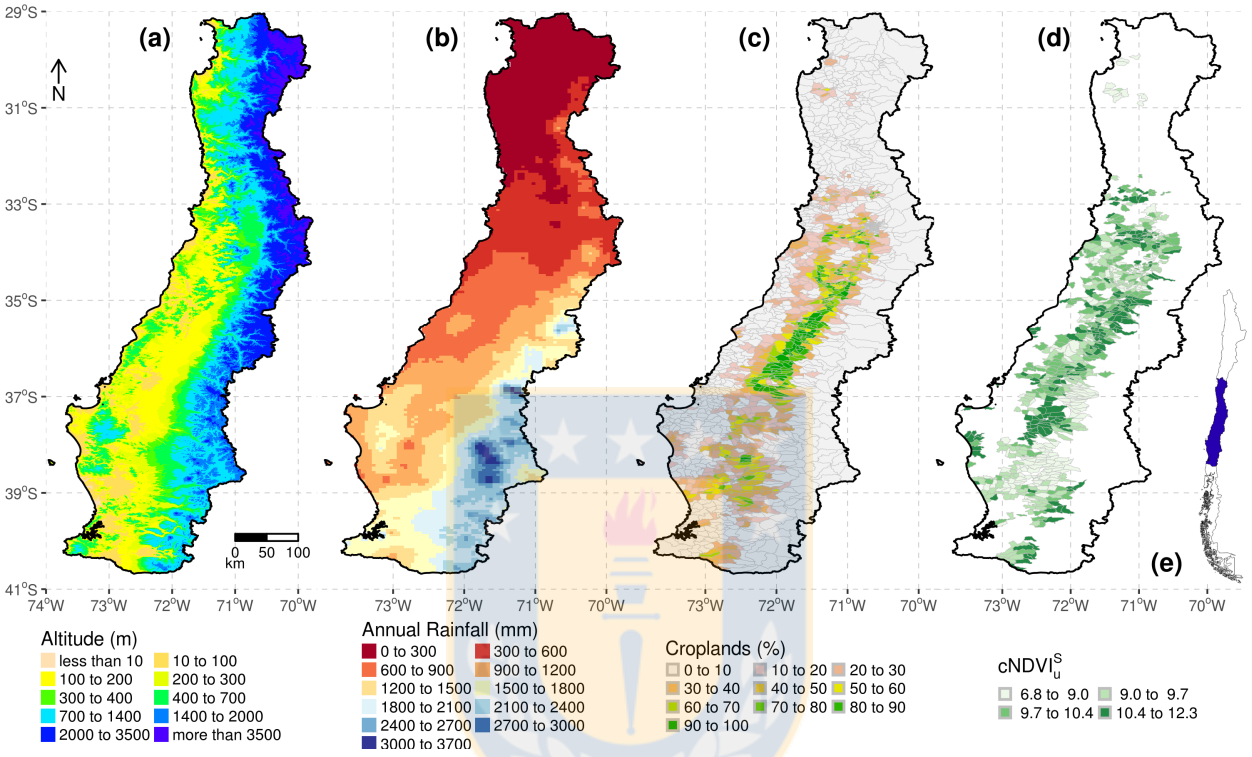

Figure 4.1: Study area with (a) elevation derived from SRTM, (b) annual rainfall derived from CHIRPS 2.0 (2000-2016 average), (c) percentage of cropland in each census unit from product MCD12Q1 scheme IGBP (2001-2013), (d) multi-annual (2000-2016) c average of NDVI cumulated over the growing season from product MOD13A1 v6, and (e) location of the study area within Chile.

4.3 Data

Data from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor were used to delimit the spatial extent of the main agricultural area of Chile, define the growing season dates, and to derive a proxy-measure of seasonal vegetation productivity anomalies as the explanatory variable. To derive a cropland mask over the study area was used the MCD12Q1 Collection 5.1 product (Friedl et al., 2010) that contains annual maps for 2001 to 2013 at 500m spatial resolution. From the MCD12Q2 product Collection 5 (Ganguly et al., 2010) was extracted start-, end-, and length-of-season (SOS, EOS, LOS, respectively), which is provided at annual intervals for 2001 to 2014 at 500m spatial resolution. To assess productivity anomalies, was used 500m resolution 16-day NDVI composites for 2000 to 2016 from the MOD13A1 Collection 6 product. To reduce remaining atmospheric noise, was smoothed the NDVI time series using a locally-weighted polynomial regression (lowess) as described by Zambrano et al. (2016). All MODIS data were obtained through the online Data Pool at the NASA Land Processes Distributed Active Archive Center (LP DAAC) and USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota.

Due to gaps in the historical station data and the low density of rain gauges, was used a satellite-derived RFE product for precipitation. Was selected the 0.05° spatial resolution, dekadal (10-day) RFE dataset of CHIRPS (Climate Hazards Group InfraRed Precipitation with Station, Funk et al. (2015)) version 2.0 for 2000 to 2016, which has recently been evaluated over Chile and shown to be a good alternative for raingauge rainfall measurements (Zambrano-Bigiarini et al., 2017; Zambrano et al., 2017).

Was used two climate oscillation indices: 1) the PDO which is often described as a long-lived El Niño-like pattern of Pacific climate variability (Zhang et al., 1997), and; 2) the MEI, which corresponds to the first component from a principal component analysis (PCA) based on six observed variables over the tropical Pacific (30°N–30°S, and 100°E–70°W), including sea-level pressure, zonal and meridional components of the surface wind, sea surface temperature, surface air temperature, and total cloud fraction of the sky (Wolter and Timlin, 2011). Both indices were provided by the National and Oceanic Atmospheric Administration (NOAA) at a monthly scale. (PDO and MEI datasets were obtained here and here)

4.4 Methods

4.4.1 Selection of census units

Because the key interest for our study is agricultural drought, was selected only census units with a predominance of croplands. Was selected the class *croplands* (class 12) from the IGBP (International Geosphere-Biosphere Programme) classification scheme contained in the MCD12Q1 data. A cropland mask was created from 14 years (2001-2014) of MCD12Q1 data (see Fig. 4.2). Per pixel, the percentage of years for which the IGBP classification scheme of MCD12Q1 indicated *cropland* was calculated as a measure of cropland intensity. Then was applied a threshold to this percentage to create a single cropland map for this study in order to consistently monitor the same areas. To choose the most appropriate threshold level that best describes actual cropland distribution, was selected 585 pixels of 500x500m resolution following a stratified random sampling over the study area. For each sample pixel, wasvisually interpreted high-resolution (<5m) imagery from Google Earth to identify agricultural fields and estimate the areal fraction of cropland (between 0% and 100%). Pixels with more than 50% cropland cover were identified as cropland pixels. The 585 cropland and no cropland samples were subsequently compared with different threshold levels of cropland intensity obtained from MODIS data. A threshold of 30% resulted in the highest global accuracy when compared with the Google Earth cropland samples (i.e., 78%), as well as when compared against a mask derived from data provided by the Ministry of Agriculture of Chile (INE, 2007) (i.e., 80%,). Given the high accuracy and the compatibility of its spatial resolution with our other input data, was decided to use the 30% intensity threshold as the cropland mask for the study.

Then was used the cropland mask to calculate the percentage of cropland area within each census unit (Fig. 4.1c). Was excluded census units from the analysis for which the cropland area did not reach 10% of its total area and/or was smaller than 750ha (30 grid cells) in order to keep only those units where agriculture is an important activity based on its spatial extent. This resulted in 785 census units.

4.4.2 Defining the growing season per census unit

To focus the analysis only on the period when crops are expected to be in the field, was estimated the average start of season (SOS_u), end of season (EOS_u) and length of season (LOS_u), for the cropland areas within each of the 785 selected units (u), using MCD12Q2 data for 2001 to 2014. To determine SOS_u was used the layer *Onset Greenness Increase* (GincO) and for EOS_u the *Onset Greenness Minimum* (GminO). Both layers have two bands, in principle allowing to identify two seasons, although was only extracted SOS and EOS for the single season that occurs in Chile, which in MCD12Q2 corresponded to the second layer for GincO and the first layer for GminO. Was then calculated the multi-annual average SOS and EOS for each cropland pixel that had at least seven years (50%) of valid data for 2001-2014. Across Chile, around 5% of the identified cropland pixels had less than seven years of retrieval, likely due to persistent clouds and

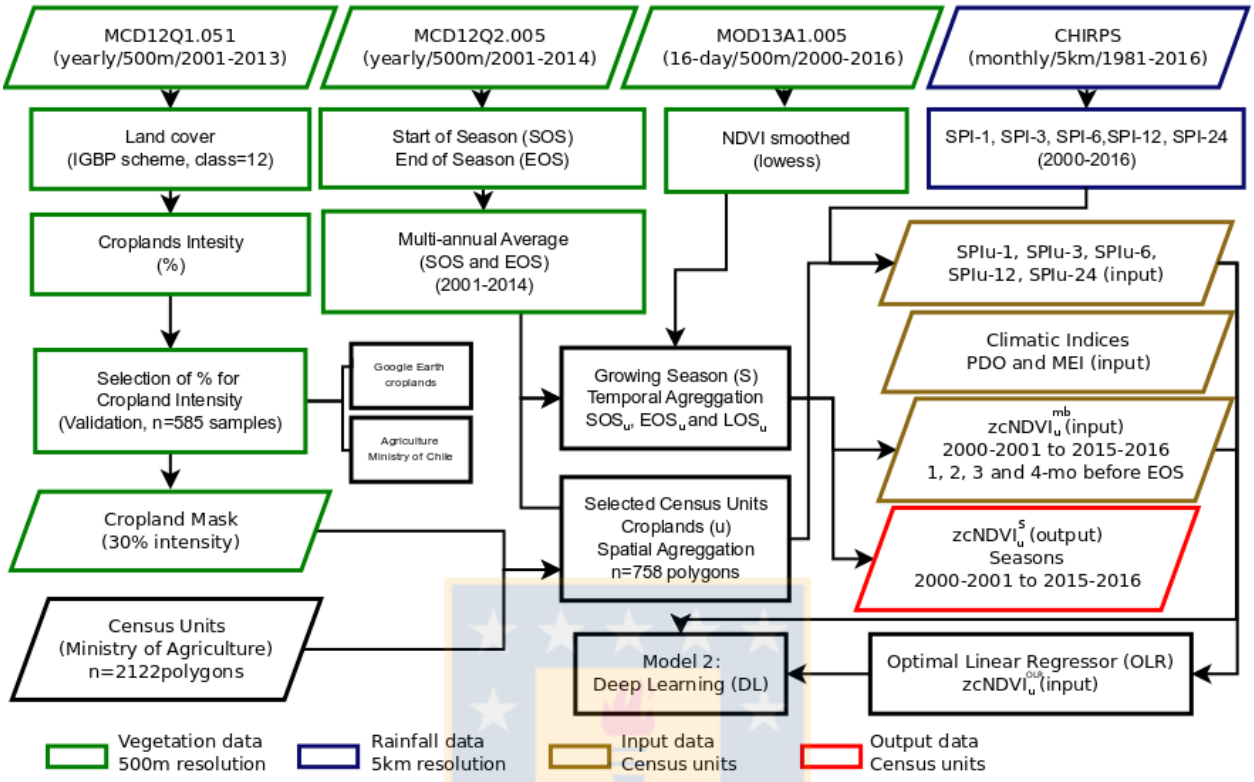

Figure 4.2: Methodological diagram to generate zcNDVI^S_u prediction models from vegetation and rainfall satellite data with climatic indices.

aerosols and possibly due to limited EVI variability in years of drought (Zhang et al., 2009). Finally, was calculated the SOS$_u$ and EOS$_u$ in each unit based on the median values for all crop pixels within each unit. LOS$_u$ was calculated as the difference between the SOS$_u$ and EOS$_u$ (Fig. 4.3). To match the SOS$_u$ with the 16-day time step of the MOD13A1 composites, the dates were rounded off to the nearest starting date, and for EOS$_u$ to the nearest end date of the 16-day period.

4.4.3 Deriving a proxy for seasonal crop biomass production

The seasonal biomass productivity of vegetation has been shown to have a strong relationship with the temporal integration of vegetation indices over the season (Jung et al., 2008; Rigge et al., 2013), and as such can be used to predict crop yields (e.g., Funk and Budde (2009)). To obtain a proxy for the net primary productivity of croplands was calculated the cumulative NDVI over the growing season (S) per unit and per year as:

$$cNDVI^S_u = \sum_{t=SOS_u}^{EOS_u} NDVI^t_u \tag{4.1}$$

Was chosen to work with NDVI rather than EVI, because: (1) was found temporal patterns to be similar and as a consequence expected very small differences in the results, and; (2) the NDVI is currently used for drought monitoring by the Ministry of Agriculture in Chile, which would likely facilitate further improvement and uptake of the approach presented in this paper.

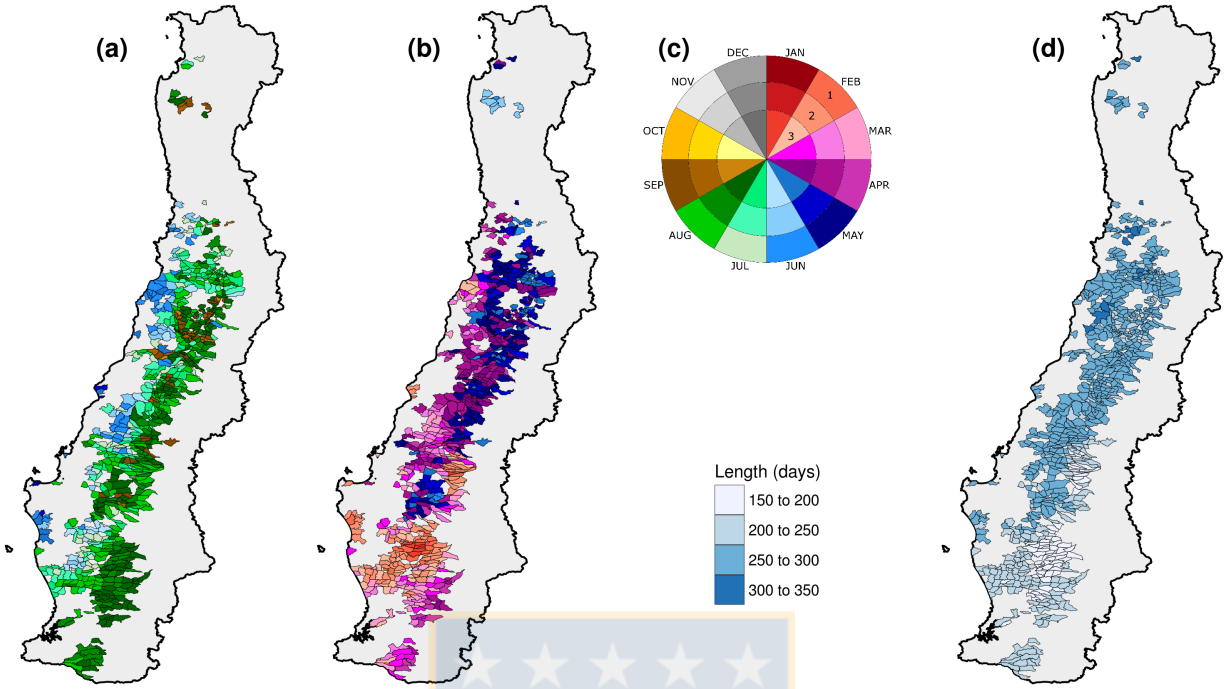

Figure 4.3: Growing Season (a) start date (SOS) and (b) end date (EOS) per census unit; (c) legend used for panels (a) and (b) following Vrieling et al (2016); the outer circle represents the first 10-day period of each month; d) growing season length per census unit. All were derived from product MCD12Q2.005.

Was transformed the into a standardized anomaly to reflect how much the cropland primary productivity is above or below normal. This can be written as:

$$zcNDVI_u^S = \frac{cNDVI_u^S - \overline{cNDVI_u^S}}{\sigma(cNDVI_u^S)} \tag{4.2}$$

where corresponds to the multi-annual average and $\overline{cNDVI_u^S}$ to the standard deviation between the 2000-2001 and 2015-2016 seasons for cNDVI$_u^S$.

4.4.4 Predictor variables

A number of predictor variables were selected based on previous studies that evaluated the predictability of drought. These include:

1. Z-scored cumulative NDVI, taking as the temporal integration period SOS until 1, 2, 3, and 4 months before (mb) EOS ($zcNDVI_u^{mb}$). Because the NDVI data are 16-day composites, effectively was analyzed the integration time periods as SOS until 32, 64, 96, and 128 days before EOS, but refer to this as *months before EOS (mb)* in this paper.
2. SPI at time-scales of 1, 3, 6, 12, and 24 months, each of which is calculated one to four month before EOS. Was refered to these as SPI$_u^{mb}$-T, with T being the time-scale and mb the month lag (1 to 4). Previous studies have shown that SPI can explain part of the variability in vegetation productivity, for example in the Sahel (Meroni et al., 2017) and in South-Central Chile (Zambrano et al., 2016). For this chapter was calculated SPI from dekadal CHIRPS data. Although the temporal aggregation was done for multiples of one month, the use of dekadal data allowed to more accurately capture the

timing of 1, 2, 3, and 4 months before EOS by pairing the EOS_u to the nearest dekad end date. For each dekad, was first calculated the cumulative rainfall over a 1, 3, 6, 12 and 24 month period before dekad end and subsequently transformed this into corresponding SPIs ending at day 10, 20, and last of each month. While short time scale SPIs (less than 9 months) are reported to relate strongest to agricultural drought (Quiring and Ganesh, 2010; Rhee et al., 2010), was included the 12- and 24-month time scales to account for a possible *memory effect* on vegetation response (Richard et al., 2008).

3. Climate oscillation indices. Following previous studies linking variability of precipitation and vegetation productivity to climate oscillations indices (see Section 4.1), was considered the MEI and the PDO oscillation indices. Was averaged these two indices for three non-overlapping time windows of three months, with a relative time lag (l) between them of 0, 3, and 6 months relative to EOS_u. Each of these resulting predictors were subsequently lagged in time to evaluate the predictive power from one to four months before EOS_u (mb). For example, for MEI can be write this as MEI_l^{mb}-3.
4. Latitude and longitude. To account for location-dependent relationships between and the predictors that would otherwise be missed in the DL model (which considers all units), was incorporated the latitude and longitude of centroid of each unit as additional predictors.

4.4.5 Prediction by OLR

Linear regression relationships were built for each unit with the purpose of finding the predictor variable that best explains the 16-year temporal variability of $zcNDVI_u^S$ per unit at one to four months before EOS_u. Optimal in OLR refers to the model with the smallest cross-validated root mean square error (RMSECV: see section 4.4.7) for the evaluated prediction timing (i.e., between one to four months before EOS_u). The 12 dependent variables included for each timing of prediction were (as described in section 4.4.4): SPI_u^{mb}-1, SPI_u^{mb}-3. SPI_u^{mb}-6, SPI_u^{mb}-12, SPI_u^{mb}-24, PDO_0^{mb}-3, PDO_3^{mb}-3, PDO_6^{mb}-3, MEI_0^{mb}-3, MEI_3^{mb}-3, MEI_6^{mb}-3 and $zcNDVI_u^{mb}$. Generalizing the notation of each predictor (P) per unit and timing to P_u^{mb}, the general linear equation can be expressed as:

$$\widehat{zcNDVI_u^S} = a \cdot P_u^{mb} + b \tag{4.3}$$

4.4.6 Prediction by DL

For each prediction timing mb, a single DL model comprising the same predictors and dependent variables for all units was trained and evaluated (LeCun et al., 2015). The DL model was implemented using the H2O platform, a scalable and open source machine learning and deep learning package that uses multi-layer feedforward (MLF) neural networks (Candel et al., 2016). MLF neural networks consist of neurons that are ordered into layers. Hidden layers connect input layers with an output layer, allowing for multiple connections with varying strength (Svozil et al, 1997). The connections are formed during the training of the network whereby the ultimate aim is to accurately predict the output layer. Having multiple levels of hidden layers and connections between them permits forming complex relationships between input and output (LeCun et al., 2015). Due to the complexity of the layer levels and their possible connections, the term *deep learning* is often used to describe such networks. Here it is describing how was trained the model, set its parameters, and how the importance of the predictors was evaluated. The reader is referred to Candel et al. (2016) and LeCun et al. (2015) for a more complete description of the MLF neural network achitecture.

DL models for each prediction timing were created for the entire study area by considering all units simultaneously. To train the model several parameters needed to be set such as "epochs" for training time, *activation* for the activation function, *hidden* for the number of hidden layers and amount of neurons, "rate" for the training rate, and others that allow to avoid overfiting such as *L1*,*L2* and *input dropout ratio*. For each prediction timing mb and validation year, the parameter values for the DL model were adjusted by the "random grid search" procedure (Candel et al., 2016; Bergstra and Bengio, 2012) to find the optimal DL model which produce the smaller mean square error. Thus, for each parameter are given a defined list

of option values (e.g, for rate were used 0.01 and 0.02) from which *random grid search* selected the optimal value given the combinated alternatives of values per parameter.

The Gedeon method (Gedeon, 1997) is used in H20 to assess the relative importance of each variable in each of the final DL models. This method uses the matrix of connection weights between layers of the trained neural network to determine which inputs are significant and thus account for variable importance. Was applied the Gedeon method to identify which variables had more significance in the final model at each prediction timing, allowing for comparison with the variables selected by the OLR method.

4.4.7 Model evaluation

To evaluate the accuracy of the predictions and compare between the OLR and the DL models, was used the R^2, RMSE and the mean absolute error (MAE). Was calculated these statistics using a leave-one-year-out cross-validation, which leaves out one year at the time allowing to compare the real observation ($zcNDVI_u^S$) with the corresponding value predicted independently from the remaining observation ($\widehat{zcNDVI_u^S}$). For example, for each census unit the $RMSE_u^{cv}$ is calculated as:

$$RMSE_u^{cv} = \sqrt{\frac{\sum (zcNDVI_u^S - \widehat{zcNDVI_u^S})^2}{n}} \tag{4.4}$$

The statistics were computed for both methods (OLR and DL) and each prediction timing (mb). Then was summarized the distribution of R_u^2cv and $RMSE_u^{cv}$ in boxplots.

To further analyze the spatial variability of the predictive power, was evaluated if the explained variability (R_u^2cv) was significantly different for units located in three zones 1) North, i.e. census units north of 32.4°S, 2) Center for units between 32.4°S and 37.42°S, and 3) South for those south of 37.42°S, given their different agro-climatic conditions (Section 2). In addition, was evaluated if season length affected the prediction accuracy by plotting R_u^2cv against LOS_u. Finally, was compared the time-series from 2000-2001 to 2015-2016 of $zcNDVI_u^S$ against the prediction ($\widehat{zcNDVI_u^S}$) made by OLR and DL in four individual characteristic census units. For this was made a four-cluster classification by k-means over the 758 units based on the multi-annual (2000-2016) average and standard deviation of $cNDVI_u$, the per-unit annual average precipitation, and LOS_u. From each cluster was randomly selected one census unit and plotted the observed $zcNDVI_u$ time series 2000-2001 to 2015-2016 seasons together with their predictions at 1, 2, 3, and 4 months before EOS for both prediction methods (OLR and DL).

4.5 Results

4.5.1 Census-specific definition of growing season

The analysis of phenology revealed a clear East-West pattern for SOS_u (Fig. 4.3a), with season start in May and June for units in the West, and predominantly July and August in the East. While a somewhat similar pattern can be observed for EOS_u, Fig. 4.3b also shows that in the South EOS_u is generally earlier (February/March) as compared to the North (April/May). As a result, shorter LOS values are found South of 38°S with less than 250 days and from here to the North the LOS have mostly more than 250 days (Fig. 4.3c). The isolated units in the extreme North part of the study area have a SOS_u of around August-September, an EOS_u between April and June and a LOS_u around 250-300 days.

4.5.2 OLR predictions

The average R_u^2cv calculated using the best predictors (minimum $RMSE_u^{cv}$) over the 758 census units was 0.94, 0.79, 0.63, and 0.51 for one, two, three, and four months before EOS_u, respectively (Fig. 4.5). As the

prediction timing approached four months before EOS the dispersion of R^2_ucv values becomes larger with the maximum interquartile range (0.49) and smallest mean (0.51) four months before EOS. The $RMSE^{cv}_u$ increases going from northern to southern units for mb between 2 and 4, while for mb=1 Center and South have a similar $RMSE^{cv}_u$ (Fig. 4.7c). Hence, census units located in the North presented higher accuracies for all prediction timings with average R^2_ucv > 0.90 and $RMSE^{cv}_u < 0.30$.

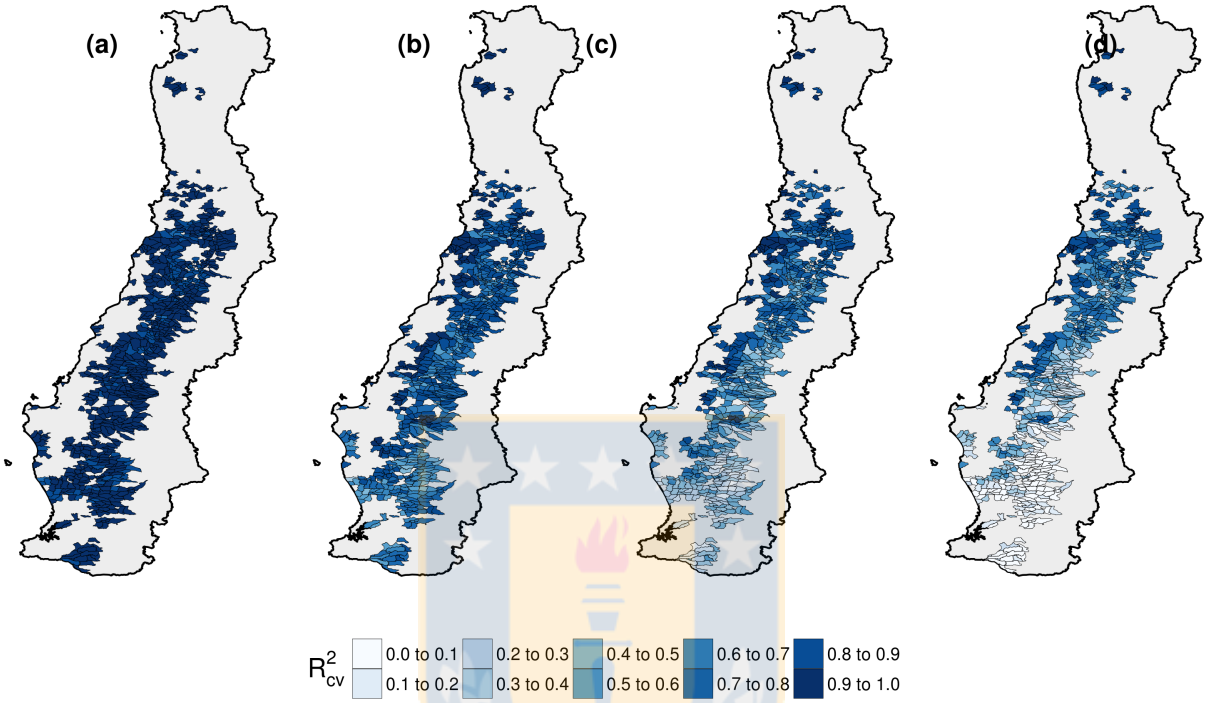

Figure 4.4: Interannual variance of season zcNDVI explained (R^2_ucv) when predicting by OLR considering per unit and time lag a single optimal predictor variable for: a) one month, b) two months, c) three months, and d) four months before EOS.

Fig. 4.5 shows that $zcNDVI^{mb}_u$ was most effective in predicting $zcNDVI^{S}_u$ across the 758 census units, explaining most of the variability and being selected for 100% of the units for one, 99.6% for two, 92% for three, and 84% for four months before EOS. Although, $zcNDVI^{mb}_u$ was the predictor selected most frequently for all prediction timings, for the South the SPI-1 was selected 48 times and SPI-3 11 times for predictions three months before EOS; for four months before EOS the best predictors after $zcNDVI^{mb}_u$ were SPI_u-1 (36 times), PDO_6-3 (21), SPI_u-12 (19), SPI_u-3 (16), and SPI_u-24 (13). However, for units where other predictors than $zcNDVI^{mb}_u$ were selected, the R^2_ucv was generally small and $RMSE^{cv}_u$ and MAE^{cv}_u large as compared to those where $zcNDVI^{mb}_u$ was selected.

4.5.3 DL predictions

Fig. 4.6 shows the spatial distribution of the R^2_ucv per unit for the four prediction timings. The average R^2_ucv calculated using the DL method over the 758 census units was 0.93 for one month, 0.79 for two months, 0.63 for three months, and 0.51 for four months before EOS (Fig. 4.7a). Accuracies were generally higher in the North zone (Fig. 4.6 and Fig. 4.7c). Accuracies decreased from North to South for prediction timings of two to four months before EOS (Fig. 4.7c).The highest accuracies were found for one month before EOS in the North zone with a R^2_ucv of 0.94 and $RMSE^{cv}_u$ of 0.25. For predictions four months before EOS, the smaller average R^2_ucv of 0.17 and larger $RMSE^{cv}_u$ of 1.02 were found in the South zone.

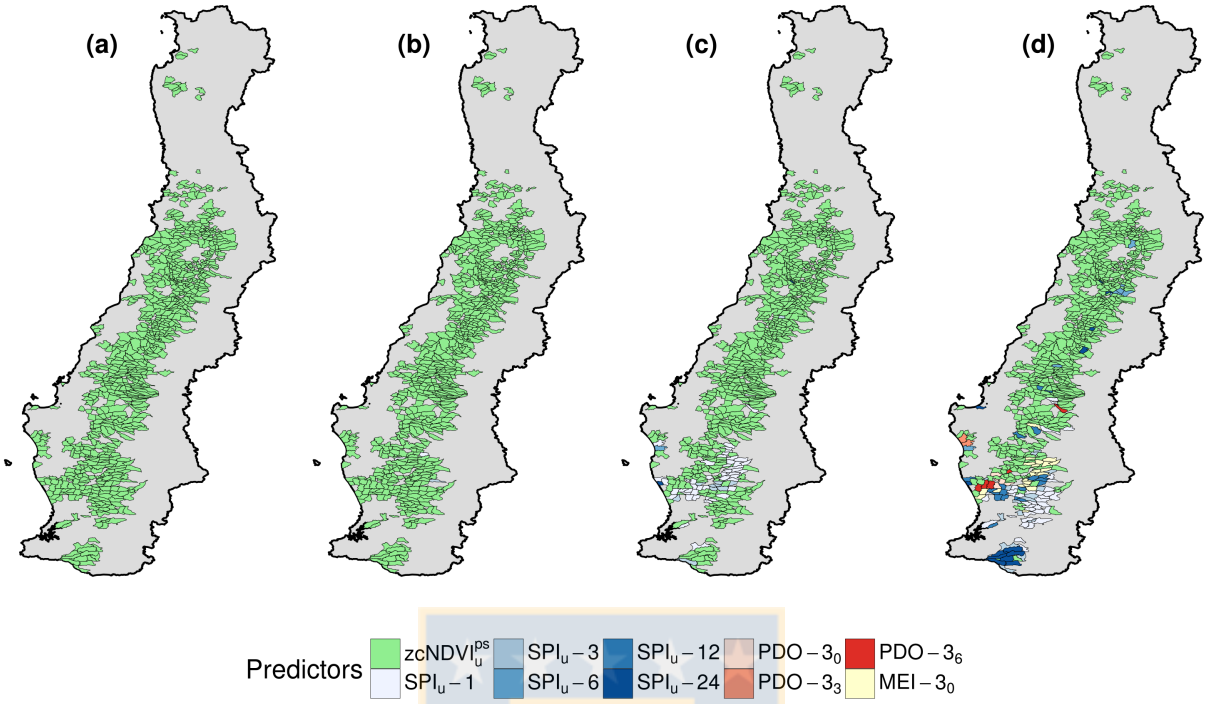

Figure 4.5: Best performing predictor based on the smallest RMSE$_u^{cv}$ when predicting the interannual variability of season zcNDVI by the OLR model considering per unit and time lag a single optimal predictor variable: a) for one month, b) two months, c) three months, and d) four months before EOS.

Table 4.1 presents the average relative importance of the predictors for the DL models. The most important variable for all prediction timings was zcNDVI$_u^{mb}$, followed by the climatic indices, PDO and MEI, mainly lagged at 3 and 6 months, though latitude was the third and second most important predictor for two and four months before EOS, respectively. The SPIs presented lower relative importance for the DL model for all prediction timings.

4.5.4 Model evaluation and comparison

On average over the 758 census units, the DL and OLR models had very similar R$_u^2$cv and RMSE$_u^{cv}$ values for all prediction timings (Fig. 4.7a). The average R$_u^2$cv difference both models was less than 0.014, and the average RMSE$_u^{cv}$ difference was less than 0.05 (Fig. 4.7a and Fig. 4.7b). For the North zone, OLR predictions were more accurate than those predicted by DL (Fig. 4.7c). Only for the South zone, for one month before EOS, OLR's RMSE$_u^{cv}$ was 12% larger than that of DL. The largest differences for the Center zone were found for one month (RMSE$_u^{cv}$ of DL 13% larger) and three months before EOS (17%). For the South zone, the largest difference between OLR and DL was for four months before EOS, with the RMSE$_u^{cv}$ for DL 14% larger than that of OLR (Fig. 4.7c).

Fig. 4.8 shows the time series of observed and predicted zcNDVI$_u^S$ for each growing season from 2000-2001 to 2015-2016, for the four selected units from the cluster analysis. The census unit *Pan de Azucar* is located in the most arid part of the study area and has a long season and a small; only irrigated crops are grown here. The *Tubul* unit represents units with intermediate levels of annual rainfall and, combined with a fairly long season (282 days in this case). The *Quintrilpe* unit has largest annual rainfall, but lower due to the shorter LOS. The last unit analyzed was Cocule, presenting large amount of annual rainfall and an average with shorter season (less than 25% of all units). Overall, both models represent reasonably well the temporal

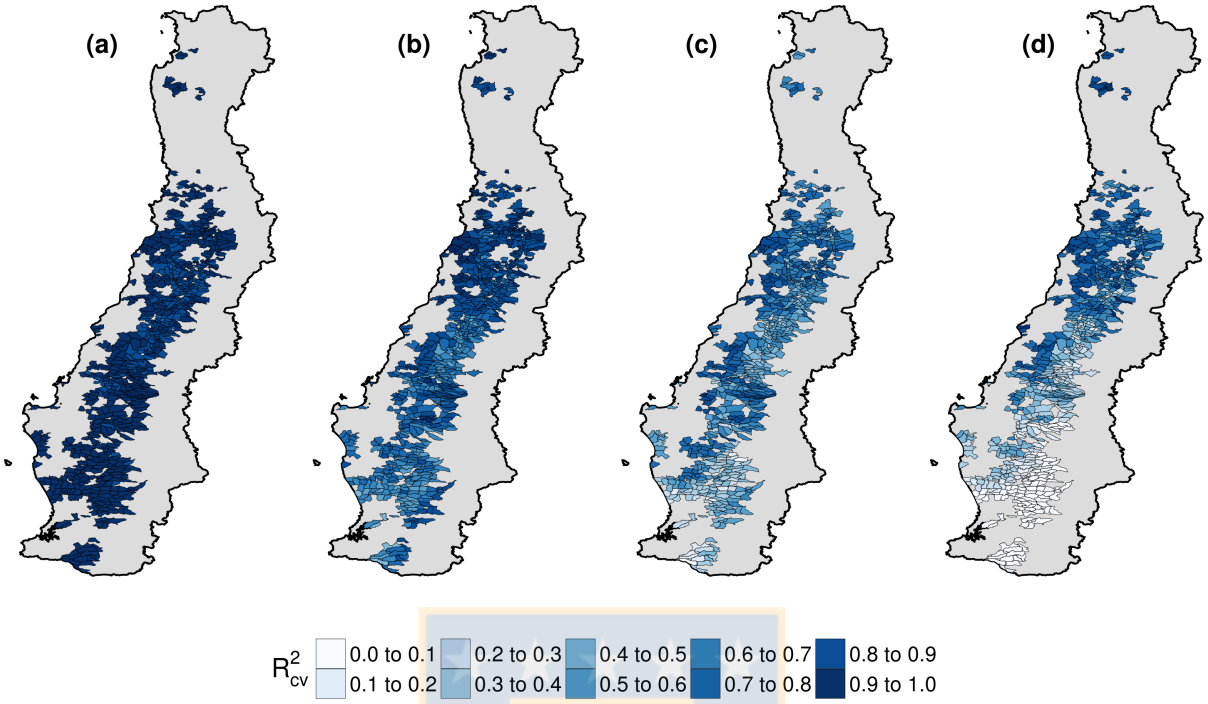

Figure 4.6: Interannual variance of season cNDVI explained (R^2_ucv) with the deep learning model using data: a) for one month, b) two months, c) three months, and d) four months before EOS.

pattern of zcNDVI^S_u, which in general was best predicted with gradual poorer performance towards four months before EOS. The predictions were most accurate for the unit *Pan de Azucar* with both models. The accuracy decreased going southwards via *Tubul* and *Quintrilpe* to *Cocule*. This latitude-dependency corresponds with the R^2_ucv pattern showed in Fig. 4.4 and Fig. 4.6. Predictions made short before EOS (mb=1) are in all cases better than those made longer before EOS. Out of these four units, OLR outperformed DL only for *Pan de Azucar* for all prediction timings, whereas for the other three units DL predictions were always more accurate for two to four months before EOS.

Fig. 4.9 plots unit-level R^2_ucv as a function of LOS$_u$ for OLR and DL methods. It shows that the accuracy of the predictions decreases for shorter LOS$_u$ and this effect becomes stronger as increase the number of months before EOS used for the prediction. The effect was similar for both models.

4.6 Discussion

This chapter shows that the selected proxy for seasonal productivity (zcNDVI^S_u) could be accurately predicted using a simple OLR approach. The more complex DL approach offered little to no advantage over OLR in this study. One explanation for this could be that the DL model attempted to fit a single model to a spatially heterogeneous study area. Further attempts to increase DL accuracies could include the clustering of the study area in groups of more homogeneous units (e.g. using the k-means applied in this study) and develop separate DL models for each, or alternatively, incorporating predictors that better reflect this heterogeneity (e.g., LOSu, or average annual rainfall). For example, Brown et al. (2008) used maps on soil, land cover, and elevation within neural networks in an attempt to relate multiple sources of NDVI data. As the addition of such predictors could account for differences in relationships between sub-groups, a similar approach could benefit the DL application of our study. Despite the good performance of OLR, potentially a principal component analysis (PCA) on all predictors per census unit could assist in better summarizing

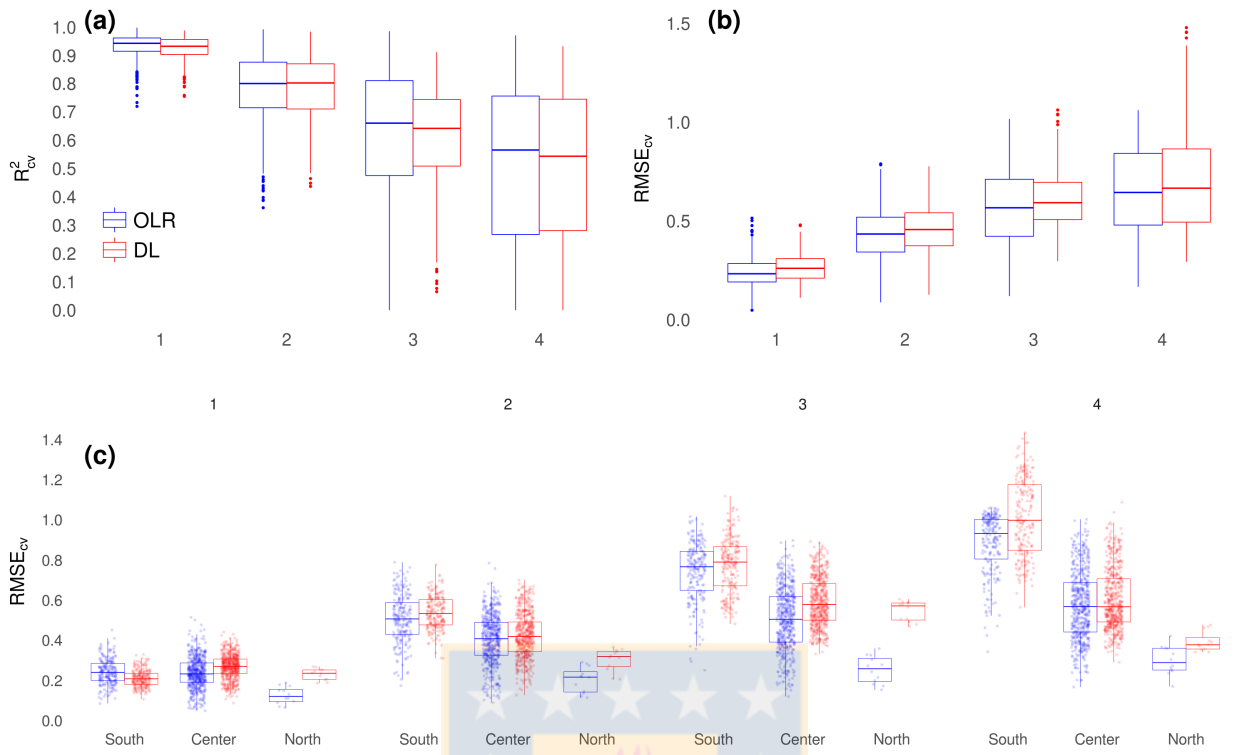

Figure 4.7: Comparison of accuracy measures between DL and OLR for the 758 census units by boxplots of: a) R^2_ucv, b) $RMSE^{cv}_u$; and c) spatial agreggation into North (29°-32.4° S), Center (32.4°-37.4° S) and South (37.4°-40.7° S) zones; for 1, 2, 3 and 4 months before EOS.

the temporal variability within each unit; the first PCA component could then be used in the OLR, even if the interpretability of the predictor variable may suffer from this.

Despite the potential for further improving our models, in this chapter, was reached an average prediction accuracy (R^2cv) between 0.51 and 0.94 depending on the timing of the prediction. In other studies about the prediction of cumulative vegetation indices through the growing season, Meroni et al. (2017) evaluated the per grid selection of SPI timescale to predict z-score cumulative FAPAR at different timing respect to the vegetation season's progress showing that SPI could explain about 40% of the variability. Also, Vrieling et al. (2016) evaluated prediction of cumulative NDVI over the season for forage insurance purpose by cumulating NDVI over shorter periods, and their results indicate that the R^2cv was above 0.75 for a large part of the study area. Other authors, study the prediction of monthly NDVI, in this regard Asoka and Mishra (2015) who used as predictor NDVI, soil moisture and seas surface temperature, obtained R^2 values between 0.14 predicting for July three months before, and 0.89 for November, one month before. Also, Tadesse et al. (2014) who used multiple predictor variables including SPI, landcover, digital elevation models, and four climatic oscillation indices, found that predicting monthly NDVI for September with 3 month before reached a R^2 of 0.50, and was larger improved by predicting one month before which had a R^2 of 0.9.

Despite the potential for further improvement, in this chapter was reached an average prediction accuracy (R^2) between 0.51 and 0.94 depending on the timing of the prediction. In other studies, Asoka and Mishra (2015) showed specific-month accuracies between 0.14 (in July with mb=3) and 0.89 (in November with mb=1). Also, Tadesse et al. (2014) found that predicting monthly NDVI for September with mb=3 reach a R^2 of 0.50, which is larger improved by predicting the NDVI values of September one month before EOS which had a R^2 of 0.9.

Was analyzed the drought prediction skill of $zcNDVI^S_u$ for the prediction timing of one, two, three, and four months before EOS at census unit level. However, as LOS varies between the units, the prediction timing

Table 4.1: Variable relative importance derived from deep learning models, mean ($\bar{X}$) for 16 models (2000-2001 and 2015-2016 growing seasons) for 1, 2, 3 and 4 month before EOS (mb).

Predictor	mb=1	mb=2	mb=3	mb=4
$zcNDVI_u^{mb}$	1.00	0.98	1.00	0.87
PDO_3^{mb}-3	0.58	0.51	0.56	0.55
MEI_3^{mb}-3	0.57	0.40	0.64	0.69
MEI_6^{mb}-3	0.57	0.64	0.61	0.63
PDO_6^{mb}-3	0.57	0.55	0.61	0.62
SPI_u^{mb}-1	0.57	0.28	0.47	0.34
SPI_u^{mb}-3	0.56	0.32	0.43	0.30
PDO_0^{mb}-3	0.56	0.54	0.57	0.57
SPI_u^{mb}-24	0.56	0.30	0.41	0.34
Latitude	0.56	0.56	0.53	0.71
Longitude	0.56	0.38	0.45	0.45
SPI_u^{mb}-12	0.55	0.33	0.41	0.33
SPI_u^{mb}-6	0.55	0.39	0.44	0.35
MEI_0^{mb}-3	0.54	0.51	0.59	0.67

could represent a different progress of the season. For example, for an LOS_u of eight months, already half of the season has passed at mb=4, so more information on that season performance is available as compared to an LOS_u of five months. For operational applications, a question is now whether it is more sensible to have a prediction at a fixed number of months before EOS, or alternatively if the prediction should relate to a relative measure of season progress. For example, Meroni et al. (2017) predicted zcNDVI from SPI for 0, 25, 50, and 75% of season progress. Was noted that in our current set-up, the DL model could have suffered from the mixing of areas with different levels of information on season performance (due to varying LOS_u) available. LOS decreases from North to South (Fig. 4.3c), following a similar pattern to the variability explained (R_u^2cv) for each model (Fig. 4.4 and 4.6), resulting in a strong relationship between LOS and R_u^2cv (Fig. 4.9). Although $zcNDVI_u^{mb}$ was in general the best predictor of $zcNDVI_u$, the prediction accuracy strongly depended on the prediction timing for both OLR and DL with poor predictive power early during the season.

In an attempt to further improve the prediction skill, additional spatio-temporal predictors could be incorporated in this study, many of which are being retrieved near real-time using satellite imagery as input. Examples include estimates of soil moisture, evapotranspiration, existing multi-scalar drought indices. For example, satellite-derived soil moisture estimates have been used to predict NDVI anomalies (Asoka and Mishra, 2015, Tadesse et al. (2014)). Asoka and Mishra (2015), showed that using soil moisture in addition to NDVI improves the prediction accuracy particularly when predicting early in the season. Another key variable relating to crop development is actual evapotranspiration, which can be estimated from satellite data. Currently operational evapotranspiration products exist at 1km spatial resolution (Mu et al., 2007) with a reported relatively good accuracy (R^2=0.70) over cropland, when compared to eddy covariance data of the United States (Velpuri et al., 2013). Another index that combines the effect of temperature and precipitation is the SPEI (Vicente-Serrano et al., 2010), which is currently spatially derived and available online (http://spei.csic.es/database.html) but has the potential for calculation at finer spatial resolution using satellite data (e.g., Mu et al. (2007), Funk et al. (2015), and Ashouri et al. (2015)).

In analogy to various other studies (e.g., Meroni et al. (2017) and Vrieling et al. (2016)), $zcNDVI_u^S$ was selected as a proxy for seasonal biomass productivity of cropland. However, due to a lack of accurate spatio-temporal data on crop production in Chile, it could not assess the accuracy of this proxy vis-à-vis actual production. In that regard it is recommended for the Ministry of Agriculture in Chile to put resources in place to collect such information on the basis of more regular and centrally organized collections of crop production information and store it in national publicly-accessible databases. This could further benefit the prediction of agricultural drought and improve Chile's resilience to future climatic shifts. Possibilities to

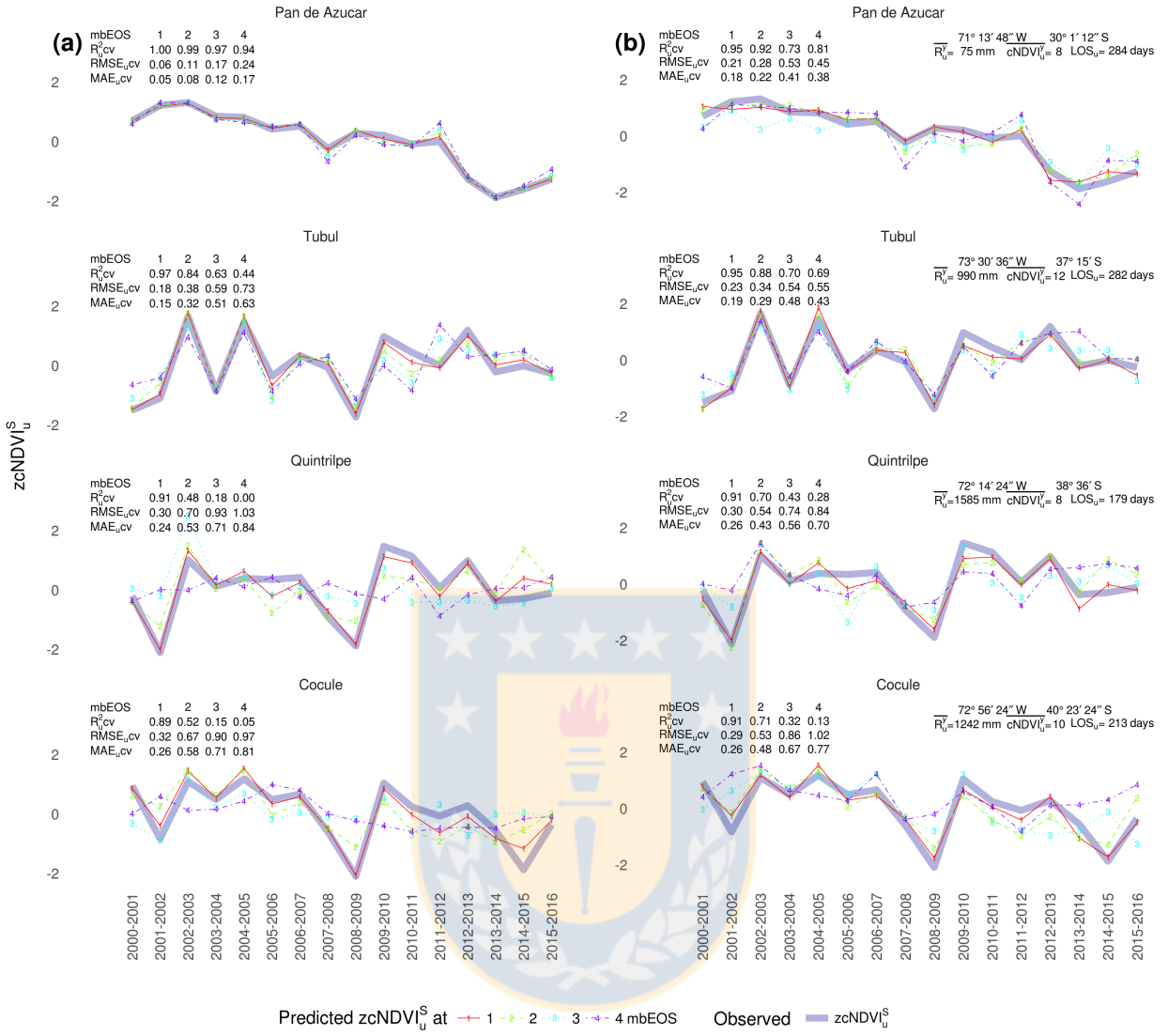

Figure 4.8: Comparison of the prediction of zcNDVI for four census units and for 1, 2, 3 and 4 months before EOS (mbEOS) obtained by a) optimal linear regression model (OLR), and b) deep learning model (DL); against the observed zcNDVI (tick blue line). In the top righ of (b) per each census unit are presented the location (Latitude and Longitude), average annual rainfall ($\overline{R^y_u}$), average cumulative ($\overline{cNDVIu}$), and the length of sesons (LOS_u).

collect such information include area-frame sampling in combination with crop cutting exercises or farmer- or food industry-reporting on yield and/or production.

In addition, was noted that the season definitions contained in the MCD12Q2 product may generally not be tailored to the precise season when crops are in the field. Because many of the retained MODIS pixels contain some level of non-agricultural vegetation, it is likely that such vegetation has an earlier SOS due to later emergence of crops, and delayed EOS due to crop harvesting before the other vegetation has come to complete senescence. The MCD12Q2 product determines SOS/EOS through curvature change of a modelled logistic EVI series, resulting in a relatively early SOS and late EOS (Zhang et al., 2006). Other agricultural drought studies have brought the EOS forward to better reflect the end of grain filling (Rojas et al., 2011). In this regard, spatially explicit crop calendar information for Chile may help to better define seasonality,

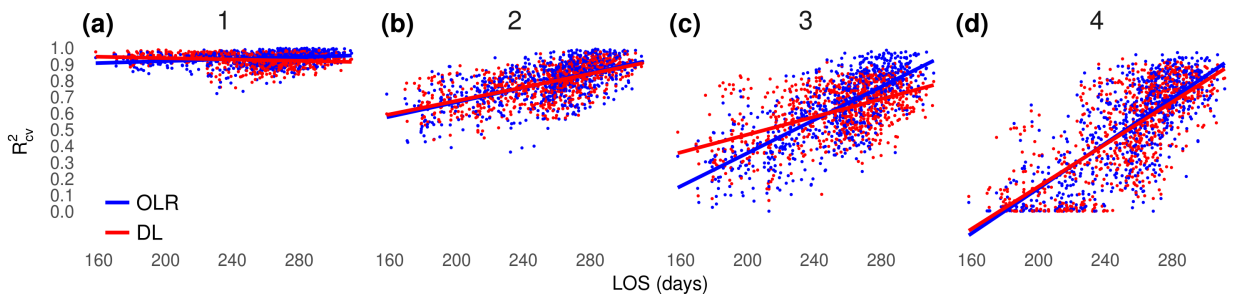

Figure 4.9: Variance of season zcNDVI explained (R^2_ucv) for the OLR and DL prediction models as function of LOS_u for: a) one month, b) two months, c) three months, and d) four months before EOS.

possibly allowing for 1) better proxies of agricultural productivity, and 2) more relevant timing of drought prediction.

The prediction models used in this chapter can assist in establishing an operational early warning system for Chile to achieve better preparedness for upcoming drought events. An operational system requires all input data to be available in near-real time. While this is generally true for NDVI (only a four-day delay after the last date of the compositing period), the CHIRPS data used here is only released during the third week of the following month. However, CHIRPS has a rapid response product available that is distributed two days after the end of each five-day period, though this is available for restricted areas (excluding Chile). The extension of this rapid response product to Chile or globally will improve the dataset's aptness for the envisaged warning system. While alternative precipitation datasets exist, their accuracy may vary (Derin et al., 2016) and would need to be evaluated over Chile before use. Although the climatic indices PDO an MEI have the same issue of a one-month delay, in this chapter (4.5.3) was showed that climatic indices with lag times of three and six months are most relevant for our predictions, and so this delay not an issue. While a few challenges remain, the models presented in this study can constitute a solid basis for improved operational drought warnings by the Chilean government.

4.7 Conclusions

Anticipating agricultural drought is major research challenge. In this chapter, was showed that publicly-available satellite data and OLR had high skill in the prediction of agricultural drought in a country where agricultural drought has been barely studied. The global coverage of the data supports the testing of the method in other drought prone regions.

Was demonstrated that cumulative NDVI during the growing season, as a proxy for crop productivity, could be accurately predicted with a similar performance of OLR and DL models. This accuracy decreased from north to south, particularly during predictions made earlier in the season (i.e., four months before EOS). This gradient matches the gradient in season length, which could be explained by the fact that more information on the season performance was available for units with a longer LOS at the timing of prediction (one to four months before EOS). Although scope exists to further improve the prediction models, the drought prediction tool developed in this chapter constitutes an important step for Chile for increased drought preparedness.

Chapter 5

Conclusions

Through this thesis has been analyzed the use of multiple time-series of satellite data from last decades, for the monitoring and prediction of agricultural drought. The country of Chile was selected as study area, mostly because its heterogeneity in climatic and agricultural condition, and also due to the increased droughts of last years and the climatic variation forecasted.

The overall achievement of this thesis as response to the hypothesis stated in Section 1.1 has been the assessment that satellite data is a good source that allows characterizing agricultural drought by defining the cropland area and growing season, and also for the deriving of a reliable proxy of agricultural drought. This proxy could be anticipated time before the end of the season using satellite-derived vegetation and precipitation, together with climatic oscillation indices. Also, was showed that the prediction accuracy decrease as the prediction timing going farther from EOS, and this accuracy is highly dependent on the length of the season. Also, in Section 4 was discussed that the models used has the potential to be improved further.

Based on the findings in chapter 2, was concluded that VCI is useful for monitoring agricultural drought in the BioBío Region and is closely correlated with SPI-3 during the modified growing season (Nov. to Apr.), which indicate that rainfall deficit beginning on September it is when has a larger impact on vegetation health, this would be related with crops types in the region, what it should be evaluated in future studies. This makes it a relevant indicator for agricultural drought monitoring and response plans. Further research is needed to associate the remote sensing values observed at high resolution (250m) with the measured crop yield (Seiler et al., 2007) and individually identify more detailed crop responses. This identification will gradually construct an effective drought management tool for the agricultural sector in Chile.

From Chapter 3, was concluded that the long-record (more than 30 years) precipitation satellite datasets evaluated in this chapter, the PERSIANN-CDR and the CHIRPS 2.0, were found to be viable options for precipitation information for countries such as Chile, which have limited in-situ precipitation measurements both in number and historical length. These datasets provide alternative data options for the scientific community to improve and extend the hydro-meteorological models and analysis in data poor countries and regions of the world. For drought object, future research must be done to evaluate the use of these datasets to derive drought indices spatially distributed with a higher resolution than traditional based drought maps that are produced from the spatial interpolation of in situ measurements. Like the SPI used here, these satellite-based precipitation data sets could also be used to derive the Precipitation Condition Index (Du et al., 2013) and the Standardized Precipitation Evapotranspiration Index (Vicente-Serrano et al., 2010). Collectively, the increased use of these remotely sensed precipitation data sets has the potential to improve the drought monitoring and early warning tools in Chile and other parts of the world.

In Chapter 4, was demonstrated that cumulative NDVI during the growing season, as a proxy for crop productivity, could be accurately predicted with a similar performance of OLR and DL models. This accuracy decreased from north to south, particularly during predictions made earlier in the season (i.e., four months before EOS). This gradient matches the gradient in season length, which could be explained by the fact

that more information on the season performance was available for units with a longer LOS at the timing of prediction (one to four months before EOS). Although scope exists to further improve the prediction models, the drought prediction tool developed in this chapter constitutes an important step for Chile for increased drought preparedness.

Finally, the development of satellite products are increasing heavily over time, and the quality of them is also improved. Nowadays, there are a huge amount of satellite data which could be used for the study of climate impact on vegetation at scale, both in spatial extension as well as temporal. Many variables are being derived from the satellite such as the ones used here for vegetation and precipitation, but also there other such as soil moisture, evapotranspiration, snow, groundwater, and a few more. However, for the processing and analysis of such amount of data new technologies has been developed in the field of *Data Science* that highly facilitate the exploration toward new insights, that otherwise would make it hard to achieve.

Appendix A

Other useful satellite data

Abstract

Agricultural drought in Chile has been studied using the Normalized Difference Vegetation Index (NDVI) and the Vegetation Condition Index (VCI), derived from NDVI. In addition, satellite-derived Standardized Precipitation Index (SPI) has been used as an indirect measure. However, there are other remote sensing data, such as Land Surface Temperature (LST), Phenology Dynamics, Evapotranspiration (ET), and Soil Moisture (SM), which provide useful information to carry out more detailed analyses. Moreover, a main challenge to improve scientific findings is to collect data about crops such as type, seasonality, and yield, in order to generate a national database, which along with satellite indices, will be essential to develop spatial crop models; thereby, broadening knowledge regarding the impact of agricultural drought in Chile.

A.1 Introduction

IPCC (2013) forecasts precipitation decrease in the central-southern region of Chile that would increase drought frequencies and intensities. Future scenarios predict that by 2050, wheat and corn yields will decrease by 5% to 20% (IPCC, 2013; Meza and Silva, 2009); however, impacts on agricultural production could be greater if taking onboard effects on cropping frequency and area (Cohn et al., 2016). Nowadays, there has been an increment in the number of remote sensing products for climate (Yang et al., 2013) and agriculture applications (Brown, 2015) that provide great opportunities to carry out drought studies (AghaKouchak et al., 2015), especially for developing countries like Chile.

A.1.1 Vegetation and precipitation indices

Since the drought of 2008, remote sensing data has been used in Chile as useful information to monitor agricultural drought. Indices as the Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1974) and the Vegetation Condition Index (VCI) (Kogan, 1995b) have been adopted by the Chilean Agroclimatic Observatory (OAC for its acronym in Spanish) (www.climatedatalibrary.cl), along with the Standardized Precipitation Index (SPI) at short-time scales, based on rain gauges and satellite, used as indirect measure of agricultural drought. Zambrano et al. (2016) proved that VCI is a useful tool for monitoring agricultural drought in the central-southern region of Chile, comparing this index with the well known Standardized Precipitation Index (SPI) (McKee et al., 1993) obtained from weather stations, and with governmental declarations regarding the emergency drought between 2000 and 2015.

Zambrano et al. (2016) used remote sensing products from the Moderate Resolution Imaging Spectroradiometer (MODIS), specifically, MOD13Q1 (Didan, 2015) product regarding vegetation indices that provides data

every 16 days at 250-meter spatial resolution and MCD12Q1 (Friedl et al., 2010)-product that provides data regarding land cover yearly at 500-meter spatial resolution-used to generate a cropland mask and then isolate the agricultural area. Satellite-derived SPIs between 1981 and 2015, at time-scale of 1, 3, and 6 months, were evaluated for its application on Chile during agricultural drought (Zambrano et al., 2017), using 278 weather station to assess the agreement between satellite and in-situ SPIs. Climate Hazards Group Infra Red Precipitation with Station (CHIRPS) (Funk et al., 2014) product, at 5-kilometer spatial resolution, showed good results in the central region of Chile, unlike in the southern and northern regions. Therefore, this product is a viable alternative for data sources spatially distributed rainfall in Chile.

A.1.2 Recent satellite data and drought indices

Besides vegetation and precipitation satellite data for application in agricultural drought, there are new remote sensing datasets, such as Evapotranspiration (ET) (Mu et al., 2007, 2011), Soil Moisture (SM) (Dorigo et al., 2015), Land Surface Temperature (LST), and phenology dynamics (Ganguly et al., 2010), which could be evaluated for its application in Chile. As indicated by Mishra et al. (2015), the agricultural drought will differ between crops because of two major factors: demand and supply, which could be estimated using satellite data.

Water demand depends on climate and specific crop characteristics. Perhaps, the most relevant measure of crop water demand is ET, which has been used in the formulation of drought indices like the Drought Severity Index (DSI) (Mu et al., 2013) and the Standardized Precipitation Evapotranspiration Index (SPEI) (Vicente-Serrano et al., 2010). The effect of temperature in agriculture could be analyzed through LST, used to derive the Temperature Condition Index (TCI) (Kogan, 1995a), which determines stress on vegetation caused by temperatures. Another main characteristic of crop water demand is the growing season, period during the agricultural drought indices should be considered. To achieve this analysis of NDVI, time-series has been generally used to estimate the seasonality (Atkinson et al., 2012; Vrieling et al., 2011) also product MCD12Q2 for land cover dynamics (Ganguly et al., 2010) from MODIS provide seasonality dates globally as a useful dataset.

Soil moisture (SM) is a water supply measure which will be used for plant development and to know the impact on yield well; this should be considered in the root zone depth Mishra et al. (2015). Two of the latest satellite soil moisture products are:

A. Climate Change Initiative (CCI) (Dorigo et al., 2015), part of the ESA Programme on Global Monitoring of Essential Climate Variables (ECV), which has 25-kilometer spatial resolution and daily frequencies and
B. Soil Moisture Active Passive (SMAP) (Colliander et al., 2017), which has 9-kilometer spatial resolution and frequencies of 7 days and the advantage that considers the soil moisture for the root zone depth. From soil moisture data, there are several drought indices, such as the Drought Severity Index (DSI) (Cammalleri et al., 2015) and the Soil Moisture Deficit Index (SMDI) (Narasimhan and Srinivasan, 2005).

A.2 Conclusion

In order to study agricultural drought in Chile, satellite indices that measure vegetation response, such as NDVI and VCI (Zambrano et al., 2016), have been used. In addition to satellite-derived precipitation indices like SPI (Zambrano et al., 2017). These indices have allowed generating a baseline in the understanding of agricultural drought process in Chile.

Future works should consider the use of new satellite data, such as LST, SM, ET, and phenology, in order to improve agricultural drought analyses. These data, along with new analysis techniques, will be useful to develop a model that helps to anticipate agricultural drought.

In addition, there is a gap in the studies presented in Chile regarding to the use of crop data to validate satellite data, mainly because it implies a lot of work to collect. A considerable challenge is to generate a national database about crops which will be crucial to advance in the knowledge of agricultural drought impact at the level of crop type through Chile.

Bibliography

Abbot, J. and Marohasy, J. (2012). Application of artificial neural networks to rainfall forecasting in Queensland, Australia. *Advances in Atmospheric Sciences*, 29(4):717–730.

Abbot, J. and Marohasy, J. (2014). Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks. *Atmospheric Research*, 138:166 – 178.

AghaKouchak, A., Farahmand, A., Melton, F. S., Teixeira, J., Anderson, M. C., Wardlow, B. D., and Hain, C. R. (2015). Remote sensing of drought: Progress, challenges and opportunities. *Reviews of Geophysics*, 53:452–480.

Aldunce, P. and González, M. (2009). *Desastres asociados al clima en la agricultura y medio rural en Chile*. Departamento de Ciencias Ambientales y Recursos Naturales Renovables de la Facultad de Cs. Agronómicas de la Universidad de Chile.

Ali, I., Greifeneder, F., Stamenkovic, J., Neumann, M., and Notarnicola, C. (2015). Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. *Remote Sensing*, 7(12):16398–16421.

Alley, W. M. (1984). The palmer drought severity index: Limitations and assumptions. *Journal of Climate and Applied Meteorology*, 23(7):1100–1109.

Almazroui, M. (2011). Calibration of TRMM rainfall climatology over Saudi Arabia during 1998-2009. *Atmospheric Research*, 99:400–414.

Amin, Z., Rehan, S., Bahman, N., and Faisal, K. (2011). A review of drought indices. *Environmental Reviews*, 19:333–349.

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P. (2015). PERSIANN-CDR: Daily precipitation climate data record form multisatellite observations for hydrological and climate studies. *Bulletin of the American Meteorological Society*, 96:69–83.

Asoka, A. and Mishra, V. (2015). Prediction of vegetation anomalies to improve food security and water management in india. *Geophysical Research Letters*, 42(13):5290–5298. 2015GL063991.

Atkinson, P. M., Jeganathan, C., Dash, J., and Atzberger, C. (2012). Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. *Remote Sensing of Environment*, 123:400 – 417.

Bajgiran, P. R., Darvishsefat, A. A., Khalili, A., and Makhdoum, M. F. (2008). Using AVHRR-based vegetation indices for drought monitoring in the Northwest of Iran. *Journal of Arid Environments*, 72(6):1086 – 1096.

Beguería, S. and Vicente-Serrano, S. M. (2013). *SPEI: Calculation of the Standardised Precipitation-Evapotranspiration Index*, r package version 1.6 edition.

Bergstra, J. and Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, pages 281–305.

Bharti, V. and Singh, C. (2015). Evaluation of error in TRMM 3B42V7 precipitation estimates over Himalayan region. *Journal of Geophysical Research: Atmospheres*, 120(12):458–473.

Bhuiyan, C., Singh, R., and Kogan, F. (2006). Monitoring drought dynamics in the Aravalli region (India) using different indices based on ground and remote sensing data. *International Journal of Applied Earth Observation and Geoinformation*, 8(4):289 – 302.

Boisier, J. P., Rondanelli, R., Garreaud, R. D., and Muñoz, F. (2016). Anthropogenic and natural contributions to the Southeast Pacific precipitation decline and recent megadrought in central Chile. *Geophysical Research Letters*, 43(1):413–421. 2015GL067265.

Bontemps, S., Defourney, P., Van Bogaert, E., Arino, O., Kalogirou, V., and Perez, J. (2011). GLOBCOVER 2009: Products description and validation report. Technical report, Université catholique de Louvain (UCL) & European Space Agency (esa).

Brown, M. E. (2015). Satellite remote sensing in agriculture and food security assessment. *Procedia Environmental Sciences*, 29:307 –.

Brown, M. E., de Beurs, K., and Vrieling, A. (2010). The response of African land surface phenology to large scale climate oscillations. *Remote Sensing of Environment*, 114(10):2286 – 2296.

Brown, M. E., Lary, D. J., Vrieling, A., Stathakis, D., and Mussa, H. (2008). Neural networks as a tool for constructing continuous NDVI time series from AVHRR and MODIS. *International Journal of Remote Sensing*, 29(24):7141–7158.

Caccamo, G., Chisholm, L., Bradstock, R., and Puotinen, M. (2011). Assessing the sensitivity of MODIS to monitor drought in high biomass ecosystems. *Remote Sensing of Environment*, 115(10):2626 – 2639.

Cammalleri, C., Micale, F., and Vogt, J. (2015). A novel soil-moisture-based drought severity index (DSI) combining water deficit magnitude and frequency. *Hydrological Processes*, 30:289–301.

Campbell, B., Vermeulen, S., Aggarwal, P., Corner-Dolloff, C., Girvetz, E., Loboguerrero, A., Ramirez-Villegas, J., Rosenstock, T., Sebastian, L., Thornton, P., and Wollenberg, E. (2016). Reducing risks to food security from climate change. *Global Food Security*, 11:34–43.

Candel, A., Parmar, V., LeDell, E., and Arora, A. (2016). Deep learning with H2O.

Cleveland, W. S. (1981). LOWESS: A Program for Smoothing Scatterplots by Robust Locally Weighted Regression. *The American Statistician*, 35(1).

Cohn, A. S., VanWey, L. K., Spera, S. A., and Mustard, J. F. (2016). Cropping frequency and area response to climate variability can exceed yield response. *Nature Climate Change*, page 6.

Colliander, A., Jackson, T., Bindlish, R., Chan, S., Das, N., Kim, S., Cosh, M., Dunbar, R., Dang, L., Pashaian, L., Asanuma, J., Aida, K., Berg, A., Rowlandson, T., Bosch, D., Caldwell, T., Caylor, K., Goodrich, D., al Jassar, H., Lopez-Baeza, E., Martínez-Fernández, J., González-Zamora, A., Livingston, S., McNairn, H., Pacheco, A., Moghaddam, M., Montzka, C., Notarnicola, C., Niedrist, G., Pellarin, T., Prueger, J., Pulliainen, J., Rautiainen, K., Ramos, J., Seyfried, M., Starks, P., Su, Z., Zeng, Y., van der Velde, R., Thibeault, M., Dorigo, W., Vreugdenhil, M., Walker, J., Wu, X., Monerris, A., O'Neill, P., Entekhabi, D., Njoku, E., and Yueh, S. (2017). Validation of {SMAP} surface soil moisture products with core validation sites. *Remote Sensing of Environment*, 191:215 – 231.

Cook, J., Nuccitelli, D., Green, S. A., Richardson, M., Winkler, B., Painting, R., Way, R., Jacobs, P., and Skuce, A. (2013). Quantifying the consensus on anthropogenic global warming in the scientific literature. *Environmental Research Letters*, 8(2):024024.

Cook, J., Oreskes, N., Doran, P. T., Anderegg, W. R. L., Verheggen, B., Maibach, E. W., Carlton, J. S., Lewandowsky, S., Skuce, A. G., Green, S. A., Nuccitelli, D., Jacobs, P., Richardson, M., Winkler, B., Painting, R., and Rice, K. (2016). Consensus on consensus: a synthesis of consensus estimates on human-caused global warming. *Environmental Research Letters*, 11(4):048002.

Cunha, A., Alvalá, R., Nobre, C., and Carvalho, M. (2015). Monitoring vegetative drought dynamics in the Brazilian semiarid region. *Agricultural and Forest Meteorology*, 214-215:494–505.

Dai, A. (2012). Increasing drought under global warming in observations and models. *Nature Climate Change*, 3(1):52–58.

Deo, R. C. and Şahin, M. (2015). Application of the extreme learning machine algorithm for the prediction of monthly Effective Drought Index in eastern Australia. *Atmospheric Research*, 153:512–525.

Derin, Y., Anagnostou, E., Berne, A., Borga, M., Boudevillain, B., Buytaert, W., Chang, C.-H., Delrieu, G., Hong, Y., Hsu, Y. C., Lavado-Casimiro, W., Manz, B., Moges, S., Nikolopoulos, E. I., Sahlu, D., Salerno, F., Rodríguez-Sánchez, J.-P., Vergara, H. J., and Yilmaz, K. K. (2016). Multiregional satellite precipitation products evaluation over complex terrain. *Journal of Hydrometeorology*, 17(6):1817–1836.

Didan, K. (2015). MOD13Q1 MODIS/Terra vegetation indices 16-day L3 Global 250m SIN Grid V006. Technical report, NASA EOSDIS Land Processes DAAC.

Dinku, T., Cecato, P., Grover-Kopec, E., M.Lemma, Connor, S., and Ropelewski, C. (2007). Validation of satellite rainfall products over East Africa's complex topography. *International Journal of Remote Sensing*, 28(7):1503–1526.

Dinku, T., Ruiz, F., Connor, S. J., and Ceccato, P. (2009). Validation and inetrcomparison of satellite rainfall estimates over Colombia. *Journal of Applied Meteorology and Climatology*, 49:1004–114.

Dore, M. H. (2005). Climate change and changes in global precipitation patterns: What do we know? *Environment International*, 31(8):1167 – 1181.

Dorigo, W., Gruber, A., Jeu, R. D., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R., and Kidd, R. (2015). Evaluation of the {ESA} {CCI} soil moisture product using ground-based observations. *Remote Sensing of Environment*, 162:380 – 395.

Dowle, M., Srinivasan, A., Short, T., with contributions from R Saporta, S. L., and Antonyan, E. (2015). *data.table: Extension of Data.frame*. R package version 1.9.6.

Du, L., Tian, Q., Yu, T., Meng, Q., Jancso, T., Udvardy, P., and Huang, Y. (2013). A comprehensive drought monitoring method integrating MODIS and TRMM data. *International Journal of Applied Earth Observation and Geoinformation*, 23:245 – 253.

Duan, Z. and Bastiaanssen, W. (2013). First results from version 7 TRMM 3B43 precipitation product in combination with a new downscaling-calibration procedure. *Remote Sensing of Environment*, 131:1–13.

Dwyer, J. and Schmidt, G. (2006). The MODIS reprojection tool. In Qu, J., Gao, W., Kafatos, M., Murphy, R., and Salomonson, V., editors, *Earth Science Satellite Remote Sensing*, pages 162–177. Springer: Berlin, Germany.

Enenkel, M., Steiner, C., Mistelbauer, T., Dorigo, W., Wagner, W., See, L., Atzberger, C., Schneider, S., and Rogenhofer, E. (2016). A combined satellite-derived drought indicator to support humanitarian aid organizations. *Remote Sensing*, 8(4):340.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., and Huang, X. (2010). MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114(1):168 – 182.

Fritz, S., Bartholome, E., Belward, A., Hartley, A., Stibig, H.-J., Eva, H., Mayaux, P., Bartalev, S., Latifovic, R., Kolmert, S., and et al. (2003). Harmonisation, mosaicing and production of the Global Land Cover 2000 database. Technical report, Joint Research Center, EC.

Funk, C. and Budde, M. E. (2009). Phenologically-tuned {MODIS} NDVI-based production anomaly estimates for Zimbabwe. *Remote Sensing of Environment*, 113(1):115 – 125.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Rowland, J., Romero, B., Husak, G., Michaelsen, J., and Verdin, A. (2014). A quasi-global precipitation time series for drought monitoring. Technical report, U.S. Geological Survey Data Series 832.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, 2(150066):21.

Ganguly, S., Friedl, M. A., Tan, B., Zhang, X., and Vermab, M. (2010). Land surface phenology from MODIS: Characterization of the collection 5 global land cover dynamics product. *Remote Sensing of Environment*, 114(8):1805–1816.

Garreaud, R. (2009). The Andes climate and weather. *Advances in Geosciences*, 7(1-9).

Garreaud, R., Alvarez-Garreton, C., Barichivich, J., Boisier, J. P., Christie, D., Galleguillos, M., LeQuesne, C., McPhee, J., and Zambrano-Bigiarini, M. (2017). The 2010–2015 mega drought in Central Chile: Impacts on regional hydroclimate and vegetation. *Hydrology and Earth System Sciences Discussions*, 2017:1–37.

Garreaud, R. and Battisti, D. (1999). Interannual (ENSO) and interdecadal (ENSO-like) variability in the Southern Hemisphere tropospheric circulation. *Journal of Climate*, 12(7):2113–2123. cited By 230.

Gebrehiwot, T., van der Veen, A., and Maathuis, B. (2011). Spatial and temporal assessment of drought in the Northern highlands of Ethiopia. *International Journal of Applied Earth Observation and Geoinformation*, 13(3):309 – 321.

Gedeon, T. D. (1997). Data mining of inputs: Analysing magnitude and functional measures. *International Journal of Neural Systems*, 08(02):209–218.

Gillette, H. (1950). A creeping drought under way. *In: Water and Sewage Works*, pages 104–105.

Guo, H., Bao, A., Liu, T., Chen, S., and Ndayisaba, F. (2016a). Evaluation of PERSIANN-CDR for meteorological drought monitoring over China. *Remote Sensing*, 8(379):17.

Guo, H., Chen, S., Bao, A., Behrangi, A., Hong, Y., Ndayisaba, F., Hu, J., and Stepanian, P. (2016b). Early assessment of integrated Multi-satellite Retrievals for Global Precipitation Measurement over China. *Atmospheric Research*, 176-177:121–133.

Hansen, J., Ruedy, R., Sato, M., and Lo, K. (2010). Global surface temperature change. *Reviews of Geophysics*, 48(4):n/a–n/a. RG4004.

Hansen, J. W., Hodges, A. W., and Jones, J. W. (1998). ENSO influences on agriculture in the Southeastern United States. *Journal of Climate*, 11(3):404–411.

Hao, Z. and AghaKouchak, A. (2013). Multivariate standardized drought index: A parametric multi-index model. *Advances in Water Resources*, 57:12–18.

Hayes, M., Svoboda, M., Wall, N., and Widhalm, M. (2011). The Lincoln declaration on drought indices: Universal meteorological drought index recommended. *Bulletin of the American Meteorological Society*, 92(4):485–488.

Hengl, T. (2009). *A Practical Guide to Geostatistical Mapping*. University of Amsterdam, Amsterdam, 2nd edition.

Hijmans, R. J. (2015). *raster: Geographic Data Analysis and Modeling*, r package version 2.5-2 edition.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15):1965–1978.

Hird, J. N. and McDermid, G. J. (2009). Noise reduction of NDVI time series: An empirical comparison of selected techniques. *Remote Sensing of Environment*, 113(1):248 – 258.

Hsu, K.-L., Gao, X., Sorooshian, S., and Gupta, H. V. (1997). Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology*, 36:1176–1190.

Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., and Ferreira, L. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1â 2):195 – 213. The Moderate Resolution Imaging Spectroradiometer (MODIS): a new generation of Land Surface Monitoring.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolff, D. B. (2007). The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8(1):38–55.

Huffman, G. J., Adler, R. F., Bolvin, D. T., and Nelkin, E. J. (2010). *The TRMM Multi-Satellite Precipitation Analysis (TMPA)*. Satellite Applications for Surface Hydrology.

INE (2007). VII Censo nacional agropecuario y forestal. Instituto Nacional de Estadística (INE), Santiago, Chile.

IPCC (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK; New York, USA.

IPCC (2014). *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel of Climate Change[Field, C. B. and Barros, V. R. and Dokken, D. J and Mach, K. J. and Mastrandrea, M. D. and Bilir, T. E. and Chatterjee, M. and Ebi, K. L. and Estrada, Y. O. and Genova, R. C. and Girma, B. and Kissel, E. S. and Levy, A. N. and MacCracken, S. and Mastrandrea, P. R. and White, L. L. and Porter, J. R. and Xie, L. and Challinor, A. J. and Cochrane, K. and Howden, S. M. and Iqbal, M. M. and Lobell, D. B. and Travasso, M. I.]*. Cambridge University Press, Cambridge, UK; New York, USA, 1132 pp.

Ji, L. and Peters, A. J. (2003). Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. *Remote Sensing of Environment*, 87(1):85 – 98.

Jia, M., Tong, L., Chen, Y., Wang, Y., and Zhang, Y. (2013). Rice biomass retrieval from multitemporal ground-based scatterometer data and RADARSAT-2 images using neural networks. *Journal of Applied Remote Sensing*, 7(1):073509–073509.

Johnson, M., Hsieh, W., Cannon, A. e., Davidson, A., and Bédard, F. (2016). Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods. *Agricultural and Forest Meteorology*, 218-219:74–84.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P. (2004). CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, 5:487–503.

Julien, Y. and Sobrino, J. A. (2010). Comparison of cloud-reconstruction methods for time series of composite NDVI data. *Remote Sensing of Environment*, 114(3):618 – 625.

Jung, M., Verstraete, M., Gobron, N., Reichstein, M., Papale, D., Bondeau, A., Robustelli, M., and Pinty, B. (2008). Diagnostic assessment of european gross primary production. *Global Change Biology*, 14(10):2349–2364.

Katsanos, D., Retailis, A., and Michaelides, S. (2016). Validation of a high-resolution precipitation database (CHIRPS) oover Cyprus for a 30-year period. *Atmospheric Research*, 169:459–464.

Kenawy, A. M. E., Lopez-Moreno, J. I., McCabe, M. F., and Vicente-Serrano, S. M. (2015). Evaluation of the TMPA-3B42 precipitation product using a high-density rain gauge network over complex terrain in northeastern Iberia. *Global and Planetary Change*, 133:188–200.

Klisch, A. and Atzberger, C. (2016). Operational drought monitoring in kenya using MODIS NDVI time series. *Remote Sensing*, 8(4):267.

Knapp, A. K., Beier, C., Briske, D. D., Classen, A. T., Luo, Y., Reichstein, M., Smith, M. D., Smith, S. D., Bell, J. E., Fay, P. A., Heisler, J. L., Leavitt, S. W., Sherry, R., Smith, B., and Weng, E. (2008). Consequences of more extreme precipitation regimes for terrestrial ecosystems. *Bioscience*, 58(9):811–821.

Kogan, F., Gitelson, A. A., Zakarin, E., Spivak, L., and Lebed, L. (2003). AVHRR-Based spectral vegetation index for quantitative assessment of vegetation state and productivity: Calibration and validation. *Photogrammetric Engineering and Remote Sensing*, 69:899–906.

Kogan, F. N. (1990). Remote sensing of weather impacts on vegetation in non-homogeneous areas. *International Journal of Remote Sensing*, 11(8):1405–1419.

Kogan, F. N. (1995a). Application of vegetation index and brightness temperature for drought detection. *Advances in Space Research*, 15(11):91–100.

Kogan, F. N. (1995b). Droughts of the late 1980s in the United States as derived from NOAA polar-orbiting satellite data. *Bulletin of the American Meteorological Society*, 76(5):655–668.

Kogan, F. N. (1997). Global drought watch from space. *Bulletin of the American Meteorological Society*, 78(4):621–636.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F. (2006). World map of the Koppen-Geiger climate classificclass updated. *Meteorologische Zeitschrift*, 15(3):259–2563.

Kouadio, L. d., Newlands, N., Davidson, A., Zhang, Y., and Chipanshi, A. (2014). Assessing the performance of MODIS NDVI and EVI for seasonal crop yield forecasting at the ecodistrict scale. *Remote Sensing*, 6(10):10193–10214. cited By 10.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444.

Logan, K., Brunsell, N., Jones, A., and Feddema, J. (2010). Assessing spatiotemporal variability of drought in the U.S. central plains. *Journal of Arid Environments*, 74(2):247 – 255.

Loon, A. F. V., Gleeson, T., Clark, J., Dijk, A. I. J. M. V., Stahl, K., Hannaford, J., Baldassarre, G. D., Teuling, A. J., Tallaksen, L. M., Uijlenhoet, R., Hannah, D. M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., Rangecroft, S., Wanders, N., and Lanen, H. A. J. V. (2016). Drought in the anthropocene. *Nature Geoscience*, 9(2):89–91.

Mann, M., Zhang, Z., Hughes, M., Bradley, R., Miller, S., Rutherford, S., and Ni, F. (2007). Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proceedings of the National Academy of Sciences of the United States*.

Marj, A. F. and Meijerink, A. M. J. (2011). Agricultural drought forecasting using satellite images, climate indices and artificial neural network. *International Journal of Remote Sensing*, 32(24):9707–9719.

McKee, T. B., Doesken, N. J., and Kleist, J. (1993). The relationship of drought frecuency and duration to time scales. In *In Proceedings of the International 8th Conference on Applied Climatology. American Meteorological Society, Anaheim, CA, USA, 17-22 January*, pages 179–184.

Meroni, M., Fasbender, D., Kayitakire, F., Pini, G., Rembold, F., Urbano, F., and Verstraete, M. (2014a). Early detection of biomass production deficit hot-spots in semi-arid environment using FAPAR time series and a probabilistic approach. *Remote Sensing of Environment*, 142:57–68.

Meroni, M., Rembold, F., Fasbender, D., and Vrieling, A. (2017). Evaluation of the Standardized Precipitation Index as an early predictor of seasonal vegetation production anomalies in the Sahel. *Remote Sensing Letters*, 8(4):301–310.

Meroni, M., Verstraete, M. M., Rembold, F., Urbano, F., and Kayitakire, F. (2014b). A phenology-based method to derive biomass production anomalies for food security monitoring in the Horn of Africa. *International Journal of Remote Sensing*, 35(7):2472–2492.

Meza, F. J. (2013). Recent trends and ENSO influence on droughts in Northern Chile: An application of the Standarized Precipitation Evapotranspiration Index. *Weather and Climate Extremes*, 1:51–58.

Meza, F. J. and Silva, D. (2009). Dynamic adaptation of maize and wheat production to climate change. *Climatic Change*, 94(1):143–156.

Miao, C., Ashouri, H., Hsu, K.-L., Sorooshian, S., and Duan, Q. (2015). Evaluation of the Persiann-CDR daily rainfall estimates in capturing the behavior of extreme precipitation events over China. *American Meteorology Society*, 16:1387–1396.

Mishra, A. K., Ines, A. V., Das, N. N., Khedun, C. P., Singh, V. P., Sivakumar, B., and Hansen, J. W. (2015). Anatomy of a local-scale drought: Application of assimilated remote sensing products, crop model, and statistical methods to an agricultural drought study. *Journal of Hydrology*, 526:15 – 29.

Mishra, A. K. and Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, 391(1–2):202 – 216.

Miura, T., Yoshioka, H., Fujiwara, K., and Yamamoto, H. (2008). Inter-comparison of ASTER and MODIS surface reflectance and vegetation index products for synergistic applications to natural resource monitoring. *Sensors*, 8(4):2480.

Moazami, S., Golian, S., Hong, Y., Sheng, C., and Kavianpour, M. R. (2016). Comprehensive evaluation of four high-resolution satellite precipitation products under diverse climate conditions in Iran. *Hydrological Sciences Journal*, 61(2):420–440.

Montecinos, A. and Aceituno, P. (2003). Seasonality of the ENSO-related rainfall variability in central Chile and associated circulation anomalies. *Journal of Climate*, 16(2):281–296. cited By 160.

Moorhead, J. E., Gowda, P. H., Singh, V. P., Porter, D. O., Marek, T. H., Howell, T. A., and Stewart, B. (2015). Identifying and evaluating a suitable index for agricultural drought monitoring in the Texas High Plains. *JAWRA Journal of the American Water Resources Association*, 51(3):807–820.

Moreno, A., García-Haro, F. J., Martínez, B., and Gilabert, M. A. (2014). Noise reduction and gap filling of fAPAR time series using an adapted local regression filter. *Remote Sensing*, 6(9):8238–8260.

Morssm, R. E., Demuth, J. L., and Lazo, J. K. (2008). Communicating uncertainty in weather forecasts: A survey of the U.S. public. *Weather and Forecasting*, 23(5):974–991.

Mu, Q., Heinsch, F. A., Zhao, M., and Running, S. W. (2007). Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sensing of Environment*, 111(4):519 – 536.

Mu, Q., Zhao, M., Kimball, J. S., McDowell, N. G., and Running, S. W. (2013). A remotely sensed global terrestrial drought severity index. *Bulletin of the American Meteorological Society*, 94(1):83–98.

Mu, Q., Zhao, M., and Running, S. W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, 115:1781–1800.

Narasimhan, B. and Srinivasan, R. (2005). Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring. *Agricultural and Forest Meteorology*, 133:69–88.

Nastos, P., Kapsomenakis, J., and Philandras, K. (2016). Evaluation of the TRMM 3B43 gridded precipitation estimates over Greece. *Atmospheric Research*, 169:497–514.

Nastos, P., Paliatsos, A., Koukouletsos, K., Larissi, I., and Moustris, K. (2014). Artificial neural networks modeling for forecasting the maximum daily total precipitation at Athens, Greece. *Atmospheric Research*, 144:141 – 150. Perspectives of Precipitation Science - Part {II}.

National Research Council (2004). *Climate Data Records from Environmental Satellites: Interim Report*. The National Academies Press, Washington, DC.

Niemeyer, S. (2008). New drought indices. *Options Méditerranéennes,*, Série A: Séminaires Méditerranéens(80):267–274.

Núñez, J., Verbist, K., Wallis, J., Schaefer, M., Morales, L., and Cornelis, W. (2011). Regional frequency analysis for mapping drought events in north-central Chile. *Journal of Hydrology*, 405(3):352 – 366.

ODEPA (2015). *Chilean agriculture overview*. Office of Agricultural Policies, Trade and Information.

Palmer, W. C. (1965). Meteorological drought; research paper no. 45. *US Department of Commerce, Weather Bureau, Washington, DC, USA*.

Palmer, W. C. (1968). Keeping track of crop moisture conditions, nationwide: The new crop moisture index. *Weatherwise*, 21(4):156–161.

Panda, S. S., Ames, D. P., and Panigrahi, S. (2010). Application of vegetation indices for agricultural crop yield prediction using neural network techniques. *Remote Sensing*, 2(3):673–696.

Peel, M. C., Finlayson, B. L., and McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, 11(5):1633–1644.

Peters, A. J., Walter-Shea, Ji, L., Viña, A., and Svoboda, M. D. (2002). Drought monitoring with NDVI-based standarized vegetation index. *Photogrammetric Engineering & Remote Sensing*, 68(1):71–75.

Pipunic, R. C., Ryu, D., Costelloe, J. F., and Su, C.-H. (2015). An evaluation and regional error modelling methodology for near-real-time satellite rainfall data over Asutralia. *Journal of Geophysical Research: Atmospheres*, 120(10):767–783.

Potopová, V., Stepánek, P., Mozný, M., Türkott, L., and Soukup, J. (2015). Performance of the standarised precipitation evapotranspiration index at various lags for agricultural drought risk assessment in the Czech Republic. *Agricultural and Forest Meteorology*, 202:26–38.

Pulwarty, R. S. and Sivakumar, M. V. (2014). Information systems in a changing climate: Early warnings and drought risk management. *Weather and Climate Extremes*, 3:14 – 21. High Level Meeting on National Drought Policy.

Quiring, S. M. and Ganesh, S. (2010). Evaluating the utility of the vegetation condition index (VCI) for monitoring meteorological drought in Texas. *Agricultural and Forest Meteorology*, 150(3):330 – 339.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Reilly, J., Tubiello, F., McCarl, B., Abler, D., Darwin, R., Fuglie, K., Hollinger, S., Izaurralde, C., Jagtap, S., Jones, J., Mearns, L., Ojima, D., Paul, E., Paustian, K., Riha, S., Rosenberg, N., and Rosenzweig, C. (2003). U.S. Agriculture and climate change: New results. *Climatic Change*, 57(1):43–67.

Rembold, F., Atzberger, C., Savin, I., and Rojas, O. (2013). Using low resolution satellite imagery for yield prediction and yield anomaly detection. *Remote Sensing*, 5(4):1704.

Rembold, F., Meroni, M., Rojas, O., Atzberger, C., Ham, F., and Fillol, E. (2015). *Chapter 14. Agricultural Drought Monitoring Using Space-Derived Vegetation and Biophysical Products: A Global Perspective*. Remote Sensing of Water Resources, Disasters, and Urban Studies. 349-365.

Rhee, J., Im, J., and Carbone, G. J. (2010). Monitoring agricultural drought for arid and humid regions using multi-sensor remote sensing data. *Remote Sensing of Environment*, 114(12):2875 – 2887.

Richard, Y., Martiny, N., Fauchereau, N., Reason, C., Rouault, M., Vigaud, N., and Tracol, Y. (2008). Interannual memory effects for spring NDVI in semi-arid South Africa. *Geophysical Research Letters*, 35(13):n/a–n/a. L13704.

Rigge, M., Smart, A., Wylie, B., Gilmanov, T., and Johnson, P. (2013). Linking phenology and biomass productivity in South Dakota mixed-grass prairie. *Rangeland Ecology & Management*, 66(5):579 – 587.

Roco, L., Engler, A., Bravo-Ureta, B., and Jara-Rojas, R. (2014). Farm level adaptation decisions to face climatic change and variability: Evidence from Central Chile. *Environmental Science & Policy*, 44:86 – 96.

Rojas, O., Vrieling, A., and Rembold, F. (2011). Assessing drought probability for agricultural areas in Africa with coarse resolution remote sensing imagery. *Remote Sensing of Environment*, 115(2):343 – 352.

Romero, H., Smith, P., Mendonca, M., and Méndez, M. (2013). Macro y mesoclimas del altiplano andino y desierto de Atacama: desafíos y estrategias de adaptación social ante su vulnerabilidad. *Revista de Geografía Norte Grande*, 55:19–41.

Rouse, Jr., J. W., Haas, R. H., Schell, J. A., and Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with erts. *NASA Special Publication*, 351:309.

Salio, P., Hobouchiand, M. P., Skabarc, Y. G., and Vila, D. (2015). Evaluation of high-resolution satellite precipitation estimates over southern South America using a dense rain gauge network. *Atmospheric Research*, 163:146–161.

Sandholt, I., Rasmussen, K., and Andersen, J. (2002). A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. *Remote Sensing of Environment*, 79(Issues 2-3):213 – 224.

Seiler, R., Kogan, F., and Sullivan, J. (1998). AVHRR-based vegetation and temperature condition indices for drought detection in Argentina. *Advances in Space Research*, 21(3):481 – 484.

Seiler, R., Kogan, F., Wei, G., and Vinocur, M. (2007). Seasonal and interannual responses of the vegetation and production of crops in Cordoba - Argentina assessed by AVHRR derived vegetation indices. *Advances in Space Research*, 39(1):88 – 94.

Shafer, B. A. and Dezman, L. E. (1982). Development of a surface water supply index (SWSI) to assess the severity of drought conditions in snowpack runoff areas. In *In Proceedings of the Western Snow Conference, Fort Collins, CO, USA, 19-23 April*, pages 164–175.

Shah, H. L. and Mishra, V. (2015). Uncertainty and bias in satellite-based precipitation estimates over Indian subcontinental basins: Implications for real-time streamflow simulation and flood prediction. *Journal of Hydrometeorology*, 17:615–636.

Sheffield, J., Goteti, G., Wen, F., and Wood, E. F. (2004). A simulated soil moisture based drought analysis for the United States. *Journal of Geophysical Research: Atmospheres*, 109(D24):n/a–n/a. D24108.

Singh, R. P., Roy, S., and Kogan, F. (2003). Vegetation and temperature condition indices from NOAA AVHRR data for drought monitoring over India. *International Journal of Remote Sensing*, 24(22):4393–4402.

Skakun, S., Kussul, N., Shelestov, A., and Kussul, O. (2016). The use of satellite data for agriculture drought risk quantification in Ukraine. *Geomatics, Natural Hazards and Risk*, 7(3):901–917.

Sykes, M. T. (2001). *Climate Change Impacts: Vegetation*. John Wiley & Sons, Ltd.

Tadesse, T., Demisse, G. B., Zaitchik, B., and Dinku, T. (2014). Satellite-based hybrid drought monitoring tool for prediction of vegetation condition in Eastern Africa: A case study for Ethiopia. *Water Resources Research*, 50(3):2176–2190.

Tadesse, T., Wardlow, B. D., Hayes, M. J., Svoboda, M. D., and Brown, J. F. (2010). The vegetation outlook (VegOut): A new method for predicting vegetation seasonal greenness. *GIScience & Remote Sensing*, 47(1):25–52.

Tan, M. L., Ibrahim, A. L., Duan, Z., Cracknell, A. P., and Chaplot, V. (2015). Evaluation of six high-resolution satellite and ground-based precipitation products over Malaysia. *Remote Sensing*, 7(2):1504–1528.

Tapiador, F. J., Turk, F., Petersen, W., Hou, A. Y., García-Ortega, E., Machado, L. A., Angelis, C. F., Salio, P., Kidd, C., Huffman, G. J., and de Castro, M. (2012). Global precipitation measurement: Methods, datasets and applications. *Atmospheric Research*, 104–105:70 – 97.

Tonini, F., Lasinio, G. J., and Hochmair, H. H. (2012). Mapping return levels of absolute NDVI variations for the assessment of drought risk in Ethiopia. *International Journal of Applied Earth Observation and Geoinformation*, 18:564 – 572.

Tsakiris, G., Pangalou, D., and Vangelis, H. (2007). Regional drought assessment based on the reconnaissance drought index (RDI). *Water Resources Management*, 21:821–833.

Unganai, L. S. and Kogan, F. N. (1998). Drought monitoring and corn yield estimation in Southern Africa from AVHRR data. *Remote Sensing of Environment*, 63(3):219 – 232.

Van der Wiel, K., Kapnick, S. B., and Vecchi, G. A. (2017). Shifting patterns of mild weather in response to projected radiative forcing. *Climatic Change*, 140(3):649–658.

Van Leeuwen, W. J., Hartfield, K., Miranda, M., and Meza, F. J. (2013). Trends and ENSO/AAO driven variability in NDVI derived productivity and phenology alongside the Andes mountains. *Remote Sensing*, 5(3):1177.

Velpuri, N., Senay, G., Singh, R., Bohms, S., and Verdin, J. (2013). A comprehensive evaluation of two MODIS evapotranspiration products over the conterminous United States: Using point and gridded FLUXNET and water balance ET. *Remote Sensing of Environment*, 139:35–49.

Vicente-Serrano, S. M. (2007). Evaluating the impact of drought using remote sensing in a Mediterranean, semi-arid region. *Natural Hazards*, 40:173–208.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate*, 23(7):1696–1718.

Vicente-Serrano, S. M., Gouveia, C., Camarero, J. J., Beguería, S., RicardoTrigo, López-Moreno, J. I., Azorín-Molina, C., Pasho, E., Lorenzo-Lacruz, J., Revuelto, J., Morán-Tejeda, E., and Sanchez-Lorenzo, A. (2012a). Response of vegetation to drought time-scales across global land biomes. *Proceedings of the National Academy of Sciences of the United States of America*, 110(1):6.

Vicente-Serrano, S. M., López-Moreno, J. I., Beguería, S., Lorenzo-Lacruz, J., Azorin-Molina, C., and Morán-Tejeda, E. (2012b). Accurate computation of a streamflow drought index. *Journal of Hydrologic Engineering*, 17(2):318–332.

Vrieling, A., de Beurs, K. M., and Brown, M. E. (2011). Variability of African farming systems from phenological analysis of NDVI time series. *Climatic Change*, 109(3):455–477.

Vrieling, A., Meroni, M., Mude, A. G., Chantarat, S., Ummenhofer, C. C., and de Bie, K. (2016). Early assessment of seasonal forage availability for mitigating the impact of drought on East African pastoralist. *Remote Sensing of Environment*, 174:44–55.

Vrieling, A., Skidmore, A. K., Wang, T., Meroni, M., Ens, B. J., Oosterbeek, K., O'Connor, B., Darvishzadeh, R., Heurich, M., Shepherd, A., and Paganini, M. (2017). Spatially detailed retrievals of spring phenology from single-season high-resolution image time series. *International Journal of Applied Earth Observation and Geoinformation*, 59:19 – 30.

Wan, Z., Wang, P., and Li, X. (2004). Using MODIS land surface temperature and normalized difference vegetation index products for monitoring drought in the southern Great Plains, USA. *International Journal of Remote Sensing*, 25(1):61–72.

Wang, P.-X., Li, X.-W., Gong, J.-Y., and Song, C. (2001). Vegetation temperature condition index and its application for drought monitoring. In *In Proceeding of the IEEE 2001 International Geoscience and Remote Sensing Symposium, IGARSS '01, 2001, Sydney, NSW, Australia, 9-13 July.*, volume 1, pages 141–143.

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, 21(12):1–20.

Wickham, H. and Francois, R. (2015). *dplyr: A Grammar of Data Manipulation*. R package version 0.4.2.

Wilhite, D. A. and Glantz, M. H. (1985). Understanding the drought phenomenon: The role of definitions. *Water International*, 10(3):111–120.

Wilhite, D. A., Sivakumar, M., and Wood, D. (2000). Improving drought early warning systems in the context of drought preparedness and mitigation. *Drought Network News (1994-2001).Paper 104*.

Wilhite, D. A., Sivakumar, M. V., and Pulwarty, R. (2014). Managing drought risk in a changing climate: The role of national drought policy. *Weather and Climate Extremes*, 3(0):4 – 13. High Level Meeting on National Drought Policy.

WMO (1986). *Report on Drought and Countries Affected by Drought During 1974-1985*. WMO, Geneva, p. 118.

Wolter, K. and Timlin, M. S. (2011). El Niño/Southern oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (mei.ext). *International Journal of Climatology*, 31(7):1074–1087.

Wu, J., Zhou, L., Liu, M., Zhang, J., Leng, S., and Diao, C. (2013). Establishing and assessing the integrated surface drought index (ISDI) for agricultural drought monitoring in mid-eastern China. *International Journal of Applied Earth Observation and Geoinformation*, 23:397 – 410.

Wu, J., Zhou, L., Zhang, J., Liu, M., Zhao, L., and Zhao, F. (2010). Analysis of relationships among vegetation condition indices and multiple-time scale SPI of grassland in growing season. In *In Proceedings of the 2010 18th International Conference on Geoinformatics, Beijing, China, 18-20 June.*, pages 1–6.

Xiong, X., Chiang, K., Sun, J., Barnes, W., Guenther, B., and Salomonson, V. (2009). NASA EOS terra and aqua MODIS on-orbit performance. *Advances in Space Research*, 43(3):413 – 422.

Yang, J., Gong, P., Fu, R., Zhang, M., Chen, J., Liang, S., Xu, B., Shi, J., and Dickinson, R. (2013). The role of satellite remote sensing in climate change studies. *Nature Climate Change*, 3:875–883.

Zambrano, F., Lillo-Saavedra, M., Verbist, K., and Lagos, O. (2016). Sixteen years of agricultural drought assessment of the BioBío Region in Chile using a 250m resolution Vegetation Condition Index (VCI). *Remote Sensing*, 8(6):530.

Zambrano, F., Wardlow, B., Tadesse, T., Lillo-Saavedra, M., and Lagos, O. (2017). Evaluating satellite-derived long-term historical precipitation dataset for drought monitoring in Chile. *Atmospheric Research*, 186:26–42.

Zambrano-Bigiarini, M., Nauditt, A., Birkel, C., Verbist, K., and Ribbe, L. (2017). Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile. *Hydrology and Earth System Sciences*, 21(2):1295–1320.

Zhang, A. and Jia, G. (2013). Monitoring meteorological drought in semiarid regions using multi-sensor microwave remote sensing data. *Remote Sensing of Environment*, 134:12 – 23.

Zhang, F., Zhang, L. W., Wang, X. Z., and Hung, J. F. (2013). Detecting agro-droughts in southwest of china using MODIS satellite data. *Journal of Integrative Agriculture*, 12(1):159 – 168.

Zhang, X., Friedl, M. A., and Schaaf, C. B. (2006). Global vegetation phenology from moderate resolution imaging spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements. *Journal of Geophysical Research: Biogeosciences*, 111(G4):n/a–n/a. G04017.

Zhang, X., Friedl, M. A., and Schaaf, C. B. (2009). Sensitivity of vegetation phenology detection to the temporal resolution of satellite data. *International Journal of Remote Sensing*, 30(8):2061–2074.

Zhang, Y., Wallace, J. M., and Battisti, D. S. (1997). ENSO-like interdecadal variability: 1900–93. *Journal of Climate*, 10(5):1004–1020.