Evaluating WRF-Chem Aerosol Indirect Effects in Southeast Pacific Marine Stratocumulus During VOCALS-REx

P.E. Saide  
*University of Iowa*

S.N. Spak  
*University of Iowa*

G.R. Carmichael  
*University of Iowa*

M.A. Mena-Carrasco  
*Universidad Andrés Bello, Santiago, Chile*

Q. Yang  
*Pacific Northwest National Laboratory, Richland, WA*

*See next page for additional authors*

Follow this and additional works at: [http://repository.uwyo.edu/geology_facpub](http://repository.uwyo.edu/geology_facpub)

Publication Information
Evaluating WRF-Chem aerosol indirect effects in Southeast Pacific marine stratocumulus during VOCALS-REx


1Center for Global and Regional Environmental Research (CGRER), University of Iowa, Iowa City, Iowa, USA
2Center for Sustainability Research, Universidad Andrés Bello, Santiago, Chile
3Pacific Northwest National Laboratory, Richland, WA, USA
4Department of Oceanography, University of Hawaii at Manoa, Honolulu, USA
5Department of Atmospheric Science, University of Wyoming, Laramie, Wyoming, USA
6Department of Chemistry, Drexel University, Philadelphia, PA, USA
7Colorado State University, Department of Atmospheric Science, Fort Collins, CO, USA
8Oregon State University, College of Earth, Ocean, and Atmospheric Sciences, Corvallis, OR, USA
9Harvey Mudd College, Department of Chemistry, Claremont, CA, USA
10Centre for Atmospheric Science, University of Manchester, Manchester, M13 9PL, UK
11Brookhaven National Laboratory, USA

Correspondence to: P. E. Said (pablo-saide@uiowa.edu)

Received: 16 October 2011 – Published in Atmos. Chem. Phys. Discuss.: 4 November 2011
Revised: 9 February 2012 – Accepted: 8 March 2012 – Published: 29 March 2012

Abstract. We evaluate a regional-scale simulation with the WRF-Chem model for the VAMOS (Variability of the American Monsoon Systems) Ocean-Cloud-Atmosphere-Land Study Regional Experiment (VOCALS-REx), which sampled the Southeast Pacific’s persistent stratocumulus deck. Evaluation of VOCALS-REx ship-based and three aircraft observations focuses on analyzing how aerosol loading affects marine boundary layer (MBL) dynamics and cloud microphysics. We compare local time series and campaign-averaged longitudinal gradients, and highlight differences in model simulations with (W) and without (NW) wet deposition processes. The higher aerosol loadings in the NW case produce considerable changes in MBL dynamics and cloud microphysics. We compare local time series and campaign-averaged longitudinal gradients, and highlight differences in model simulations with (W) and without (NW) wet deposition processes. The higher aerosol loadings in the NW case produce considerable changes in MBL dynamics and cloud microphysics, in accordance with the established conceptual model of aerosol indirect effects. These include increase in cloud albedo, increase in MBL and cloud heights, drizzle suppression, increase in liquid water content, and increase in cloud lifetime. Moreover, better statistical representation of aerosol mass and number concentration improves model fidelity in reproducing observed spatial and temporal variability in cloud properties, including top and base height, droplet concentration, water content, rain rate, optical depth (COD) and liquid water path (LWP). Together, these help to quantify confidence in WRF-Chem’s modeled aerosol-cloud interactions, especially in the activation parameterization, while identifying structural and parametric uncertainties including: irreversibility in rain wet removal; overestimation of marine DMS and sea salt emissions, and accelerated aqueous sulfate conversion. Our findings suggest that WRF-Chem simulates marine cloud-aerosol interactions at a level sufficient for applications in forecasting weather and air quality and studying aerosol climate forcing, and may do so with the reliability required for policy analysis.

1 Introduction

Clouds play a major role in Earth’s radiative balance (Ramanathan et al., 1989; Cess et al., 1989). However, uncertainties in the processes that affect cloud optical properties and modify this balance are still high (Solomon et al., 2007). These processes are driven by the indirect climatic effects of aerosols (Lohmann and Feichter, 2005), which can modify cloud albedo (Twomey, 1977) and lifetime (Albrecht, 1989), evaporate clouds (Gräbl, 1979), change thermodynamics in deep convective clouds (Andronache et al., 1999), increase precipitation in ice clouds (Lohmann, 2002), and change the surface energy budget (e.g., Liepert, 2002).

Low-level marine clouds have been shown to contribute substantially to cloud radiative forcing (Ramanathan et al., 1989). However, these clouds are not well represented by contemporary models (Wyant et al., 2010). Previous work
has shown problems in the ability of global and regional models to accurately represent marine stratocumulus clouds (Vellore et al., 2007; Otkin et al., 2009; Wyant et al., 2010; Abel et al., 2010), leading to difficulties in predicting cloud cover on an operational basis (e.g., Shah et al., 2010). Some problems are thought to be related to boundary layer schemes generating insufficient vertical mixing, resulting in an unrealistically shallow cloud-topped boundary layer (Otkin and Greenwald, 2008). By comparison, Large Eddy Simulation (LES) models have been shown to more effectively describe stratocumulus clouds and their transitions (e.g., Feingold et al., 1998; Khairoutdinov and Kogan, 2000; Berner et al., 2011; Wang et al., 2010). Efforts have been made to couple models at both scales (regional and LES), obtaining accurate representation of stratocumulus (Zhu et al., 2010).

However, operational use of these coupled models for numerical weather prediction (NWP) or climate studies is not yet feasible. Cloud data assimilation has been an alternative way to improve clouds in NWP (e.g., Vellore et al. 2006; Errico et al., 2007; Michel and Auligné, 2010).

Uncertainties in modeling aerosol indirect effects diminish our capability to generate reliable climate projections, to evaluate policy questions and geo-engineering proposals, and to provide accurate weather and air quality predictions. Including indirect aerosol effects has been shown to improve cloud representations in global models (Lohmann and Lesins, 2002), and a range of approaches in modeling them have been assessed (Ghan and Easter, 2006). On the regional scale, including aerosol indirect effects tends to impact clouds optical properties (Chapman et al., 2009) and precipitation (Neelekos et al., 2009), and often produces better cloud representation by optical properties, dynamics and microphysics (Gustafson et al., 2007; Q. Yang et al., 2011). The LES scale has been able to show the effect on cloud structure by different cloud condensation nuclei (CCN) loadings, and effectively simulate the dynamics of open cells, “pockets of open cells,” and closed cell marine clouds (Wang and Feingold, 2009a, b).

Intensive measurement campaigns provide a wealth of observations that present the opportunity to evaluate models and to identify, quantify, and hopefully reduce these uncertainties. The VAMOS Ocean-Cloud-Atmosphere-Land Study Regional Experiment (VOCALS-REx, Wood et al., 2011) was an international field program designed to make observations of poorly understood but critical components of the coupled climate system of the southeast Pacific on the coast of Chile and Peru. Reactive gas and aerosol observations show a marked longitudinal gradient from elevated values close to shore due to polluted conditions to cleaner remote conditions (Allen et al., 2011), while cloud properties correlate to some extent with this gradient (Bretherton et al., 2010). Model evaluation studies emerging from the campaign have identified difficulties in accurately representing MBL and stratocumulus clouds (Abel et al., 2010; Sun et al., 2010; Andrejczuk et al., 2011) without considering aerosol feedbacks to meteorology. Q. Yang et al. (2011) present a comprehensive evaluation of the WRF-Chem system on a regional scale highlighting the effects of aerosol feedbacks, showing that the inclusion of aerosol-cloud interactions typically improve model performance in simulating cloud properties.

In this work, we build upon previous regional simulations including aerosol feedbacks using the WRF-Chem model. Several modeling studies have performed sensitivity analyses of the effects of aerosol loading on cloud properties (e.g., Chen et al., 2011). Starting from a base configuration, we find another configuration that better represents aerosol mass and number concentrations, and then analyze the impacts of these different aerosol loading on MBL dynamics and cloud microphysics, and compare them to observations and to the canonical conceptual model of warm cloud indirect effects. We perform an extensive evaluation of different aspects of the model representation, and identify areas for improvements and remaining problems.

2 Methods

For the purposes of defining representative spatial zones characterized by broadly internally similar thermodynamic aerosol and composition regimes (when averaged over the length of the VOCALS-REx campaign) we choose to use the three areas as defined by Bretherton et al. (2010) and Allen et al. (2011). These are the “coastal zone” (or “off shore”), east of 75°W, and the “remote zone” (west of 80°W), with the two regions separated by a “transition zone” near the 78°W meridian (75°W–80°W).

2.1 WRF-Chem model configuration

The WRF-Chem model simulates meteorology and atmospheric constituents, as well as their interactions (Skamarock et al., 2008; Grell et al., 2005). We configured WRF-Chem v3.3 with a combination of model structures, parametric choices, and input data to best represent marine stratocumulus conditions, atmospheric chemistry, and secondary aerosols, with the goal of future use in meteorological and air quality forecasting.

A 12 × 12 km² horizontal resolution domain is employed, covering 91°W–65°W longitude and 40°S–12°S latitude. This choice attempts to optimize between spatial resolution, critical for representing cloud dynamics, and complete coverage of the VOCALS region of the Southeast Pacific within the limitations of computing time. The domain accounts for most major Chilean and Peruvian anthropogenic sources shown in airmass trajectories (Chand et al., 2010) to impact VOCALS-REx observations, including the Andean cordillera, and covers the entire VOCALS-REx experimental domain (to ~85°W), with a margin towards the west and north to avoid excessive boundary condition influence on meteorology and atmospheric composition. Following recommendations by Wang et al. (2011), a 75-level vertical
resolution was chosen to reduce MBL and cloud height underestimation. The first few levels are as in Saide et al. (2011) with \(\sim 10\) \(\text{m}\) thickness, and the average vertical layer spacing between 60 \(\text{m}\) and 3 km is \(\sim 60\) \(\text{m}\). In preliminary testing, this resolution produced accurate MBL and cloud heights for all longitudes, which were \(\sim 100\)–300 \(\text{m}\) greater than the 39-level resolution used in Saide et al. (2011).

Model structure was configured to combine modules included in contemporary WRF-Chem public release code that best represent known aerosol, cloud, and MBL processes and their couplings. Wherever possible, the most complete representations of complex physical and chemical processes were chosen. This application requires a boundary layer closure scheme that can make use of (and maintain numerical stability at) high vertical resolution, and can accurately represent the diurnal evolution of the MBL at low wind speeds. Mellor-Yamada type schemes have generally exhibited good cloud representation under these conditions (Otkin and Greenwald, 2008; Zhu et al., 2010; Rahn and Garreaud, 2010). The MYNN level 2.5 scheme (Nakanishi and Niino, 2004) was chosen since it performed well in prior applications at this resolution over Chile (Saide et al., 2011). The Lin microphysics scheme (Chapman et al., 2009) and Goddard short wave radiation (Chou et al., 1998; Fast et al., 2006) were chosen to support aerosol direct, indirect, and semi-direct feedbacks to meteorology. Activation of aerosols from the interstitial to the cloudborne “attachment state” (Ghan and Easter, 2006) is based on a maximum supersaturation determined from a Gaussian spectrum of updraft velocities and the internally mixed aerosol properties within each aerosol size bin (Abdul-Razzak and Ghan, 2002). The updraft velocity distribution is centered in the model vertical wind component plus the subgrid vertical velocity diagnosed from vertical diffusivity. No cumulus scheme was used following the recommendation of Q. Yang et al. (2011). The RRTM longwave radiation scheme (Mlawer et al., 1997) was used. Gases and aerosols were simulated using the CBMZ gas-phase chemical mechanism (Zaveri et al., 1999; Fast et al., 2006) with dimethyl sulfide (DMS) reactions coupled to the 8-bin sectional MOSAIC (Zaveri et al., 2008) aerosol module. Seawater DMS concentration was set to 2.8 \(\text{nM}\), following the VOCA Modeling Experiment Specifications (http://www.atmos.washington.edu/~myant/vocals/model/VOCA_ModelSpec.htm) and in agreement with measurements during VOCALS-REx (Hind et al., 2011). DMS is transferred to the air using sea-air exchange as in Liss and Merlivat (1986).

We chose emissions and chemical boundary conditions to best resolve spatial and temporal variability in aerosols and their precursors, taking into account a complete range of natural and anthropogenic emissions sources. Continental emissions of biogenic trace gases (e.g., isoprene) were predicted hourly by the MEGAN algorithm (Guenter et al., 2006), and daily biomass burning locations and fuel loadings were obtained from FIRMS MODIS fire detections (Davies et al., 2009) and modeled hourly using WRF-Chem’s plume rise model (Freitas et al., 2006, 2007). Volcanic and anthropogenic emissions, including point and area sources, were taken from the VOCA inventory described in detail by Mena-Carrasco et al. (2012). Table 1 shows a summary of the sources of information for the emission inventories compiled for this research. In cases where particulate matter (PM) was not speciated, 10\%, 30\% and 70\% were associated to elemental carbon, organic carbon and crustal aerosol, respectively. Chemical boundary conditions are obtained from 6-hourly MOZART global simulations (Emmons et al., 2010). MOZART fields were found to overestimate near-shore concentrations, so the model was started from clean initial conditions and spun up for 6 days to avoid biasing results. MOZART sulfur dioxide (\(\text{SO}_2\)) boundary conditions in the free troposphere (FT) were found to be underestimated, so a global minimum background level of 30 ppt and a 50 ppt minimum for heights over 3.5 km were set, in agreement with flight profile measurements in the remote region (Allen et al., 2011; Kazil et al., 2011). Sea salt aerosol emissions were modeled following Gong et al. (1997), but resultant concentrations from the default scheme were found to substantially overestimate ship-based measurements from the NOAA RV Ronald H. Brown (Ron Brown). In order to avoid misleading indirect effects due to these biases, submicron emissions were reduced by a factor of 10 and supermicron emissions were reduced by a factor of 2, in line with campaign-averaged observations from the Ron Brown. Default WRF-Chem sea salt emissions do not consider sulfate coming from seawater, speciating sea salt as Na and Cl only. Wind-blowed dust was not modeled, due to known high biases in WRF-Chem’s online wind-blow dust emissions, concentrations, and resultant aerosol optical depth over land, and poor model representation of Andean dust composition. No organic sea emissions were considered in this study, as there was little to no evidence of these submicrometer particles during the campaign (Shank et al., 2012). Also, no secondary organic aerosols (SOA) were modeled as the fraction of SOA to total organic aerosol is thought to be low in this region (~10\%, Kanakidou et al., 2005). New particle formation is modeled by the Wexler et al. (1994) scheme.

WRF-Chem simulations covered the entire VOCALS-REx campaign period, 15 October–16 November 2008, along with the extra 15 days that Ron Brown stayed in the domain (16–30 November). The model was run with an initial “chemical” spin-up period of 6 days with meteorology re-initialized from analyses at the middle of the modeling period using the previous chemical state. We found that 3–4 days of spin up are necessary to overcome the underestimated MBL height present in the National Centers for Environmental Prediction (NCEP) Final Analysis (FNL) from GFS (Sun et al., 2010), as only increasing vertical resolution is insufficient for debiasing the offshore MBL (Andrejczuk et al., 2011).
We found that aerosol wet deposition has a large influence over the modeling results. In the present version of WRF-Chem, in- and below-cloud wet removal of gases and aerosols in CBMZ-MOSAIC are modeled following Easter et al. (2004). This mechanism assumes that the removal processes are irreversible, and does not consider aerosol resuspension due to rain evaporation. This becomes an important issue for the Southeast Pacific during Austral spring, since most of the drizzle observed during VOCALS-REx evaporated before reaching the surface (Bretherton et al., 2010), leading to a great contrast between cloud base and surface rain rates. Thus, irreversible removal of aerosol by rain might create an unrealistically strong sink, which is supported by previous modeling results (Q. Yang et al., 2011). Kazil et al. (2011) implemented wet removal considering rain evaporation, but for a different modal aerosol approach and in the context of LES simulations. To assess the importance of modeled wet removal processes, we performed simulations where wet deposition was excluded, which results in higher aerosol loadings. This represents an upper limit to below cloud aerosol, and reflects the fact that low rain rates were observed at the sea surface (0.01 mm h\(^{-1}\) on average) during the VOCALS campaign (M. Yang et al., 2011), indicating that most rain evaporated before reaching the surface (Bretherton et al., 2010), suggesting nearly zero wet deposition. Thus, by turning off wet deposition the unrealistic sink of aerosol mass generated by not considering resuspension is removed. However, the effects in terms of number concentration are uncertain due to complex interactions: one droplet can collect thousands of particles by collision-coalescence but, as some have observed (Mitra et al., 1992; Feingold et al., 1996), only one aerosol is released after evaporation. Since rain rates increase as aerosol number concentration decrease, a cloud-scavenged ultra-clean layer can be generated which can lead to conditions of particle nucleation (Kazil et al., 2011) potentially recovering the number of particles lost before. Without an aerosol module that includes reversible wet deposition, and for the sake of studying the sensitivity to different aerosol loads, both simulations were conducted for the whole period. The simulation with wet deposition turned on is hereafter referred as the base run or “W”, while the simulation without wet deposition is called “NW”. Since W represents large aerosol removal, while NW no aerosol removal, we hypothesize that a model with a correct wet deposition scheme should be bounded by these two states.

### 2.2 Observations

The observations used for comparison are provided by the VOCALS-REx airborne and marine platforms. Carbon monoxide (CO), Ozone (O\(_3\)), SO\(_2\), DMS gases; and sulfate (SO\(_4\)), nitrate (NO\(_3\)), ammonium (NH\(_4\)) and organic carbon (OC) interstitial aerosol (from AMS: non-refractory non-sea salt) observations collected by the NSF C-130, DoE G-1, FAAM BAe-146 aircrafts and Ron Brown are thoroughly described by Allen et al. (2011), Kleinman et al. (2012), M. Yang et al. (2011), and Hawkins et al. (2010), while C-130 cloud water composition measurement methods can

---

**Table 1.** Summary of emissions inventories configurations. Areas are divided in metropolitan region, which is located on Central Chile, the rest of Chile and the rest of the domain.

<table>
<thead>
<tr>
<th>Geographical area</th>
<th>Sector</th>
<th>Inventory name</th>
<th>Species</th>
<th>Base year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metropolitan region</td>
<td>Mobile sources</td>
<td>SECTRA</td>
<td>PM(<em>{10}), PM(</em>{2.5}), CO, NO(_x), SO(_x), VOCs, NH(_3)</td>
<td>2010</td>
</tr>
<tr>
<td>Metropolitan region</td>
<td>Residential sources</td>
<td>Chilean Ministry of environment DICTUC (2007)</td>
<td>PM(<em>{10}), PM(</em>{2.5}), CO, NO(_x), SO(_x), VOC, NH(_3)</td>
<td>2005</td>
</tr>
<tr>
<td>Metropolitan region</td>
<td>Point sources</td>
<td>Chilean Ministry of environment DICTUC (2007)</td>
<td>PM(<em>{10}), PM(</em>{2.5}), CO, NO(_x), SO(_x), VOC, NH(_3)</td>
<td>2005</td>
</tr>
<tr>
<td>Rest of Chile</td>
<td>Power plant emissions</td>
<td>Chilean power plant emissions standard KAS (2009)</td>
<td>PM(<em>{10}), PM(</em>{2.5}), CO, NO(_x), SO(_x), VOC, NH(_3)</td>
<td>2009</td>
</tr>
<tr>
<td>Rest of Chile</td>
<td>Smelter emissions</td>
<td>Chilean air quality standards revision Mena-Carrasco (2010)</td>
<td>PM(<em>{10}), PM(</em>{2.5}), CO, NO(_x), SO(_x), VOC, NH(_3)</td>
<td>2010</td>
</tr>
<tr>
<td>Rest of Chile</td>
<td>Mobile sources</td>
<td>SECTRA regional Corvalán et al. (2005)</td>
<td>PM(<em>{10}), PM(</em>{2.5}), CO, NO(_x), SO(_x), VOC, NH(_3)</td>
<td>2005</td>
</tr>
<tr>
<td>Rest of domain</td>
<td>Total anthropogenic excluding power and smelting EDGAR 3.2 Olivier et al. (1994)</td>
<td>PM(<em>{10}), PM(</em>{2.5}), CO, NO(_x), SO(_x), VOC, NH(_3)</td>
<td>2005</td>
<td></td>
</tr>
<tr>
<td>Rest of domain</td>
<td>Total anthropogenic</td>
<td>Bond et al. (2004)</td>
<td>Black carbon and organic carbon</td>
<td>2005</td>
</tr>
</tbody>
</table>
be found in Benedict et al. (2012). The University of Wyoming 94 GHz cloud radar (WCR) aboard C-130 provided radar reflectivities, which were then corrected (Bretherton et al., 2010) and converted to rainfall estimates using the Z-R relationship described in Comstock et al. (2004). This presents results consistent with Particle Measuring Systems (PMS) Two Dimensional Cloud Probe (2D-C) probe rainfall estimates during VOCALS (Bretherton et al., 2010). The WCR, along with an upward-pointing lidar (WCL) provided cloud top and base height estimates from the C-130 (Bretherton et al., 2010). Cloud top and base heights from Ron Brown were estimated using a millimeter-wave cloud radar (MMCR) and a Vaisala CL31 ceilometer, respectively (de Szoeke et al., 2010). Capping inversion height (CIH) was estimated as the height at which the temperature was a minimum, provided the relative humidity was at least 45 % (Jones et al., 2011) in both Ron Brown soundings (Wood et al., 2011) and aircraft vertical profiles. C-130 Gerber PVM-100 Probe cloud water content, PMS Cloud Droplet Probe (CDP) and Forward Scattering Spectrometer Probe (FSSP-100) cloud droplet number concentration, and PMS Passive Cavity Aerosol Spectrometer Probe (PCASP) accumulation mode aerosol number concentration observations used are described in Kazil et al. (2011) and Bretherton et al. (2010). BAE-146 cloud and accumulation mode aerosol measurements (Allen et al., 2011) were performed with similar instruments as in C-130 (Dropet Measurement Technologies (DMT) CDP-100, PCASP) while G-1 (Kleinman et al., 2012) used a DMT Cloud and Aerosol Sampling (CAS) probe and a PCASP, respectively. An intercomparision of the cloud microphysics probes fitted to BAE-146 and C-130 was performed on 31 October 2008 and 4 November 2008. The aircraft performed straight and level runs (of the order of 10's of km in length) through the same region of cloud approximately 5 min apart, finding that the number concentration, LWC and size distributions were similar within calibration and systematic error. However, G-1 cloud microphysics measurements showed inconsistencies compared to other probes used (Kleinman et al., 2012) probably due to shattering of drizzle on CAS inlet (McFarquhar et al., 2007). On the Ron Brown, total number of particles over 13 nm was measured with a TSI 3010 Condensation Particle Counter. Cloud optical depth (COD), cloud liquid water path (LWP) and cloud effective radius were obtained from MODIS-Aqua retrievals.

2.3 Performance statistics

Table 2. Measure of model performance using data obtained from a box and whisker plot. A “match” is defined as a model (observation) median or mean falling between the two prescribed percentiles of the observation (model) distribution. Matches ranges from 4 (perfect match) to 0 (no match). An “overlap” is defined as when modeled and observed inner quartiles (boxes) or inner deciles (whiskers) overlap. Overlaps ranges from 2 to 0. The final performance is assigned as the best of the three criteria.

<table>
<thead>
<tr>
<th></th>
<th>Matches in the 25th–75th percentile</th>
<th>Matches in the 10th–90th percentile</th>
<th>Overlaps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>4 or 3</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Very good</td>
<td>2 or 1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Good</td>
<td>–</td>
<td>4 or 3</td>
<td>2</td>
</tr>
<tr>
<td>Fair</td>
<td>–</td>
<td>1 or 2</td>
<td>1</td>
</tr>
<tr>
<td>Poor</td>
<td>–</td>
<td>–</td>
<td>0</td>
</tr>
</tbody>
</table>

Here, we introduce a measure of accuracy deduced from these plots, hereafter referred to as the Box and Whisker (BoW) metric and summarized in Table 2. We define a “match” as a model (observation) median or mean falling between the two prescribed percentiles of the observation (model) distribution. The first criterion (Table 2, column 1) uses the 25th and 75th percentiles (boxes in Fig. 1) and the second criterion (Table 2, column 2) uses the 10th and 90th percentiles (whiskers in Fig. 1). The third criterion (Table 2, column 3) is based on overlapping of the 25th and 75th (box overlapping) and 10th and 90th (whisker overlapping) percentile distributions. The net level of accuracy is determined from the sum of these three criteria, which is converted into a qualitative category: excellent, very good, good, fair and poor. Some of the advantages of the BoW method are that: it is independent of the variable being assessed; no absolute threshold of accuracy is specified for any variable; it is based on basic statistical parameters; it can easily be read from a box and whisker plot; and it transforms a quantitative measure of accuracy into a qualitative description. However, some issues could be encountered when the distributions are strongly skewed, as the mean could be found outside the inner quartile.

Statistics for cloud microphysics and aerosol number concentration were computed for aircrafts profile means instead of local point to point comparisons, since observed clouds could be in different levels than model values, generating mismatches with modeled and observed clouds both present, but at different levels. For estimating the modeled 117 nm to 3 μm PCASP aerosol number concentration, values from bins 3 to 7 (156 nm to 2.5 μm) are integrated along with 42 % of the second bin (78 to 156 nm), which corresponds to the fraction over 117 nm using the logarithmic diameter.

Aircraft modeled and observed gases, aerosol mass, cloud heights and rain statistics were computed for one minute...
average values, while Ron Brown statistics were computed for ten minute intervals. For the case of rain statistics, model results are not filtered for missing observations and vice versa as information on rain frequency can be extracted from the total sampling time in each longitude bin on the top of the box and whisker plots (e.g., Fig. 1). This creates minor inconsistencies; mainly in the 100 m height estimates, which was not measured in sub-cloud flight legs (Bretherton et al., 2010).

3 Results and discussion

We first focus on evaluating atmospheric concentrations of selected gases and aerosols for the base and NW simulations. Then, model performance is assessed for MBL dynamics and cloud microphysics. Finally, spatial and temporal variability in chemical transport and cloud effects are investigated on an episodic basis.

3.1 Trace gas and aerosol evaluation

Figure 1 shows aircrafts flight statistics for gaseous and aerosol concentration for selected species for the MBL and free troposphere (FT). For this plot, MBL concentrations were considered for heights lower than 1200 m or below cloud, and FT concentrations for heights in between 1700 and 3200 m in order to avoid cloud contamination (Allen et al., 2011). Trace gas data quality was assessed by intercomparison between aircrafts, obtaining 1.5 and 4 ppb as the uncertainties for CO and O$_3$, respectively (Allen et al., 2011). As seen in Fig. 1, MBL CO is overestimated across the entire

Fig. 1. Observed and modeled statistic for selected gaseous and aerosol species gridded into 2.5 degree longitudinal zones in between 22°S and 18°S. For each zone, centre solid (dashed) lines indicate the median (mean), boxes indicate upper and lower quartiles with upper and lower decile whiskers. The sampling time in decimal hours in each longitude bin is indicated at the top. Left column and right column are for marine boundary layer (MBL) and free troposphere (FT), respectively.
modeled longitudinal range with the exception of the close to the shore bin where highly polluted plumes were detected by the G-1. Measurement uncertainty is well below these differences pointing to a model bias. Neglecting these non-resolved plumes, close to shore overestimation is probably due to a lower MBL than observed (see Sect. 3.2), while remote zone issues are likely due to overestimation of the MBL CO MOZART boundary conditions, as these air masses often had no contact with the continent for a long time period (Allen et al., 2011). However, we cannot rule out the possibility that a combination of overestimates in Central Chile anthropogenic emissions (e.g., Jorquera and Castro, 2010; Saide et al., 2009) and too much entrainment in the model could generate MBL concentrations similar to FT concentrations. The latter case is less likely, as it would have similarly affected O₃. Remote FT CO shows very good to excellent performances driven mostly by MOZART boundary conditions over the east-central Pacific. Even though MBL CO shows poor to fair performance in BoW metrics, differences are no more than 15 ppb, and the observed longitudinal trend (decreasing towards remote zone) and spread (<10 ppb) are often well simulated, indicating that transport in the MBL is resolved. The base and NW models show very small differences, attributable to changes in entrainment and MBL heights (see Sect. 3.2).

Figure 1b shows O₃ was well simulated for both the MBL and FT, with very good to excellent BoW metrics, and with similar spread. There is no clear longitudinal trend in either the model or the observations. An important point is that, as mentioned in Allen et al. (2011), the model resolves the ~30 ppb difference between MBL and FT and also the higher variability in the FT. The lower O₃ in the MBL is the result of chemical destruction during the day, transport from FT during the night due to entrainment (M. Yang et al., 2011), and lower photolysis rates and temperatures under the cloud deck. Due to the ability of the model to correctly maintain the MBL to FT O₃ difference, we surmise that entrainment is simulated effectively. Hydroxyl radical (OH) in VOCALS MBL was estimated by Yang et al. (2009) from the DMS budget and found to have maximum diurnal values of 3–5 × 10⁶ molecules cm⁻³. WRF-Chem showed OH peaks in the lower end of this range, at ~2.5–3.5 × 10⁶ molecules cm⁻³.

Statistics for gas and aerosol components of the sulfur cycle are shown in Fig. 1c–e. The C-130 measured FT DMS (Fig. 1c, right panel) was usually below the detection limit (5 ppt), as in the model. The spikes in DMS for the 75th to 90th percentile show times where the cloud top heights were >1700 m, which is better captured by NW as explained later. In general, MBL DMS has a high bias, with poor to fair BoW scores. As discussed by Q. Yang et al. (2011), this is likely related to an overestimation of DMS emissions due to overestimation of the modeled DMS ocean:atmosphere transfer velocity. Similarly to CO, and despite the emission bias, the modeled longitudinal trend is captured very well by the model. Ron Brown atmospheric DMS measurements showed higher values, in better agreement with the model but still lower. FT SO₂ is also skillfully simulated, presenting mostly excellent BoW metrics and follows the observed longitudinal trend. FT Remote zone SO₂ is mostly affected by boundary conditions, showing the importance of setting lower thresholds for influx from MOZART (see Sect. 2). In the MBL, the model usually exhibits very good performance, but cannot maintain the ~20 ppt lower threshold observed. When looking at the modeled SO₂ diurnal cycle, the higher values are obtained after DMS photochemical destruction, but rapidly decay to values as low as 1 ppt due to cloud processing and conversion to SO₄ aerosol. Modeled conversion appears to occur at a higher rate than observed, which is investigated later in more detail. Finally, Fig. 1e shows sulfate mass concentrations statistics. MBL SO₄ clearly shows the impact of wet deposition as no remote zone W and NW model distributions overlap. The observations are typically closer to the NW results, and sometimes between the two simulations. As most of the rain evaporated before reaching the sea surface, we find that the NW results are more realistic, and that any overestimation could be due to the combination of high DMS emissions rates and high SO₂ to SO₄ cloud conversion yields. In the FT, both configurations show similar results with very good to excellent BoW performances. Other species play much smaller roles in aerosol composition. The model significantly underestimates NH₄ (possibly due to emissions, which are poorly constrained in the region), estimates of NO₃ are below detection limits as observed, and accurately predicts organic carbon (not shown). Most modeled NO₃ is found in the coarser sectional bins, as it is displaced by SO₄ in fine aerosol bins due to the low NH₃ concentrations (Seinfeld and Pandis, 2006), and therefore rarely appears in aircrafts AMS observations, where aerodynamic diameter is capped at 500–700 nm. Coarser aerosol is dominated by sea-salt, where NO₃ displaces Cl creating a chloride deficit (Q. Yang et al., 2011).

Table 3. Observed and modeled cloud chemistry statistics. Values where the observations (model) were inexistent were removed from the model (observations) statistics.

<table>
<thead>
<tr>
<th></th>
<th>Observation</th>
<th>NW model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std dev</td>
</tr>
<tr>
<td>pH</td>
<td>4.94</td>
<td>0.91</td>
</tr>
<tr>
<td>H₂O₂ (µM)</td>
<td>132.20</td>
<td>130.94</td>
</tr>
<tr>
<td>TOC (µg C L⁻¹)</td>
<td>2028.9</td>
<td>571.0</td>
</tr>
<tr>
<td>Cl⁻ (µN)</td>
<td>855.3</td>
<td>1205.0</td>
</tr>
<tr>
<td>NO₂⁻ (µN)</td>
<td>72.77</td>
<td>120.39</td>
</tr>
<tr>
<td>SO₄²⁻ (µN)</td>
<td>298.35</td>
<td>465.94</td>
</tr>
<tr>
<td>Na⁺ (µN)</td>
<td>1204.6</td>
<td>2008.6</td>
</tr>
<tr>
<td>NH₄⁺ (µN)</td>
<td>90.65</td>
<td>163.89</td>
</tr>
<tr>
<td>Ca²⁺ (µN)</td>
<td>128.60</td>
<td>203.39</td>
</tr>
</tbody>
</table>
In order to further explore $\text{SO}_2$ to sulfate conversion processes, we compare cloud chemistry observations to the NW model (Fig. 2 and Table 3), as both models show similar cloud aerosol composition. Consistent with observations (Benedict et al., 2012), model results show that bulk (summing all sizes) cloud drop ion concentrations are dominated by sea salt, followed by sulfate (Fig. 2a, b). Bulk sulfate concentrations are underestimated, since sulfate coming from seawater is not modeled. As shown by Table 3, in general the model does a good job representing the mean and variability of the ion concentrations. The most notable problems are $\text{Ca}^{+2}$, which is very low since no dust was modeled, and $\text{NH}_4^+$, which is underestimated as it was in the AMS interstitial aerosol evaluation. Model pH shows the same tendency as observations, increasing towards the remote region as sulfate aerosol is more abundant close to shore (not shown). However, model pH is always under 5, while values up to 7 were observed, leading to underprediction in mean pH (Table 3). We found that the bulk model is extremely sensitive to chloride concentrations, as a decrease in only $5\%$ in $\text{Cl}^-$ (as in observations) will increase average pH by 1 and increase single values up to 2.5 pH units. This is important, as WRF-Chem uses a bulk cloud chemistry scheme (Chapman et al., 2009) and small variations in $\text{Cl}^-$ (thus in pH) can generate a shift in the dominant mechanism of $\text{SO}_2$ to sulfate conversion, from the roughly pH independent $\text{H}_2\text{O}_2$ reaction (Martin and Damschen, 1981) to the $\text{O}_3$ reaction which increases in rate with pH (Hoffmann and Calvert, 1985), resulting in a speeding up of the $\text{SO}_2$ to sulfate conversion and even further reductions in $\text{SO}_2$ concentrations. However, as pointed out by M. Yang et al. (2011), most droplets nucleate from sulfate particles, so their pH will be acidic and dominated by the hydrogen peroxide reaction. This behavior for the majority of droplets is seen in modeled cloud water aerosol in the bins that dominate nucleation (Fig. 2c). We also see very low sea salt influence, as $\text{Na}^+$ percentage is low and $\text{Cl}^-$ is diffused into the droplet from HCl gas rather than entering the droplet as sea salt. All this implies that there is clear need for sized-resolved cloud chemistry (e.g., Fahey and Pandis, 2001), and that aqueous chemistry should be considered for nucleation and accumulation modes only (Kazil et al., 2011). Measurements considering the nature of the cloud condensation nuclei (CCN) composition and size should also be performed (Bator and Collett, 1997). Analyzing the $\text{H}_2\text{O}_2$ pathway, $\text{H}_2\text{O}_2$ concentrations are slightly overestimated by the model (Table 3), which cannot explain the $\text{SO}_2$ gap between model and observations. M. Yang et al. (2011) found that to close the $\text{SO}_2$ budget, the Martin and Damschen (1981) $\text{H}_2\text{O}_2$ rate expression yielded best results, while other reaction rates were too fast to reach mass balance. We compared these rates to the McArdle and Hoffmann (1983) rates implemented in WRF-Chem, reaching the same answer, which could explain the difference in $\text{SO}_2$. Another factor that influences increased $\text{SO}_2$ depletion is the consistent overestimation of cloud fraction, as WRF-Chem NW shows average cloud fractions of $86\%$ on the Ron Brown track, while the MMCR on board of Ron Brown (M. Yang et al., 2011) showed values of $67\%$.

### 3.2 MBL and marine Stratocumulus dynamics

Box and whisker plots for cloud base and cloud top for the VOCALS-REx period are shown in Fig. 3a, b. The WRF-Chem NW model shows up to 200 m higher mean and median cloud top and base heights than the base model, bringing it closer to observations. The largest differences are found in the remote zone. The higher accumulation mode aerosol load obtained by NW allows for a less broken cloud deck with smaller droplet radius and less precipitation (see Sect. 3.3), which affects the MBL energy budget by decreasing average downward shortwave radiation (SW), upward surface heat flux and top of the atmosphere (TOA) outgoing longwave radiation (LW) by $50-60\text{ W m}^{-2}$, $\sim 3\text{ W m}^{-2}$ and $1-3\text{ W m}^{-2}$, respectively. Precipitation, SW absorption and LW cooling of clouds are the main drivers of entrainment in a cloud topped boundary layer, which in turn determines its cloud height. As shown by Pincus and Baker (1994),
when number concentration of droplets increase, precipitation decreases, which increases entrainment. However, this also generates extra cloud water that produces thicker clouds that absorb more shortwave radiation (lower model downward SW), heating the layer and decreasing entrainment. Also, when clouds rise, cloud top temperature tends to decrease, decreasing LW cooling (model TOA outgoing LW decrease) and thus reducing entrainment. An overall increase in entrainment is achieved which cause cloud heights to rise (Pincus and Baker, 1994), in agreement with our results.

As aerosol loads increase for NW, the direct and semi-direct effects are also expected to change. However, semi-direct effects should not play an important role as BC observations (Shank et al., 2012) and model results show very low concentrations. A simulation where the aerosol radiation feedbacks were turned off using the NW configuration shows small differences for cloud top pressure (<±1%), cloud fraction (<±10%) and water content (<±10%), implying that indirect effects dominate under clean conditions like those observed during VOCALS-REx, where aerosol radiative effects become more important under heavily polluted conditions (Koren et al., 2008).

As seen in Fig. 3a, b, close to shore both simulations have large cloud height negative biases (fair to good BoW), as the coarse resolution is unable to resolve the steep topography and land-sea transition (Wang et al., 2011). In the remote zone the NW heights captured well the observations (mostly excellent BoW classifications) while the base model underestimated the heights (fair to excellent BoW scores). The NW model also better represents observed temperature and water vapor profiles in the remote zone both from aircraft profiles and ship-based soundings (not shown), as the typical MBL structure approaches the observed vertical profile. Even so, there are still some periods where the model does not simulate the very high cloud heights observed in the remote zone, as depicted by the 95th extremes of the observed distribution and as seen in the Ron Brown time series in Fig. 4, which are responsible for the lower model means. However, these periods of poor performance appear episodic, and there are periods where WRF-Chem NW does reach the observed heights (e.g., RF03 and RF05 on Fig. 4). Episodic underestimation of cloud heights is thought to come from meteorological boundary condition issues, as the model is unable to represent the high clouds condition occurring over several days (e.g., 19–23 November). The model shows good agreement with observations for a range of very different conditions: a 20S/POC drift flight (RF02) with very thin clouds, two flights to 85 W (RF03 and RF05) with different longitudinal cloud trends, and a coastal pollution survey flight (RF12) capturing the latitudinal gradient in cloud height.

In order to explore the model representation of MBL dynamics in more detail, a decoupling measure was computed. Jones et al. (2011) showed that an effective decoupling indicator can be calculated as the difference in total water mixing ratio ($q_t$, water vapor plus cloud water) between two levels: 25% and 75% of the capping inversion height (CIH), considering a value below (above) 0.5 g kg$^{-1}$ as a coupled (decoupled) MBL. Observed decoupling index and capping inversion height were obtained here from aircrafts vertical profiles following the method of Jones et al. (2011), and modeled values were obtained mapping the profiles and computing a modeled CIH and decoupling index. Figure 3c shows longitudinal statistics for all aircrafts flights. Both simulations represent several basic aspects of the observed decoupling. The modeled decoupling index is accurately predicted everywhere (very good to excellent BoW performance) but in the transition zone, where performance is lower but still good. On average, areas west of 78°W are decoupled while areas east of 78°W are coupled both in the observations and the model, and better represented by the NW simulation. The spread of the decoupling index on each zone shows longitudinal transition from coupled to decoupled MBLs. Observations show a sharp longitudinal transition from coupled to decoupled MBLs, which is also represented by the model.

3.3 Cloud microphysics

Figure 5 shows statistics for cloud properties and for aerosol number concentration. Model cloud water representation is
very good to excellent (in BoW metrics) for both models (Fig. 5a), but the NW model typically shows higher amounts of cloud water than the base model as clouds are more permanent and thicker (consistent with the Twomey effect found by Albrecht, 1989). A clear difference is seen when analyzing number of droplets (Fig. 5b) where the increase in cloud albedo is more evident (Twomey, 1974) and modeled inner quartiles do not overlap over the remote region. This is a result of the difference in sub-cloud aerosol number concentration (Fig. 5c), where not even the modeled deciles overlap. Comparing observed and model droplet number and aerosol number concentration in the remote zone, the NW model presents excellent results while the base model is biased low (poor to good BoW scores), showing vast improvements in cloud microphysics by increasing the sub-cloud aerosol to near observed levels. We found consistency in the results, as when aerosol loads are relatively close to observations, the number of droplets also becomes closer to the observations. We thus conclude that the activation routine in WRF-Chem is consistent and reliable. The inability of NW to represent the lower end and spread of the cloud droplet and aerosol number distributions can be related to not considering wet deposition, as aerosol number is expected to decrease for intense sub-cloud rain events since a single droplet can collect a large amount of particles and release just one when evaporation occurs (Kazil et al., 2011). Close to shore overestimation of droplet number concentration by both simulations may be explained by the slight overestimation of aerosol number and also by the fact that the model finds that the aerosol number concentration in the 1st bin (40 to 78 nm in diameter) is an important contributor to activated particles. The latter is not captured by the PCASP aerosol number concentrations, as it only measures aerosol diameters over 117 nm. Free troposphere aerosol number concentration (Fig. 5d) follows the same trend as in the MBL, with good to excellent BoW accuracy and few differences between the two simulations.

Rain estimates extracted from radar reflectivities and model statistics are shown in Fig. 6. Focusing on the cloud top (a) and cloud base (b) plots, it can be seen that the model captures in-cloud rain stratification, showing lower mean and median values for the cloud top. WRF-Chem also represents the longitudinal gradient in rain rate (Bretherton et al., 2010). The NW model tends to exhibit better agreement with observations, showing lower mean and median rainfall rates. The NW model has higher concentrations of activated particles and smaller effective droplet radius, which decrease autoconversion and suppress precipitation (Albrecht, 1989). The equilibrium reached in the base simulation is off since as more precipitation is produced, more particles are scavenged (and not recovered after evaporation), further reducing the number of particles and leading to even more precipitation in a reinforcing feedback. Besides showing higher precipitation rates, the base run also shows higher precipitation occurrence along the flight track (sampling time on the top of each plot), while the NW results tend to agree more closely with observations. At lower altitudes (500 and 100 m), observed precipitation occurrence decreases, which is also captured by both models, with the NW model always showing lower occurrence. At 100 m, the base model shows better agreement with remote observed rain rates and NW overestimates the mean and medium. Since the NW rain is scarcer, in agreement with observations, heavy drizzle events tend to skew
Fig. 5. Box and whisker plots for selected cloud properties and aerosol number concentration as in Fig. 1. (a) Profile mean cloud water content. (b) Profile mean number of droplets concentration. (c) and (d): marine boundary layer (MBL) and free troposphere (FT) mean profile aerosol number concentration. Number of profiles is indicated at the top of each longitude bin.

Fig. 6. Box and whisker plots for radar derived and modeled rain at different heights as in Fig. 1. (a) and (b) corresponds to rain rates just below the cloud top and at the cloud base while (c) and (d) corresponds to rain rates at fixed heights of 500 and 100 m. The numbers above each zone represent sampling time in decimal hours.

the distribution. However, modeled rain range given by the outer deciles agrees with the observations.

While episodic comparisons with in-situ observations are critical, it is also important to consider model performance for regional climatology, as the model should represent monthly mean values and their spatial features. Figure 7 shows COD and LWP for MODIS and both WRF-Chem simulations. Model COD was computed by first computing the effective radius as in Martin et al. (1994) and then COD as proposed by Slingo (1989) for the 0.64–0.69 µm band, as the MODIS reference wavelength for this retrieval is 0.65 µm (King et al., 2006). The base WRF-Chem model usually underestimates COD, while the NW model is closer to the observations. Several features are well represented: close to shore hotspots of COD around 17°S and 26°S, a nearshore local COD minimum around 36S, and an increase in COD around 20°S from 80°W to 75°W. In the remote zone (83°W to 90°W), observed COD tends to fall between both models but closer to NW, for the same reasons presented before to explain episodic performance: the base model is unable to generate a thick enough cloud layer and drizzles too much, while the NW clouds do not dissipate when moving westwards, thereby increasing cloud lifetime (Albrecht, 1989). LWP path shows a different behavior, as both models underestimate MODIS LWP, probably due to a model bias in the Lin microphysics parameterization, as the Morrison scheme (Morrison and Pinto, 2005) generates higher LWP (Q. Yang et al., 2011), as discussed further in the text. However, NW model results consistently show higher values and a better agreement with observations.

3.4 Aerosol feedbacks and relation to sources

The Ron Brown provided a unique platform for continuous point measurements. Not only does the vessel have a much longer residence time in each regional model grid-cell than research aircraft, it also records the complete diurnal cycle at each location. One of the points where Ron Brown sampling efforts were focused was at (20°S, 75°W), where it spent approximately 8 days, 4 days on each leg of the cruise. This area was found to be affected by coastal sources (Allen
et al., 2011; Bretherton et al., 2010; Hawkins et al., 2010). To evaluate model performance and aerosol interactions, Fig. 8a compares total observed sulfate to the base and NW models. Observed values are closer to the NW model, but both models resolve most of the periods where SO$_4^{2-}$ concentrations increase over 1 µg m$^{-3}$. As seen in Fig. 8b, the SO$_4^{2-}$ episodes are well correlated with aerosol number concentrations over 13 nm in diameter, a relationship also represented by the model. Observed aerosol number concentration is in the high range of the models because the lowest bin modeled is 40 nm, and does not include the 13 to 40 nm window. These sulfate episodes do not follow any diurnal pattern and are a constant factor affecting aerosol concentrations in this zone. Model results, including prior modeling by Spak et al. (2010) that only included anthropogenic sulfur emissions, clearly indicate that these peaks can be attributed to continental sources, usually coming from Central Chile. As an example, Fig. 8c shows the evolution of second bin (78 to 156 nm in diameter) SO$_4^{2-}$ (main contributor to aerosol number concentration) in a distinct pollution plume from the time emitted in Central Chile until it reaches the Ron Brown, 2 days after. When fresh, the maximum value of the plume is found on model level 17, around 650 m above sea level. At this height, it is transported by southeasterly trades (Rahn and Garreaud, 2010) until it makes contact with the MBL, where it starts entraining and SO$_2$ to SO$_4^{2-}$ conversion is enhanced in clouds. Once in the MBL, lower wind speeds result in the plume taking a longer time to reach the Ron Brown location. In the MBL, the plume receives additional SO$_4^{2-}$ contributions from DMS, as a near-shore DMS emissions hot spot is found off central Chile (26° S-36° S) due to wind shear generated by the subtropical low-level jet (Garreaud and Muñoz, 2005; Muñoz and Garreaud, 2005). By performing a simulation without DMS initial conditions and emissions, we estimate the DMS contribution to sulfate to be from 15% to 25% in mass (which could be overestimated as shown in previous sections) by the time the plume reaches the Ron Brown location for the case analyzed, showing that these episodes are generated mainly by anthropogenic sources. The ability of these plumes to reach this zone is thought to be determined by the position of the surface pressure maximum of the Southeast Pacific Anticyclone (Spak et al., 2010).

Distant sources from Central Chile often have a visible footprint in SO$_4^{2-}$ mass and aerosol number concentration over the study domain, and these enhanced aerosols participate in cloud feedbacks such as drizzle suppression. Figure 9 shows curtain plots of radar reflectivity (proxy for precipitation) and aerosol number concentration for C-130 RF05 flight, showing very marked and correlated longitudinal gradients both on aerosol load and precipitation for NW and observed values. The base model (W) shows the same gradient (not shown), with higher rain rates in the remote region. Bretherton et al. (2010) argued that lack of drizzle near the coast is not just a microphysical response to high droplet concentrations; but other aspects such as lower LWP and thinner clouds (related to shallower and coupled MBL) can be comparably important. However, synoptic conditions present during RF05 flight were such that MBL height differences between off shore and remote zone weren’t significant (cloud top and base heights differences less than 250 m and 200 m respectively, Fig. 4) and the remote region was not decoupled (no vertical gradients on aerosol concentrations inside the MBL,
also verified by flight vertical profiles as in Fig. 3) but we still see very high precipitation gradients. Thus, differences in aerosol load might be playing a more important role than previously thought. The enhanced aerosols also participate in cloud feedbacks visible in satellite retrievals of cloud properties. Figure 10a shows MODIS-Aqua cloud effective radius for an overpass on a day with a thick cloud deck, where aerosol feedbacks are more pronounced. Figure 10b, c show model results for effective radius and second bin sulfate surface concentrations. Model cloud effective radius clearly decreases when the MBL is dominated by high accumulation mode sulfate concentrations, following a similar shape to the plume, which can also be observed in the MODIS overpass. The scene shows two distinct plumes coming from Central Chile: an older one between 23°S and 20°S and a fresher one in between 29°S and 25°S, both showing a decrease in cloud effective radius in both model and the observations. These findings highlights the need to consider aerosol interactions and transport from far-away sources in high-resolution studies and NWP applications over the region and similar persistent coastal stratocumulus in eastern boundary tropical and subtropical areas.

3.5 Assessing differences due to model configuration

WRF-Chem is a community model with several choices of parameterizations to represent various processes (Skamarock et al., 2008). These choices result in different model configurations, which can produce different predictions. For example, there are two microphysics schemes that can be used to study aerosol indirect effects in WRF-Chem v3.3: the Lin scheme (used in this study); and the Morrison scheme (Q. Yang et al., 2011; Morrison and Pinto, 2005). Figure 11 shows results from a sensitivity column study where both schemes were run until reaching stable conditions as a function of number of droplets. Significant differences of over an order of magnitude are found in rain rates between both schemes. For the VOCALS-REx case of study, the Lin scheme rain rates showed good performance (Fig. 6), while the Morrison scheme showed under-prediction (Q. Yang et al., 2011), in accordance with Fig. 11. The higher rain rates in our study can also explain the larger under-prediction of sulfate mass and aerosol number concentration by the base model (Figs. 1 and 3) compared to the Q. Yang et al. (2011) study. We hypothesize that the main cause of the rain mismatch is the different autoconversion (cloud droplets to rain conversion) parameterization used in both schemes. The Morrison scheme uses Khairoutdinov and Kogan (2000) parameterization, which uses regressed coefficients (multiplicative and power laws) from cloud drop size spectra predicted by LES simulations, which shows a linear behavior in the log scale (Fig. 11a); while the Lin scheme uses Liu et al. (2005) which introduces a threshold function which depends on droplet number concentration that is responsible for the curve shape and rain suppression in Fig. 11a. On the
Fig. 9. Curtain plots for radar reflectivity (Z, in dBZ) and accumulation mode aerosol number concentration (# cm\(^{-3}\)) for C-130 flight RF05 on 25 October. (a) and (b) show radar observed and NW model Z while (c) shows NW model as the curtain and one minute average PCASP observations as colored circles. Observed Z and PCASP aerosol are 1 min averages. Model Z is computed according to Appendix A. Solid lines represent flight track with the line becoming segmented on (c) every time there is a PCASP observation. For all panels, bottom scale is time in hours and top scale is longitude in degrees.

Fig. 10. Horizontal plots of cloud effective radius (µm, a and b) and first level, second bin (78–156 nm aerosol diameter) SO\(_4\) concentration (µg m\(^{-3}\), c). (a) shows MODIS-AQUA cloud effective radius for 16 October 17:00 UTC overpass while (b) and (c) shows NW model results for the same time. Model cloud effective radius is computed for the cloud top.
other hand, the Morrison scheme shows higher liquid water content (Fig. 11b), which is not completely explained by the lower precipitation, as LWP differences are still present when rain rates are similar for low number of droplets. This is also coherent with the fact that Q. Yang et al. (2011) shows better agreement to MODIS LWP than our study, where this configuration under-predicts it.

Full double moment microphysics schemes (Lin scheme is double moment for cloud water only) are necessary to improve process representation in models (e.g., Morrison et al., 2009). As autoconversion seems to be generating low performance in rain rates, we propose to implement and test the Liu et al. (2005) parameterization in the Morrison scheme. The implementation has to come along with the inclusion of aerosol re-suspension due to rain evaporation on the wet deposition scheme to avoid the MBL aerosol biases seen in Figs. 1 and 3.

4 Conclusions

There is an imperative need for reducing uncertainty and improving the atmospheric models used in studies of aerosol-cloud interactions at scales needed for NWP, air quality predictions, and policy assessments. In this context, several intensive measurement campaigns have been carried out to improve our understanding of aerosol and cloud interactions and they provide an extensive data base for use in testing and improving models. In this work we test the regional model WRF-Chem for the VOCALS-REx campaign which focused on studying the persistent stratocumulus deck on the South East Pacific, off the shore of Chile and Peru. Starting from the fact that the inclusion of aerosol cloud interactions in the model are important to represent processes in this region (Q. Yang et al., 2011), we perform model simulations designed to address the questions: what are the effects on cloud dynamics and microphysics from changing the sub-cloud aerosol loads? And do these effects bring model results closer to observations when aerosol loads are in better agreement to measurements? To address these questions results from two model simulations, with (base) and without wet deposition (NW) were analyzed. Both runs represent an incomplete modeling picture, as the base run lacks aerosol re-suspension (which is important in drizzling stratocumulus), and excluding wet deposition means neglecting a known removal process. These simulations produce significant differences in aerosol amounts, particularly in the remote zone where sulfate mass and accumulation mode aerosol number distributions do not overlap with each other and can be one order of magnitude different. Observed aerosol mass and number are usually closer to the NW results, because the model wet deposition process irreversibly removes aerosol even for evaporating rain. Little surface rain was observed during the campaign, so evaporation of drizzle drops is a likely source of sub-cloud aerosols. The increase in aerosol number in NW generate a significant difference between the models in terms of marine boundary layer (MBL) dynamics and cloud microphysics, in accordance to warm clouds aerosol indirect effects. These include an increased number of cloud droplets (Twomey, 1974) showing no overlap of the inner quartiles from the two models in the remote zone; increased MBL and cloud heights (Pincus and Baker, 1994) reaching up to 200 m differences; drizzle suppression on average concentrations and on number of detections; increased liquid water content and increased cloud lifetime (Albrecht, 1989); which helps answer the first question. MBL dynamics and cloud microphysics observed values are usually closer to NW or at least fall in between both models showing that better aerosol statistical performance lead to changes in the right direction, which helps answer the second question. This study demonstrates the capabilities of the WRF-Chem model to simulate aerosol/cloud interactions, particularly regarding the activation routine, which simulates number of droplet concentrations more accurately when sub-cloud aerosol loads more closely match observations. However, the model needs further improvements to address issues such as aerosol re-suspension in rain wet removal, overestimation in oceanic DMS and sea-salt emissions, increased cloud driven SO2 to sulfate conversion and move from bulk to sectional/modal aqueous chemistry. Also, an assessment of model differences when using distinct WRF-Chem configurations shows these seem to be related to the microphysics schemes, specifically to different autoconversion parameterizations which can generate over an order of magnitude disagreement on rain rates predictions for the same conditions.

Besides performing campaign averaged comparisons, we quantified local model performance for stratocumulus properties and their hourly evolution against ship-based

Fig. 11. Results from column study for comparing Lin and Morrison microphysics schemes for a profile on (80°W, 20°S) at 00:00 UTC on 28 October 2008. (a) shows maximum rain rate per profile while (b) shows liquid water path (LWP) per profile. Each profile is run with a different droplet number concentration using a 12 s time step for enough time to reach stable conditions. For (a), missing points means rain rate equal to zero.
measurements (NOAA Ron Brown), three aircraft observations (NSF C-130, DoE G-1, FAAM BAe-146) and satellite retrievals (MODIS) using them to explain how aerosols and model processes affect system response. For instance, hourly evolution of cloud heights was evaluated showing a good model performance for the diurnal cycle and different synoptic conditions with the exception of periods where the model is not able to recover from the underestimated MBL height found on the boundary conditions. Also, an episodic study was performed showing that anthropogenic sources from Central Chile substantially changed aerosol mass and number, rain and cloud optical properties over the ocean both in modeled and observed values, showing that indirect effects might be playing a more important role in modulating cloud properties and dynamics than stated in previous studies.

In our analysis we have attempted to perform a complete multi-platform evaluation for a regional simulation of clouds and aerosols, where we included VOCALS-REx observations which were not compared to models previously, such as decoupling state, trace gas concentrations (carbon monoxide, ozone), cloud aerosol composition, cloud water ionic balance and radar reflectivities. These are all crucial to fully quantifying regional model performance in this tightly coupled system. In order to provide quantification to this evaluation, we introduced a new metric for assessing model performance that uses box and whisker plots. This metric is independent of the variable being analyzed, thus allows doing performance cross-comparison in between different models and variables.

Together with improving model issues already mentioned, future work should be focused on continuing to validate models with aerosol and cloud interactions from measurement campaigns in other locations, as conditions for each region vary extensively. Also, several observation platforms such as close to shore flights (NERC Dornier 228 and CIRPAS Twin Otter) and inland measurements (Iquique, Paposo and Paranal sites) were not considered as the modeling was too coarse for their use (12 km² grid cells). Thus, finer resolution studies for the same area are needed to exploit these data (4–1 km² grid cells). These studies help to better quantify the uncertainties in models, that need to be considered when these models are used as tools for policy makers and for weather and air quality forecasts.

Appendix A

Model radar reflectivity

The Lin microphysics scheme implemented in WRF-Chem uses an exponential distribution for rain droplets (Chen and Sun, 2002):

\[ N(D) = N_0 \exp(-\Lambda D) \]  

(A1)

\[ \Lambda = \left( \frac{\pi p_q N_0}{\rho q_r} \right)^{1/4} \]

Where \( N_0 \) is the intercept parameter, \( \Lambda \) the slope parameter, \( p_q \) and \( \rho \) water and air densities respectively, and \( q_r \) the rain water content. Radar reflectivity \( (Z) \) is computed as the sixth moment of the rain drop distribution as:

\[ Z = \int_0^\infty D^6 N(D) dD \]

which is only valid for Rayleigh scattering regime. As the radar used in this study is W-band (94 GHz frequency, ~3 mm wavelength) then Rayleigh scattering might not be valid for droplets over 0.5 mm (O’Connor et al., 2004). Instead, we use Mie calculations to obtain \( Z \) (Arai et al., 2005):

\[ Z_{Mie} = \frac{\lambda^4}{\pi^3 K^2} \int_0^\infty \sigma_{Mie}(D) dD \]  

(A2)

Where \( \lambda \) is the radar wavelength, \( K \) the absorption coefficient of water and \( \sigma_{Mie} \) the backscattering cross-section, which is a function of droplet diameter and radar wavelength. Then, as \( N_0 \) is fixed for the Lin scheme \( 8e+6m^4 \), \( Z_{Mie} \) can be computed as a function of rain water content by numerically integrating this equation. For each diameter, \( \sigma_{Mie} \) is computed using Mätzler (2002) code which is based on the Appendix of Bohren and Huffman (1983).

Acknowledgements. We thank all people and organizations that participated in the VOCALS-REx campaign, generating and allowing us to use the complete and comprehensive data set. Special thanks to Timothy Bates, Paquita Zuidema and Ludovic Bariteau for helping interpret and compare observational data, to William Gustafson and Jerome Fast for comments on modeling and to two anonymous reviewers for their constructive comments. Ron Brown, C-130 and G-1 measurements were obtained from the VOCALS data archive NCAR/EOL, which is sponsored by the National Science Foundation (NSF). MODIS data was obtained from the NASA Langley Research Center Atmospheric Science Data Center. Special thanks to the staff of the NERC Research Aviation Facility for supporting the deployment of the C130 and the staff of the UWYO King Air national facility for enabling the deployment of the WCR and WCL onboard the C130 during VOCALS-REx. We also thank the UK National Environment Research Council (NERC) for funding the VOCALS UK contingent to the project (grant ref: NE/F019874/1) and the NERC Facility for Airborne and Atmospheric Measurement (FAAM) and Direct Flight and Avalon for operational support of the BAe-146 aircraft. This work was carried out with the aid of NSF grants 0745986, grant number ULR024979 from the National Center for Research Resources (NCRR), a part of the National Institutes of Health (NIH), FONDECYT Initiation grant 11090084, and Fulbright-PUCIT scholarship number 15093810. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the founding institutions.

Edited by: R. Wood
References


Fast, J. D., Gustafson Jr., W. I., Easter, R. C., Zaveri, R. A.,


