UNIVERSIDAD DE CHILE
Facultad de Ciencias Forestales y de la Conservación de la Naturaleza
Magíster en Áreas Silvestres y Conservación de la Naturaleza

MONITORING ANDEAN HIGH ALTITUDE WETLANDS IN CENTRAL CHILE WITH SEASONAL OPTICAL DATA: A COMPARISON BETWEEN WORLDVIEW-2 AND SENTINEL-2 IMAGERY

Proyecto de grado presentado como parte de los requisitos para optar al grado de Magíster en Áreas Silvestres y Conservación de la Naturaleza.

ROCÍO ALEJANDRA ARAYA LÓPEZ
Ingeniera Agrónomo

SANTIAGO, CHILE
2017

UNIVERSIDAD DE CHILE
Facultad de Ciencias Forestales y de la Conservación de la Naturaleza
Magíster en Áreas Silvestres y Conservación de la Naturaleza

MONITORING ANDEAN HIGH ALTITUDE WETLANDS IN CENTRAL CHILE WITH SEASONAL OPTICAL DATA: A COMPARISON BETWEEN WORLDVIEW-2 AND SENTINEL-2 IMAGERY

Proyecto de grado presentado como parte de los requisitos para optar al grado de Magíster en Áreas Silvestres y Conservación de la Naturaleza.

ROCÍO ALEJANDRA ARAYA LÓPEZ
Ingeniera Agrónomo

SANTIAGO, CHILE
2017

Tesis presentada como parte de los requisitos para optar al grado de Magíster en Áreas Silvestres y Conservación de la Naturaleza

Profesor Guía	Nombre	Jaime Hernández Palma
	Nota	
	Firma	
Profesor Co Guía	Nombre	Fabian Fassnacht
	Nota	
	Firma	
Profesor Consejero	Nombre	Juan Pablo Fuentes Espoz
	Nota	
	Firma	
Profesor Consejero	Nombre	Gustavo Cruz Madariaga
	Nota	
	Firma	

AGRADECIMIENTOS

Los datos satelitales WV-2 y de humedad de suelo fueron cedidos por el Proyecto “Seguimiento de Vegas Altoandinas con Imágenes Satelitales de Alta Resolución Espacial” financiado a través de convenio ALTO MAIPO SpA y la Facultad de Ciencias Forestales y de la Conservación de la Naturaleza, U. de Chile (AM-CO292).

TABLA DE CONTENIDO

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

Página

RESUMEN

En la cuenca del Río Maipo, ubicada en la zona central de Chile, la actividad minera es el principal factor que afecta vegas altoandinas, a través del consumo de agua y explotación de la tierra. Debido a que estos humedales son altamente vulnerables y susceptibles a cambios en el suministro de agua, alteraciones y modificaciones en el régimen hidrológico tienen efectos directos sobre la condición ecofisiológica y la cobertura de la vegetación. El objetivo principal de este trabajo fue evaluar el potencial de las imágenes de alta resolución Worldview-2 y Sentinel-2 para el mapeo de vegas altoandinas a través de clasificadores de una clase, en específico "*Bias support vector machine*" (BSVM), y la estimación del estado hídrico de vegas mediante el uso del algoritmo de regresión '*Partial least square regression*'.

Los resultados de esta investigación demuestran que la combinación de un limitado número de observaciones y datos de teledetección reducen significativamente el esfuerzo de campo y permiten la obtención de mapas que detallan de manera precisa vegas altoandinas. BSVM logró clasificar áreas de vegas con una precisión general de más del 85% con ambos sensores. Además, se logró estimar la humedad superficial del suelo (primeros 20 cm) dentro de las áreas clasificadas como vegas con datos de teledetección y algoritmos de regresión simple. La predicción de la humedad del suelo alcanzó valores de R^2 hasta 63% y error cuadrático medio normalizado entre 11% y 18% con Sentinel-2, mientras que las estimaciones con Worldview-2 arrojaron resultados no satisfactorios.

El enfoque presentado es particularmente valioso para monitorear áreas de humedales de alta montaña con acceso limitado, como en los sectores húmedos de la Cordillera de los Andes.

ABSTRACT

In the Río Maipo watershed, situated in the central of Chile, mining activities are impacting high altitude Andean wetlands through the consumption and exploitation of water and land. As wetlands are vulnerable and particularly susceptible to changes of water supply, alterations and modifications in the hydrological regime have direct effects on their eco-physiological condition and vegetation cover. The aim of this study was to evaluate the potential of Worldview-2 and Sentinel-2 sensors to identify and map Andean wetlands through the use of the one-class classifier Bias support vector machine (BSVM), and then to estimate soil moisture content of the identified wetlands during the summer period using partial least square regression.

The results obtained in this research showed that the combination of remote sensing data and a small sample of ground reference measurements enable to map Andean high wetlands with high accuracies. BSVM was capable to classify the meadow areas with an overall accuracy of over ~85% for both sensors. Our results also indicate that it is feasible to estimate surface soil moisture (first 20 cm) with optical remote sensing data and simple regression approaches. Surface soil moisture estimates reached R^2 values of up to 63%, and standard mean square errors between 11% and 18% using Sentinel-2 data, while Worldview-2 estimates resulted in non-satisfying results. The presented approach is particularly valuable for monitoring high-mountain wetland areas with limited accessibility such as in the Andes.

1. INTRODUCTION

Andean high altitude wetlands (‘*vegas’*; wet meadows) are wetlands usually located in humid areas of the Andes in north-central Chile (Squeo et al., 2006a; Squeo et al., 2006b). These wetlands are important key components of the Andean ecosystems, providing several ecosystem services like water supply and habitat for wildlife and livestock (Ahumada & Faúndez, 2009; Contreras, 2007; García & Otto, 2015). However, Andean high altitude wetlands have suffered a strong degradation and transformation in the last decades, especially as a consequence of agriculture and mining activities (Cepeda-Pizzaro & Pola, 2013; Squeo et al., 2006).

In the Maipo River Basin located in central Chile, mining activity is the main factor impacting Andean wetlands, through the consumption and exploitation of water and land (GORE-RMS, 2013). Andean wetlands are particularly susceptible to changes of water supply, alterations and modifications in the hydrological regime which were found to have direct effects on vegetation cover (Ahumada et al., 2011; Ahumada & Faúndez, 2009). In order to understand these ecosystems better, as well as for conservation planning and an efficient management of resources, there is a strong need for a spatially explicit and up-to-date inventory of Andean wetland ecosystems. This need is contrasted by a lack of baseline dataset and limited knowledge on these habitats (Muñoz-Schick et al., 1997).

Remote sensing (RS) data has proven to be a useful source of information for mapping and monitoring wetland ecosystem at different temporal and spatial

scales (e.g. Adam et al., 2010; Ozesmi & Bauer, 2002; Rundquist et al., 2001), and has been considered a valuable tool for their conservation (Baker et al., 2006; Davranche et al., 2013; MacAlister & Mahaxay, 2009; Whiteside & Bartolo, 2015). RS allows obtaining objective information in remote and isolated areas such as the high Andean basins. Accurate and up-to-date mapping, inventory and assessment of wet meadows is extremely important to assess changes in wetlands' vegetation composition and cover and water content due to natural or anthropogenic disturbances over time (Ozesmi & Bauer, 2002; Rundquist et al., 2001; Wang et al., 2015).

Soil moisture (SM) is one of the key parameters estimated by RS data and closely related to wetland ecology. SM is usually estimated by means of passive microwaves sensors such as SMOS (e.g., Kerr et al., 2012), that could be disaggregated to higher resolution with thermal data (Merlin et al., 2012b; Molero et al., 2016). Another alternative is the use active radar systems, normally with VV polarization (e.g. Zribi et al., 2005; Zribi & Decharme, 2008) However, approaches using passive microwave and radar sensors mostly deliver SM information at coarse resolution (~100−4000 m). Reason why there is an interest in the development of spectral indices and physical models based on free optical sensors, such as Sentinel-2. First studies applying vegetation indices (e.g. NDVI) in conjunction with short-wave infrared (SWIR) information have already been presented (e.g. Sadeghi et al., 2017). Such approaches may enable the estimation of SM at local scales, which is key when mapping non-extensive isolated wetlands such as the Andean wetlands which are located in the highlands of the Andes and usually do not occupy large extensions of lands as they depend on special micro-relief characteristics which only occur locally.

The aim of this study was to analyse the potential of high spatial resolution optical data, i.e. Worldview-2 and Sentinel-2, for mapping and monitoring Andean wetlands of central Chile. To achieve this, we propose the use of a one-class-classifier to map the occurrences of Andean wetlands using sparse ground data. Such approach has been applied earlier for other ecosystems (e.g. Mack et al., 2016; Stenzel et al., 2017) and we deem this approach particularly valuable in our study area which is geographically isolated and where field data is hard to acquire. With the suggested work-flow we aim to create baseline data to complement the Chilean Wetland National Inventory, which is known to underestimate the extent of Andean wetlands (MMA-Centro de Ecología Aplicada, 2011), which hampers their proper conservation and management.

Secondly, we tested the suitability of the two satellite systems to estimate SM within the areas identified as Andean wetlands during the first step by applying empirical regression algorithms. Such an approach may become possible because in the Andes wetlands the SM gradient is usually related to a vegetation cover gradient which should be detectable by optical remote sensing data. The combination of both approaches could be the basis for a monitoring system that is able to detect early responses of Andean wetlands to natural or anthropogenic disturbances.

2. OBJECTIVES

2.1 General objective

Analyze the potential of high spatial resolution optical data, in specific, WorldView-2 and Sentinel-2 satellite images for mapping and monitoring Andean high altitude wetlands of central Chile

2.2 Specific objectives

Evaluate and compare the accuracy in the identification of Andean high altitude wetlands of central Chile through WorldView-2 and Sentinel-2 satellite images

Evaluate and compare the suitability of WorldView-2 and Sentinel-2 sensors to estimate soil moisture content within the areas identified as Andean wetlands.

3. MATERIALS AND METHODS

3.1. Study area

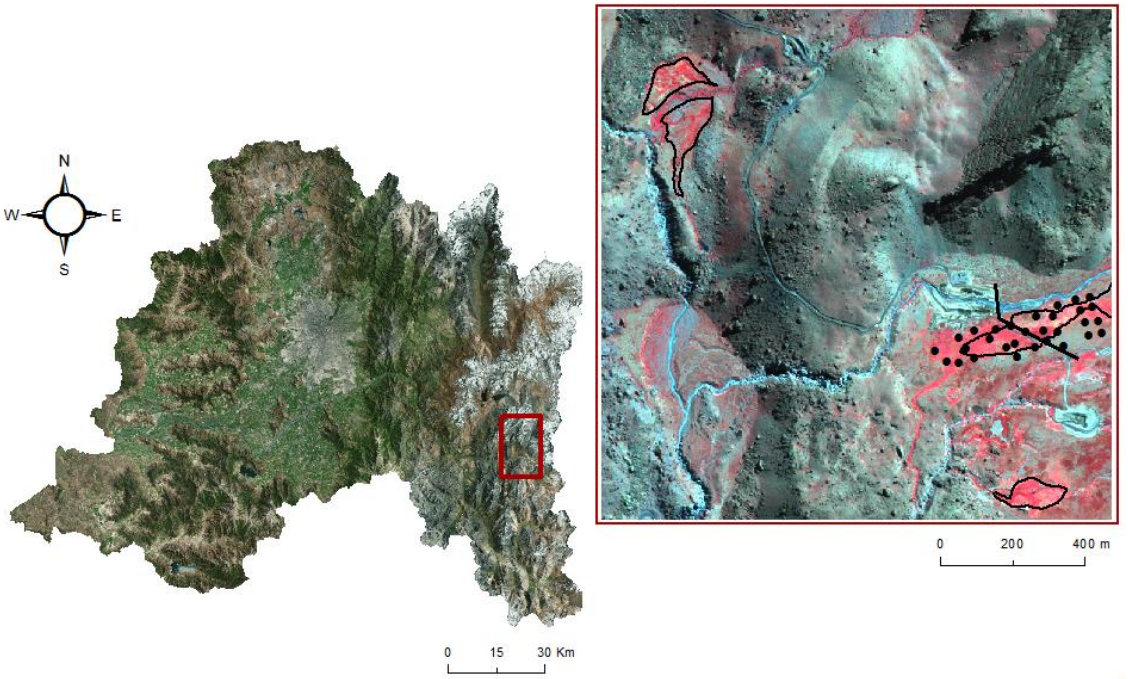

Fig. 1 Study area in San José del Maipo, Chile. Areas delineated with black lines correspond to the locations of Andean wetlands as listed in the national wetland inventory of the year 2011 (MMA-Centro de Ecología Aplicada, 2011). Dark dots show the location of the soil moisture sampling locations during the summer season.

The study area is situated in the high-alpine areas of the Maipo River Basin in the Metropolitana region of Chile (Fig. 1). Fieldwork was conducted on the Estero La Engorda sub basin (6259600 S−407100 W) which covers an area of 60,000 m². This area is mainly characterized by arid Mediterranean climate, with

precipitations occurring mainly as snow during the winter. Summers are dry, with occasional rain-storms. Currently, the wetlands located in La Engorda are being disturbed through the construction and installation of weirs and temporary roads supporting a hydroelectric installation nearby.

Fig. 2. High Andean vegetation of the Estero La Engorda.

Wet meadows are dominated by species with rhizome growth form and very small grasses (less than 40 cm in height), with dominant species from the genera Carex and Scirpus. This type of vegetation is called zonal (local) vegetation and it is determined by factors like precipitation, altitude and slope (Ahumada & Faúndez, 2009) but also depends on soil properties and humidity (Fig.2).

3.2 Ground reference data

During the field survey soil moisture was measured at 30 stratified random point samples during the summer period of 2015-2016, with repetitive measurements taken in December, January, February and March.

Table 1. Descriptive statistics of the soil moisture samples (g/g) during the summer season.

	2015	2016		
	December	January	February	March
Minimum	0.03	0.01	0.01	0.01
Median	0.27	0.34	0.27	0.23
Mean	1.02	1.11	1.06	1.08
Maximum	8.41	6.72	8.65	11.45
Coefficient of variation (%)	181	163	182	215
Shapiro-Wilk (P-value)	0.00000003	0.0000002	0.00000002	0.000000005

At each sampling point, soil moisture measurements were obtained from soil samples taken by a cylinder with a diameter of 52 mm (first 15 cm of soil). The soil samples were stored in plastic bags inside coolers to avoid moisture loss by evapotranspiration. In the laboratory, we followed the procedure of Sadzawka et al. (2006) to estimate the gravimetric soil moisture (g/g), where the soil samples were dried at 105 ± 5 °C for 48 h. Table 1 depicts the descriptive statistics of each field campaign. To evaluate the temporal variability of the soil moisture, a non-parametric Kruskal-Wallis test was performed, as the soil moisture samples did not reach the normality assumption according to the Shapiro-Wilk test; Table 1).

3.3 Remote sensing data

S We acquired WorldView-2 (WV-2) imagery from December 20th and March 20th. The data were delivered in radiance units, and were successively converted to reflectance at the top of the atmosphere using ENVI® (Exelis Visual Information Solutions, Boulder, CO, version 4.7) and the method proposed by Updike & Comp (2010). WorldView-2 collects multispectral information in 8 bands, covering the spectral ranges between 400−450 nm (coastal), 450−510 nm (blue), 510−581 nm (green), 585−625 nm (yellow), 630−690 nm (red), 705−745 nm (red edge), 770−895 nm (near-infrared 1; NIR 1) and 860−1040 nm (NIR 2), with a pixel size of 2 m.

Sentinel-2 (S-2) imagery was acquired through the “Copernicus Scientific Data Hub” webpage. We acquired imagery for the four dates most closely matching the dates of the field campaigns. S-2 has 13 bands with different spatial and spectral resolutions: 398−594 nm (coastal), 515−605 nm (green), 626−702 nm (red) and 690−980 nm (NIR) at 10 m resolution; four bands in the near-infrared (683−722, 722−758, 755−810 and 831−897 nm) and two bands in the short-wave infrared (1470−1756 and 1960−2444 nm; SWIR) at 20 m resolution; and three bands at 60 m resolution, at the coastal (415−740 nm), NIR (919−971 nm) and SWIR (1298−1447 nm). We used only the 10 and 20 m bands due to the small distance between field sampling points. The bands at 20 m were resampled to 10 m using the Sentinel-2 Toolbox of SNAP (Sentinel Application Platforms, version 5.0).

Atmospheric corrections were carried out for all images by applying the dark object subtraction method (Chavez, 1988). Furthermore, geometric corrections were performed in ENVI® using a SRTM digital elevation model (DEM) of 90 m. Finally, all images were brightness normalized using the method of Feilhauer et al. (2010) to maximize spectral differences of vegetated areas.

3.4 Mapping Andean high wetlands

3.4.1 Classification of Andean wetlands

To map the current extent of Andean wetlands we applied the one-class classifier biased support vector machines (BSVM; Liu et al., 2003). BSVM is a special form of binary SVM, that labels only the class of interest (positive class) while all other classes remain unlabelled ('background' class; Li & Guo, 2014; Mack et al., 2014). This reduces the amount of required in situ information significantly (Li et al., 2011; Song et al., 2016). This algorithm has been used successfully to map grasslands and wetland ecosystems using remote sensing data before (e.g. Mack et al., 2016; Stenzel et al., 2017).

We used the Chilean wetland inventory of the year 2011 (MMA-Centro de Ecología Aplicada, 2011) to obtain 200 samples for training the BSVM. The positions of the training samples were randomly selected from inside the areas indicated as wetland in the inventory data. 200 more random pixels were selected in the whole study area to provide samples for the 'background' class. This process based on a completely random selection, hence, the background class is likely to contain some pixels from the positive class. A 70% of the

selected pixels were used for training while the remaining 30% were consecutively used for validation.

We applied BSVM with a radial basis kernel (e.g. Mack et al., 2016; Stenzel et al., 2017) and the parameters ‘sigma’, ‘Cv’ and ‘Cb’ were tuned by a 10-fold cross-validation to find the best model. ‘Sigma’ is a function related to the kernel, ‘Cv’ is the penalty associated to classification errors and ‘Cb’ is the relative cost of the errors inside the ‘background’ class. The classifier was trained with 8 predictors in case of the WV-2 images and 10 predictors for the S-2 images (i.e. spectral bands).

To select the best model we consider two accuracy metrics often used in combination with one-class-classifiers: a) the probability of positive prediction (PPP), and b) the true positive rate (TPR; Mack et al., 2014). The model which maximized the TPR while minimized the PPP was selected. The final outputs of the one-class classifier are maps showing the probability of occurrence of the positive class (i.e. wetlands). The analysis was performed using the R-package ‘OneClass’ (Mack et al., 2014).

3.4.2 Predicting soil moisture

Soil moisture (SM) was modelled using WV-2 and S-2 imagery separately. In case of WV-2, only the field data obtained in December and March were used, while for S-2 two sets of data were used: 1) December and March (to compare with WV-2), and 2) using all available field data (December, January, February and March). Furthermore, to compare also the radiometric characteristic of both

sensors, a homogenization of the spatial and spectral resolution was performed, where: 1) the resolution of WV-2 was resampled (‘bilinear’) to 10 m to match S-2, and 2) only 6 bands of S-2 (the ones closest to the wavelength of WV-2) were used to match the spectral characteristics of WV-2.

We used partial least square regression (PLSR; Wold et al., 2001) trained with the reflectance values and the field-measured SM estimates to predict SM across all areas identified as wetlands during the preceding mapping procedure (see section 2.4.3.). PLSR is a well-known algorithm that summarizes the predictors’ information into fewer less correlated components (Wold et al., 2001). The models were optimized by a backward-selection of predictors (Martens & Martens, 2000), while the number of selected components (to minimize the possibility of overfitting) were selected by the leave-one-out cross-validated root mean square error (RMSEval); soil moisture was modelled with an increasing number of components as predictors, and the model with lower RMSEval was selected for each sensor.

Once the best model was selected (in terms of RMSEval), a bootstrap procedure with 500 iterations was applied for further validation. In each iteration, n observations were randomly selected with replacement from the n available samples. From this selection, approximately a 36.8% of the samples were not chosen in each iteration, and were used as holdout samples for the independent validation (Lopatin et al., 2016). Model performances were compared based on differences in the coefficients of determination (r^2; calculated as the squared Pearson's correlation coefficient), the normalized root mean square error (nRMSE) and the bias between predicted and observed LVs of the holdout

samples in the bootstrap. nRMSE is calculated as $nRMSE = (\sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}/[max(SM) - min(SM)]) * 100$, where n is the number of samples and SM is soil moisture. Likewise, the bias of prediction was measured as one minus the slope of a regression without intercept of the predicted versus observed values (Lopatin et al., 2016).

We tested other regression algorithms, such as random forest, support vector machines and K-nearest neighbours, which did not result in an improvement of the model; hence their results are not presented in the manuscript. The R-package 'autopls' was applied for the analysis (Schmidtlein et al., 2012).

Finally, we applied a one-sided bootstrapping test (Lopatin et al., 2016) to check for significant differences in the models ($\alpha = 0.05$; in terms of R^2 and nRMSE). We tested for significant differences between the classification results of the models trained with different number of observations (S-2 models using 120 and 60 observation), as well as for the sensors (homogenized models of WV-2 and S-2.)

3.4.3 Mapping soil moisture

Predictive maps of SM were calculated for each available image of the best obtained model. We used the probability maps depicted from the BSVM classification to mask out all non-wetlands areas (i.e. probabilities < 80%; threshold selected to only include high probability areas). Additionally, bootstrapping with 100 iterations was performed during the SM map predictions

to obtain the median and the coefficient of variation (CV, given in %) of SM values for each pixel. High values of CV indicate a less stable SM prediction.

4. RESULTS

4.1 Identification of Andean wetlands

The BSVM classifications resulted in probability maps showing the extent and location of all detected wetlands (Fig. 3). The confusion matrices (Table 2) show a general tendency for higher performances in March than in December. S-2 obtained an overall accuracy (OA) of ~90%, a producer accuracy (PA) of ~99% and a user accuracy (UA) of 77% for both images, while WV-2 depicted slightly lower accuracies (OA ~ 85%, PA ~ 94% and UA ~71%). Omission errors (i.e. exclusion) were small for both sensors, with values of 2% for S-2 and values between 4% and 8% for WV-2. Likewise, commission errors (i.e. false positives) were of 30% for WV-2 and 25% for S-2, showing a tendency for overestimating the wetlands presence.

Table 2. Classification results of the Biased Support Vector Machine classifications with Worldview-2 and Sentinel-2.

	Worldview-2		Sentinel-2	
	December	March	December	March
User accuracy (%)	71	72	77	77
Producer accuracy (%)	92	96	98	100
Commission error (%)	29	28	23	22
Omission Error (%)	8	4	2	2
Overall accuracy (%)	85	86	90	91

From all pixels classified as wetlands, only ~55% obtained probabilities of over 80% for both sensors. WV-2 registers an increase of the wetland area for Mach

in comparison to December (from 30 ha to 33 ha), while on the contrary S-2 accounted a decrease of the wetland area from December to March (from 25 ha to 23 ha) when considering only the pixels with a probability of over 80% (Table 3).

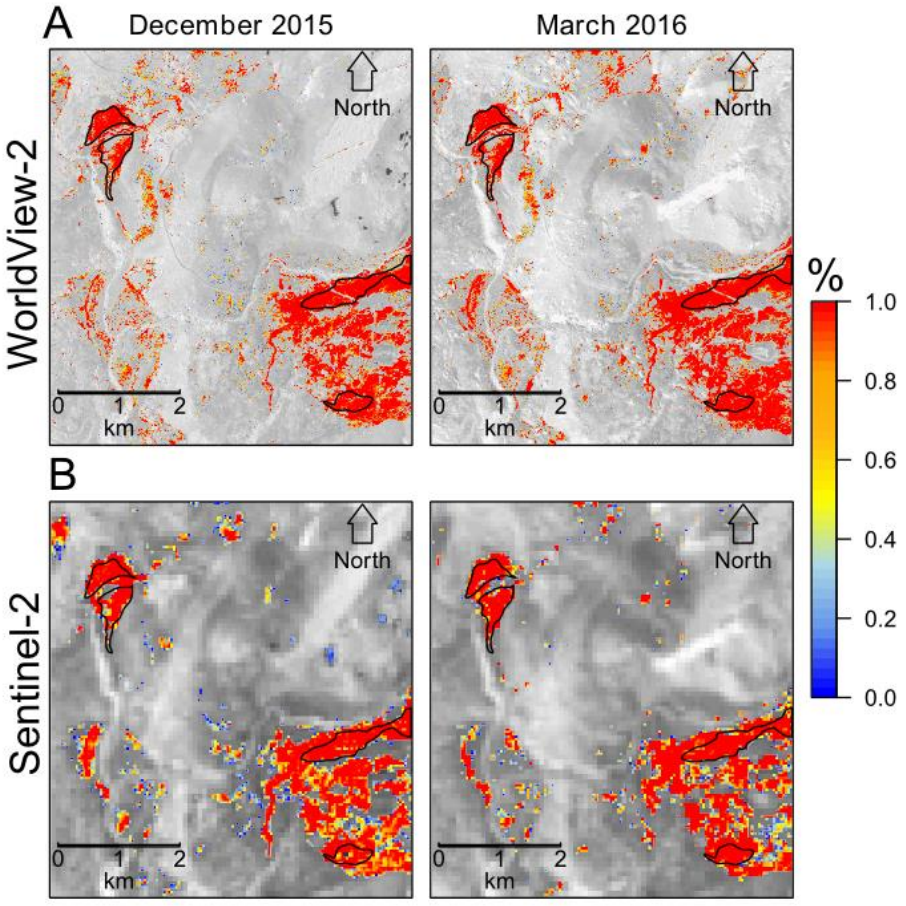

Fig. 3. Probability of occurrence (OCC) of Andean wetlands in La Engorda valley using WorldView-2 and Sentinel-2 imageries from the beginning (December) and ending (March) of the summer. Black lined polygons show the wetlands presented in the Wetland National Inventory of 2011 (MMA-Centro de Ecología Aplicada, 2011). Areas with zero OCC were masked out.

Table 3. Resulting wetland cover estimates obtained with BSVM and the sensors WV-2 and S-2.

Sensor	Date	Total area (ha)	Classified area as meadow (ha)	Classified area as meadow with probability > 80% (ha)
WV-2	December	241,54	58	30
	March		58	33
S-2	December	241,24	41	25
	March		35	23

4.2 Soil moisture prediction

In Table 4 the result of all tested models using PLSR and leave-one-out cross-validation are presented. Generally, models using S-2 imagery showed higher performances. The best model was obtained by including samples from all four available field measurement dates (R^2val = 0.58 and nRMSEval = 15.30%). The accuracy of the WV-2 model increased when enlarging the pixel side (to 10 m to match S-2). Nevertheless, when matching the two sensor characteristics (i.e., both with 10 m pixel and 6 bands with similar wavelengths), S-2 models still obtained significantly ($\alpha = 0.05$) higher fits than the WV-2 models (R^2val = 0.40 and nRMSEval = 18.20%). In Fig. 4, the bootstrapped distribution of accuracies using the best model (i.e. S-2 with 120 observations) can be seen, where the model reaches median values of $R^2 = 0.63$, nRMSE = 11.1% and bias = 0.28. The model shows that small SM values tend to be slightly overestimated, while middle and high SM values tend to be underestimated.

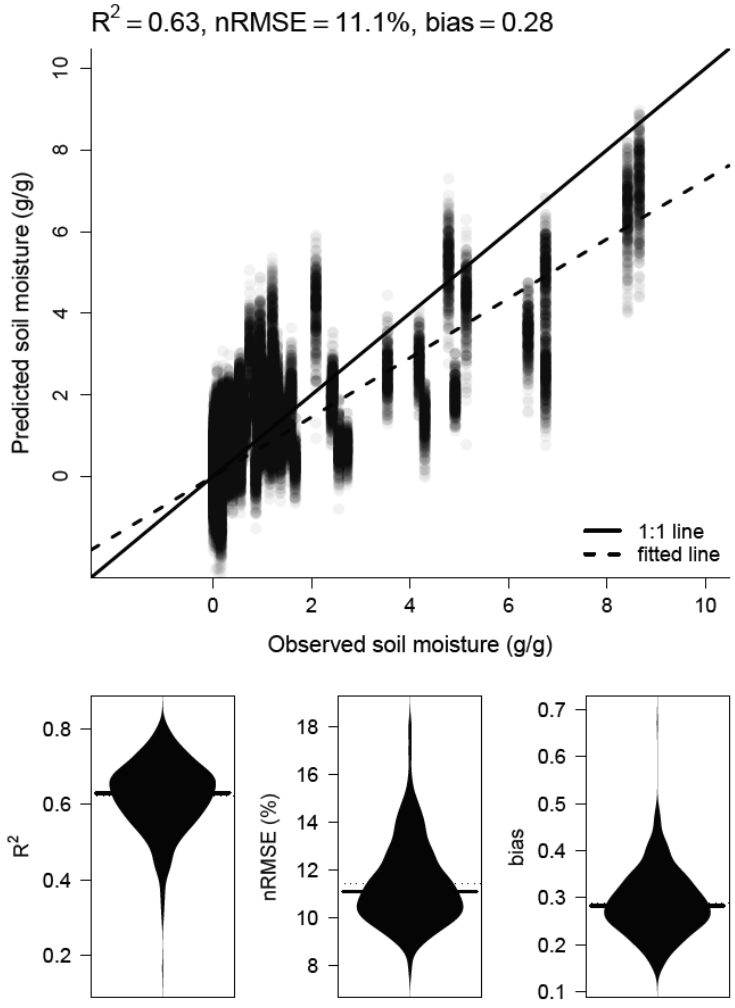

Fig. 4. Bootstrapping validation for the models using observation from all summer season. On the top, a scatterplot of predicted against observed values of soil moisture is presented while on the bottom the bootstrapped distribution of the accuracies are displayed.

Table 4. Summary of soil moisture (g/g) predictions by WV-2 and S-2 with PLSR. * denote a significant improvement (α = 0.05) of the model accuracy.

Model	Sensor	Dataset	Number of observation	Spatial resolution (m)	Number of bands	R^2val	nRMSE val (%)
Original information							
1	WV-2	Dec/Mar	60	2	6	0.11	22.16
2	S-2	Dec/Mar	60	10	10	0.51	16.39
3	S-2	Dec/Jan/ Feb/Mar	120	10	10	0.58	15.30
Homogenized information							
4	WV-2	Dec/Mar	60	10	6	0.25	20.44
5	S-2	Dec/Mar	60	10	6	0.40*	18.20*

Finally, median and coefficient of variation (CV) prediction maps for December and March are presented in Fig. 5. The patterns agree with the field observations, where the central areas and areas closest to the road intervention are highly saturated with water while the exterior areas show less soil moisture. Moreover, areas depicting higher SM values were more stable during bootstrapping (CV < 12%), while driest areas were highly variable (CV > 12%).

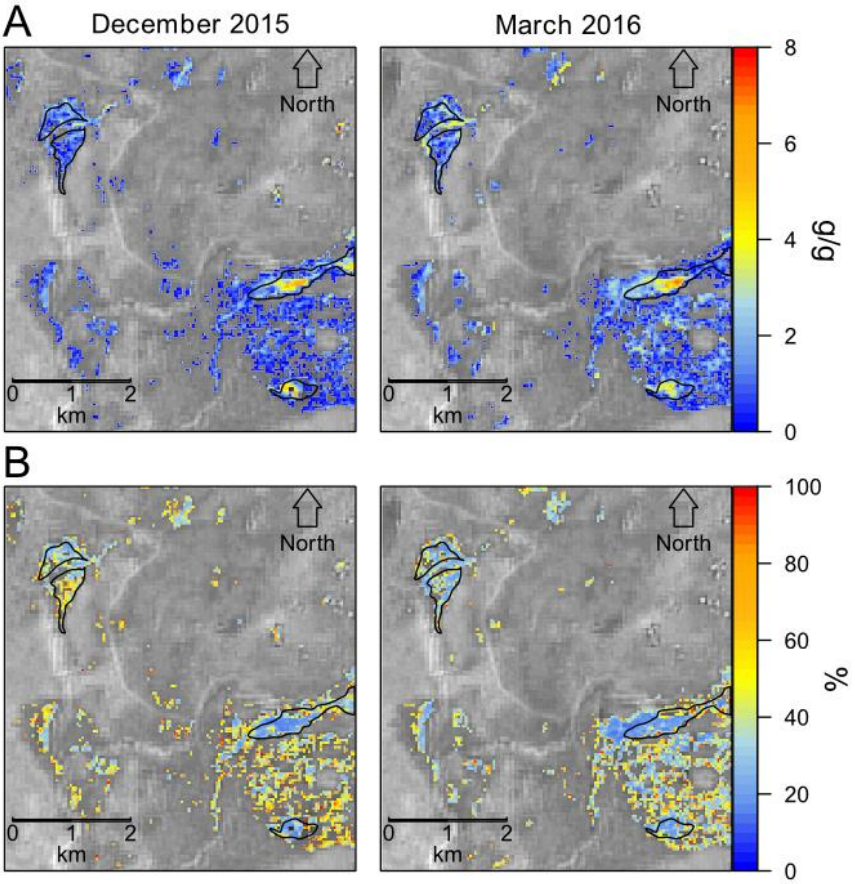

Fig. 5. Soil moisture prediction maps for the beginning (December) and end (March) of the summer season. In A, the mean bootstrap value per pixel is showed; while in B the coefficient of variation (CV in %) per pixel is presented. Black lined polygons shows the wetlands presented in the Wetland National Inventory of 2011 (MMA-Centro de Ecología Aplicada, 2011). Predictions were carried out only within areas with a wetland occurrence probability of over 80% according to the BSVM classification.

5. DISCUSSION

5.1 Classification of Andean wetlands

The results showed that it is feasible to accurately map the presence and extent of Andean wetlands by optical RS with few field efforts. The BSVM classification resulted in high performances with both sensors, with a slight tendency to overestimate the wetland area (commission errors of ~30%). In our case this commission error is likely to at least partly be explainable by the incompleteness of the reference dataset which is known to underestimate the number and the extent of Andean wetlands (MMA-Centro de Ecología Aplicada, 2011; see section 4.3 for an explanation of how these areas were mapped in the reference dataset). On the other hand, such errors can also be attributed to the structure of the so called 'background' class, where a random selection of pixels (including from the positive class) is performed. This process has been reported to lead to commission errors in earlier studies. For example, Skowronek et al. (2017) reported commission errors of 34% when mapping the invasive species Phalaris aquatica. Moreover, Stenzel et al. (2014) mapped alkaline fens and raised bogs with similar commission errors, while Mack et al. (2016) obtained high levels of overestimation of raised bogs (~75%). This underlines the importance of the 'background' class when working with the BSVM classifier. Recently, Mack et al. (2016) developed an iterative process to 'clean' the 'background' class as much as possible of pixels of the positive class. This procedure improved their classification results for mapping raised bogs by ~26%. In our case, the separation of the positive class (i.e. wetlands) from the rest of the landscape was accurate enough without any previous iterative

process (i.e. overall accuracies over ~85%), as the spectral characteristics of the non-wetlands are clearly separable from the wetland areas. Not performing the iterative step saves a considerable amount of CPU process time (Mack et al., 2016), which is an important factor to account for when monitoring large areas.

One advantage of BSVM is that the classification maps (Fig. 3) can be displayed as continuous probability of occurrence maps (OCC given in %). Contrary to other binary classifiers that deliver result with sharp boundaries, BSVM hence allows for the detection of continuous patterns in the landscape, which agree more with the graduate gradients of water, nutrient and plant covers in nature (Reschke & Hüttich, 2014; Tiner, 2015).

Here, we used each monthly field samples independently for the classification. Similar approaches were presented by e.g. (Lee & Yeh (2009) and McCarthy et al., (2015), which used optical RS of single dates to classify mangroves and coastal marshes respectively. Contrary, most of the previous studies focusing on wetland mapping used seasonal information to enhance the classification (i.e. higher range of photosynthetic activity; e.g. Adam et al., 2010; Davranche et al., 2010; Ozesmi & Bauer, 2002; Townsend et al., 2001). Seasonal information is particularly valuable if the target wetland ecosystem is difficult to classify from the other landscape types occurring in the background. In our case, the high altitudinal Andes wetlands had a spectral signature which differed notably from the landscape which allowed us obtain good classification accuracies (overall accuracies > 85%) without adding seasonal information (e.g. McCarthy et al., 2015).

Concerning spatial resolution, Frohn et al. (2012) and McCarthy et al. (2015) agreed that high spatial resolution imagery is a key component to take into account when classifying wetlands at a local scale. Here, both sensors achieved similar accuracies, which indicates that for mapping Andean wetlands, a spatial resolution of 10 m as provided by Sentinel-2 is adequate. Nevertheless, differences in the areas classified as wetlands by the two sensors were up to ~5-10 ha. WV-2 showed systematically larger estimates of wetland areas than S-2. One reason for this could be that WV-2 is capable to identify smaller isolated wetland areas because of its higher spatial resolution compared to Sentinel-2.

5.2. Prediction of soil moisture by optical sensors

Soil moisture (SM) values for ground samples showed a high variability (CV > 100%, Table 1) across small distances with hardly any clear gradient. Furthermore, the samples were highly skewed (i.e. high presence of small values and few large ones), probably due to the small sample size (Hills & Reynolds, 1969). This resulted in a low range of observed SM values with the potential presence of outliers (Kaleita et al., 2005) in water-saturated areas. In our case, the elimination of outliers did not translate into any improvement of the models, hence we kept all samples in the model.

Soil moisture is also known to be influenced by seasonal effect. Hills & Reynolds (1969) pointed out four decades ago the importance of an adequate sample size and the selected seasonal time. Extreme sample periods (e.g. beginning and

end of summer) will enhance the sample variability on such environments as Andean highlands (Western et al., 2003).

From all tested SM models, S-2 data with all 120 observation reached highest accuracies. Applying reference data from all four field sampling dates instead of just the beginning and the end of the season improved the accuracies, especially in terms of R^2 (from 0.51 to 0.58), while the difference in terms of nRMSE was not pronounced (non-significant differences were found for the sample size differences; residuals of the best model can be seen in Fig. A1). S-2 showed significantly better performance than WV-2 for predicting SM ($\alpha = 0.05$). We relate this to S-2's two bands in the SWIR which were found to have a high importance for estimating SM (see Table A1 for variable importance of PLSR). The importance of the SWIR region to estimate SM has been stated in many earlier studies (e.g. Chen et al., 2014; Harris et al., 2006; Kaleita et al., 2005; Wang et al., 2008). Surprisingly, even when using only bands in the VIS and NIR spectral regions, S-2 still obtained higher performances than WV-2. This is maybe due the higher spectral resolution of S-2 which features bands with a narrower spectral range than WV2, and hence might be able to depict more detailed spectral information related to SM.

The bootstrap validation of the selected model showed normal distributions or R^2, nRMSE and bias values, with a systematic tendency to underestimate higher values of SM. This behaviour is usual in this kind of models, where optical sensors are not able to differentiate properly gradients of high or saturated values of SM. Some earlier studies consequently masked these areas out for the analysis (e.g. Sadeghi et al., 2017). Fig. 5 shows that the highest

uncertainties in terms of a high coefficient of variation of SM predictions were observed in areas with low SM values. We assume that this relates to the fact that in the examined Andean wetlands, the vegetation cover is the key-indicator for the estimation of SM with areas showing a higher vegetation cover typically being located at more humid areas. Nevertheless, the significant difference between S-2 and WV-2 shows that the SWIR information is key when accurately estimating SM in areas with similar plant cover but differing SM values. According to Muller & Decamps (2011) and Kaleita et al. (2005), drier soils tend to present erratic reflectance that complicate an accurate SM prediction. This effect may have been amplified in our case by the bias toward lower SM values in the field observations which is likely to also increase the variability of the observed reflectance matching a low SM situation in the field. In future studies, it should be ensured that the full range of SM values is sampled in a more or less balanced way.

5.3 Potential of the examined sensor systems for operational monitoring of Andean wetlands in the framework of the Chilean National Wetland Inventory.

Chile has an historical deficit of ecological information about its wetlands, with Andean high altitude wetlands being one of the least studied wetland ecosystems (Cepeda-Pizzaro & Pola, 2013; Squeo et al., 2006). Even though Chile joined the Ramsar convention already in 1981, the first general classification and detection of national wetlands within the framework of the National Wetland Inventory (MMA-Centro de Ecología Aplicada, 2011) was only conducted in 2011. This inventory was carried using Landsat imagery, where water bodies, vegetated areas and vegetation with visible water were classified

by using the normalized difference water index (NWI) and the normalized vegetation index (NDVI). Validation was carried out by means of two sources of information: 1) a water bodies layer elaborated by the military geographical institute (Instituto Geográfico Militar (IGM)), with a scale of 1:50,000, and 2) with GIS information created by the Ministry of Environment in the northern area, by using Landsat imagery and field information. Hence, the current inventory methodology is not exhaustive and it does not use reliable validation data, which may hamper a proper monitoring and conservation of these ecosystems.

Here, we proposed the use of optical Sentinel-2 data has a reliable alternative for Andean high wetlands monitoring for several reasons:

Sentinel 2 data is free of cost: Hence, the data can be used for long term monitoring in conjunction with few sampling efforts, which is an important issue for governmental management. The high temporal resolution of Sentinel 2 furthermore allows for the implementation of an operational monitoring system. Reference data could be obtained with a sparse number of field campaign and additional high-resolution RGB images (e.g. GoogleEarth) which are nowadays freely available.

Spatial resolution: Sentinel-2 has higher spatial resolution than other freely available optical RS data with SWIR information, such as Landsat and MODIS, and clearly outcompetes other sensor systems used for SM estimations (e.g. using passive microwaves with or without thermal information (e.g. Merlin et al., 2012a; Zribi et al., 2005) and active radar (e.g. Gao et al., 2017)) with regard to

spatial detail. For the Andean wetlands this is a key requirement due to their small size and often patchy occurrences.

Spectral properties: Sentinel-2 was found to deliver sufficient spectral details to adequately identify Andean wetlands from all other land-cover classes in this study. The direct comparison to the commercial WorldView-2 data underlined the very high radiometric quality of the sensor system.

All these characteristics turn S-2 into a robust alternative for monitoring Andean wetlands. The presented classification approach may allow the establishment of new datasets to actualize the National Wetland Inventory. The monitoring of soil moisture gradients may furthermore support the early detection of impacts caused by anthropogenic impacts or global warming.

6. CONCLUSIONS

Optical remote sensing was found to be a reliable source of information to map Andean wetlands in complicated terrain. Models based on Sentinel-2 data showed high performances for mapping the occurrence and extent of Andean wetlands and for estimating soil moisture content within the wetlands. According to the aims of the study we can conclude:

The one-class classifier Bias Support Vector Machines was found to deliver accurate results for the delineation of Andean wetlands when using single-date imagery and minimal field data. Both WorldView-2 and Sentinel-2 imagery deliver similar results demonstrating that the separation of the wetlands from the landscape background is possible without information from the short-wave infrared. This is important when the monitoring of detailed areas is needed. In this regard, the use of new generations of very-high resolution imagery, such as WorldView-2 and 3, Pleiades or the new Venus, is especially interesting.

To predict soil moisture SWIR information is fundamental, especially in areas with similar plant cover but differing soil moisture values. In our study Sentinel-2 allowed the monitoring of Andean wetlands with sufficient accuracy. It could hence be used for monitoring and management purposes, particularly in environments where the size of some of the wetlands are too small for the use of traditional passive microwaves-thermal and active radar approaches.

The presented approach and results may be of high importance for the monitoring and management of Andean wetlands in Chile in a local and regional

scale, and particularly in areas where the anthropogenic pressures over water resources is elevated due to the mining industry.

Appendix A. Best model information

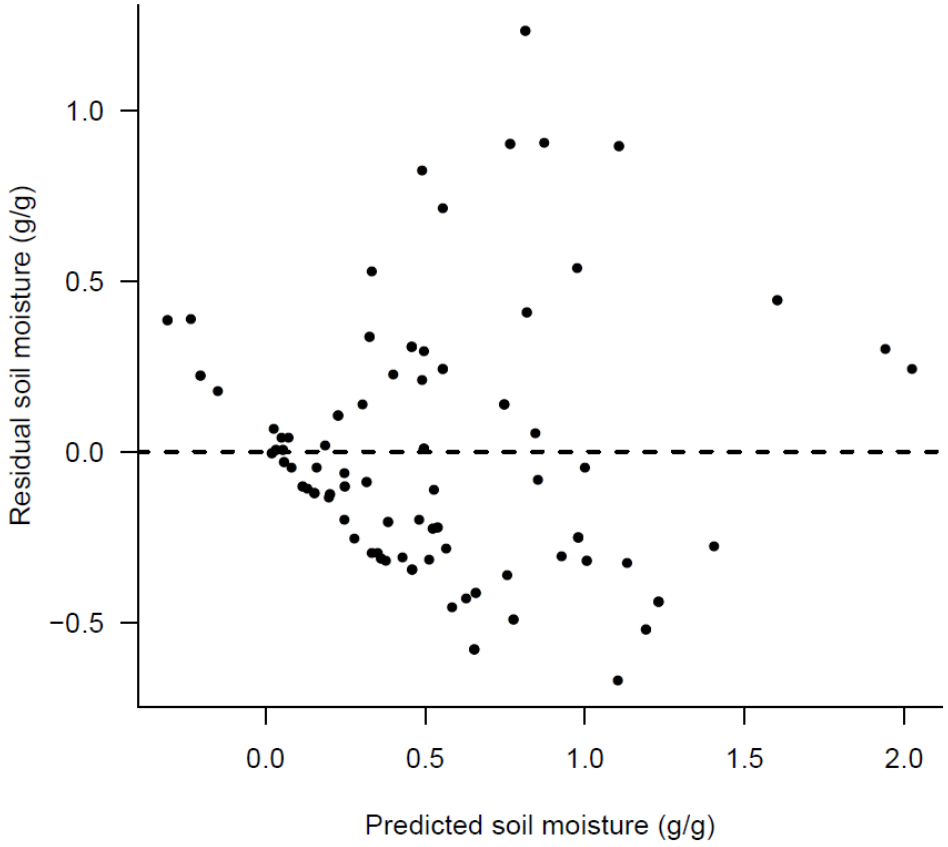

Fig. A1. Model residuals of the model using all summer season Sentinel-2 data.

Table A1. Variable importance of the best model using partial least square regression (PLSR). The importances correspond to the model coefficients.

Band	Wavelength (nm)	Resolution (m)	Description	P-value	Sig. level
B2	490	10	Blue	0.001	*
B3	560	10	Green	0.139	-
B4	665	10	Red	1.097e-05	***
B5	705	20	IR	7.495e-05	***
B6	740	20	IR	5.499e-05	***
B7	783	20	IR	0.121	-
B8	842	10	IR	0.155	-
B8a	865	20	IR	0.0006	*
B11	1610	20	SWIR	0.032	*
B12	2190	20	SWIR	0.648	-

Significance level: '***' 0.001; '**'; 0.01; '*' 0.05 '.

7. BIBLIOGRAPHY

Adam, E., Mutanga, O. & Rugege, D. (2010). Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: A review. Wetl. Ecol. Manag. 18, 281–296. doi:10.1007/s11273-009-9169-z

Ahumada, M., Aguirre, F., Contreras, M. & Figueroa, A. (2011). Guía para la Conservación y Seguimiento Ambiental de Humedales Andinos. Minist. del Medio Ambient. Santiago, Chile.

Ahumada, M., & Faúndez, L. (2009). Guía descriptiva de los sistemas vegetacionales azonales hídricos terrestres de la ecorregión altiplánica (SVAHT). Minist. Agric. Chile, Serv. Agrícola y Ganad. Santiago, Chile.

Baker, C., Lawrence, R., Montagne, C. & Patten, D. (2006). Mapping wetlands and riparian areas using Landsat ETM+ Imagery and decision-tree-based models. Soc. Wetl. Sci. Wetl. 26, 465–474.

Cepeda-Pizzaro, J., & Pola, M.L. (2013). Relaciones de abundancia de órdenes de hexápodos terrestres en vegas altoandinas del desierto-transicional de Chile. Idesia 31, 31–39. doi:10.4067/S0718-34292013000200005

Chavez, P.A.T.S. (1988). An Improved Dark-Object Subtraction Technique for Atmospheric Scattering Correction of Multispectral Data. Remote Sens. Environ. 24, 459–479.

Chen, C.F., Valdez, M.C., Chang, N. Bin, Chang, L.Y. & Yuan, P.Y. (2014). Monitoring spatiotemporal surface soil moisture variations during dry seasons in central america with multisensor cascade data fusion. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7, 4340–4355. doi:10.1109/JSTARS.2014.2347313

Contreras, R. (2007). Uso de vegas y bofedales de la zona cordillerana y precordillerana de la región de Atacama. Universidad de Chile.

Davranche, A., Lefebvre, G. & Poulin, B. (2010). Wetland monitoring using classification trees and SPOT-5 seasonal time series. Remote Sens. Environ. 114, 552–562. doi:10.1016/j.rse.2009.10.009

Davranche, A., Poulin, B. & Lefebvre, G. (2013). Mapping flooding regimes in Camargue wetlands using seasonal multispectral data. Remote Sens. Environ. 138, 165–171. doi:10.1016/j.rse.2013.07.015

Feilhauer, H., Asner, G.P., Martin, R.E. & Schmidtlein, S. (2010). Brightness-normalized Partial Least Squares Regression for hyperspectral data. J. Quant. Spectrosc. Radiat. Transf. 111, 1947–1957. doi:10.1016/j.jqsrt.2010.03.007

Frohn, R., D'Amico, E., Lane, C., Autrey, B., Rhodus, J. & Liu, H. (2012). Multi-temporal sub-pixel Landsat ETM+ classification of isolated wetlands in Cuyahoga County, Ohio, USA. Wetlands 32, 289–299.

Gao, Q., Zribi, M., Escorihuela, M.J. & Baghdadi, N. (2017). Synergetic Use of Sentinel-1 and Sentinel-2 Data for Soil Moisture Mapping at 100 m Resolution. Sensors 17.

García, E., & Otto, M. (2015). Caracterización ecohidrológica de humedales alto andinos usando imágenes de satélite multitemporales en la cabecera de cuenca del Río Santa, Ancash, Perú. Ecol. Apl. 14, 115–125.

GORE-RMS (2013). Estrategia Regional para la Conservación de la Biodiversidad en la Región Metropolitana de Santiago 2015-2025.

Harris, A., Bryant, R.G. & Baird, A.J. (2006). Mapping the effects of water stress on Sphagnum: Preliminary observations using airborne remote sensing. Remote Sens. Environ. 100, 363–378. doi:10.1016/j.rse.2005.10.024

Hills, R.C., & Reynolds, S.G. (1969). Illustrations of soil moisture variability in selected areas and plots of different sizes. J. Hydrol. 8, 27–49.

Kaleita, a L., Tian, L.F. & Hirschi, M.C. (2005). Relationship between soil moisture content and soil surface reflectance. Trans. Asae 48, 1979–1986. doi:doi: 10.13031/2013.19990

Kerr, Y.H., Waldteufel, P. & Richaume, P. (2012). The SMOS Soil Moisture Retrieval Algorithm. IEEE Trans. Geosci. Remote Sens. 50, 1384–1403.

Lee, T., & Yeh, H. (2009). Applying remote sensing techniques to monitor shifting wetland vegetation: A case study of Danshui River estuary mangrove communities, Taiwan. Ecol. Eng. 35, 487–496.

Li, W., & Guo, Q. (2014). A new accuracy assessment method for one-class remote sensing classification. IEEE Trans. Geosci. Remote Sens. 52, 4621–4632. doi:10.1109/TGRS.2013.2283082

Li, W., Guo, Q. & Elkan, C. (2011). A Positive and Unlabeled Learning Algorithm for One-Class Classification of Remote-Sensing Data. IEEE Trans. Geosci. Remote Sensing2 49, 717–725.

Liu, B., Dai, Y., Li, X., Lee, W.L. & Yu., P.S. (2003). Building Text Classifiers Using Positive and Unlabeled Examples, in: Third IEEE International Conference on Data Mining, Melbourne, Nov 19-22. pp. 179–186.

Lopatin, J., Dolos, K., Hernández, H.J., Galleguillos, M. & Fassnacht, F.E. (2016). Comparing Generalized Linear Models and random forest to model vascular plant species richness using LiDAR data in a natural forest in central Chile. Remote Sens. Environ. 173, 200–210. doi:10.1016/j.rse.2015.11.029

MacAlister, C., & Mahaxay, M. (2009). Mapping wetlands in the Lower Mekong Basin for wetland resource and conservation management using Landsat ETM images and field survey data. J. Environ. Manage. 90, 2130–2137. doi:https://doi.org/10.1016/j.jenvman.2007.06.031

Mack, B., Roscher, R., Stenzel, S., Feilhauer, H., Schmidtlein, S. & Waske, B. (2016). Mapping raised bogs with an iterative one-class classification approach. ISPRS J. Photogramm. Remote Sens. 120, 53–64. doi:10.1016/j.isprsjprs.2016.07.008

Mack, B., Roscher, R. & Waske, B. (2014). Can i trust my one-class classification? Remote Sens. 6, 8779–8802. doi:10.3390/rs6098779

Martens, H., & Martens, M. (2000). Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). Food Qual. Prefer. 11, 5–16. doi:http://dx.doi.org/10.1016/S0950-3293(99)00039-7

McCarthy, M.J., Merton, E.J. & Muller-Karger, F.E. (2015). Improved coastal wetland mapping using very-high 2-meter spatial resolution imagery. Int. J. Appl. Earth Obs. Geoinf. 40, 11–18. doi:10.1016/j.jag.2015.03.011

Merlin, O., Escorihuela, M.J., Mayoral, M.A., Hagolle, O., Al Bitar, A. & Kerr, Y.H. (2012a). Self-calibrated evaporation-based disaggregation of SMOS soil moisture: An evaluation study at 3 km and 100 m resolution in Catalunya, Spain. Remote Sens. Environ. 10.1016/j.rse.2012.11.008. doi:10.1016/j.rse.2012.11.008

Merlin, O., Rudiger, C., Bitar, A. Al, Richaume, P., Walker, J.P. & Kerr, Y.H. (2012b). Disaggregation of SMOS Soil Moisture in Southeastern Australia. IEEE Trans. Geosci. Remote Sens. 50, 1556–1571.

MMA-Centro de Ecología Aplicada (2011). Diseño del inventario nacional de humedales y el seguimiento ambiental.

Molero, B., Merlin, O., Malbéteau, Y., Bitar, A. Al, Cabot, F., Stefan, V., Kerr, Y., Bacon, S., Cosh, M.H., Bindlish, R. & Jackson, T.J. (2016). SMOS disaggregated soil moisture product at 1 km resolution: Processor overview and first validation results. Remote Sens. Environ. 180, 361–376. doi:https://doi.org/10.1016/j.rse.2016.02.045

Muller, E., & Decamps, H. (2011). Modeling soil moisture reflectance. Remote Sens. Environ. 76, 173–180.

Muñoz-Schick, M., Núñez, H. & Yáñez, J. (1997). Libro rojo de los sitios prioritarios para la conservación de la biodiversidad en Chile. Ambient. y Desarro. 13, 90–99.

Ozesmi, S.L., & Bauer, M.E. (2002). Satellite remote sensing of wetlands. Wetl. Ecol. Manag. 10, 381–402. doi:10.1023/A:1020908432489

Reschke, J., & Hüttich, C. (2014). Continuous field mapping of Mediterranean wetlands using sub-pixel spectral signatures and multi-temporal Landsat data. Int. J. Appl. Earth Obs. Geoinf. 28, 220–229. doi:10.1016/j.jag.2013.12.014

Rundquist, D.C., Narumalani, S. & Narayanan, R.M. (2001). A review of wetlands remote sensing and defining new considerations. Remote Sens. Rev. 20, 207–226. doi:10.1080/02757250109532435

Sadeghi, M., Babaeian, E., Tuller, M. & Jones, S.B. (2017). The optical trapezoid model: A novel approach to remote sensing of soil moisture applied to Sentinel-2 and Landsat-8 observations. Remote Sens. Environ. 198, 52–68.

Sadzawka, A., Carrasco, M.A., Grez, R., Mora, M.L., Flores, P. & Neaman, A. (2006). M{é}todos de an{á}lisis de suelos recomendados para los suelos de Chile. Instituto de Investigaciones Agropecuarias, Santiago, Chile.

Schmidtlein, S., Feilhauer, H. & Bruelheide, H. (2012). Mapping plant strategy types using remote sensing. J. Veg. Sci. 23, 395–405. doi:10.1111/j.1654-1103.2011.01370.x

Skowronek, S., Asner, G.P. & Feilhauer, H. (2017). Performance of one-class classifiers for invasive species mapping using airborne imaging spectroscopy. Ecol. Indic. 37, 66–76.

Song, B., Li, P., Li, J. & Plaza, A. (2016). One-Class Classification of Remote Sensing Images Using Kernel Sparse Representation. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9, 1613–1623. doi:10.1109/JSTARS.2015.2508285

Squeo, F., Cepeda, J., Olivares, N.C. & Arroyo, M.T.K. (2006). Interacciones ecológicas en la alta montaña Del Valle Del Elqui, in: GEOECOLOGÍA de Los ANDES Desérticos. La Alta Montaña Del Valle Del Elqui. Ediciones Universidad de La Serena, La Serena, pp. 69–103.

Squeo, F., Ibacache, E., Warner, B. & Espinoza, D. (2006a). Productividad y Diversidad Florística De La Vega Tambo, in: GEOECOLOGÍA de Los ANDES Desérticos. La Alta Montaña Del Valle Del Elqui. Ediciones Universidad La Serena, La Serena, pp. 325–351.

Squeo, F., Warner, B., Aravena, R. & Espinoza, D. (2006b). Bodedales: high altitude peatlands of the central Andes. Rev. Chil. Hist. Nat. 79, 245–255. doi:10.4067/S0716-078X2006000200010

Stenzel, S., Fassnacht, F.E., Mack, B. & Schmidtlein, S. (2017). Identification of high nature value grassland with remote sensing and minimal field data. Ecol. Indic. 74, 28–38. doi:http://dx.doi.org/10.1016/j.ecolind.2016.11.005

Stenzel, S., Feilhauer, H., Mack, B., Metz, A. & Schmidtlein, S. (2014). Remote sensing of scattered natura 2000 habitats using a one-class classifier. Int. J. Appl. Earth Obs. Geoinf. 33, 211–217. doi:10.1016/j.jag.2014.05.012

Tiner, R. (2015). Classification of wetland types for mapping and large-scale inventories, in: Remote Sensing of Wetland. Applications and Advances. Taylor & Francis Group, Boca Raton, Florida, pp. 19–41.

Todd Updike, & Comp, C. (2010). Radiometric Use of WorldView-2 Imagery Technical Note 1 WorldView-2 Instrument Description.

Townsend, P. a, Walsh, S.J., Ecology, S.P., Dec, N., Townsend, a, Stephen, J. & J, S. (2001). Remote Sensing of Forested Wetlands : Application of

Multitemporal and Multispectral Satellite Imagery to Determine Plant Community Composition and Structure in Southeastern Stable URL : http://www.jstor.org/stable/20051167 and sensing of forested wetland. Plant Ecol. 157, 129–149. doi:10.1023/A:1013999513172

Wang, B., Chen, Y. & Lü, C. (2015). Evaluating flood inundation impact on wetland vegetation FPAR of the Macquarie marshes, Australia. Environ. Earth Sci. 74, 4989–5000.

Wang, L., Qu, J.J., Hao, X. & Zhu, Q. (2008). Sensitivity studies of the moisture effects on MODIS SWIR reflectance and vegetation water indices. Int. J. Remote Sens. 29, 7065–7075. doi:10.1080/01431160802226034

Western, A., Zhou, S., Grayson, B.R., McMahon, A.T., Blöschl, G. & Wilson, J.D. (2003). Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes. J. Hydrol. 286, 113–134.

Whiteside, T.G., & Bartolo, R.E. (2015). Mapping aquatic vegetation in a tropical wetland using high spatial resolution multispectral satellite imagery. Remote Sens. 7, 11664–11694.

Wold, S., Sjöström, M. & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. Chemom. Intell. Lab. Syst. 58, 109–130. doi:http://dx.doi.org/10.1016/S0169-7439(01)00155-1

Zribi, M., Baghdadi, N., Holah, N. & Fafin, O. (2005). New methodology for soil surface moisture estimation and its application to ENVISAT-ASAR multi-incidence data inversion. Remote Sens. Environ. 96, 485–496. doi:https://doi.org/10.1016/j.rse.2005.04.005

Zribi, M., & Decharme, B. (2008). A Method for Soil Moisture Estimation in Western Africa Based on the ERS Scatterometer. IEEE Trans. Geosci. Remote Sens. 46, 438–448.