



1 Forest Fire Aerosol – Weather Feedbacks over Western North 2 America Using a High-Resolution, Fully Coupled, Air-Quality 3 Model

4 Paul A. Makar¹, Ayodeji Akingunola¹, Jack Chen¹, Balbir Pabla¹, Wanmin Gong¹, Craig Stroud¹,
5 Christopher Sioris¹, Kerry Anderson², Philip Cheung¹, Junhua Zhang¹, Jason Milbrandt³

6 ¹Air Quality Research Division, Atmospheric Science and Technology Directorate, Environment and Climate Change
7 Canada, 4905 Dufferin Street, Toronto, M3H 5T4, Canada

8 ²Natural Resources Canada (emiritus), Summerland, British Columbia, Canada.

9 ³Meteorological Research Division, Atmospheric Science and Technology Directorate, Environment and Climate
10 Change Canada, 2121 Trans Canada Highway, Montreal, Canada

11 *Correspondence to:* Paul A. Makar (paul.makar@canada.ca)

12 **Abstract.** The influence of both anthropogenic and forest fire emissions, and their and subsequent chemical and
13 physical processing, on the accuracy of weather and air-quality forecasts, was studied using a high resolution, fully
14 coupled air-quality model. Simulations were carried out for the period 4 July through 5 August 2019, at 2.5-km
15 horizontal grid cell size, over a 2250 x 3425 km² domain covering western Canada and USA, prior to the use of the
16 forecast system as part of the FIREX-AQ ensemble forecast. Several large forest fires took place in the Canadian
17 portion of the domain during the study period. A feature of the implementation was the incorporation of a new on-
18 line version of the Canadian Forest Fire Emissions Prediction System (CFFEPSv4.0). This inclusion of
19 thermodynamic forest fire plume-rise calculations directly into the on-line air-quality model allowed us to simulate
20 the interactions between forest fire plume development and weather.

21 Incorporating feedbacks resulted in improvements in most metrics of both air-quality and meteorological model
22 forecast performance, through comparison of no-feedback and feedback simulations with surface, radiosonde, and
23 satellite observations. For the meteorological simulations, these improvements occurred at greater than the 90%
24 confidence level. Relative to the climatological cloud condensation nuclei and aerosol optical properties used in the
25 no-feedback simulations, the fully coupled model's aerosol indirect and direct effects were shown to result in feedback
26 loops characterized by increased surface temperatures, decreased lower troposphere temperatures, and increased lower
27 troposphere cloud droplet and raindrop number densities. The aerosol direct and indirect effect reduced oceanic cloud
28 droplet number densities and increased oceanic rain drop number densities, relative to the no-feedback climatological
29 simulation. The aerosol direct and indirect effects were responsible for changes to the aerosol concentrations at greater
30 than the 90% confidence level throughout the model domain, and to NO₂ and O₃ concentrations within forest fire
31 plumes.

32 The simulations show that incorporating aerosol direct and indirect effect feedbacks can significantly improve the
33 accuracy of weather and air quality forecasts, and that forest fire plume rise calculations within a fully coupled model
34 changes the predicted fire plume dispersion and emissions, the latter through changing the meteorology driving fire
35 behaviour and growth.



36 1 Introduction

37 Atmospheric aerosol particles may be emitted (primary particles) or result from the condensation of the products of
38 gas-phase oxidation reactions (secondary aerosol). With increasing transport time from emission sources, the
39 processes of coagulation (colliding particles stick adhere creating larger particles) and condensation (low volatility
40 gases condense to particle surfaces) tend to result in particles which have a greater degree of internal mixing (internal
41 homogeneous mixtures). Primary and near-source particles are more likely to have a single or a smaller number of
42 chemical constituents (external mixtures).

43 Atmospheric particles also modify weather through well-established pathways. Under clear sky conditions, the
44 particles may absorb and/or scatter incoming light, depending on their size, shape, mixing state (internal, external or
45 combinations) and their composition. The presence of the particles themselves may thus affect the radiative budget
46 of the atmosphere, resulting in either positive or negative climate forcing (i.e. the absorption of a greater amount of
47 incoming solar radiation versus increased scattering reflection of that radiation back out into space, a process known
48 as the Aerosol Direct Effect; ADE). Aerosols can also alter the atmospheric radiative balance through interactions
49 with clouds, this influence being referred to as the Aerosol Indirect Effect (AIE). Three broad classes of categories
50 by which cloud/aerosol interactions take place (Oreopoulos *et al.*, 2020) include the first indirect effect, where higher
51 aerosol loadings resulting in increasing numbers of cloud droplets with smaller sizes, hence increasing cloud albedo
52 (Twomey *et al.*, 1977), the second indirect effect, where higher aerosol loadings suppress the collision-coalescence
53 activity of the smaller droplets, reducing precipitation/drizzle, changing cloud heights, and changing cloud lifetime in
54 warm clouds (Albrecht, 1989), and aerosol “invigoration” of storm clouds, where higher aerosol loadings may result
55 in delayed glaciation of cloud droplets, in turn leading to greater latent heat release and stronger convection (Rosenfeld
56 *et al.*, 2018).

57 The uncertainties associated with the ADE and particularly AIE account for a large portion of the uncertainties in
58 current climate model predictions for radiative forcing between 1750 and 2011 (Mhyre *et al.*, 2013). Carbon dioxide
59 is believed to have a positive (warming) global radiative forcing of approximately $1.88 \pm 0.20 \text{ W m}^{-2}$, while the direct
60 and indirect effects both have nominal values of approximately -0.45 W m^{-2} , with uncertainty ranges encompassing -
61 0.94 to $+0.07$ and -1.22 to 0.0 W m^{-2} respectively. These uncertainties have spurred research designed to better
62 characterize the ADE and AIE, and reduce these uncertainties, through both observations and atmospheric modelling.
63 Observational studies of the ADE have established its large impact; for example, high aerosol loading over Eurasian
64 boreal forests has been found to double the diffuse fraction of global radiation (i.e. increased scattering), a change
65 sufficient to affect plant growth characterized via gross primary production (Ezhova *et al.*, 2018). Aerosol assimilation
66 of Geostationary Ocean Color Imager Aerosol Optical Depth (AOD) observations into a coupled meteorology-
67 chemistry model showed that South Korean AOD values increased by as much as 0.15 with the use of assimilation;
68 these increases corresponded to a local -31.39 W m^{-2} reduction in solar radiation received at the surface, and reductions
69 in planetary boundary layer height, air temperature, and surface wind speed over land, and a deceleration of vertical
70 transport (Jung *et al.*, 2019). Other studies in East Asia have shown ADE decreasing local shortwave reaching the
71 surface by -20 W m^{-2} (Wang *et al.*, 2016), as well as significant changes in surface particulate matter and gas
72 concentrations in response to these radiation changes.



73 However, one commonality amongst the recent studies of the ADE for air-quality models is a tendency towards
74 negative biases in predicted aerosol optical depths, potentially indicating systematic under-predictions in aerosol mass,
75 aerosol size, and/or inaccuracies in the assumptions for shape and/or mixing state. Mallet *et al.* (2017) noted this
76 negative bias for regional climate model AOD predictions associated with large California forest fires compared to
77 OMI and MRIS satellite observations. Palacios-Pena *et al.* (2018) noted that high AOD events associated with forest
78 fires were under-predicted by most models in a study employing a multi-regional-model ensemble. The chosen AOD
79 calculation methodology and mixing state assumptions employed in models also plays a role in systematic biases:
80 Curci *et al.* (2015) compared aerosol optical depths, single scattering albedos, and asymmetry factors at different
81 locations to observations, varying the source model for the aerosol composition, as well as the mixing state
82 assumptions used in generating aerosol optical properties, for Europe and North America. AODs were biased low by
83 a factor of two or more, regardless of model aerosol inputs or mixing state assumptions at 440 nm, single scattering
84 albedos were biased low by up to a factor of two, with the poorest performance for “core-shell” approaches, while
85 asymmetry factor estimates showed no consistent bias relative to observations. However, the assumed mixing state
86 was clearly a controlling factor in the negative biases; the AOD predictions closest to the observations at 440 nm
87 assumed an external mixture with particle sulphate and nitrate assumed to grow hygroscopically as pure sulphuric
88 acid, lowering their refractive index with increasing aerosol size. This mixing state assumption and the different
89 homogeneous mixture assumptions gave the best fit for single scattering albedo relative to observations. While not
90 commenting on aerosol direct effect implications, Takeishi *et al.* (2020) noted that forest fire aerosols increase particle
91 number concentrations but reduce their water uptake (hygroscopicity) relative to anthropogenic aerosols, with the
92 latter effect reducing the resulting cloud droplet numbers by up to 37%. Mixing state and hygroscopicity properties
93 of aerosols were thus shown to have a controlling influence on the ADE.

94 The AIE has often been shown to be locally more important for the radiative balance than ADE in terms of magnitude
95 of the radiative forcing and response of predicted weather to AIE and ADE (Makar *et al.*, 2015(a); Jiang *et al.*, 2015;
96 Nazarenko *et al.*, 2017). Several recent studies have attempted to characterize the relative importance of the AIE with
97 the use of multi-year satellite observations, sometimes making use of models and data assimilation. Saponaro *et al.*
98 (2017) used MODIS/Aqua linked observations of aerosol optical depth and Ångström exponent to various cloud
99 properties, noting that the cloud fraction, cloud optical thickness, liquid water path, and cloud top height all increased
100 with increasing aerosol loading, while cloud droplet effective radius decreased, with the effects dominating at low
101 levels (between 900 to 700 hPa). Zhao *et al.* (2018) examined 30 years of cloud and aerosol data (1981-2011), and
102 found that increasing aerosol loading up to $AOD < 0.08$ increased cloud cover fraction and cloud top height, while
103 further increases in aerosol loading (AOD from 0.08 to 0.13) resulted in higher cloud tops, and larger cloud droplets.
104 In polluted environments ($AOD > 0.30$) cloud droplet effective radius, optical depth and water path; cloud droplet
105 effective radius increased with increasing AOD. The first ADE was most sensitive to AOD in the AOD range 0.13 to
106 0.30; and the reduction of precipitation efficiency associated with the second aerosol indirect effect occurred for AODs
107 between 0.08 and 0.4, in oceanic areas downwind of continental sources.

108 However, sources of uncertainty in AIE estimates persist, in part due to the number of poorly understood processes
109 contributing to the atmospheric response to the presence of aerosols. Nazarenko *et al.* (2017) showed that short-term



110 atmospheric radiative changes were reduced in magnitude when sea-surface temperature and sea-ice coupling was
111 included in climate change simulations. Suzuki *et al.* (2019) showed that the vertical structure of atmospheric aerosols,
112 as well as their composition, had a significant influence on radiative forcing. Penner *et al.* (2018) and Zhu *et al.* (2020)
113 examined the impact of aerosol composition on cirrus clouds via ice nucleation, finding negative forcings for most
114 forms of soot, but a contrary impact of secondary organic aerosols. Rothenburg *et al.* (2018) noted that tests of aerosol
115 activation schemes carried out under current climate conditions had little variability, but had much greater variability
116 for pre-industrial simulations, implying that the available data for evaluation using current conditions may poorly
117 constrain ADE and AIE parameterizations used in simulating in past climates.

118 Forest fires are of key interest for improving the understanding and representation of ADE and AIE in models, due to
119 the large amount of aerosols released during these biomass burning events. Forest fire emissions and interactions with
120 weather are also of interest due to the expectation that the meteorological conditions resulting in forest fires may
121 become more prevalent in the future under climate change (Hoegh-Guldberg *et al.*, 2018). Observations of aerosol
122 optical properties during long-range transport events of North American forest fire plumes to Europe showed 500 nm
123 AOD values of 0.7 to 1.2 over Norway, with Ångström exponents exceeding 1.4 and absorbing angstrom exponents
124 ranging from 1.0 to 1.25, along with single scattering albedos greater than 0.9 at the surface and up to 0.99 in the
125 column over these sites (Markowicz *et al.*, 2016). Biomass burning was shown to have a specific set of optical
126 properties relatively independent of fuel type for three different types of biomass burning in China (cropland), Siberia
127 (mixed forest) and California (needleleaf forest). The increase in upward radiative forcing at the top of the atmosphere
128 due to fires being linearly correlated to AOD (R from 0.48 to 0.68), with slopes covering a relatively small range from
129 20 to 23 W m⁻² unit AOD⁻¹. O'Neill *et al.* (2001) showed that forest fires have a profound impact on aerosol optical
130 depth in western Canada, accounting for 80% of the summer AOD variability in that region, with a factor of three
131 increase in AOD levels from clear-sky to forest fire plume conditions. O'Neill *et al.* (2001)'s analysis of TOMS
132 AVHRR and GOES imagery suggested that forest fire aerosols increase in size with increasing downwind distance,
133 due to secondary aerosol aging and condensation chemistry. We note here that reanalyzing the data presented in
134 O'Neill *et al.* (2001) results in a linear relationship between fine mode particle effective radius (r_{eff} , μm) and the base
135 10 logarithm of distance from the fires (D, km) of $r_{eff} = 0.0106 \log_{10}(D) + 0.1163, R^2 = 0.18$). Mallet *et al.*
136 (2017) simulated AODs in the range 1 to 2 for biomass burning events, and also noted changes in direct radiative
137 forcing at the top of the atmosphere from positive to negative in both model results and simulations, with increasing
138 downwind distance from the sources. Lu *et al.* (2017) carried out simulations with 5-km horizontal grid spacings for
139 the eastern Russia forest fires of 2002 assuming an internal mixture for emitted aerosols with the WRF-CHEM model,
140 and noted impacts on cloud formation for two different periods. The first period was characterized by high cloud
141 droplet and small ice nuclei numbers, where the fire plumes reduced cloud rain and snow water content, large scale
142 frontal system dynamics were altered by smoke, and precipitation was delayed by a day. The second period was
143 characterized by high numbers for cloud droplets and ice nuclei, where the fire plumes reduced rain water content,
144 increased snow water content, and precipitation locations changed locally across the simulation domain. Russian
145 forest fire simulations for 2010 with suites of fully coupled air-quality models (Makar *et al.*, 2015; Palacios-Pena *et*



146 *al.*, 2018; Baro *et al.*, 2017) showed substantial local impacts, such as reductions in average downward shortwave
147 radiation of up to 80Wm^{-2} and temperature of $-0.8\text{ }^{\circ}\text{C}$ (Makar *et al.*, 2015(a)).

148 Given the above developments in direct and indirect parameterizations, and the increasing amount of information
149 available for estimating forest fire emissions, the impact of forest fires on weather, in the context of weather
150 forecasting, is worthy of consideration. Air-quality model predictions of forest fire plumes have been provided to the
151 public under operational forecast conditions of time- and memory-space limited computer resources (e.g. Chen *et al.*,
152 2019; James *et al.*, 2018; Ahmadov *et al.*, 2019, Pan *et al.*, 2017). These simulations make use of satellite retrievals
153 of forest fire hot-spots, climatological data on the extent of area burned by land use type, databases of fuel type linked
154 to emission factors, and an *a priori* weather forecast to provide the meteorological inputs required to predict forest
155 fire plume rise. The latter point is worthy of note in the context of the direct and indirect feedback studies noted above
156 – both climate and weather simulations with prescribed forest fire emissions have consistently resulted in large
157 perturbations of weather patterns in the vicinity of the forest fires. However, the approaches for predicting forest fire
158 plume rise up until now have relied on weather forecast information lacking those meteorological feedback effects.
159 The connection of the ADE and AIE within an air-quality and weather forecast model is referred to as “coupling”,
160 with such a model being “fully-coupled”. However, we note that in the field of very high resolution forest fire
161 behaviour modelling, coupling of biomass burning with the atmosphere has also been defined as the interaction of
162 dynamic meteorology with the heat released by the fire, where the initial meteorology determines fire spread on the
163 landscape. This in turn, modifies the temperature and wind fields, in turn affecting future fire spread (Clark *et al.*,
164 1996, Linn *et al.*, 2002). The coupling presented in the present paper refers to that between the aerosols released by
165 fires and other sources to meteorology through the ADE and AIE, with the resulting changes in meteorology in turn
166 influencing fire behaviour (fire intensity, fuel consumption, etc.), in turn influencing emissions height and distribution,
167 closing this feedback loop.

168 A key consideration in parameterizing the AIE (via aerosol-cloud interaction) is the manner in which the cloud
169 condensation process is represented in the meteorological component of the modelling system. In numerical weather
170 prediction (NWP) models, clouds and precipitation are represented by a combination of physical parameterizations
171 that are each targeted at a specific subset of moist processes. These include “implicit” (subgrid-scale) clouds generated
172 by the boundary layer and the convection parameterization schemes (e.g Sundqvist, 1988), and “explicit” clouds from
173 the grid-scale condensation scheme (Milbrandt and Yau, 2005(a,b), Morrison and Milbrandt, 2015, Milbrandt and
174 Morrison, 2016). Depending on the model grid these “moist physics” schemes vary in their relative importance.

175 However, regardless of the horizontal grid cell size, the grid-scale condensation scheme plays a crucial role in
176 atmospheric models, though to different degrees and using different methods, depending on the grid spacing and the
177 corresponding relative contributions of the implicit schemes. A grid-scale condensation scheme will in general consist
178 of the following three components: 1) a subgrid cloud fraction parameterization (CF, or cloud “macrophysics”
179 scheme); 2) a microphysics scheme; and 3) a precipitation scheme (Jouan *et al.*, 2020). The cloud fraction (CF) is the
180 percentage of the grid element that is covered by cloud (and is saturated), even though the grid-scale relative humidity
181 may be less than 100%. The microphysics parameterization computes the bulk effects of a complex set of cloud
182 microphysical processes. If precipitating hydrometeors are advected by the model dynamics, the precipitation is said



183 to be *prognostic*; if precipitation is assumed to fall instantly to the surface upon production, it is considered *diagnostic*.
184 The precipitation “scheme” is not a separate component per se, since it simply reflects the level of detail in the
185 microphysics parameterization, but it is a useful concept to facilitate the comparison of different grid-scale
186 condensation parameterizations.

187 With a wide range of grid cell sizes in current NWP models, there is a wide variety of types of condensation schemes
188 and degrees of complexity in their various components. For example, cloud-resolving models (with grid spacing on
189 the order of 1 km or less) have typically used detailed bulk microphysics schemes (BMSs), with prognostic
190 precipitation, and no diagnostic or prognostic CF component (i.e. the CF is either 0 or 1). Large-scale global models
191 use condensation parameterizations, sometimes referred to as “stratiform” cloud schemes, typically with much simpler
192 microphysics and diagnostic precipitation, but with more emphasis on the details of the CF. However, with continually
193 increasing computer resources and decreasing grid spacing (both in research and operational prediction systems), the
194 distinction between schemes designed for specific ranges of model resolutions is disappearing and condensation
195 schemes are being designed or modified to be more versatile and usable across a wider range of model resolutions
196 (e.g. Milbrandt and Morrison, 2016).

197 Aerosol-cloud interactions and feedback mechanisms are difficult to represent in grid-scale condensation schemes
198 with very simple microphysics components. For example, to benefit from the predicted number concentrations of
199 cloud condensation nuclei and ice nuclei, the microphysics needs to be double-moment (predicting both mass and
200 number) for at least cloud droplets and ice crystals, respectively. Until recently, detailed BMSs were only used at
201 cloud resolving scales, hence requiring these relatively high resolutions to be recommended in feedback modelling.
202 In recent years, multi-moment BMSs have been used in operational NWP for model grid spacings of 2-4 km (e.g.
203 Seity *et al.*, 2010, Pinto *et al.*, 2015, Milbrandt *et al.*, 2016). Further, condensation schemes with detailed microphysics
204 are starting to use non-binary CF components (e.g. Chosson *et al.*, 2014, Jouan *et al.*, 2020), thereby allowing detailed
205 microphysics to be used at larger scales, and hence allowing the same indirect feedback parameterizations to be used
206 at multiple scales. Nevertheless, the expectation is that detailed parameterization will provide a more accurate
207 representation of cloud formation at the near cloud-resolving scales, without the complicating aspect of a diagnostic
208 CF, motivating the use of km-scale grid spacing for feedback studies.

209 The formation of secondary aerosols from complex chemical reactions are another key consideration in feedback
210 forecast implementation, given the impact of aerosol composition on aerosol optical and cloud formation properties,
211 as described above.

212 In the sections which follow, we describe our high resolution, fully coupled air-quality model with on-line forest fire
213 plume rise calculations, which was created as part of the FIREX-AQ air-quality forecast ensemble
214 (<https://www.esrl.noaa.gov/csl/projects/firex-aq/>), to address the following questions:

215 (1) Will a fully coupled model of this nature provide improved forecasts of *both* weather and air-quality, using
216 standard operational forecast evaluation tools, techniques and metrics of forecast confidence? That is, despite the
217 uncertainties in the literature as described above, are these processes sufficiently well described in our model that
218 their use results in a formal improvement in forecast accuracy?



219 (2) Are the changes in forest fire plume rise associated with implementing this process directly within a fully coupled
220 model sufficient to result in significant perturbations to weather predictions and to chemistry? What are these
221 perturbations?

222 We employ our fully coupled model with 2.5-km grid cell size domain covering most of western North America, and
223 compare model results to surface meteorological and chemical observations, and to vertical column observations of
224 temperature and aerosol optical depth (AOD), in order to quantitatively evaluate the effect of feedback coupling of
225 the ADE and AIE on model performance. We then compare feedback and no-feedback simulations to show the
226 impacts of the ADE and AIE feedbacks on cloud and other meteorological predictions, and on key air quality variables
227 (particulate matter, nitrogen dioxide, and ozone). We begin our analysis with a description of our modelling platform.

228 **2 Model Description**

229 **2.1 GEM-MACH**

230 The Global Environmental Multiscale – Modelling Air-quality and CHEMistry (GEM-MACH) model in its fully
231 coupled configuration has been described elsewhere (Makar *et al.*, 2015a,b; Gong *et al.*, 2015, 2016). Briefly, the
232 model combines the Environment and Climate Change Canada Global Environmental Multiscale weather numerical
233 weather prediction model (GEM, Cote *et al.*, 1998, Girard *et al.*, 2014) with gas and particle process representation
234 using the on-line paradigm, with options for climatological versus full coupling between meteorology and chemistry.
235 In GEM-MACH's climatological coupling configuration, prescribed, invariant climatological values for aerosol
236 optical properties and cloud condensation nuclei (CCN) are employed within the model's radiative transfer and cloud
237 microphysics modules. In the full coupling configuration, the ADE is simulated using GEM-MACH's predicted
238 aerosol loading and Mie scattering using a binary water-dry aerosol homogeneous mixture assumption, at the 4
239 wavelengths employed by GEM's radiative transfer algorithms, and at additional wavelengths for diagnostic purposes.
240 The full coupling also includes the AIE by simulating aerosol-cloud interaction via explicit droplet nucleation using
241 the algorithm of Abdul-Razzak and Ghan (2002) along with on-line aerosol composition and size (Gong *et al.*, 2015).
242 This droplet nucleation replaces the decoupled model's existing droplet nucleation calculation in the Predicted Particle
243 Properties (P3) microphysics scheme (Morrison and Milbrandt, 2015, Milbrandt and Morrison, 2016). The latter
244 assumes an invariant aerosol population of a single lognormal size distribution (with a geometric mean diameter of
245 100 nm and total aerosol number concentration of 300 cm^{-3} , consisting of pure ammonium sulphate; Morrison and
246 Grabowski, 2008). The prognostic cloud droplet number and mass mixing ratios from the P3 microphysics are then
247 transferred back to the chemistry module for using in cloud processing of gases and aerosols (cloud scavenging and
248 chemistry) calculations, completing the AIE feedback process loop (Gong *et al.*, 2015).

249 The chemistry modules of GEM-MACH also include process representation for gas-phase chemistry (ADOMII
250 mechanism, 42 gas species, Stockwell *et al.*, 1989), cloud processing including aqueous chemistry, scavenging of
251 gases and aerosols, below-cloud removal and wet deposition (Gong *et al.*, 2015), particle microphysics employing a
252 sectional size distribution and 8 chemical species (Gong *et al.*, 2003), particle inorganic thermodynamics (Makar *et al.*,
253 *et al.*, 2003), the formation of secondary organic aerosols using a modified yield approach (Stroud *et al.*, 2018), process



254 representation for surface fluxes as a boundary condition on the solution for vertical diffusion, and semi-Lagrangian
255 advection for transported chemical tracers.

256 The specific base model version employed in these simulations is GEMv4.9.8/GEM-MACHv2 (Moran *et al.*, 2018),
257 incorporating the following additional improvements in addition to those noted above: (a) the AIE parameterization
258 was modified for use with the P3 cloud microphysics scheme; (b) forest canopy shading and turbulence was
259 parameterized following Makar *et al.*, (2017); (c) anthropogenic plume rise was parameterized through calculating
260 residual buoyancy of the rising plume (Akingunola *et al.*, 2018); (d) emissions of crustal material undergo a
261 meteorological modulation with crustal material emissions being inhibited when the soil water content is predicted to
262 be greater than 10%; (e) emissions/deposition of NH₃ are implemented using a bidirectional flux parameterization
263 (Whaley *et al.*, 2018; Zhang *et al.*, 2003); (f) CH₄ is treated as a reactive and emitted tracer; (g) the KPP-generated
264 RODAS-3 solver (Sandu and Sander, 2006) is used for the solution of ADOMII gas-phase mechanism (Stockwell
265 and Lurmann, 1989); (g) MODIS retrievals were used to create monthly leaf area index values for use in the model's
266 biogenic emissions, forest canopy shading and turbulence, and deposition algorithms; (h) a parameterization for the
267 impacts of vehicle-induced turbulence on vertical diffusive transport was employed (Makar *et al.*, 2020).

268 Simulations were carried out with a 2.5-km horizontal grid cell spacing over a 900 x 1370 grid cell domain, covering
269 most of western Canada and the USA (Figure 1). The meteorological boundary conditions for the simulation were a
270 combination of 10-km GEM forecasts updated hourly (themselves originating in data assimilation analyses of real-
271 time weather information), and 2.5-km GEM simulations employing, in the northern portion of the 2.5-km domain,
272 the Canadian Land Data Assimilation System (Carrera *et al.*, 2015), to better simulate surface conditions. Both
273 “feedback” and “no feedback” simulations were carried out on a 30-hour forecast cycle (Figure 2). Following the
274 usual practice for weather forecasts, the analysis-driven meteorological forecasts at 10 km resolution were updated
275 operationally every 24 hours at 12 UT. These 10 km resolution forecasts were followed by a 30-hour meteorology-
276 only forecast at 2.5-km resolution on the high resolution domain of Figure 1, the last 24 hours of which were used as
277 meteorological initial and boundary conditions for the 24-hour 2.5-km fully-coupled GEM-MACH simulation. These
278 two stages of meteorology-only simulations were carried out to prevent chaotic drift from the observed meteorology,
279 and to allow spin-up time for the cloud fields of that meteorology to reach equilibrium (6-hour timeframe). Chemical
280 lateral boundary conditions were taken from climatologies based on ECMWF MACC-II global atmospheric chemical
281 composition modelling and reanalysis (Inness *et al.*, 2013). Chemical initial concentrations for each consecutive
282 forecast within the 2.5- km GEM-MACH model domain were “rolled over” or “daisy-chained” between subsequent
283 forecasts without chemical data assimilation. Forecast performance scores presented here are for the inner 2.5-km
284 domain from this set of linked 24 forecast simulations, mimicking operational forecast conditions.

285 **2.2 CFFEPS Version 4.0: On-line forest-fire plume rise calculations**

286 In addition to the above algorithm improvements relative to GEM-MACH implementations, this model system setup
287 has incorporated the first on-line calculation of forest-fire plume-rise by energy balance driven using on-line
288 meteorology, in a new version of the Canadian Forest Fire Emissions Prediction System (CFFEPS). The algorithms



289 of CFFEPSv2.03 are described in detail and evaluated elsewhere (Chen *et al.*, 2019), but will be outlined briefly here,
290 as well as subsequent modifications to this forest fire emissions processing module.

291 CFFEPS combines near-real-time satellite detection of forest fire hotspots with national statistics of burn areas by
292 Canadian province and by specific fuel type across North America. CFFEPS assumes persistence fire growth in the
293 subsequent 24- to 72-hour forecasts with hourly fuel consumed calculated (kg m^{-2}), based on GEM forecast
294 meteorology and predicted fire behaviour in grid cells representing fire locations. The modelled fire fuel consumption
295 is then linked with combustion-phase specific emission factors (g kg^{-1}) for fire specific emissions and chemical
296 speciation. Fire energy associated with the modelled combustion process is also estimated, and is used in conjunction
297 with *a priori* forecasts of meteorology within the column to determine plume rise. In its off-line/non-coupled
298 configuration (Chen *et al.*, 2019), CFFEPS carries out residual buoyancy calculations at five preset pressure levels
299 (surface, 850, 700, 500, 250 mb). CFFEPS predicts plume injection heights, which are in turn used to redistribute the
300 mass emissions below the plume top to the model hybrid levels. This approach employed in CFFEPSv2.03 provided
301 a substantial improvement in forecast accuracy relative to the previous approach employing modified Briggs (Briggs,
302 1965, Pavlovic *et al.*, 2016) plume rise formulae in the offline GEM-MACH forecast system (Chen *et al.*, 2019).

303 However, other work has shown the substantial impact of large forest fires on regional weather (Makar *et al.*, 2015a;
304 Palacios-Pena *et al.*, 2018, Baro *et al.*, 2017), including changes to the surface radiative balance and atmospheric
305 stability. These findings imply that plume rise calculations employing an *a priori* weather forecast lacking the impact
306 of fire plumes via the ADE and AIE may not accurately predict the weather conditions critical to subsequent forest
307 fire plume rise prediction. In order to study this possibility, and to allow forest fire plumes to influence weather and
308 hence subsequent fire spread/growth, several changes were made to CFFEPS implementation, resulting in version 4.0
309 of CFFEPS, used here. The process flow within CFFEPSv2.03 versus CFFEPSv4.0 are compared in Figure 3. The
310 original C language CFFEPSv2.03 code was converted to FORTRAN90, and following successful off-line
311 comparisons to the original code, was then integrated as an on-line subroutine package within GEM-MACH itself,
312 with the near-real-time satellite hotspot data and location fuel parameters being read into GEM-MACH directly
313 (CFFEPSv4.0 is this new on-line package). A key advantage of the CFFEPSv4.0 subroutine integration within GEM-
314 MACH is that the residual buoyancy calculations for plume injection heights are now carried out over the model
315 hybrid model layers, rather than the five coarse resolution, prescribed pressure levels of CFFEPSv2.03, making
316 complete use of GEM-MACH's detailed vertical structure. Additionally, CFFEPSv4.0 allows plume rise calculations
317 to be updated during model runtime. When GEM-MACH is run in fully-coupled mode, the ADE and AIE
318 implementations allow model-generated aerosols to modify the predicted meteorology, in turn influencing predicted
319 fire emissions and plume rise, closing these feedback loops. The on-line implementation of CFFEPSv4.0 thus allows
320 us to investigate the effects of meteorology on subsequent forest fire plume development, the changes to modelled
321 aerosol compositions, and, ultimately, the feedbacks to weather.

322 The formation of particles from forest fires affects meteorology on the larger scale via the ADE and AIE, in turn
323 modifying the regional scale atmospheric features affecting fire growth, such as the temperature profiles below forest
324 fire plumes. However, we note that the local scale weather modifications due to the addition of forest fire heat to the
325 atmosphere are not yet incorporated into fire spread or GEM microphysics. Specifically, when the feedback version



326 of GEM-MACH incorporating CFFEPSv4.0 is used in its fully coupled configuration, CFFEPSv4.0 calculates forest
327 fire plume rise using the meteorological predictions which include the ADE and AIE generated by forest fire aerosols
328 on atmospheric stability from the current fully-coupled model timestep. The resulting added aerosol mass due to the
329 fire in turn affects the meteorology through ADE and AIE, closing this feedback loop. To the best of our knowledge,
330 this is the first implementation of a dynamic forest fire plume injection height scheme incorporated into a fully coupled
331 high-resolution, air quality forecast modelling system. The impact of this feedback on both weather and air-quality
332 can be substantial, as we show in the following sections.

333 The locations of the daily forest hotspots detected during the study period, and the corresponding magnitude of the
334 daily PM_{2.5} emissions generated by CFFEPS for each hotspot are shown in Figure 4. Individual hotspots with the
335 highest magnitude emissions are located in the state of Nevada (Figure 4(a), southern boxed region). However, the
336 largest ensemble emissions from a suite of hotspots occurs in northern Alberta (Figure 4(a), northern boxed region).
337 Expanded views of the northern Alberta and Nevada hotspots are shown in Figure 4(b,c) respectively – the use of
338 smaller symbols shows that the Alberta hotspots are groups representing large spreading fires, which overlapped in
339 Figure 4(a), while the Nevada hotspots indicate single fires of small spatial extent and duration rather than larger
340 spreading fires. The Alberta fires are thus the most significant sources of forest fire emissions in the study domain for
341 the period analyzed here.

342

343 **2.2 Feedback and No-Feedback Simulations**

344 Two simulations were carried out for the period July 4th through August 5th 2019; a “feedback” (ADE and AIE
345 feedbacks enabled – fully coupled model) and a “no-feedback” simulation (ADE and AIE make use of GEM’s
346 climatological aerosol radiative and CCN properties – the decoupled model). During this period, five large forest fires
347 took place in the northern portion of the modelling domain. The two parallel combined meteorology and air-quality
348 forecasts in the fully coupled model with/without ADE and AIE coupling were evaluated using the US EPA AIRNOW
349 data (<https://www.airnow.gov>) and Environment and Climate Change Canada’s EMET and ARCAD operational
350 forecast evaluation systems, respectively. Following evaluation, the simulation mean values of hourly meteorological
351 and chemical tracer predictions were compared to analyze the impact of fully coupled ADE and AIE feedbacks on
352 both sets of fields.

353 **3 Model Evaluation**

354 **3.1 Meteorology Evaluation**

355 Surface meteorological conditions were evaluated at three-hour intervals from the start of both of the two sets of paired
356 24-hour forecasts using standard metrics of weather forecast performance including mean bias (MB), mean absolute
357 error (MAE), root mean square error (RMSE), correlation coefficient (R) and standard deviation (σ). In all
358 comparisons, a 90% percent confidence level assuming a normal distribution was used to identify statistically different
359 results between forecast simulations. Note that 90% confidence levels are commonly used in meteorological forecast



360 evaluation, with values of 80% to 85% recommended (Pinson and Kariniotakis, 2004) and up to 90% used (Luig *et*
361 *al.*, 2001) for variables such as wind speed, rather than the 95% or 99% confidence levels in other fields, in recognition
362 of the difficulties inherent in prognostic forecasts of the chaotic weather system. Here, the confidence range
363 formulation of Geer (2014) has been applied using a 90% confidence level in model predictions, with the statistical
364 measures considered different at the 90% confidence level when the 90% confidence ranges do not overlap. The
365 surface meteorological evaluations shown here only include those variables and metrics where results were
366 significantly different at the 90% confidence level.

367 Several model forecast output variables were evaluated and the surface variables showing statistically significant
368 differences relative to observations at the 90% confidence level included: 2 m temperature, surface pressure, 2 m
369 dewpoint temperature, 10 m wind speed, sea-level pressure, and accumulated precipitation (the latter in 3 different
370 metrics). The comparisons are shown as time series in three-hourly intervals as a function of forecast hour prediction
371 time forward from forecast hour 0, for grid cells corresponding to measurement locations in Figures 5, 6, 7, 8, 9, 10,
372 and 11 for each of these quantities, respectively. Note that these statistics measure domain-wide performance, across
373 all of the reporting stations within the model domain, during the sequence of 24-hour forecasts comprising the
374 simulation period. The duration of the time series in these comparison figures is thus a function of the duration of the
375 contributing forecasts.

376 Figure 5 shows an example analysis for surface temperature bias for the entire model domain. Figure 5(a) shows the
377 average model MB time series across all stations and all forecasts at the given forecast hours, while Figure 5(b) shows
378 the corresponding difference in the MB absolute values. The difference plot in Figure 5(b) shows the feedback – no-
379 feedback scores, such that scores below the zero line indicate superior performance of the feedback forecast, while
380 those above the zero line indicate superior performance of the no-feedback forecast. Here, the feedback forecast was
381 statistically superior at forecast hours 3, 6, 15 and 18 at the 90% confidence level at these forecast hours, and both
382 simulations were at par (differences below the 90% confidence level) at hours 9, 12, 21 and 24. The feedback forecast
383 thus has superior performance, at greater than 90% confidence, over half of the forecast hours evaluated within the
384 domain, during the simulation period.

385 All of the metrics for which surface temperature forecast performance differed at the 90% confidence level are shown
386 in Figure 6. In addition to MB, the scores for MAE, and RMSE showed superior forecast performance for the feedback
387 relative to the no-feedback case at the 90% confidence level for hours 6 through 15, while the improvement for the
388 correlation coefficient was only reached the 90% confidence level at hours 6 and 12.

389 Model 10-m windspeed forecasts were also improved with the incorporation of feedbacks for hour ranges between
390 hour 3 and hour 12, depending on the metric, with the longest duration improvement for MB, MAE, and RMSE, and
391 the shortest duration for correlation coefficient and standard deviation (Figure 9). A marginal performance
392 degradation of the feedback forecast at close to 90% confidence at hours 15-18 can also be seen for root mean square
393 error, correlation coefficient, and standard deviation in this Figure.

394 Precipitation forecast performance from the two simulations varied depending on the metric chosen (Figure 11). The
395 metrics in this case are based on the number of coincident precipitation “events” versus “non-events” as shown in
396 contingency Table 1.



397 The Heidke skill score { $HSS = 2(AD - BC) / [(A + C)(C + D) + (A + B)(B + D)]$ } measures the fractional
398 improvement of the forecast over the number correct by chance. The Frequency Bias { $FB = (A + B) / (A + C)$ }
399 measures the frequency of event over-forecasts ($FB > 1$) versus event under-forecasts ($FB < 1$). The Equitable Threat
400 score { $ETS = (A - \tilde{A}) / (A + C + B - \tilde{A})$, where $\tilde{A} = (A + B)(A + C) / (A + B + C + D)$ } measures the observed
401 and/or forecast events that were correctly predicted. Following standard practice at Environment and Climate Change
402 Canada, the HSS is used as a measure of total precipitation accumulated over a 6-hour interval, with no lower limit
403 on the amount of precipitation defining an “event”, while FB and ETS define precipitation “events” as being those
404 with greater than 2mm / 6 hours – consequently FB and ETS have a smaller number of data points for comparison
405 than HSS.

406 Figure 11 shows improvements to the fully coupled precipitation forecast at the 90% confidence level were seen for
407 the HSS 6-hour accumulated metric by hours 18 and 24, while the frequency bias index of 6-hour accumulated
408 precipitation showed degradation at hours 6 and 24, and the equitable threat score of 6-hour accumulated precipitation
409 showed degradation at hour 24. As is noted above, the latter two metrics employed a minimum 6-hour precipitation
410 threshold of 2 mm prior to comparisons (this is the reason for the reduced number of points available for comparison
411 in Figure 11(b,c) relative to Figure 11(a)). These findings suggest that the fully coupled model’s trend towards
412 improved total precipitation over time (Figure 11(a)) is the result of improved performance for relatively low-level
413 precipitation events ($< 2\text{mm } 6\text{hr}^{-1}$), offsetting a degradation of performance for higher level precipitation events.
414 Precipitation events have thus become more frequent, but “lighter” with the use of the feedback parameterizations.
415 The meteorological forecast performance metrics with statistically significant differences for surface pressure,
416 dewpoint temperature, and sea-level pressure are shown in Figures 7, 8, and 10 respectively. The model performance
417 differences in these three Figures show a similar pattern: a degradation in performance with the use of feedbacks at
418 hour 3, with the differences between the two forecasts either dropping below the 90% confidence level, or the feedback
419 forecast showing an improvement by hour 6, followed by several hours in which the feedback forecast has a superior
420 performance. The duration of this latter period varies between the metrics, from up to 18 hours for MAE for surface
421 pressure (Figure 7(b)) to 6 hours for the standard deviation of dew-point temperature (Figure 8(d)).

422 We believe that the initial loss of performance for the feedback forecast may represent a form of “model spin-up” that
423 may be unique to fully coupled models, but may be affected or improved with further adjustments to the forecast
424 cycling setup for the chemical species. As noted earlier (Figure 2), in order to prevent chaotic drift from observed
425 meteorology, we made use of a 30-hour 2.5-km resolution analysis-driven weather forecast to update our fully coupled
426 model’s initial meteorology at hour zero of each 24 hour forecast. The cloud fields provided as initial conditions at
427 hour zero include observation analysis for the 6 hours prior to hour zero - these have reached a quasi-equilibrium in
428 the high-resolution weather forecast (Figures 2(b,e)) by the time they are used as initial and boundary conditions in
429 the fully coupled model (Figure 2(c,f)). However, the fully coupled model’s *aerosol* fields at hour zero, used to
430 initialize the subsequent forecast (Figure 2, dashed blue arrow), still reflect the locations of aerosol-cloud interactions
431 in the previous fully coupled simulation. We believe that the initial three to six hours of feedback forecast degradation
432 represents the time required for the fully coupled model to reach a new equilibrium consistent between both its aerosol
433 and the cloud fields.



434 One possible solution for this model spin-up inconsistency would be to eliminate the intermediate driving 2.5-km
435 meteorological simulation in favour of a longer 30-hour fully coupled forecast with the first six hours removed as
436 spin-up (i.e. extend the duration of steps (c) and (f) in Figure 2 to 30 hours, starting at UT hour 6). The duration of
437 the forecast experiments carried out here was limited to 24 hours due to limited computational resources, and, more
438 importantly, the operational requirement for an on-time forecast delivery for the purpose of the FIREX-AQ field
439 campaign. The 24-hour forecast simulations carried out in Figure 2 (c,f) each required nearly 3 hours of
440 supercomputer processing time; longer simulation periods were not possible within the operational window available
441 for forecasting.

442 The amalgamated observations and model pairs of vertical temperature profile data from 39 radiosonde sites in western
443 North America are shown in Figures 12 and 13. Improvements in the forecasted temperature vertical profile with
444 increasing forecast time are evident at 850 hPa in the 12th hour forecast (Figure 12) and at 925 and 1000 hPa in the
445 24th hour (Figure 13) forecast. The forecast simulation with aerosol feedbacks enabled also showed slight
446 improvements in the 10, and 50 hPa 12th hour temperatures and 50 hPa 24th hour temperatures, while 500 hPa 24th
447 hour temperature performance degraded slightly. There are larger differences between the 1000 hPa forecasts, though
448 these also have the least number of contributing stations (i.e. only those located close to sea-level contribute to the
449 lowest level temperature biases). Other levels of the atmosphere showed no statistically significant change at the 90%
450 confidence level in temperature profile forecast performance with the use of feedbacks.

451 3.2 Chemistry Evaluation

452 Chemistry forecast quality is usually evaluated using standard statistical metrics against hourly observations collected
453 from surface measurement stations. Both simulations' performance for ozone (O₃), nitrogen dioxide (NO₂) and
454 particulate matter with diameters less than 2.5 μm (PM_{2.5}) were evaluated using hourly AIRNOW data (USA: AQS
455 network: <https://www.epa.gov/aqs>; Canada: NAPS network: <http://maps-cartes.ec.gc.ca/rnspa-naps/data.aspx>). The
456 summary performance metric scores for the two simulations grouped, according to contributing measurement network,
457 are shown in Table 2, with boldface values indicating the better score for the given simulation case. With respect to
458 this table, we note that:

- 459 (a) The feedback simulation generally outperforms the no-feedback simulation (more bold-face scores in the
460 "feedback" columns, with a few notable exceptions, discussed below).
- 461 (b) In some evaluation metrics, the feedback simulation showed substantial quantitative improvements over the no-
462 feedback simulation (e.g. feedback PM_{2.5} MB is reduced by over a factor of 3 relative to its no-feedback
463 counterpart over Western Canada, the region of wildfire activity).
- 464 (c) For cases when the no-feedback simulation outperforms the feedback simulation, the relative magnitude of the
465 performance difference is smaller than the feedback simulation's improvements (e.g. Western USA PM_{2.5} and
466 Western Canada NO₂ mean bias degradations of 9.0 and 27.8 percent relative to the PM_{2.5} improvement of a
467 factor of 3 noted above).



468 (d) Both simulations have negative mean biases for O₃ of -3.5 to -3.7 ppbv throughout the model domain, negative
469 biases for PM_{2.5} in the Western USA, positive biases for PM_{2.5} in Western Canada, negative biases for NO₂ in
470 Western Canada and positive biases for NO₂ in the Western USA.

471 One possible cause for the overall model biases noted in (d) may reside in inadequate chemical boundary conditions
472 used for the forecasting setup. An unprecedented large high-resolution (2.5-km) model domain and on-time delivery
473 of forecasts in support of FIREX-AQ put a significant constraint on available computational resources. 2.5-km
474 chemical boundary conditions for these simulations were taken from seasonal climatologies from ECMWF global
475 model analyses, rather than coarse resolution model simulations. During the time simulated, large forest fires occurred
476 both within the domain (in northern Alberta and Saskatchewan in Canada), and outside of this model domain (in
477 Alaska north of the Panhandle). These out-of-domain sources were thus not available as lateral boundary conditions
478 on the 2.5-km domain, with possible impacts on model performance for the three species thus resulting in overall
479 forecast biases in both the feedback and the no-feedback simulations, particularly in Western Canada.

480 The impact of lateral boundary conditions on model predictions can be seen when comparing MODIS retrievals of
481 aerosol optical depth (AOD) with model predictions (Figure 14). AOD is a function of both the particle's abundance
482 and optical properties, integrated throughout the vertical column. However, direct comparisons between satellite and
483 model-predicted AOD values must be undertaken with some care, due to the nature of the satellite retrievals quality
484 assurance and control procedures, the motion of the orbiting spacecraft, and the scan time of the instrument. For a
485 polar-orbiting instrument such as MODIS, the time at which overpasses occur varies with location, and valid satellite
486 retrievals may not occur when the location being scanned is obscured by clouds. Observed averages may be built up
487 over multiple valid scans over time, but the number of valid scans contributing to the local average at any given
488 location will vary, due to the time and space variation in cloud cover. Here, individual valid Collection 6.1
489 MODIS/Aqua (MYD04_L2 AOD_550_Dark_Target_Deep_Blue_Combined) 10 km resolution 550 nm AODs were
490 matched in time and space to the nearest model 2.5-km grid cell and output frequency hour. Levy *et al.*, (2013)
491 contains details on the MODIS combined AOD product. Both model and observed values were then locally averaged
492 across the simulation period, to a common resolution 0.1° latitude-longitude grid co-located with the model grid, in
493 order to generate the comparison shown in Figure 14. Individual pixels of this image thus incorporate a spatially
494 varying number of values in the local averages, but these local averages are matched in time, space, and local averaging
495 period. The region over which comparisons were made thus corresponds to the high resolution model domain, and
496 white areas within the images correspond to regions where satellite data used in the averaging were excluded due to
497 QA/QC constraints, such as the presence of clouds, surface ice, etc.

498 The MODIS-observed local average AOD values of Figure 14 (a,b) are generally higher than the model predictions
499 (c,d), with the exception of the region co-located with large forest fires in Northern Alberta and Saskatchewan (Figure
500 14 (d), yellow regions). In addition, as noted above, other satellite imagery has shown the presence of several large
501 forest fires occurring in Alaska, outside of the modelling domain, with smoke plumes extending from these sources
502 down through the northern and western coastal boundaries of the model domain (yellow regions in Figure 14(a)). It
503 is likely that at least some of the biases in PM_{2.5} and NO₂ and consequently in the production of secondary ozone,
504 reflect the absence of these sources. Scatterplots of all paired AOD values in Figure 15 (a,b) and for the northern



505 portion of Alberta (Figure 15(c,d)) show that the overall negative bias is due to a large number of underestimated
506 “background” values of AOD in the model simulations, while in the immediate vicinity of the Alberta / Saskatchewan
507 forest fires, model values are considerably higher than observations.

508 The local model positive biases in AOD within Northern Alberta might partly be attributed to overestimates in emitted
509 particulate matter mass in the CFFEPS module, or from inaccurate treatment of fire plume centerline dispersion
510 downwind of primary emissions over large fire sources. Previous work with CFFEPS by Chen *et al.* (2019) for the
511 2017 fire season has shown similar PM_{2.5} positive biases for western Canada, with MB of +5.8 $\mu\text{g m}^{-3}$ (88 stations)
512 and for Western USA with MB of +8.6 $\mu\text{g m}^{-3}$ (221 stations). These positive biases (Chen *et al.*, 2019) were higher
513 specific to sub-regions closer to areas of active fires (MB of +12 $\mu\text{g m}^{-3}$ for the sub-region including the provinces of
514 Alberta and British Columbia, and +29 $\mu\text{g m}^{-3}$ for the sub-region comprising the states of Idaho, Montana, Oregon
515 and Washington, respectively). Our analysis here and these earlier results suggest a positive bias in CFFEPS’ PM_{2.5}
516 emissions or insufficient dilution/vertical extent of the predicted fire plumes.

517 The local AOD positive biases could also be the result of the mixing state assumptions of the Mie code used here for
518 generating aerosol optical properties. These assumptions may also account for negative AOD biases over much of
519 the remainder of the model domain. We have used a mass-weighted homogeneous mixture approach, with the
520 complex refractive index values for the 8 particle species being calculated for pure water-dry component homogeneous
521 mixtures at each of the 12 particle size bins, followed by mass weighting to generate values for each of the model
522 components. As noted earlier, this overall negative bias of AOD predictions is a common problem in air-quality
523 models and may be due to assumptions regarding the model mixing state (Curci *et al.*, 2015). That comparison of
524 multiple mixing state assumptions on AOD with observations for European and North American modelling domains
525 (Curci *et al.*, 2015), showed a typical factor of two model under-prediction of 440 nm North American AOD across
526 all mixing state assumptions, with European AOD negative biases ranging from unbiased to a factor of 2. For the
527 latter group, those models employing an assumption of external mixing, with hygroscopic growth factors for sulphate
528 and nitrate assumed to be similar to those of sulphuric acid, had the highest AODs and hence closest values compared
529 to observations at 440nm. However, in that investigation, the latter method also sometimes resulted in AOD over-
530 predictions by a factor of 2 at 870 nm. These earlier findings along with overestimates at forest fire plumes with our
531 current homogeneous mixture approach at 550nm suggest that the hygroscopic growth may be overestimated for forest
532 fire particles, in turn overestimating forest fire AODs locally, while external mixing assumptions may be required to
533 improve model AOD performance elsewhere in the model domain.

534 3.3 Model Evaluation Summary

535 Overall, the incorporation of feedbacks in this study has resulted in improvements in weather and air-quality forecast
536 accuracy, albeit with some caveats. Weather forecast variables showed improvements at the 90% confidence level
537 for several fields, and vertical profiles showed improvements, particularly close to the surface, and with increasing
538 forecast lead time. Total precipitation scores also improved. A previously unexpected spin-up issue specific to fully
539 coupled models was noted: the impact of fully coupled particulate matter on cloud variables was sufficiently strong
540 that cloud field adjustment in the first 6 hours of the forecast was required prior to some weather forecast variable



541 improvements to be apparent (surface pressure, dewpoint temperature, sea-level pressure). While the current forecast
542 cycling duration was constrained by operational requirements, this suggests that forecast cycling should include both
543 air-quality and meteorological variables during fully-coupled forecast spin-up periods. That is, the model tracer
544 concentrations 6 hours prior to the current forecast start-up could also be used during the initial meteorological spin-
545 up period, thus allowing chemistry and cloud formation to spin-up simultaneously. Scores for surface PM_{2.5}, NO₂,
546 and O₃ also generally improved with the incorporation of feedbacks, with some metrics showing large improvements.
547 In comparison to satellite-based AOD values, the current model's AOD values were generally biased low, with
548 exceptions being in the regions of Alberta and Saskatchewan with active forest fires where AOD was biased high.
549 The latter comparison also showed that large fires off-domain in Alaska likely had a large impact on AODs in the
550 eastern and northern section of the model domain due to missing boundary condition contributions. These sources
551 were missing due to operational limitations in the model simulations shown here.

552

553 **4 Effects of Feedbacks on Selected Simulation-Period Average Variables**

554 In this section, we compare time averages of the entire study period for the two simulations, both at the surface and in
555 vertical cross-sections through the model domain, to illustrate some of the changes in both weather and air-quality
556 associated with the incorporation of feedbacks. We have found differences at greater than 90% confidence between
557 the predicted meteorological and chemical forecasts in the vicinity of the Alberta/Saskatchewan forest fires, as well
558 as in contrasting changes between land and sea. We note again here that the “no-feedback” simulation makes use of
559 time and spatially invariant aerosol CCN and optical properties, within the meteorological portion of the model. The
560 comparisons thus show the differences associated with the use of climatological constant aerosol properties, and the
561 fully coupled model-generated aerosols.

562 As in the meteorological evaluation, we have made use of 90% confidence levels in order to gauge the level of
563 significance of the differences between the feedback and no-feedback simulations in the following analysis. At each
564 model grid cell the values of the standard deviation about the mean for each respective simulation was calculated.
565 The difference between the means becomes significant at a given confidence level c if the regions defined by $M_f \pm$
566 $z^* \frac{\sigma_f}{\sqrt{N}}$ and $M_{nf} \pm z^* \frac{\sigma_{nf}}{\sqrt{N}}$ do not overlap, where N is the number of points averaged, M_f is the feedback mean value, M_{nf}
567 is the no-feedback mean value, σ_f and σ_{nf} are the corresponding standard deviation, and z^* is the value of the \sqrt{c}
568 percentile point for the fractional confidence interval c of the normal distribution ($z^* = 1.645$ at $c = 0.90$). Grid cell
569 values where this overlap does not occur (i.e. where the mean values differ at or above the 90% confidence level) may
570 be defined via the following equation¹:

¹ Note that for cases where $M_{nf} > M_f$, significance at the confidence level associated with z^* occurs when the range
of standard errors about the mean do not overlap, i.e. $M_{nf} - z^* \frac{\sigma_{nf}}{\sqrt{N}} > M_f + z^* \frac{\sigma_f}{\sqrt{N}}$, or

$(M_{nf} - M_f) / \left(\frac{z^*}{\sqrt{N}} (\sigma_{nf} + \sigma_f) \right) > 1$. Similarly, for cases where $M_f > M_{nf}$, significance at the confidence level
associated with z^* occurs when $(M_f - M_{nf}) / \left(\frac{z^*}{\sqrt{N}} (\sigma_{nf} + \sigma_f) \right) > 1$. Equation (1) may thus be used to describe both
conditions.



571
$$\frac{|M_{nf}-M_f|}{\frac{z^*}{\sqrt{N}}(\sigma_f+\sigma_{nf})} > 1 \quad (1)$$

572 The differences in the mean grid cell values between the simulations for which the above quantity is greater than unity
573 thus differ at or greater than the 90% confidence level. Differences in the mean values, as well as the value of the
574 above ratio, are thus reported in the following section.

575 **4.1 Effects of Feedbacks on Time-Averaged Meteorology**

576 The feedback – no-feedback differences in the simulation-period average cloud droplet number density (number kg⁻¹
577 of air) and mass density (g water kg⁻¹ of air) along centred cross-sections spanning the length and width of the model
578 domain are shown in Figure 16 (the locations of the cross-sections are shown in Figure 1). The “Ocean”, “Land”, and
579 “Forest Fire” regions identified are with reference to the approximate locations of these features along these cross-
580 sections. Figure 16 also shows the confidence ratio values as described above – regions where the predicted mean
581 values differ at or above the 90% confidence level are shown in red, while those differences below the 90% confidence
582 interval are shown in blue. Feedbacks increase the cloud droplet number density over the northern part of the domain,
583 including the region impacted by the Alberta/Saskatchewan forest fires, from the surface up to about 600 mb (roughly
584 equivalent to hybrid level 0.600), and decrease further aloft and along the length of the model domain into the western
585 USA (Figure 16(a)). Cloud droplet numbers also decrease over the ocean, but increase eastwards over the land (Figure
586 16(b)). The latter is unrelated to the forest fires; this is an indication that the modelled aerosol number concentration
587 over the ocean is much lower than the single climatological aerosol population assumed in the no-feedback run,
588 resulting in lower cloud droplet number concentrations. In both cases, the differences are significant at the 90%
589 confidence level from the surface up to hybrid level 0.87 and in isolated regions at hybrid level 0.550 along the south
590 to north cross-section (Figure 16(c)), and over the ocean in the west to east cross-section (Figure 16(d)). Higher-than-
591 climatology aerosol loadings, a large portion of which are due to the forest fires, resulted increased cloud droplet
592 number densities in the lower troposphere, while decreasing them in the mid-to-upper troposphere. This impact of
593 feedbacks is in accord with the satellite observations of Saponaro *et al.* (2017), and was also seen in Takeishi *et al.*
594 (2020). In contrast, cloud droplet mass density (i.e. cloud liquid water content) largely decreases across the domain
595 along the north-south cross-section (Figure 16(e)), as well as over the ocean, with a varying pattern over the land in
596 the east-west cross-section (Figure 16(f)). The magnitudes and significance levels for the average change in cloud
597 droplet mass are lower than for cloud droplet number, with the most significant differences occurring over the ocean
598 (Figure 16(g,h)).

599 Consistent with the cloud droplet number changes, rain droplet numbers and mass mixing ratios increase with the
600 feedback simulation over both the forest region impacted by the forest fires (Figure 17(a,e)) and over the ocean (Figure
601 17(b,f)), with a varying impact over the land and more distant from the forest fire sources (Figure 17(f)). The changes
602 are significant at the 90% confidence level for rain droplet number in these regions (compare Figure 17(a) with 17(c);
603 17(b) with 17(d)), while the rain droplet mass changes approach but are below the 90% confidence level.

604 These results suggest that relative to the no-feedback simulation, which employs climatological aerosol CCN
605 properties, the AIE in the feedback simulation is causing significant change in hydrometeor numbers, and a less



606 significant increase in hydrometeor mass. In the forest fire-impacted region, the ADE and AIE in the feedback
607 simulation significantly increase the number of cloud droplets near the surface and decrease the number of cloud
608 droplets in the middle to upper troposphere (Figure 16(a,c)). The rain drop number in the middle troposphere (Figure
609 17(a,c)) also increases significantly between hybrid levels 0.90 to 0.70 (Figure 17(e,g)). Near-surface rain drop
610 number and rain drop mass differences throughout the cross sections (Figure 17(e,f)) fall below the 90% confidence
611 level (Figure 17(g,h)).

612 Over the oceans, water droplet number and mass both decrease (Figure 16(b,f)), and raindrop number and mass
613 increase (Figure 17(b,f)); more atmospheric water is converted to rain drops as a result of the feedbacks, relative to
614 the climatology in the no-feedback simulation. However, these changes are more significant aloft than at the surface,
615 with the difference in both rain drop number and mass falling below the 90% confidence level near the surface. We
616 interpret these changes as a shift in over-ocean liquid hydrometeor numbers and to a lesser degree the water mass aloft
617 from cloud droplets to rain drops due to the AIE in the feedback setup relative to the climatology of the no-feedback
618 simulation. The changes occur at the 90% confidence level aloft, but the near-surface changes are smaller and are
619 usually below the 90% confidence level.

620 Differences in the average precipitation flux and the confidence ratio values are shown in Figure 18. Changes in
621 average precipitation (Figure 18(a)) appear random, though locally these differences are significant at the 90%
622 confidence level (Figure 18(b)). Both the magnitude of the differences and the frequency in their reaching the 90%
623 confidence level increase westwards. Given the local and episodic nature of rainfall events the high level of
624 significance in this case probably results from the presence or absence of individual rainfall events between the two
625 simulations affecting the local average and standard deviations.

626 Several systematic changes in the average values of the model's meteorological output fields were noted due to the
627 use of feedbacks relative to aerosol property climatologies (Figure 19), although all fall below the 90% confidence
628 level for the difference in the mean values between the two simulations (Figure 20). Surface air temperature generally
629 increased (Figure 19(b)) though less so in the region most affected by forest fires, dewpoint temperature decreased
630 (Figure 19(c)) implying a decrease in relative humidity with feedbacks. Surface pressure increased, particularly in
631 the region downwind of the Alberta / Saskatchewan fires (Figure 19(d)). Planetary boundary layer height increased
632 over the land (Figure 19(e)), consistent with decreased atmospheric stability. The friction velocity also increased with
633 the use of feedbacks (Figure 19(f)); this is consistent with a decrease in stability and an increase in turbulent energy
634 The air temperature increases are limited to the lowest part of the atmosphere (Figure 21 (a,b)), usually below hybrid
635 level 0.914 (approximately 1km above the surface). The feedbacks decrease temperatures between hybrid levels 0.914
636 and 0.721. Feedbacks thus increase or do not affect temperatures near the surface, and decrease temperatures in the
637 lower free Troposphere. However, all of these features, despite their large geographic range, fall below the 90%
638 confidence level, reflecting the large variability in surface temperatures contained within each simulation. Longer
639 time simulations than carried out here are required in order to improve confidence in the temperature predictions
640 across all forecast hours. However, these results, particularly for surface temperature, may be contrasted with Figures
641 6(a), 12 and 13, which suggests that temperature differences rise above the 90% confidence interval at specific forecast
642 times at surface and upper atmosphere measurement sites.



643 **4.2 Effects of Feedbacks on Time-Averaged Chemistry**

644 In the previous meteorological impacts section, changes in aerosol loading relative to the climatology, dominated by
645 forest fires, were shown to have a significant impact on cloud formation and atmospheric temperatures through ADE
646 and AIE. These might be expected to in turn influence and be influenced by particulate matter emitted by the forest
647 fires, with the plume rise of the forest fires dependent on the meteorological changes. Air temperatures increase
648 slightly in the model surface layer (Figure 19(b), +0.01 to +0.05 °C) but decrease at greater magnitudes through the
649 rest of the lower Troposphere (hybrid level 0.980 up to 0.667, Figure 21), with a maximum decrease of -0.5°C between
650 hybrid levels 0.871 and 0.824. The increase in air temperatures near the surface and decreases aloft (a stronger negative
651 vertical gradient in temperature) implies a decrease in atmospheric stability associated with feedbacks, given that the
652 overall temperature gradient from the surface has become more negative. This is echoed by the response of the
653 concentration fields to the near-surface stability change, as can be seen through comparisons of the PM_{2.5}, NO₂ and
654 O₃ surface concentrations changes (Figure 22) and as vertical cross-sections (Figures 23, 24, 25), respectively.

655 For all three surface fields, changes above the 90% confidence level occur near the forest fires themselves (red regions,
656 near top of model domain, Figure 22(a,b,c)). The differences in particulate matter concentrations are also significant
657 at the 90% confidence level throughout the model domain (Figure 22(a)). Note that while the PM_{2.5} mean values are
658 significantly different at the 90% confidence level throughout the model domain, the magnitude of those differences
659 are sometimes small, particularly in the upper atmosphere, where the aerosol concentrations are relatively small.
660 However, the regions with the larger magnitude regional differences in PM_{2.5} concentrations also occur at greater
661 than the 90% confidence level (compare spatial locations of coloured regions in Figure 22(a,b) to red regions in Figure
662 22(c,d)).

663 Near-surface PM_{2.5} decreases in the regions downwind of the forest fires (Figure 22(a), Figure 23(a), note the large
664 blue region and more intense blue region near surface in 23(a)), suggesting less PM_{2.5} mass is present near the surface
665 due to the feedbacks. This could reflect a change in injection height of the plumes in addition to other transport
666 changes associated with the decrease in atmospheric stability. Lower troposphere decreases in PM_{2.5} over the ocean
667 and increases over the land at or greater than the 90% confidence level may also be seen (Figure 23(b,d)).

668 Feedbacks result in an increase in near-surface NO₂ in several inland urban centers and less NO₂ at surface level
669 downwind (Figure 22(b), though these differences are only significant at the 90% confidence level within the forest
670 fire plumes (Figure 24(a,c)). Ocean versus land NO₂ differences remain below the 90% confidence level.

671 Feedbacks decreased surface O₃ near the forest fires (Figure 22(c), Figure 25(a)), while decreasing increasing O₃ aloft.
672 The forest fires are also the only area where the differences in between mean ozone forecasts have greater than the
673 90% confidence.

674 Overall, the most significant effects of the feedbacks were: (1) changes in PM_{2.5} concentrations throughout the model
675 domain, and (2) changes in NO₂ and O₃ within the forest fire plumes.

676 The feedback-induced changes in primary and secondary pollutants in the forest fire regions are consistent with the
677 decrease in atmospheric stability noted above – a greater proportion of the primary particulate matter and NO₂ resulting
678 from near-surface forest fire emissions of NO are carried upwards with the addition of feedbacks. The decrease in
679 surface ozone and increase further aloft in the fire region (Figure 25(a)) spatially matches the changes in NO₂ (Figure



680 23(a)) – this implies that the changes associated with feedbacks occur in NO_x-limited environments, i.e., with
681 relatively high VOC/NO_x ratios. In such environments, decreases in NO_x emissions may lead to decreases in the rate
682 of secondary O₃ formation.

683 Our analysis thus suggests enhanced upward transport occurs in forest fire plumes due to feedbacks, and that this
684 transport is linked to feedback-induced: (1) decreases in local atmospheric stability (Figure 21(a)); (2) increases in
685 cloud droplet numbers near the surface (Figure 16(a)); and (3) increases in rain drop numbers aloft (Figure 17(a)).
686 This combination suggests the presence of an AIE feedback loop – decreased stability results in higher forest fire
687 plume rise, in turn lofting proportionally more particles to higher levels in the atmosphere where they may act as cloud
688 condensation nuclei, increasing cloud droplets aloft (Figure 16(a)). This in turn results in increased lower middle
689 troposphere cooling, through the 1st AIE (increase in cloud droplet numbers aloft leading to increased cloud albedo
690 and cooling of the atmosphere below the cloud tops) while the corresponding decreases in particles and cloud
691 condensation nuclei at lower levels results in a smaller near-surface impact on the AIE and ADE, hence relatively
692 minor changes on near-surface temperatures (Figure 21(a)). This combination maintains a feedback-induced less
693 stable temperature gradient, relative to the no-feedback simulation employing aerosol property climatologies.

694 Similarly, over the oceans, the feedback-induced decrease in surface PM_{2.5} (Figure 23(a)) is accompanied by lower-
695 middle troposphere cooling (Figure 21(b), note the dark blue band over ocean between hybrid levels 0.914 and 0.824),
696 implying a decrease in stability. While aerosol increases aloft are not large, the lower level PM_{2.5} act as CCN,
697 resulting in increased cloud formation and convection, in turn increasing cloud droplet and rain drop numbers aloft,
698 which through the 1st AIE, maintains the slightly less stable temperature profile. We acknowledge that these changes
699 in temperature fall below the 90% confidence level for the averages over all times, though note that differences in
700 mean bias relative to observations for the two simulations became significantly different at specific times of day in
701 the forecasts (Figure 6(a), hours 3, 6, 15 and 18, corresponding to 15, 18, 3 and 6 UT, or 9 AM, 12 noon, 9 PM, and
702 midnight MDT), implying that the temperature changes at these specific times reach a higher level of significance.

703 **4.3 Summary, Differences in Forecast Simulation-Period Averages**

704 Relative to the no-feedback simulation employing an aerosol climatology, the AIE feedback as simulated here is
705 associated with decreases in stability over both ocean and forest-fire influenced land areas. Over oceans, near-surface
706 particulate matter is removed as cloud condensation nuclei, resulting in increased cloud droplet numbers, maintaining
707 the temperature gradient through the 1st aerosol indirect effect. In the vicinity of forest fires, decreases in stability
708 result in increased transport of PM_{2.5} aloft, increasing the availability of cloud condensation nuclei aloft, increasing
709 cloud droplet numbers aloft, hence also maintaining the increased temperature gradient through the 1st aerosol indirect
710 effect. We note that the ADE may also play a weak role, particularly in the southern part of the domain, where lower
711 atmosphere temperature gradient increases are not accompanied by significant changes in cloud droplet numbers
712 (Figure 16(a), southern half of the cross-section), but are accompanied by significant though small magnitude increases
713 in PM_{2.5} in the lower atmosphere (Figure 23(a), southern half of cross-section), and temperature profile changes
714 (Figure 21) below the 90% confidence level.

715



716 5 Conclusions

717 The work carried out here suggests that the answers to our two research questions (“Can fully coupled models improve
718 both air-quality and meteorological forecasts?” and “Are the changes in forest fire forecasts associated with
719 implementing forest fire emissions within a fully coupled model sufficient to significantly perturb weather and
720 meteorology?”) are both a qualified “yes”.

721 The simulations analyzed here were conducted in preparation for an experimental forecast carried out as part of the
722 FIREX-AQ campaign, and hence were limited by operational time constraints to a sequence of nested 24-hour
723 forecasts. However, the high resolution domain size employed was sufficiently large to result in improvements in
724 weather forecast performance for both surface and profile variables at or above the 90% confidence level.
725 Improvements in model performance for PM_{2.5}, NO₂ and O₃ were also found, across most statistical measures. The
726 differences between feedback and no-feedback simulations occurred at or above the 90% confidence level throughout
727 the model domain for PM_{2.5}, though were limited at the 90% confidence level for NO₂ and O₃ to the immediate
728 vicinity of the forest fires. There, increased vertical transport associated with feedbacks lowered near-surface NO₂
729 concentrations and increased them aloft, resulting in reduced surface O₃ in the NO_x-limited regions of the forest fire
730 plume.

731 The simulations suggest that the homogeneous mixture approach for aerosol optical properties results in a general
732 under-prediction of aerosol optical depths, in accord with Curci *et al.* (2015), and external mixture approaches are
733 recommended in further study. However, this general negative bias in simulated AOD is locally offset by positive
734 biases in the vicinity of forest fires. This suggests that forest fire plumes have significantly different optical properties,
735 and may be less hygroscopic than industrial aerosols of comparable size. Special / separate treatment of forest fire
736 CCN and optical properties are therefore also recommended in future work.

737 Fully coupling forest fire plume rise calculations with the weather parameters was shown to have a significant impact
738 on the height of primary pollutants reached by forest fires, the formation of near-surface ozone near the forest fires,
739 and on particulate matter throughout much of the three-dimensional model domain. These changes were largely driven
740 by the AIE, which maintains an increased temperature gradient (reduced stability) over the forest-fire-influenced and
741 oceanic portions of the region studied. Weak evidence for the influence of the ADE was shown in the southern part
742 of the domain, where increases in particulate matter were also accompanied by decreases in stability between the
743 surface and the lower-middle troposphere (the differences were at a lower than 90% confidence level for these
744 comparisons of temperatures averaged over all model times).

745 Relative to the no-feedback aerosol climatology for CCN and aerosol optical properties, the simulations carried out
746 here suggested that in the vicinity of forest fires feedbacks significantly increase cloud droplet near the surface,
747 increase cloud droplet number aloft, and significantly increase rain drop number densities aloft, relative to forecasts
748 driven by climatological aerosol properties. Over the oceans, feedbacks decreased cloud droplet number density and
749 increased rain drop number density aloft. Cloud droplet mass increased to a lesser degree (with smaller regions above
750 the 90% confidence level), as did rain drop mass (the mean differences for which for the most part remained below
751 the 90% confidence level). This provides some evidence for a shift in atmospheric water mass associated with
752 feedbacks from cloud water to rain over the oceans relative to the no-feedback climatology, though this shift occurred



753 largely within the variability of the cloud fields within each simulation. Longer simulations may be needed to achieve
754 higher confidence in this finding.

755

756 **Data Availability**

757 The datasets used here for model evaluation are available from the publicly accessible websites (AQS network)
758 <https://www.epa.gov/aqs> and (NAPS network) <http://maps-cartes.ec.gc.ca/rnspa-naps/data.aspx>.

759 **Author Contribution**

760 PAM: experiment design, conceptualization, analysis, writing of manuscript drafts; AA: model code and run script
761 design and implementation, statistical analysis of model results, model analysis graphics; JC: forest fire emissions
762 processing system design and coding, manuscript contributions, draft review and assistance; BP: forecast system
763 simulations and design; Wanmin Gong¹: indirect effect updates, advice on P3 implementation, manuscript
764 contributions, manuscript review; CStroud: code version contributions, manuscript review; CSioris: AOD analysis,
765 manuscript contributions and review; KS: forest fire emissions processing system design and coding, manuscript
766 contributions and review; PC: forecast system simulations and design; JZ: emissions processing and input field
767 assistance, manuscript review and contributions; J.M.: indirect effect updates and advice on implementing AIE in the
768 P3 scheme, manuscript review and contributions.

769

770 **Competing Interests**

771 The authors declare that they have no conflict of interest.

772

773 **References**

- 774 Abdul-Razzak, H., and Ghan, S.J.: A parameterization of aerosol activation. 3. Sectional representation, *J. Geophys.*
775 *Res. Atm.*, 107 (D3), <http://dx.doi.org/10.1029/2001JS000483>, 6 pp, 2002.
- 776 Ahmadov, R., James, E., Grell, G., Alexander, C., Benjamin, S., McKeen, S., Pereira, G., Freitas, S., Csiszar, I.,
777 Tsidulko, M., Kondragunta, S., Xu, C., Wong, K.Y., and Albers, S.: Forecasting smoke, visibility and smoke-
778 weather interactions using a coupled meteorology-chemistry modeling system: Rapid Refresh and High-
779 Resolution Rapid Refresh coupled with Smoke (RAP/HRRR-Smoke), EGU 2019 Conference presentation,
780 Geophysical Research Abstracts, Vol. 21, EGU2019-18605, 2019.
781 <https://meetingorganizer.copernicus.org/EGU2019/EGU2019-18605.pdf>, 2019.
- 782 Akingunola, A., Makar, P.A., Zhang, J., Darlington, A., Li, S.-M., Gordon, M., Moran, M.D., and Zheng, Q.: A
783 chemical transport model study of plume-rise and particle size distribution for the Athabasca oil sands, *Atm.*
784 *Chem. Phys.*, 18, 8667-8688, 2018.
- 785 Albrecht, B.A.: Aerosols, cloud microphysics, and fractional cloudiness, *Science*, 245(4923), 1227-1230,
786 <https://doi.org/10.1126/science.245.4923.1227>, 1989.



- 787 Baro, R., Palacios-Pena, L., Baklanov, A., Balzarini, A., Brunner, D., Forkel, R., Hirtl, M., Honzak, L., Perez, J.L.,
788 Pirovano, G., San Jose, R., Schroder, W., Werhahn, J., Wolke, R., Zabhar, R., Jimenez-Guerrero, P.: Regional
789 effects of atmospheric models on temperature: an evaluation of an ensemble of online coupled models, *Atm.*
790 *Chem., Phys.*, 17, 9677-9696, 2017.
- 791 Briggs, G.A.: A plume rise model compared with observations, *J. Air. Poll. Cont. Assoc.*, 15, 433-438,
792 <https://doi.org/10.1080/000022470.1965.10468404>, 1965.
- 793 Briggs, G.A.: Plume rise and buoyancy effects, atmospheric sciences and power production, in: DOE/TIC-
794 27601(DE84005177), edited by Randerson, D., Technical Information Center, U.S. Dept. of Energy, Oak Ridge,
795 TN, USA, 327-366, 1984.
- 796 Carrera, M.L., Belair, S., Bilodeau, B.: The Canadian Land Data Assimilation System (CalDAS): description and
797 synthetic evaluation study, *J. Hydrometeor.*, 16, 1293-1314, 2015.
- 798 Chen, J., Anderson, K., Pavlovic, R., Moran, M.D., Englefield, P., Thompson, D.K., Munoz-Alpizar, R., and Landry,
799 H.: The FireWork v2.0 air quality forecast system with biomass burning emissions from the Canadian Forest
800 Fire Emissions Prediction System v2.03, *Geosci. Model Dev.*, 12, 3283–3310, 2019.
- 801 Clark, T.L., Coen, J. and Latham, D.: Description of a coupled atmosphere–fire model. *International Journal of*
802 *Wildland Fire*, 13(1), pp.49-63, 2004.
- 803 Clark, T.L., Jenkins, M.A., Coen, J.L., Packham, D.R.: A coupled atmosphere-fire model: role of convective froude
804 number and dynamic engineering at the fireline, *Int. J. Wildland Fire*, 6, 177-190, 1996.
- 805 Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M., Staniforth, A.: The operational CMC/MRB global
806 environmental multiscale (GEM) model. Part I: design considerations and formulation. *Mon. Wea. Rev.*, 126,
807 1373-1395, 1998.
- 808 Curci, G., Hogrefe, C., Bianconi, R., Im, U., Balzarini, A., Baro, R., Brunner, D., Forkel, R., Giordano, L., Hirtl, M.,
809 Honzak, L., Jimenez-Guerrero, P., Knote, C., Langer, M., Makar, P.A., Pirovano, G., Perez, J.L., San Jose, R.,
810 Syrakov, D., Tuccella, P., Werhahn, J., Wolke, R., Zabkar, R., Zhang, J., and Galmarini, S.: Uncertainties of
811 simulated aerosol optical properties induced by assumptions on aerosol physical and chemical properties: an
812 AQMEII2 perspective, *Atm. Env.*, 115, 541-522, 2015.
- 813 Ezhova, E., Ylivinkka, I., Kuusk, J., Komsaare, K., Vana, M., Krasnova, A., Noe, S., Archinov, M., Belan, B., Park,
814 S-B., Lavric, J.V., Heimann, M., Petaja, T., Vesala, T., Mammarella, I., Kolari, P., Back, J., Rannik, U.,
815 Kerminen, V.-M., and Kulmala, M.: Direct effect of aerosols on solar radiation and gross primary production in
816 boreal and hemiboreal forests, *Atm. Chem. Phys.*, 18, 17863-17881, 2018.
- 817 Geer, A.J.: Significance of changes in medium-range forecast scores, *Tellus A.*, 68, 30229,
818 <http://dx.doi.org/10.3402/tellusa.v68.30229>, 2016.
- 819 Girard, C., Plante, A., Desgagne, M., McTaggart-Cowan, R., Cote, J., Charron, M., Gravel, S., Lee, V., Patoine, A.,
820 Qaddouri, A., Roch, M., Spacek, L., Tanguay, M., Vaillancourt, P.A., and Zadra, A.: Staggered vertical
821 discretization of the Canadian Environmental Multiscale (GEM) model using a coordinate of the log-
822 hydrostatic-pressure type. *Mon. Wea. Rev.*, 142, 1183–1196, 2014.



- 823 Gong, S.L. Barrie, L.A., Blanchet, J.-P., von Salzen, K., Lohmann, U., Lesins, G., Spacek, L., Zhang, L.M., Girard,
824 E., Lin, H., Leaitch, R., Leighton, H., Chylek, P., Huang, P.” Canadian Aerosol Module: a size-segregated
825 simulation of atmospheric aerosol processes for climate and air quality models. 1. Module development. J.
826 Geophys. Res., 108, 4007, <http://dx.doi.org/10.1029/2011JG002002>, 2003.
- 827 Gong, W., Beagley, S.R., Zhang, J., Cousineau, S., Chen, J., Sassi, M., Munoz-Alpizar, R., Morrison, H., Lyons, L.,
828 and Bellavance, P.: Modelling regional air quality in the Canadian Arctic: impact of North American wildfire
829 and Arctic shipping emissions, in: Air Pollution Modelling and Its Application, XXIV, edited by Steyn, D.G.
830 and Chaumerliac, N., Springer, Switzerland, 301-306, 2016.
- 831 Gong, W., Makar, P.A., Zhang, J., Milbrandt, J., Gravel, S., Hayden, K.L., Macdonald, A.M., Leaitch, W.R.:
832 Modelling aerosol-cloud-meteorology interaction: a case study with a fully coupled air quality model (GEM-
833 MACH), *Atm. Env.*, 115, 695-715, 2015.
- 834 Hoegh-Guldberg, O., Jacob, D., Taylor, M., Bindi, Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K.L.,
835 Engelbrecht, F., Guiot, J., Hijioka, Y., Mehrotra, S., Payne, A., Seneviratne, S.I., Thomas, A., Warren, R., and
836 Zhou, G.: Impacts of 1.5°C Global Warming on Natural and Human Systems. In: *Global Warming of 1.5°C. An*
837 *IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global*
838 *greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate*
839 *change, sustainable development, and efforts to eradicate poverty* [Masson-Delmotte, V., Zhai, P., Pörtner, H.-
840 O., Roberts, D., Skea, J., Shukla, P.R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S.,
841 Matthews, J.B.R., Chen, Y., Zhou, X., Gomis, M.I., Lonnoy, E., Maycock, T., Tignor, M., and Waterfield, T.
842 (eds.)]. In Press, 2018.
- 843 Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R.J.,
844 Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser,
845 J.W., Kapsomenakis, J., Lefever, K., Leitao, J., Razinger, M., Richter, A., Schultz, M.G., Simmons, A.J., Suttie,
846 M., Stien, O., Thepaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., The MACC team: The MACC reanalysis:
847 an 8 yr data set of atmospheric composition, *Atm. Chem. Phys.*, 13, 4073-4109, [http://dx.doi.org/10.5194/acp-](http://dx.doi.org/10.5194/acp-13-4703-2013)
848 [13-4703-2013](http://dx.doi.org/10.5194/acp-13-4703-2013), 2013.
- 849 James, E., Ahmadov, R., and Grell, G.: Realtime Wildfire Smoke Prediction in the United States: the HRRR-Smoke
850 Model, in: EGU General Assembly Conference Abstracts, vol. 20 of EGU General Assembly Conference
851 Abstracts, p. 19526, 2018
- 852 Jiang, Y., Yang, X.-Q., and Liu, X., Seasonality in anthropogenic aerosol effects on East Asian climate simulated
853 with CAM5, *J. Geophys. Res. Atmos.*, 120, 10837-10861, doi:10.1002/2015JD023451, 2015.
- 854 Jouan, C., J.A. Milbrandt, P. Vaillancourt, F. Chosson, H. Morrison: Adaptation of the Predicted Particles Properties
855 (P3) microphysics scheme for large-scale numerical weather prediction. *Wea. Forecasting* (under review),
856 2020.
- 857 Jung, J., Souri, A.H., Wong, D., Lee, S., Jeon, W., Kim, J., Choi, Y.: The impact of the direct effect of aerosols on
858 meteorology and air quality using aerosol optical depth simulation during the KORUS-AQ campaign, *J.*
859 *Geophys. Res. Atm.*, 124, 8303-8319, <https://doi.org/10.1029/2019JD030641>, 2019.



- 860 Linn, R., Reisner, J., Colman, J.J. and Winterkamp, J.: Studying wildfire behavior using FIRETEC. International
861 journal of wildland fire, 11(4), pp.233-246, 2002.
- 862 Lu, Z., and Sokolik, I.N.: Examining the impact of smoke on frontal clouds and precipitation during the 2002 Yakutsk
863 wildfires using the WRF-Chem-SMOKE model and satellite data, J. Geophys. Res. Atm., 122, 12765-12785,
864 <https://doi.org/10.1002/2017JD027001>, 2017.
- 865 Luig, A., Bofinger, S., Beyer, H.G.: Analysis of confidence intervals for the prediction of regional wind power
866 output, Proc. European Wind Energy Conference, Copenhagen, pp 725-728, 2001. Available at
867 <https://www.researchgate.net/publication/228793040> Analysis of confidence intervals for the prediction o
868 f regional wind power output, last accessed August 23, 2020.
- 869 Makar, P.A., Gong, W., Milbrandt, J., Hogrefe, C., Zhang, Y., Curci, G., Zabkar, R., Im, U., Balzarini, A., Baro, R.,
870 Bianconi, R., Cheung, P., Forkel, R., Gravel, S., Hirtl, H., Honzak, L., Hou, A., Jimenz-Guerrero, P., Langer,
871 M., Moran, M.D., Pabla, B., Perez, J.L., Pirovano, G., San Jose, R., Tuccella, P., Werhahn, J., Zhang, J.,
872 Galmarini, S.: Feedbacks between air pollution and weather, part 1: Effects on weather. Atmospheric
873 Environment, 115, 442-469, 2015a.
- 874 Makar, P.A., Gong, W., Hogrefe, C., Zhang, Y., Curci, G., Zabkar, R., Milbrandt, J., Im, U., Balzarini, A., Baro, R.,
875 Bianconi, R., Cheung, P., Forkel, R., Gravel, S., Hirtl, H., Honzak, L., Hou, A., Jimenz-Guerrero, P., Langer,
876 M., Moran, M.D., Pabla, B., Perez, J.L., Pirovano, G., San Jose, R., Tuccella, P., Werhahn, J., Zhang, J.,
877 Galmarini, S.: Feedbacks between air pollution and weather, part 2: Effects on chemistry. Atmospheric
878 Environment, 115, 499-526, 2015b.
- 879 Makar, P.A., Moran, M.D., Schultz, M.T., Taylor, A.: Speciation of volatile organic compound emissions for
880 regional air quality modeling of particulate matter and ozone, J. Geophys. Res., 108,D2, 4041,
881 doi:1029/2001JD000797, 2003.
- 882 Makar, P.A., Staebler, R.M., Akingunola, A., Zhang, J., McLinden, C., Kharol, S.K., Pabla, B., Cheung, P. and
883 Zheng, Q.: The effects of forest canopy shading and turbulence on boundary layer ozone, Nature
884 Communications, 8, art. no. 15243, doi: 10.1038/ncomms15243, 2017.
- 885 Makar, P.A., Stroud, C., Akingunola, A., Zhang, J., Ren, S., Cheung, P., and Zheng, Q.: Vehicle induced turbulence
886 and atmospheric pollution, Proceedings of the National Academy of Sciences of the USA (under review), 2020.
- 887 Mallet, M., Solmon, F., Roblou, L., Peers, F., Turquety, S., Waquet, F., Jethva, H., Torres, O.: Simulation of optical
888 properties and direct and indirect radiative effects of smoke aerosols over marine stratocumulus clouds during
889 summer 2008 in California with the regional climate model RegCM, J. Geophys. Res. Atm., 122, 10312-10337,
890 <https://doi.org/10.1002/2107JD026905>, 2017.
- 891 Markowicz, K.M., Pakszys, R., Ritter, C., Zielenski, T., Udisti, R., Cappelletti, D., Mazzola, M., Shobara, M., Xian,
892 P., Zawadzka, O., Lisok, J., Petelski, T., Makuch, P., Karasinski, G.: Impact of North American intense fires on
893 aerosol optical properties measured over the European Arctic in July 2015, J. Geophys. Res. Atm., 121, 14487-
894 14512, doi:10.1002/2016JD025310, 2016.



- 895 Milbrandt, J.A. and Morrison, H.: Parameterization of Cloud Microphysics Based on the Prediction of Bulk Ice
896 Particle Properties. Part III: Introduction of Multiple Free Categories. *J. Atmos. Sci.*, 73, 975-995, doi:
897 <https://doi.org/10.1175/JAS-D-15-0204.1>, 2016.
- 898 Milbrandt, J. A., S. Bélair, S., Faucher, M., Vallée, M., Carrera, M.L., and Glazer, A.: The pan-Canadian High
899 Resolution (2.5 km) Deterministic Prediction System. *Wea. Forecasting*, 31, 1791-1816, 2016.
- 900 Milbrandt, J.A., and Yau, M.K.: A multimoment bulk microphysics parameterization. Part I: analysis of the role of
901 the spectral shape parameter, *J. Atmos. Sci.*, 62, 3051-3064, 2005a.
- 902 Milbrandt, J.A., and Yau, M.K.: A multimoment bulk microphysics parameterization. Part II: a proposed three-
903 moment closure and scheme, *J. Atmos. Sci.*, 62, 3065-3081, 2005b
- 904 Moran, M.D., Pavlovic, R., and Anselmo, D.: Regional air quality deterministic prediction system (RAQDPS):
905 update from version 019 to version 020, Montreal, available at:
906 [http://collaboration.cmc.ec.gc.ca/cmc/CMOI/product_guide/docs/tech_note/technote_raqdps-](http://collaboration.cmc.ec.gc.ca/cmc/CMOI/product_guide/docs/tech_note/technote_raqdps-v2.0_20180918_e.pdf)
907 [v2.0_20180918_e.pdf](http://collaboration.cmc.ec.gc.ca/cmc/CMOI/product_guide/docs/tech_note/technote_raqdps-v2.0_20180918_e.pdf), last access: 1 December 2018.
- 908 Morrison, H., and Milbrandt, J.A.: Parameterization of Cloud Microphysics Based on the Prediction of Bulk Ice
909 Particle Properties. Part I: Scheme Description and Idealized Tests. *J. Atmos. Sci.*, 72, 287-311, doi:
910 <https://doi.org/10.1175/JAS-D-14-0065.1>, 2015.
- 911 Morrison, H. and W. W. Grabowski, W.W.: Modeling supersaturation and subgrid-scale mixing with two-moment
912 warm bulk microphysics. *J. Atmos. Sci.*, 65, 792-812, doi: <https://doi.org/10.1175/2007JAS2374.1>, 2008.
- 913 Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D.
914 , Mendoza, B. , Nakajima, T., Robock, A., Stephens, G., Takemura, T. and Zhang, H.: Anthropogenic and
915 Natural Radiative Forcing. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working*
916 *Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D.
917 Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds)].
918 Cambridge University Press, Cambridge, United Kingdom and New York, USA, 2013.
- 919 Nazarenko, L., Rind, D., Tsigaridis, K., Del Genio A.D, Kelly, M., and Tausnev, N.: Interactive nature of climate
920 change and aerosol forcing, *J. Geophys. Res. Atm.*, 122, 3457-3480, doi:10.1002/2016JD025809, 2017.
- 921 Oeropoulos, L., Cho, N., and Lee, D.: A global survey of apparent aerosol-cloud interaction signals, *J. Geophys. Res.*
922 *Atm.*, 125, e2019JD031287, <https://doi.org/10.1029/2019JD031287>, 21pp, 2020.
- 923 O'Neill, N.T., Eck, T.G., Holben, B.N., Smirnov, A., Royer, A., and Li, Z.: Optical properties of boreal forest fire
924 smoke derived from Sun photometry, *J. Geophys. Res.*, 107, 4125, doi 10.1029/2001JD000877, 21 pp., 2002
- 925 Palacios-Pena, L., Baro, R., Baklanov, A., Balzarini, A., Brunner, D., Forkel, R., Hirtl, M., Honzak, L., Lopez-
926 Romero, J.M., Motavez, J.P., Perez, J.L., Pirovano, G., San Jose, R., Schroder, W., Werhahn, J., Wolke, R.,
927 Zabkar, R., and Jimenez-Guerrero, P.: An assessment of aerosol optical properties from remote-sensing
928 observations and regional chemistry-climate coupled models over Europe, *Atm. Chem. Phys.*, 18, 5021-5043,
929 2018.
- 930 Pan, L., Kim, H. C., Lee, P., Saylor, R., Tang, Y., Tong, D., Baker, B., Kondragunta, S., Xu, C., Ruminski, M. G.,
931 Chen, W., McQueen, J., and Stajner, I.: Evaluating a fire smoke simulation algorithm in the National Air Quality



- 932 Forecast Capability (NAQFC) by using multiple observation data sets during the Southeast Nexus (SENEX)
933 field campaign, *Geosci. Model Dev. Discuss.*, <https://doi.org/10.5194/gmd-2017-207>, 2017.
- 934 Pavlovic, R., Chen, J., Anderson, K., Moran, M.D., Beaulieu, P.A., Davignon, D. and Cousineau, S.: The FireWork
935 air quality forecast system with near-real-time biomass burning emissions: Recent developments and evaluation
936 of performance for the 2015 North American wildfire season. *Journal of the Air & Waste Management*
937 *Association*, 66(9), 819-841, 2016.
- 938 Penner, J.E., Zhou, C., Garnier, A., Mitchell, D.L.: Anthropogenic aerosol indirect effects in cirrus clouds, *J.*
939 *Geophys. Res. Atm.*, 123, 11652-11677, <https://doi.org/10.1029/2018JD029204>, 2018.
- 940 Pinson, P., and Kariniotakis, G.: On-line adaptation of confidence intervals based on weather stability for wind power
941 forecasting, in *Proc. Of the Global Wind Power Conference*, Chicago, USA, 2004. Available at [https://hal-](https://hal-mines-paristech.archives-ouvertes.fr/hal-00529488/document)
942 [mines-paristech.archives-ouvertes.fr/hal-00529488/document](https://hal-mines-paristech.archives-ouvertes.fr/hal-00529488/document), last accessed August 23, 2020.
- 943 Pinto, J. O., Grim, J.A., Steinter, M.: Assessment of the High-Resolution Rapid Refresh model's ability to predict
944 mesoscale convective systems using object-based evaluation. *J. Atmos. Sci.*, 30, 892-913, 2015.
- 945 Rothenberg, D., Avramov, A., Wang, C.: On the representation of aerosol activation and its influence on model-
946 derived estimates of the aerosol indirect effect, *Atm. Chem. Phys.*, 18, 7961-7983, 2018.
- 947 Sandu, A. and Sander, R.: Technical note: Simulating chemical systems in Fortran90 and Matlab with the Kinetic
948 PreProcessor KPP-2.1, *Atmos. Chem. Phys.*, 6, 187–195, <https://doi.org/10.5194/acp-6-187-2006>, 2006
- 949 Saponaro, G., Kolmonen, P., Sogacheva, L., Rodriguez, E., Virtanen, T., and de Leeuw, G.: Estimates of the aerosol
950 indirect effect over the Baltic Sea region derived from 12 years of MODIS observations, *Atm. Chem. Phys.*, 17,
951 3133-3143, 2017.
- 952 Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-
953 France convective-scale operational model. *Mon. Wea. Rev.*, 139, 876-913, 2010.
- 954 Stockwell, W.R. and Lurmann, F.W.: Intercomparison of the ADOM and RADM gas-phase chemical mechanisms,
955 Electric Power Institute Topical Report, Electric Power Institute, Palo Alto, California, 323pp, 1989.
- 956 Stroud, C.A., Makar, P.A., Zhang, J., Moran, M.D., Akingunola, A., Li, S.-M., Leithead, A., Hayden, K., Siu, M.:
957 Improving air quality model predictions of organic species using measurement-derived organic gaseous and
958 particle emissions in a petrochemical-dominated region. *Atmospheric Chemistry and Physics*, 18 (18), pp.
959 13531-13545, 2018.
- 960 Sundqvist, H.: Parameterization of condensation and associated clouds in models for weather prediction and general
961 circulation simulation. In: Schlesinger M.E. (eds) *Physically-Based Modelling and Simulation of Climate and*
962 *Climatic Change*. NATO ASI Series (Series C: Mathematical and Physical Sciences), vol 243. Springer,
963 Dordrecht, 1988.
- 964 Suzuki, K., and Takemura, T.: Perturbations to global energy budget due to absorbing and scattering aerosols, *J.*
965 *Geophys. Res. Atm.*, 124, 2194-2209, <https://doi.org/10.1029/2018JD029808>, 2019.
- 966 Takeishi, A., Storerlvm, T., and Fierce, L.: Disentangling the microphysical effects of fire particles on convective
967 clouds through a case study, *J. Geophys. Res. Atm.*, doi:10.1029/2019JD031890, 2020.
- 968 Twomey, S.: The influence of pollution on the shortwave albedo of clouds, *J. Atmos. Sci.*, 34, 1149-1152, 1977.



- 969 Wang, J., Allen, D.J., Pickering, K.E., Li, Z., and He, H.: Impact of aerosol direct effect on East Asian air quality
970 during the EAST-AIRE campaign, *J. Geophys. Res. Atm.*, 121, 6534-6554, doi:10.1002/2016JD025108, 2016.
- 971 Whaley, C.H., Makar, P.A., Shephard, M.W., Zhang, L., Zhang, J., Zheng, Q., Akingunola, A., Wentworth, G.R.,
972 Murphy, J.G., Kharol, S.K., and Cady-Periera, K.E.: Contributions of natural and anthropogenic sources to
973 ambient ammonia in the Athabasca oil sands and north-western Canada, *Atm. Chem. Phys.*, 18, 2011-2034,
974 2018.
- 975 Zhang, L., Brook, J.R., and Vet, R.: A revise parameterization for gaseous dry deposition in air-quality models,
976 *Atm. Chem. Phys.*, 3, 2067-2082, <https://doi.org/10.5194/acp-3-2067-2003>, 2003.
- 977 Zhao, X., Liu, Y., Yu, F., and Heidinger, A.K.: Using long-term satellite observations to identify sensitive regimes
978 and active regions of aerosol indirect effects for liquid clouds over global oceans, *J. Geophys. Res. Atm.*, 123,
979 457-472, <https://doi.org/10.1002/2017JD027187>, 2018.
- 980 Zhu, J., and Penner, J.E.: Indirect effects of secondary organic aerosol on cirrus clouds, *J. Geophys., Res. Atm.*, 125,
981 e2019JD032233, <https://doi.org/10.1029/2019JD032233>, 19 pp., 2020.
- 982
- 983



984 **Tables:**

985

Event Forecast	Event Observed	
	Yes	No
Yes	A	B
No	C	D

986

987

988

989

990

Table 1. Event versus non-event contingency table. A = number of events forecast and observed; B=number of events forecast but not observed; C=number of events observed but not forecast; D = number of cases where events were neither forecast nor observed.



991 **Table 2: Summary performance metrics for ozone, nitrogen dioxide, and PM2.5. Bold-face indicates the simulation with**
 992 **the better performance score for the given metric, chemical species and sub-region, italics indicate a tied score, and regular**
 993 **font the simulation with the lower performance score. FO2: fraction of scores within a factor of 2. MB: Mean Bias. MGE:**
 994 **Mean Gross Error. R: Correlation Coefficient. RMSE: Root Mean Square Error. COE: Coefficient of Error. IOA:**
 995 **Index of Agreement.**
 996

Chemical	Region	Simulation	FO2	MB	MGE	NMGE	R	RMSE	COE	IOA
PM2.5	Western Canada	No	0.451	0.777	4.335	0.878	0.278	7.642	-0.669	0.165
		Feedback	0.453	0.236	4.215	0.829	0.219	6.976	-0.534	0.233
	Western USA	No	0.536	-1.639	3.793	0.605	0.358	6.061	-0.166	0.417
		Feedback	0.524	-1.786	3.773	0.602	0.361	5.978	-0.162	0.419
O₃	Western Canada	No	0.773	-3.683	7.76	0.347	0.634	9.996	0.138	0.569
		Feedback	0.78	-3.553	7.693	0.344	0.635	9.854	0.145	0.573
	Western USA	No	0.879	-3.584	9.667	0.257	0.763	12.534	0.319	0.66
		Feedback	0.881	-3.456	9.607	0.256	0.763	12.458	0.322	0.661
NO₂	Western Canada	No	0.528	-0.417	2.29	0.594	0.519	3.353	0.053	0.527
		Feedback	0.52	-0.533	2.274	0.591	0.513	3.314	0.055	0.527
	Western USA	No	0.428	0.669	2.278	0.759	0.585	3.917	-0.161	0.419
		Feedback	0.427	0.578	2.24	0.746	0.581	3.858	-0.141	0.429

997



Figures:

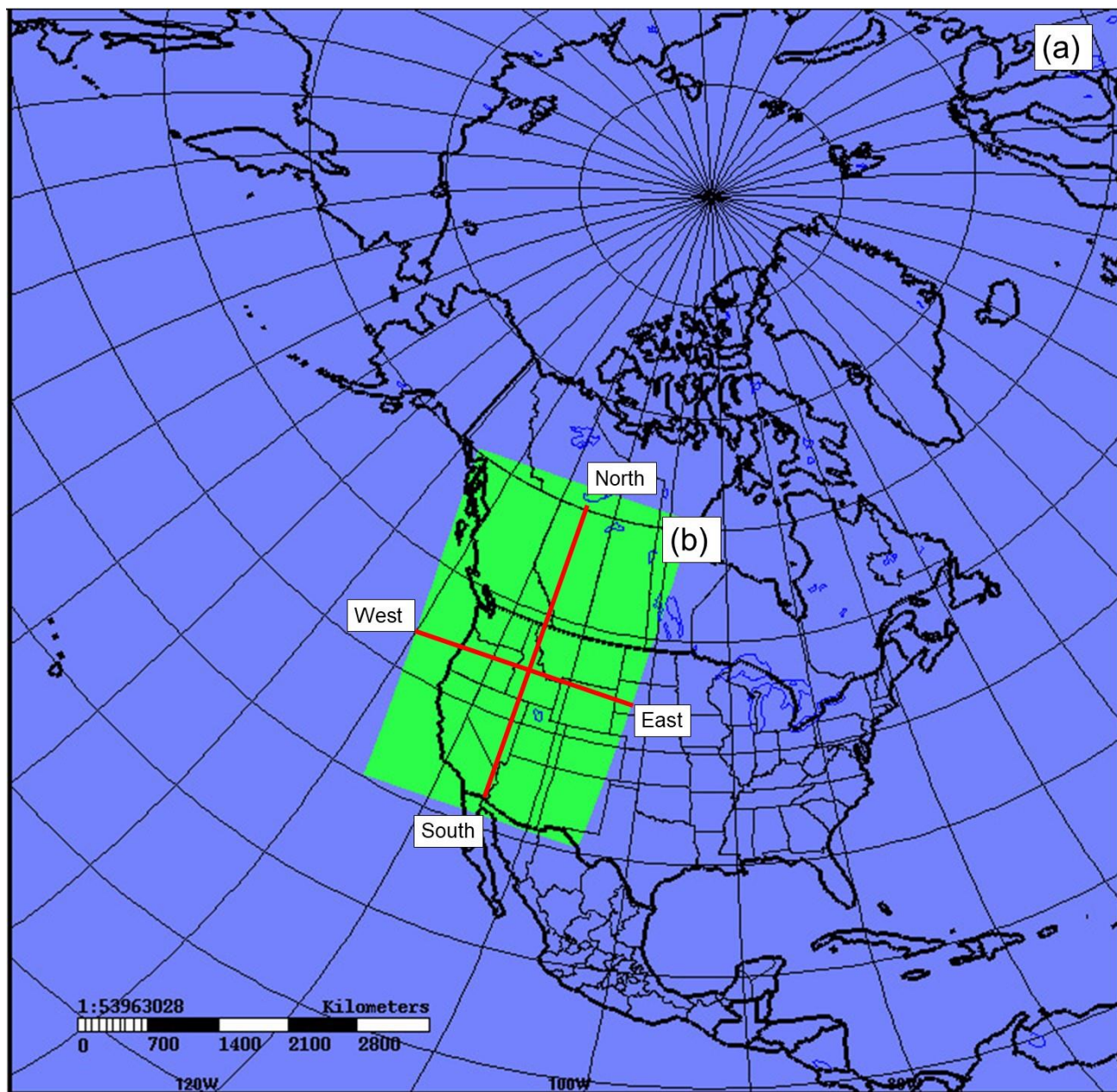


Figure 1: GEM-MACH domains: (a) Operational GEM 10km resolution forecast domain. (b) Experimental 2.5-km grid cell size forecast domain used here. Red lines indicate locations of illustrative South to North and West to East cross-sections appearing in subsequent analysis in the text.

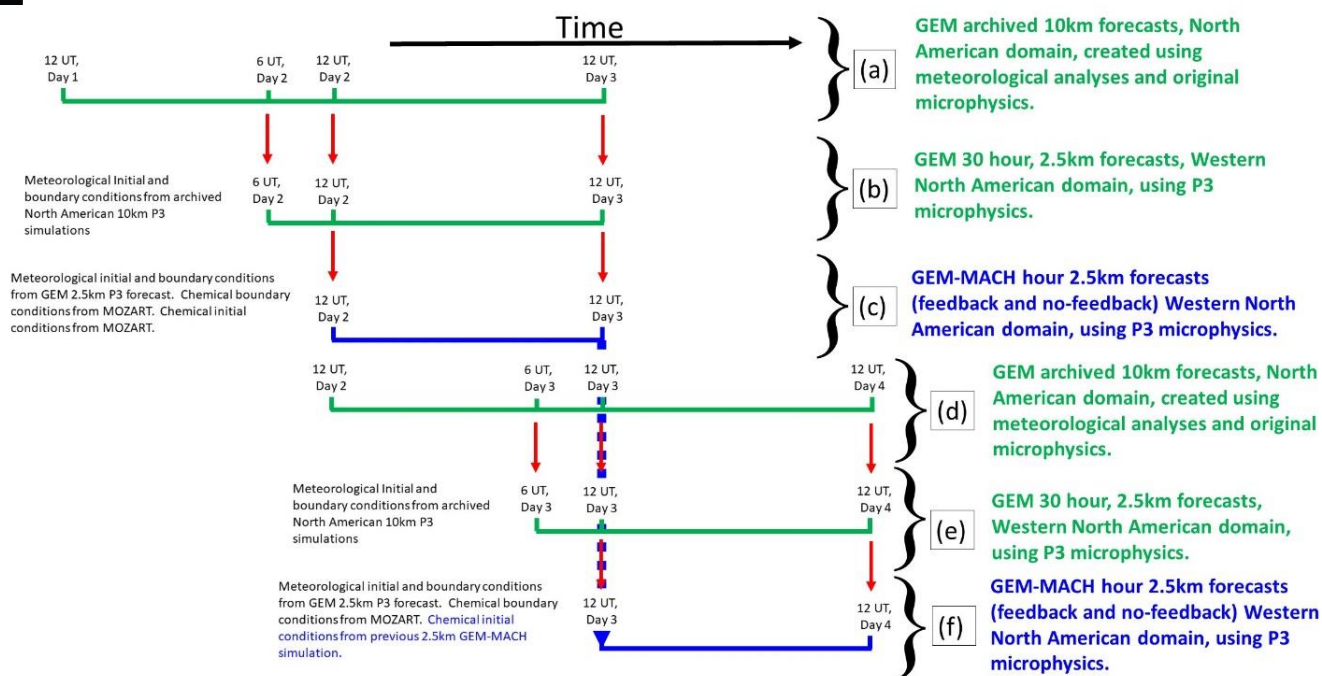


Figure 2: Example time sequencing of model simulations used to generate the 2.5-km GEM-MACH simulations carried out here. Green lines and print indicate GEM (weather forecast only) simulations), blue lines and print indicate 2.5-km GEM-MACH simulations. Steps (a) through (f) illustrate the sequence of forecasts used to generate two consecutive days of 2.5km GEM-MACH simulations.

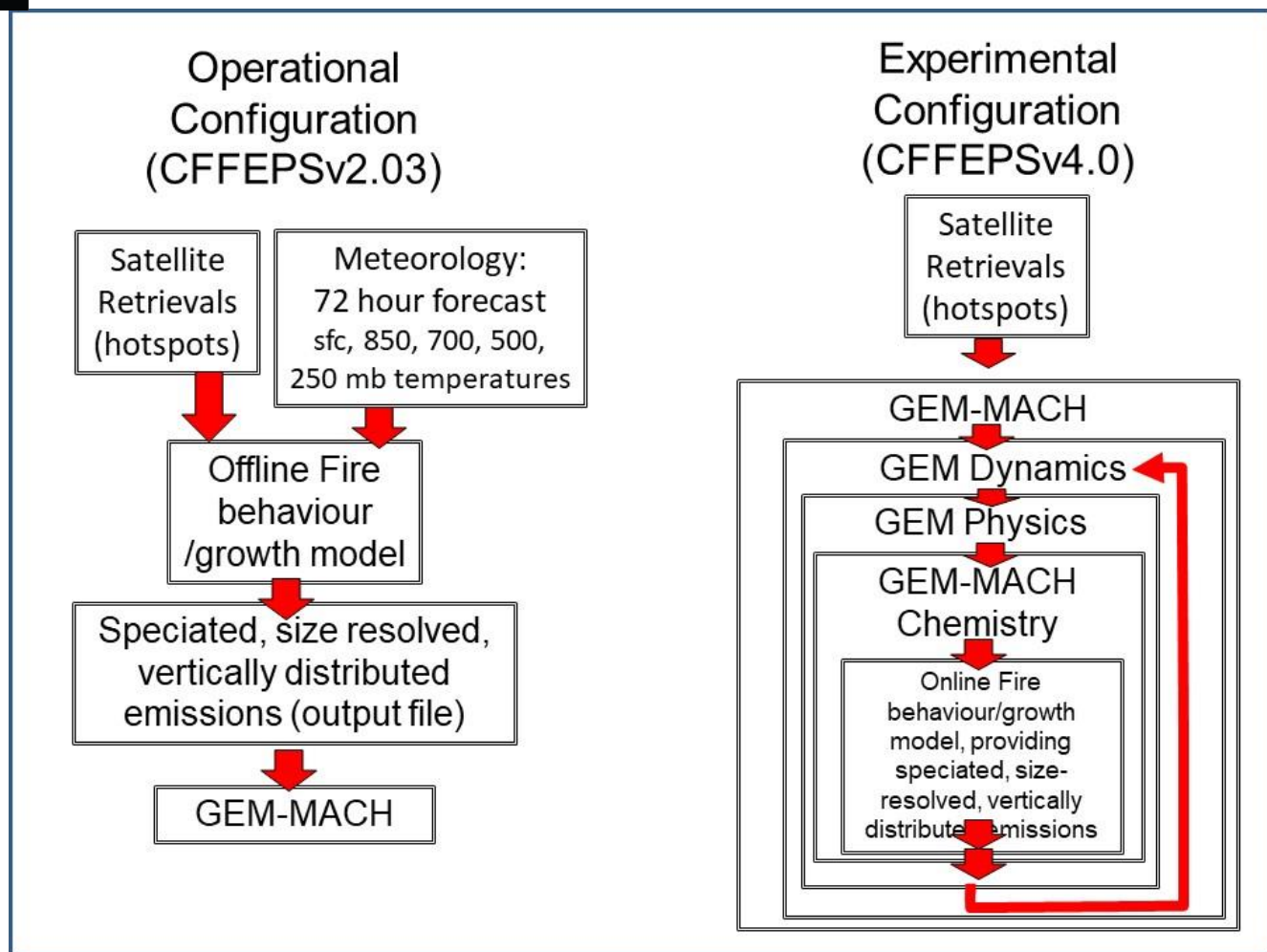


Figure 3: Process comparison between original (CFEFSv2.03, left) and on-line (CFEFSv4.0, right) forest fire emissions and vertical plume distribution algorithms.

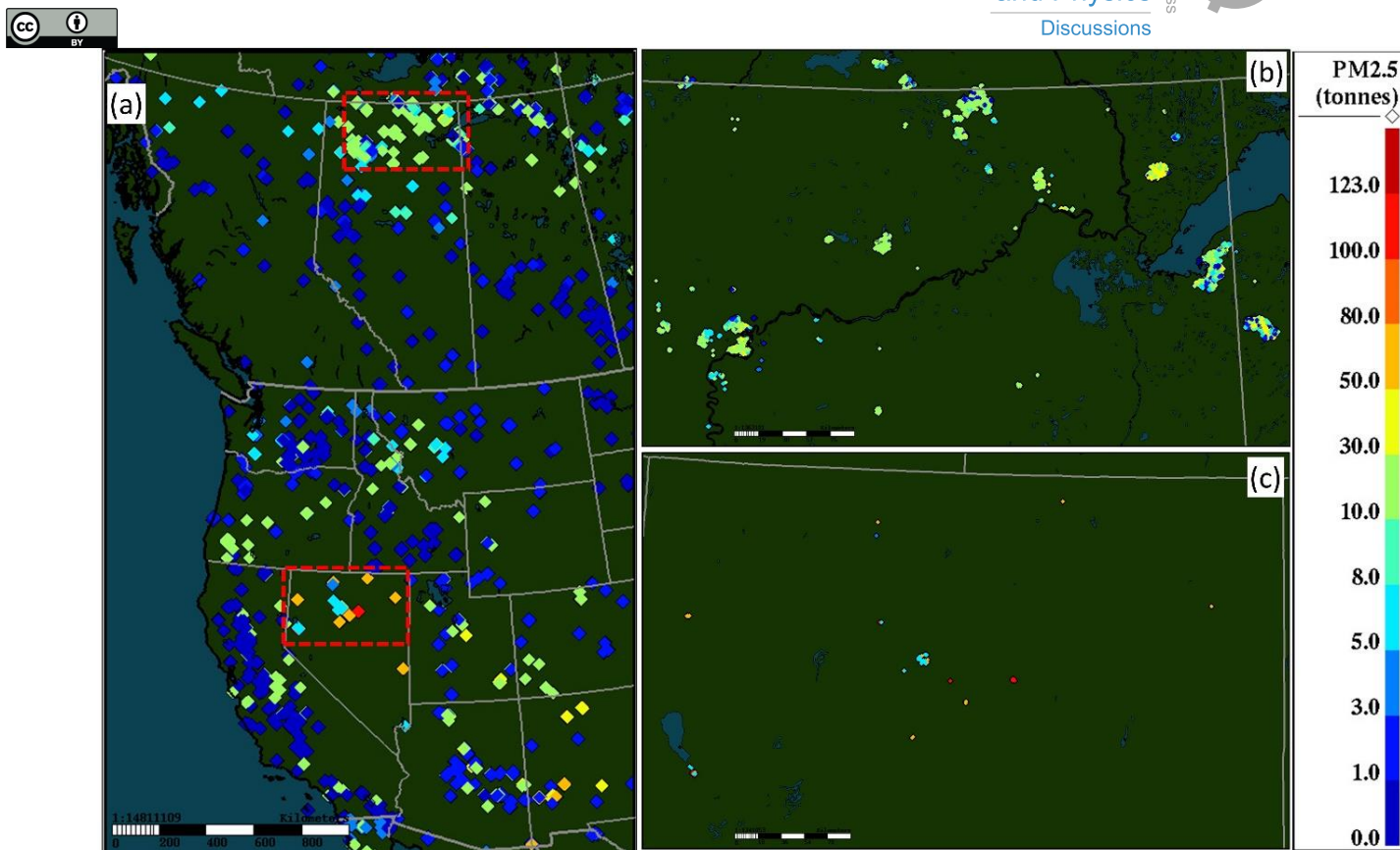


Figure 4: Hotspot locations during the study period, colour-coded by daily total tonnes PM_{2.5} emitted. (a) Entire model 2.5-km domain, with northern Alberta and northern Nevada sub-regions as red dashed boxes; (b) northern Alberta zoom, with smaller symbols for individual hotspots showing the large fire regions; (c) northern Nevada zoom, to the same scale as (b), showing isolated hotspots with high emissions.

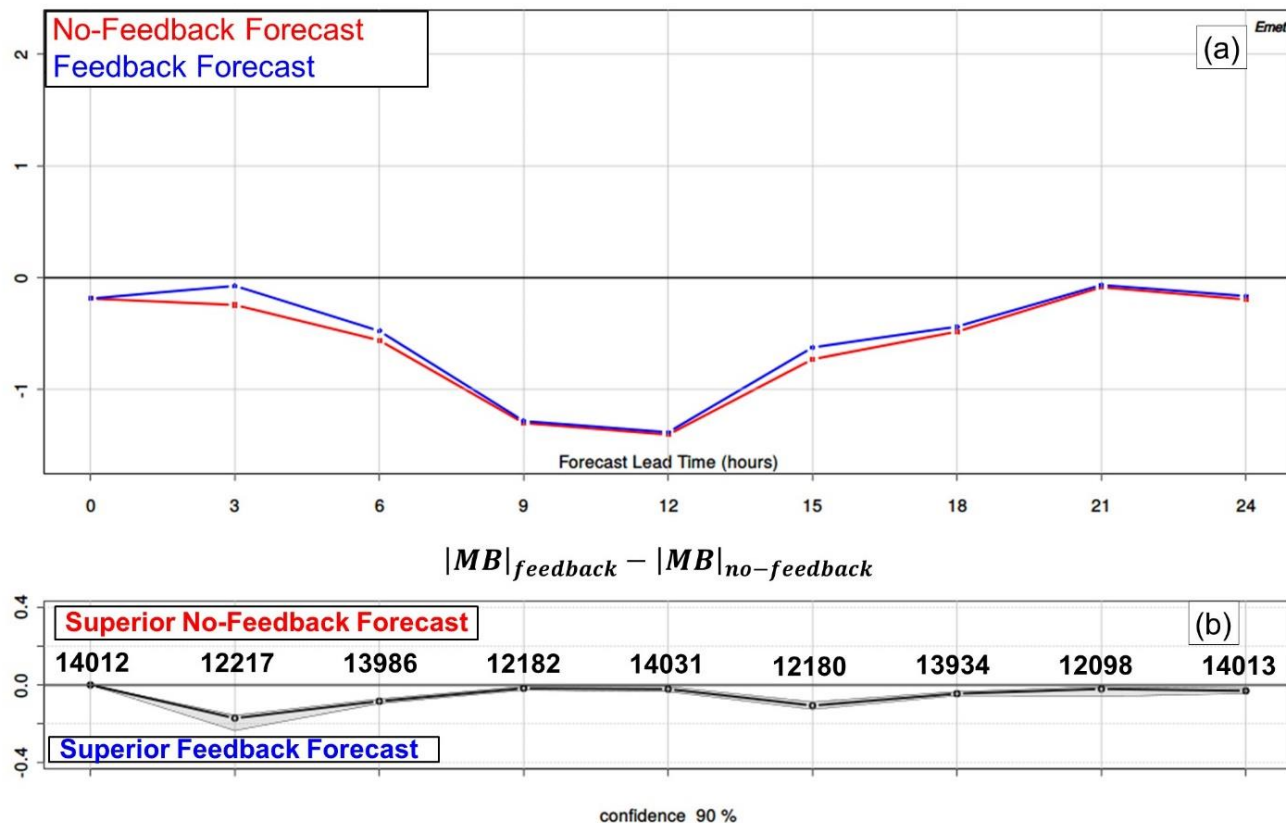


Figure 5: Mean bias in surface temperature (°C) at forecast hours starting at 0 UT. (a) Red line: no-feedback forecast values; blue line: feedback forecast values. (b) Difference in absolute value of mean bias between the two forecasts ($|MB|_{feedback} - |MB|_{no-feedback}$), with the region below 90% confidence level shown shaded grey. Mean values above/below the '0' line, and outside of the shaded region thus indicate differences in the mean between the two forecasts which differ at or above the 90% confidence level. Values of the difference which appear below/above the zero line and outside of the grey area thus indicate superior domain average performance for the feedback/no-feedback forecasts at each of the 3-hourly intervals, respectively. Numbers appearing above the metric differences are the number of observations contributing to the calculated metrics.

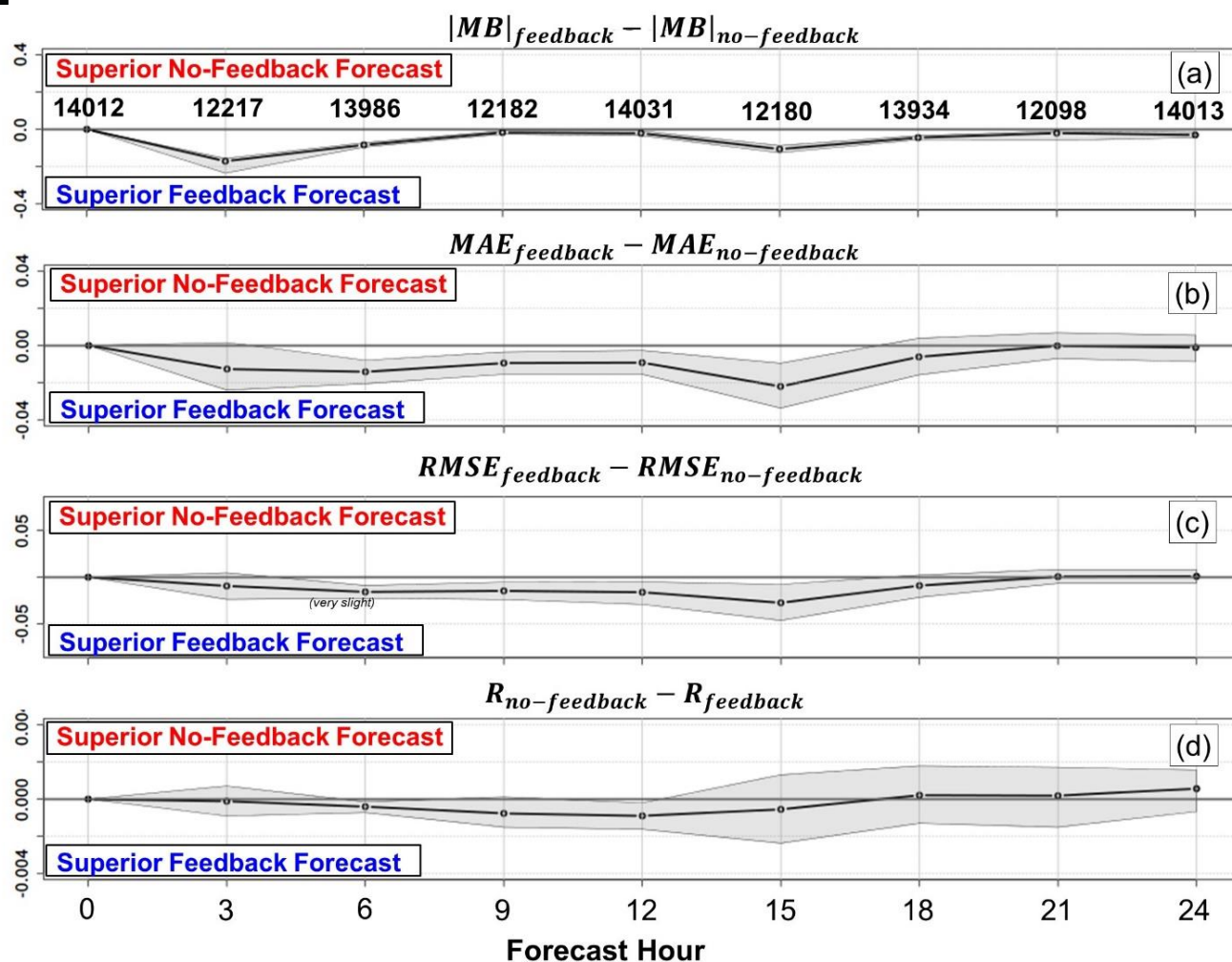


Figure 6: Summary meteorological performance comparison for surface temperature (C). (a) mean bias, (b) mean absolute error, (c) root mean square error and (d) Pearson correlation coefficient. 90% confidence level shown in grey. Numbers appearing above the absolute mean bias differences are the number of stations contributing to the calculated metrics.

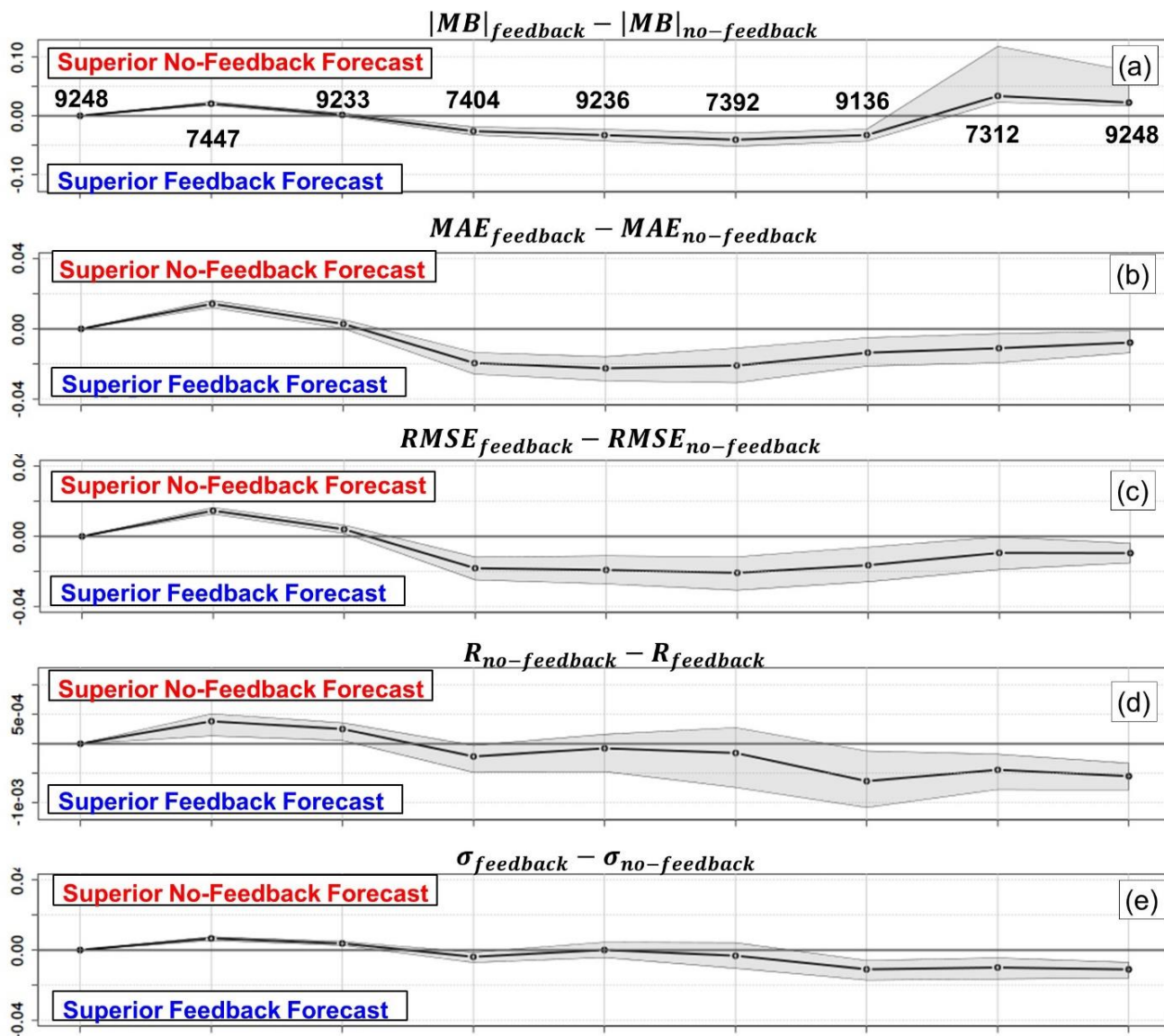


Figure 7: Summary meteorological performance comparison for surface pressure (hPa). (a) mean bias, (b) mean absolute error, (c) root mean square error, (d) Pearson correlation coefficient, and (e) standard deviation. 90% confidence level shown in grey. Numbers appearing above the absolute mean bias differences are the number of stations contributing to the calculated metrics.

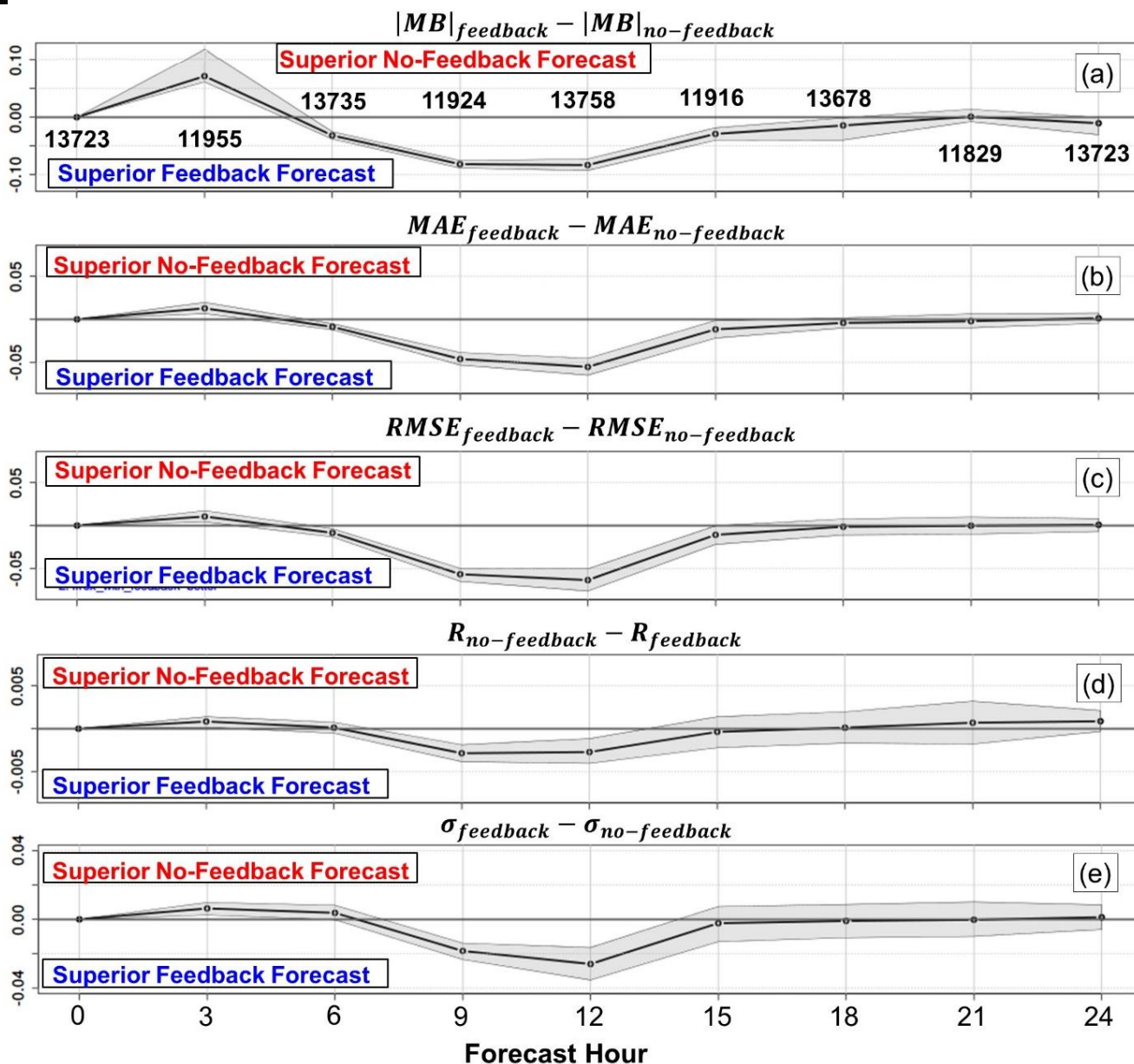


Figure 8: Summary meteorological performance comparison for dewpoint temperature (C). (a) mean bias, (b) mean absolute error, (c) root mean square error, (d) Pearson correlation coefficient, and (e) standard deviation. 90% confidence level shown in grey. Numbers appearing above the absolute mean bias differences are the number of stations contributing to the calculated metrics.

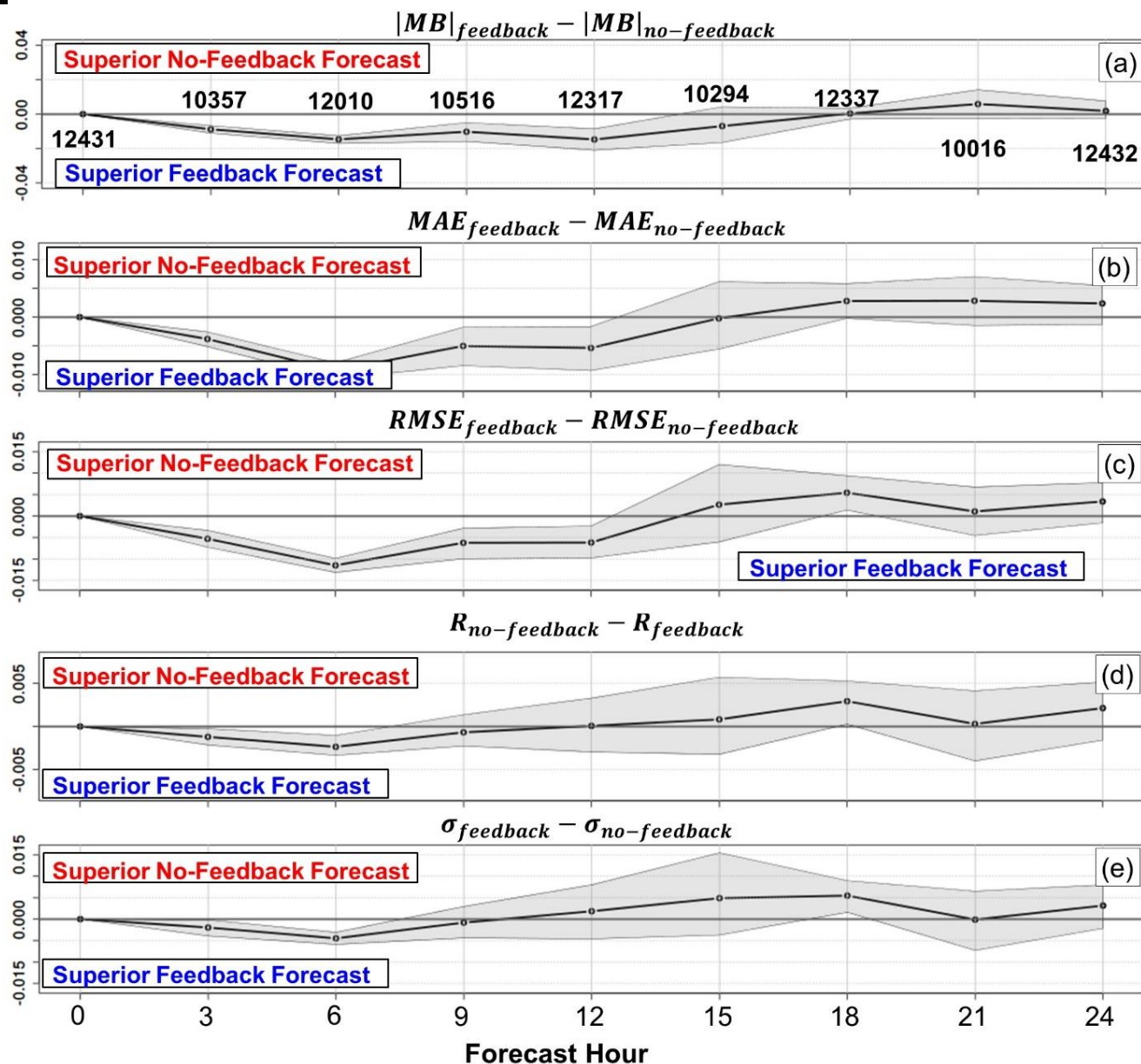


Figure 9: Summary meteorological performance comparison for 10m windspeed (m s⁻¹). (a) mean bias, (b) mean absolute error, (c) root mean square error, (d) Pearson correlation coefficient, and (e) standard deviation. 90% confidence level shown in grey. Numbers appearing above the absolute mean bias differences are the number of stations contributing to the calculated metrics.

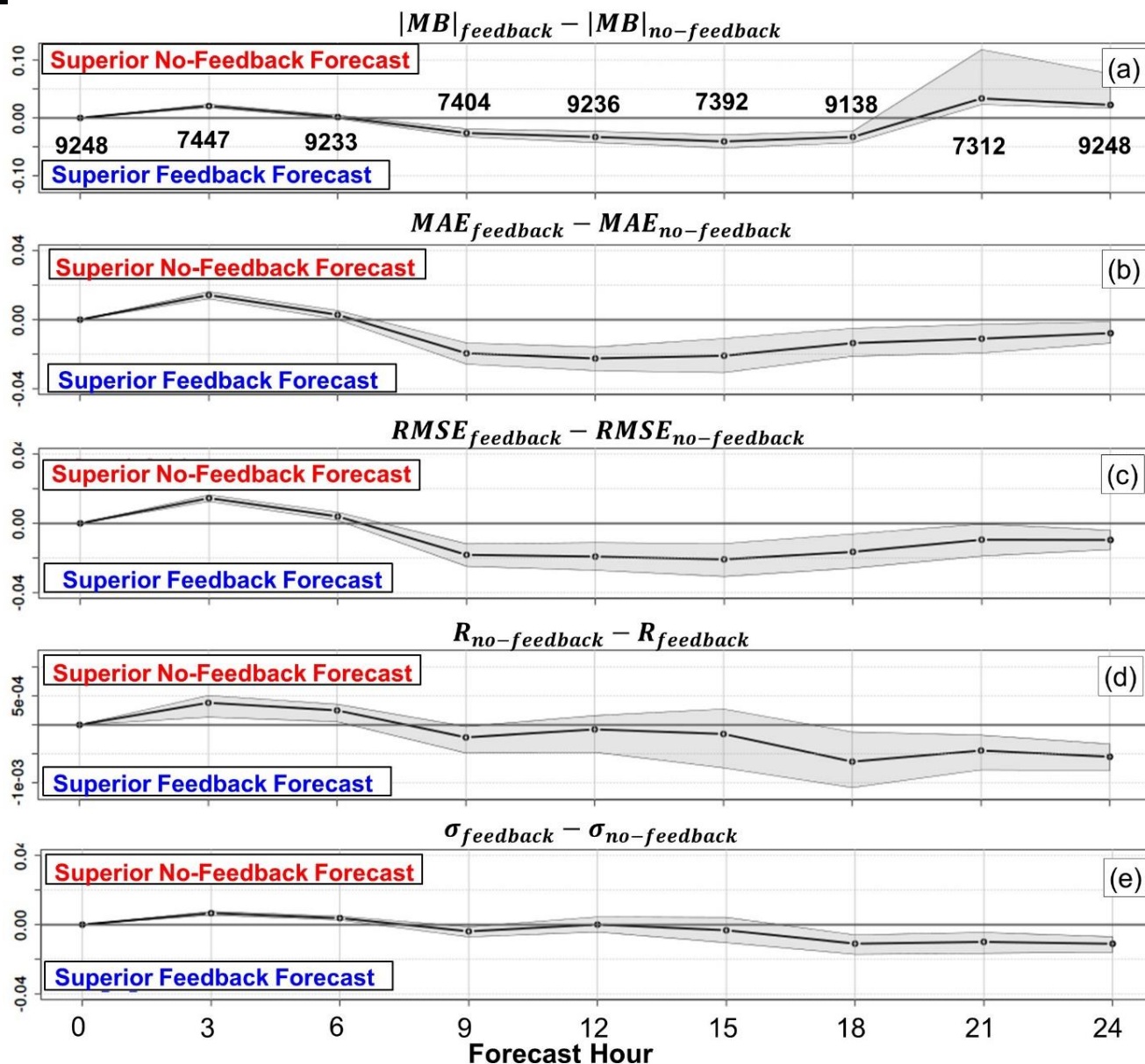


Figure 10: Summary meteorological performance comparison for sea-level pressure (hPa). (a) mean bias, (b) mean absolute error, (c) root mean square error, (d) Pearson correlation coefficient, and (e) standard deviation. 90% confidence level shown in grey. Numbers appearing above the absolute mean bias differences are the number of stations contributing to the calculated metrics.

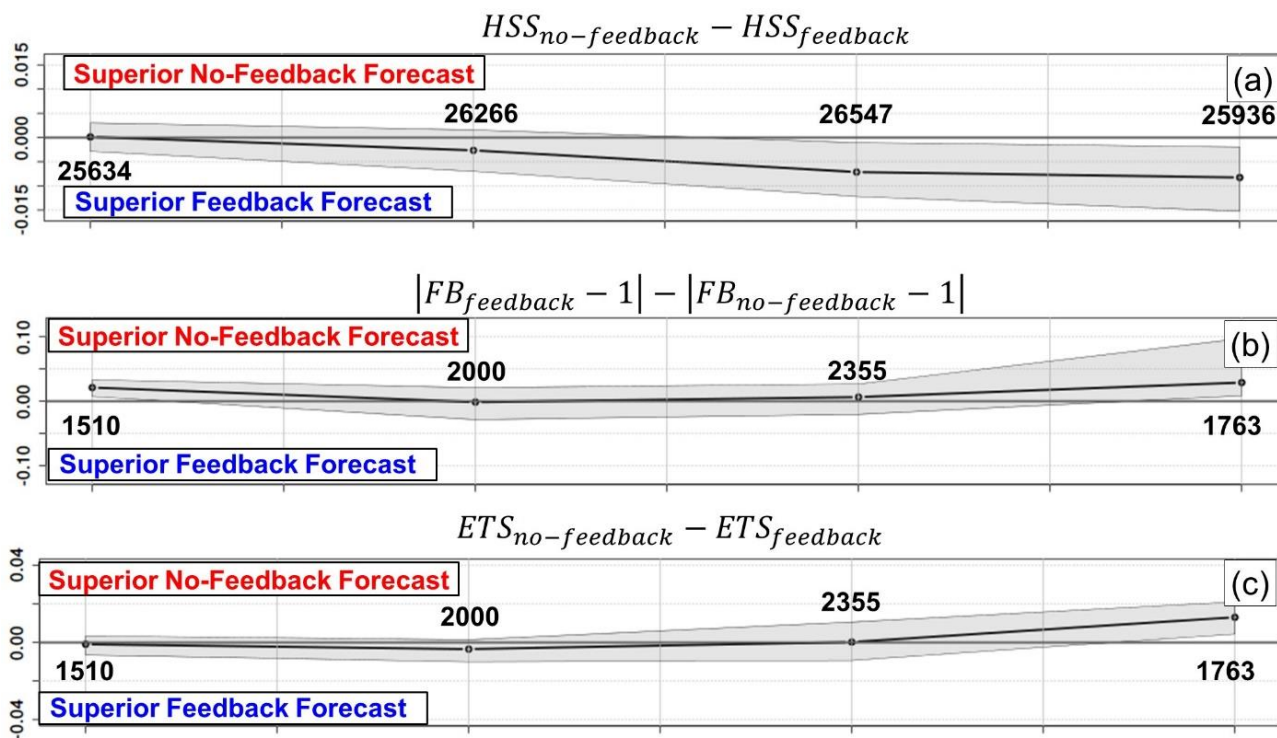


Figure 11: Precipitation performance evaluation (mm precipitation). (a) Heike skill score of 6-hour accumulated precipitation (No-Feedback – Feedback). (b) Frequency bias index of 6-hour accumulated precipitation (threshold of 2 mm, No-Feedback – Feedback). (c) Equitable Threat Score of 6-hour accumulated precipitation (threshold of 2 mm, No-Feedback – Feedback).

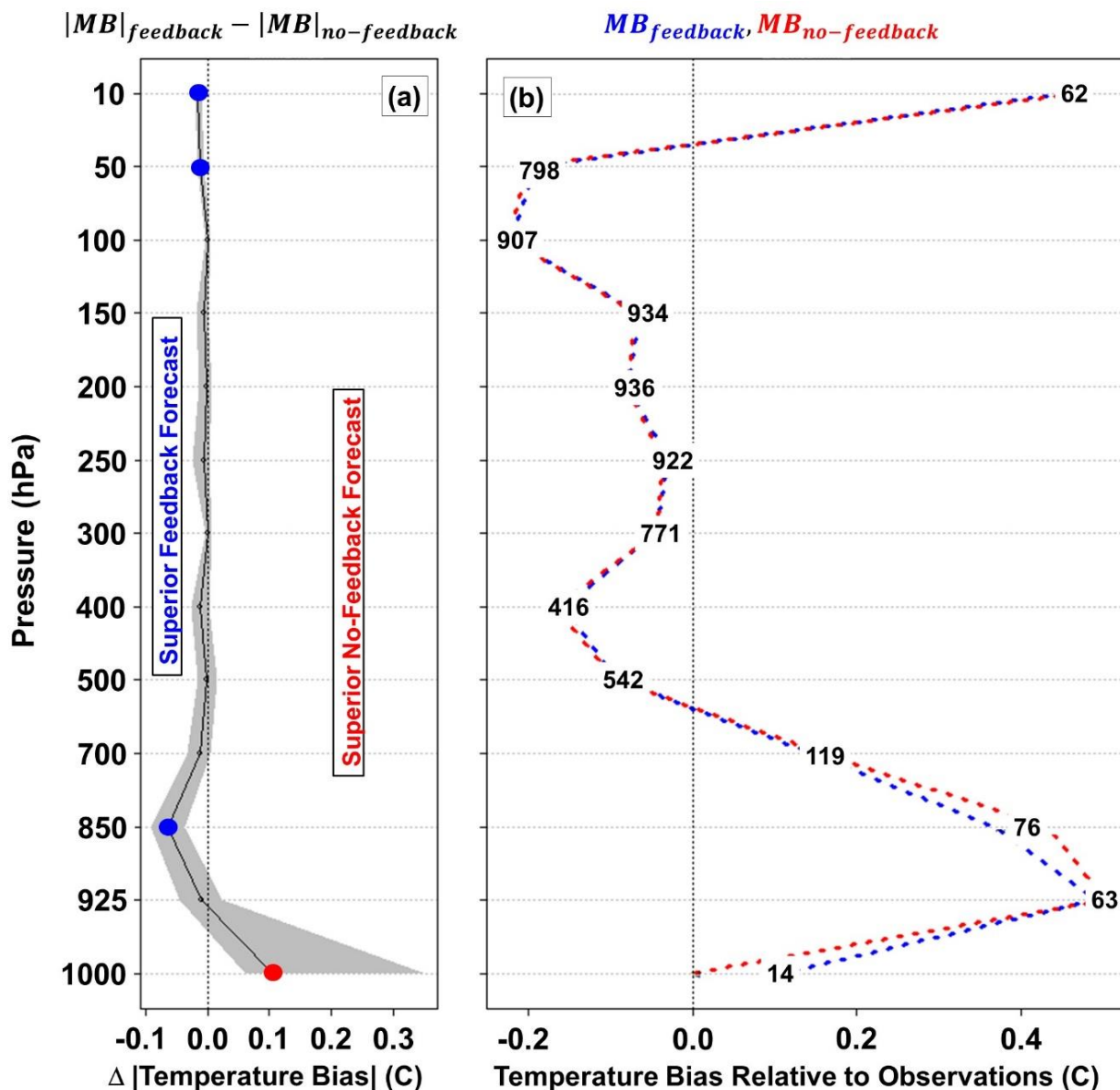


Figure 12: Forecast hour 12 (0 UT) summary upper air temperature performance comparison for air temperature (mean bias, C). (a) Difference in absolute value of mean bias in temperature, (feedback forecast – no-feedback forecast). Grey regions represent 90% confidence levels, blue symbols: pressure levels at which the feedback mean bias outperforms the no-feedback mean-bias at > 90% confidence. Red symbols: pressure levels at which the no-feedback mean bias outperforms the feedback mean bias at > 90% confidence. 90% confidence level shown in grey. (b) Mean bias in upper air temperature for feedback (blue) and no-feedback (red) (C). Numbered values on the profiles indicate the number of observed data-model pairs at each pressure level.

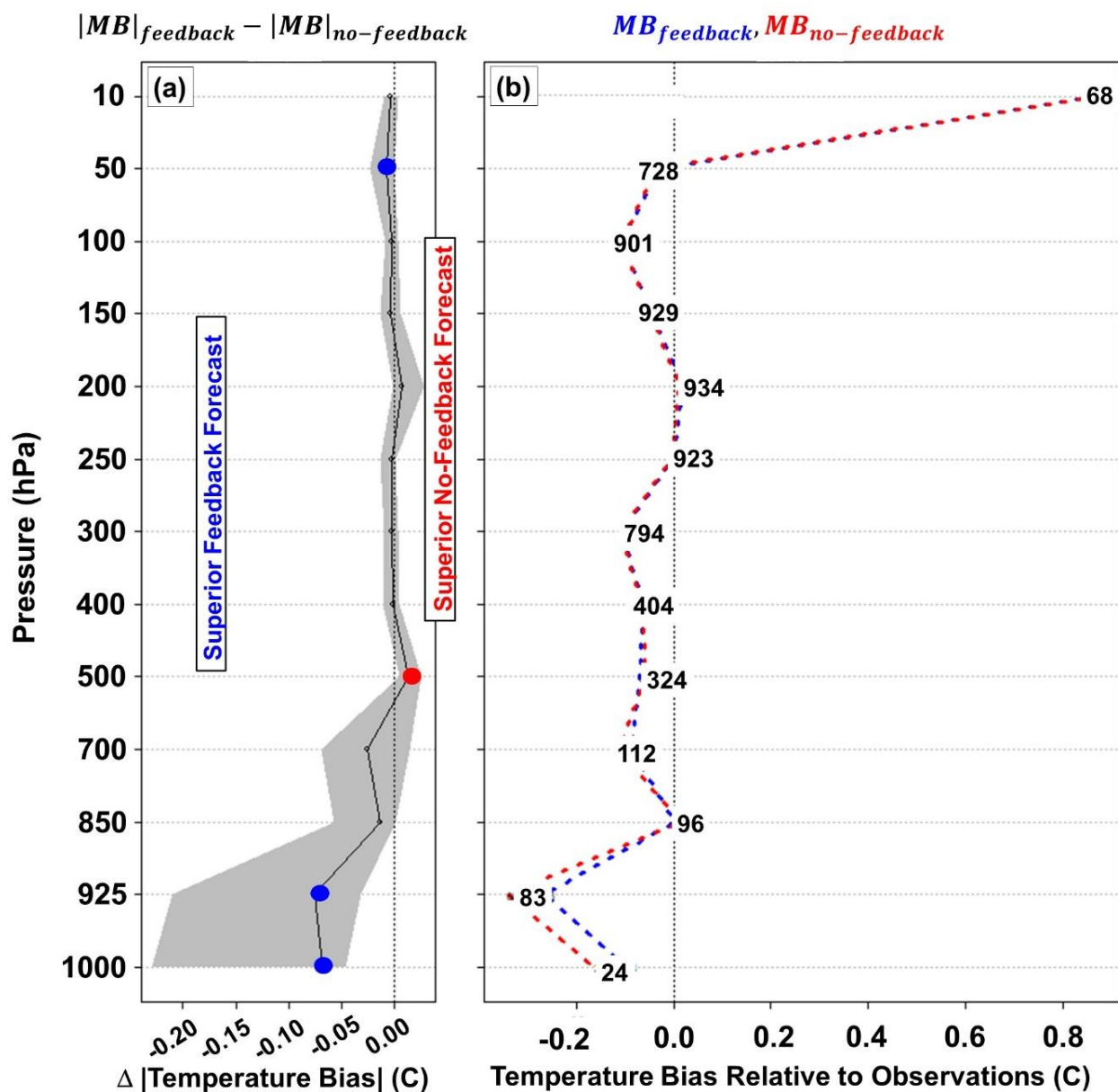


Figure 13: Forecast hour 24 (12 UT) summary upper air temperature performance comparison for air temperature (mean bias, C). (a,b) as in Figure 12.

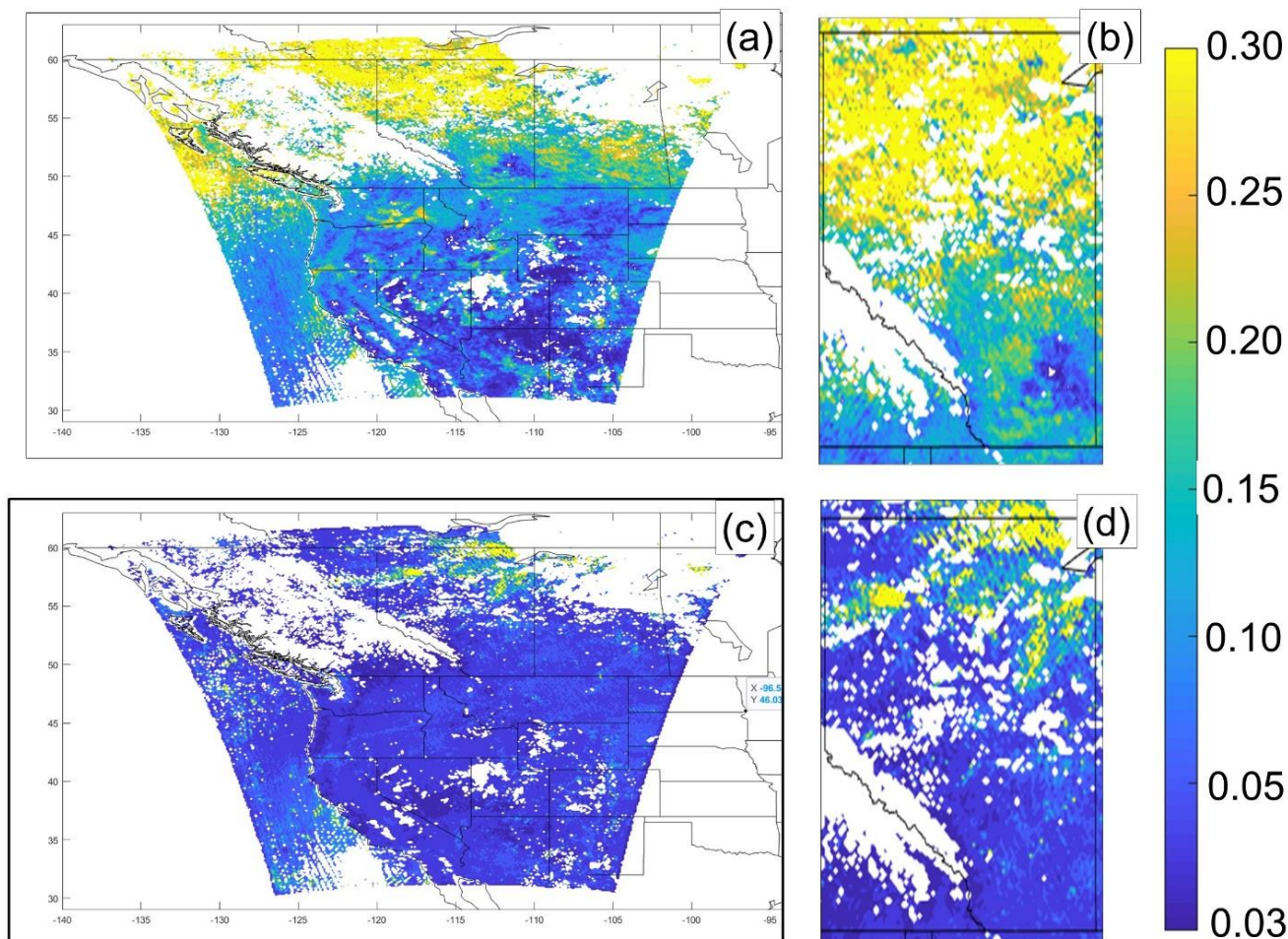


Figure 14: 550nm AOD comparison. (a) MODIS observations spatiotemporally averaged and sampled over the model domain and forecast duration and (b) zoom-in on the province of Alberta. (c,d) GEM-MACH nearest output hour AOD values corresponding to satellite sensor data availability for the same regions.

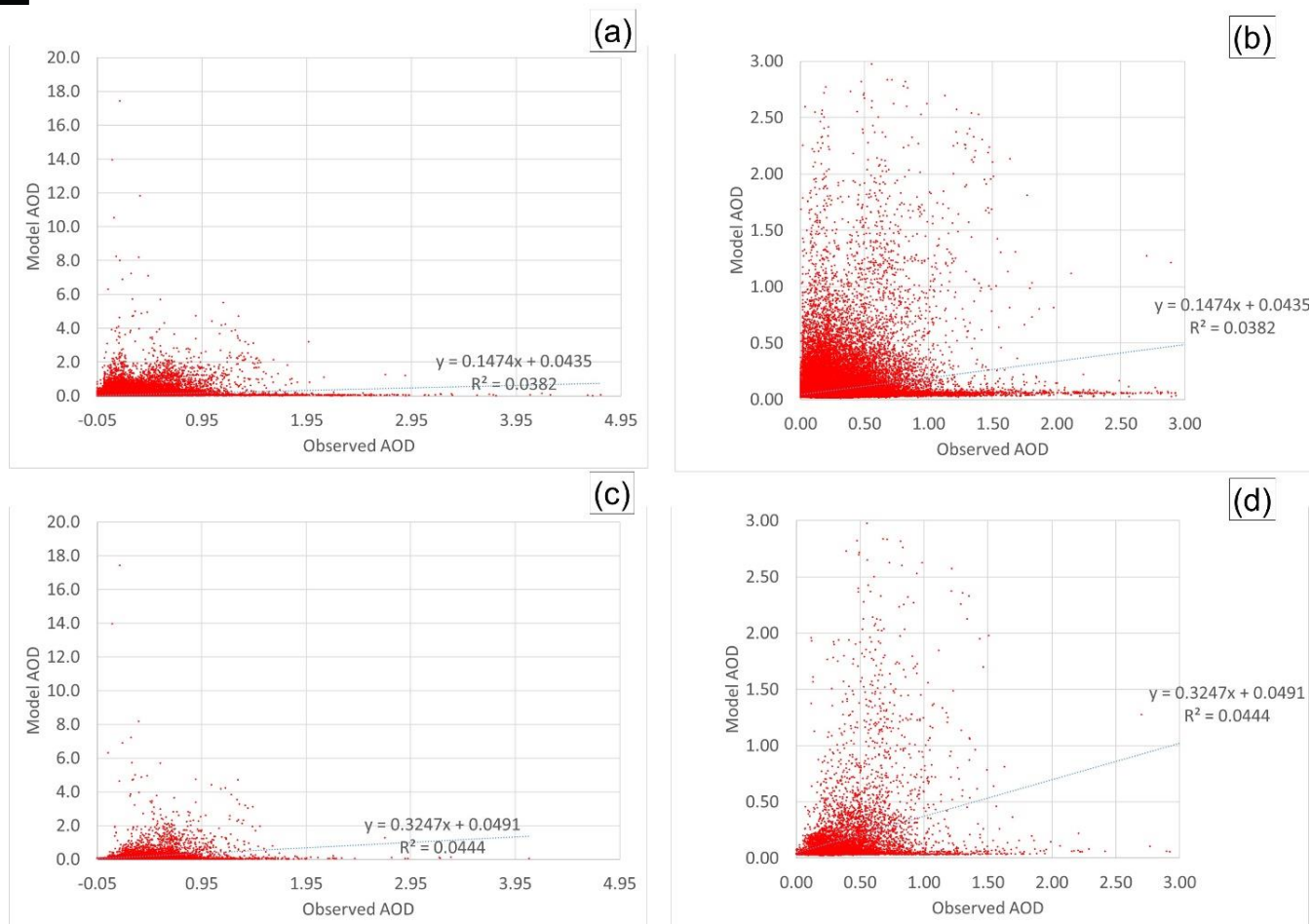


Figure 15: Scatterplots comparing model and observed AOD values. (a) All model-observation pairs. (b) As in (a), rescaled to focus on AODs in the 0-3 range. (c) Model-observation pairs in the region bounded by 56N, 60N, -110W and -120 W (northern Alberta). (d) As in (c), rescaled to focus on AODs in the 0-3 range.

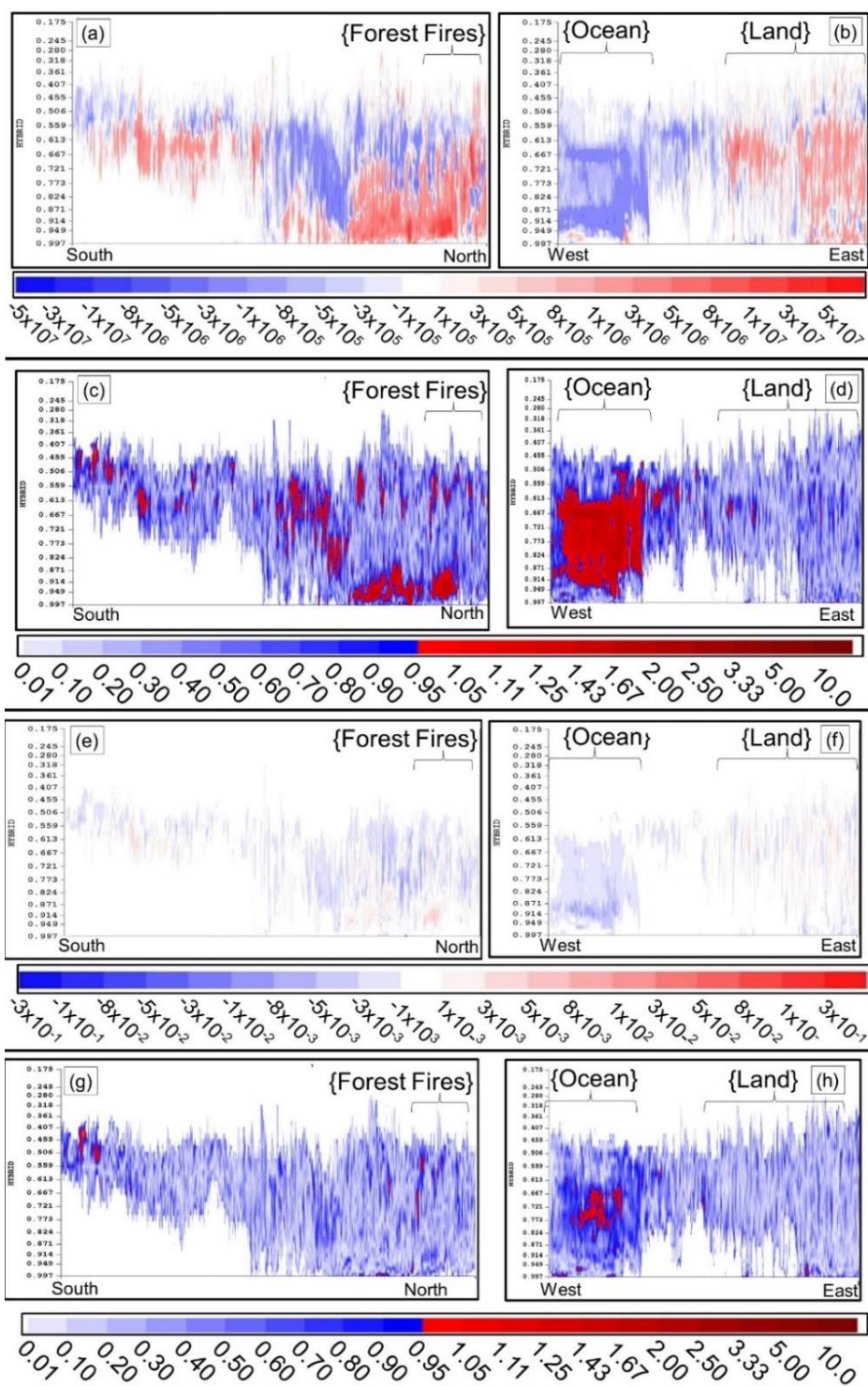


Figure 16. (a,b) Difference in mean (Feedback – No-Feedback) cloud droplet number simulations along south to north and east to west cross-sections through the middle of the model domain. (c,d) Corresponding significance level of mean cloud droplet number



differences using the confidence ratio defined in equation (1) – red areas indicate ratio values greater than unity, i.e., significance at or above the 90% confidence level. (e,f) Difference in mean cloud droplet mass (g kg^{-1}) (g,h) Corresponding significance level of mean cloud droplet mass difference. *Note: the vertical axis in hybrid coordinates does not show all model levels for clarity; the model has much finer resolution in the lower part of the atmosphere than shown, and the portion of the vertical domain shown encompasses only the lower half of the levels in the model.*

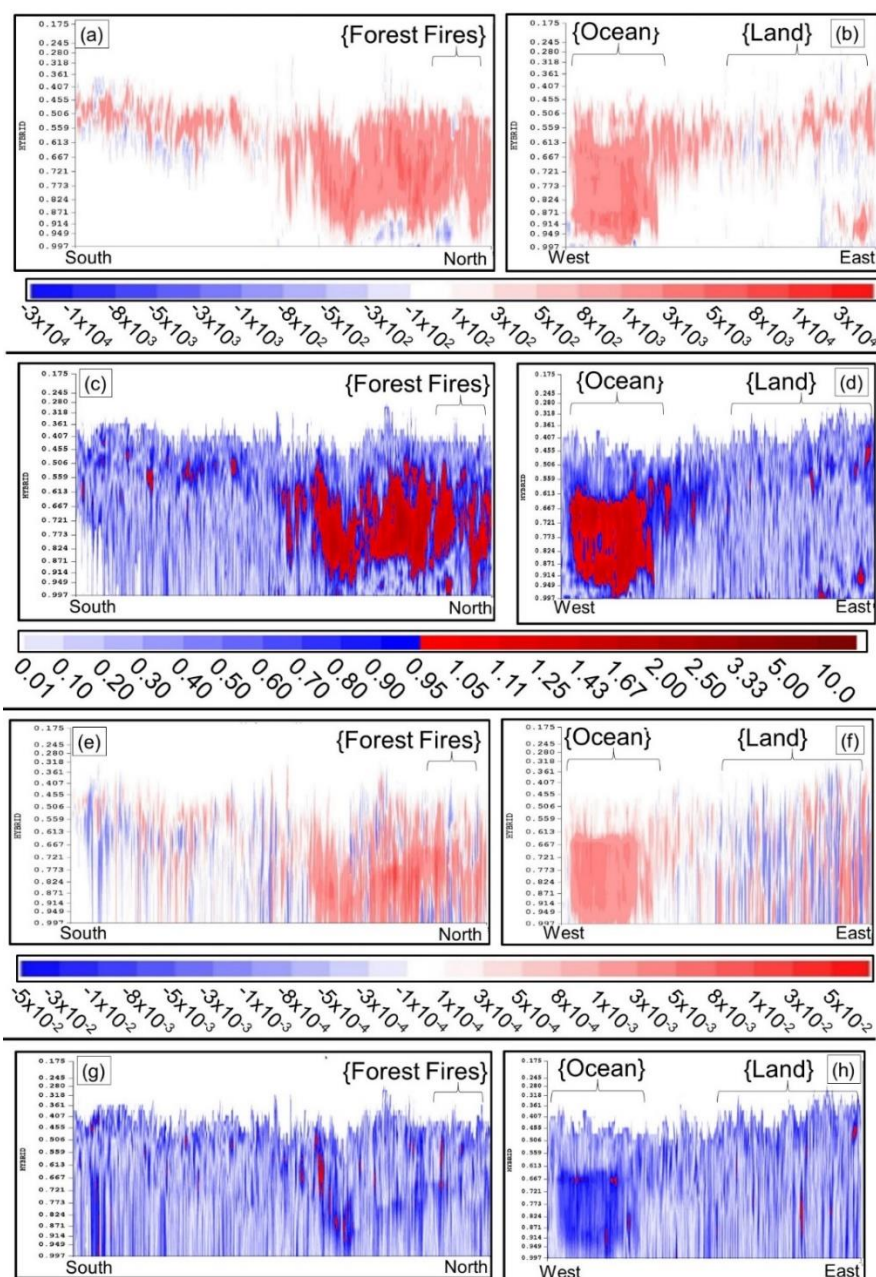


Figure 17. (a,b) Difference in mean (Feedback – No-Feedback) rain drop number simulations along south-to-north and east-to-west cross-sections through the middle of the model domain. (c,d) Corresponding significance level of mean rain drop number differences



using the confidence ratio defined in equation (1) – red areas indicate ratio values greater than unity, i.e., significance at or above the 90% confidence level. (e,f) Difference in rain cloud drop mass (g kg^{-1}) (g,h) Corresponding significance level of mean rain drop mass difference.

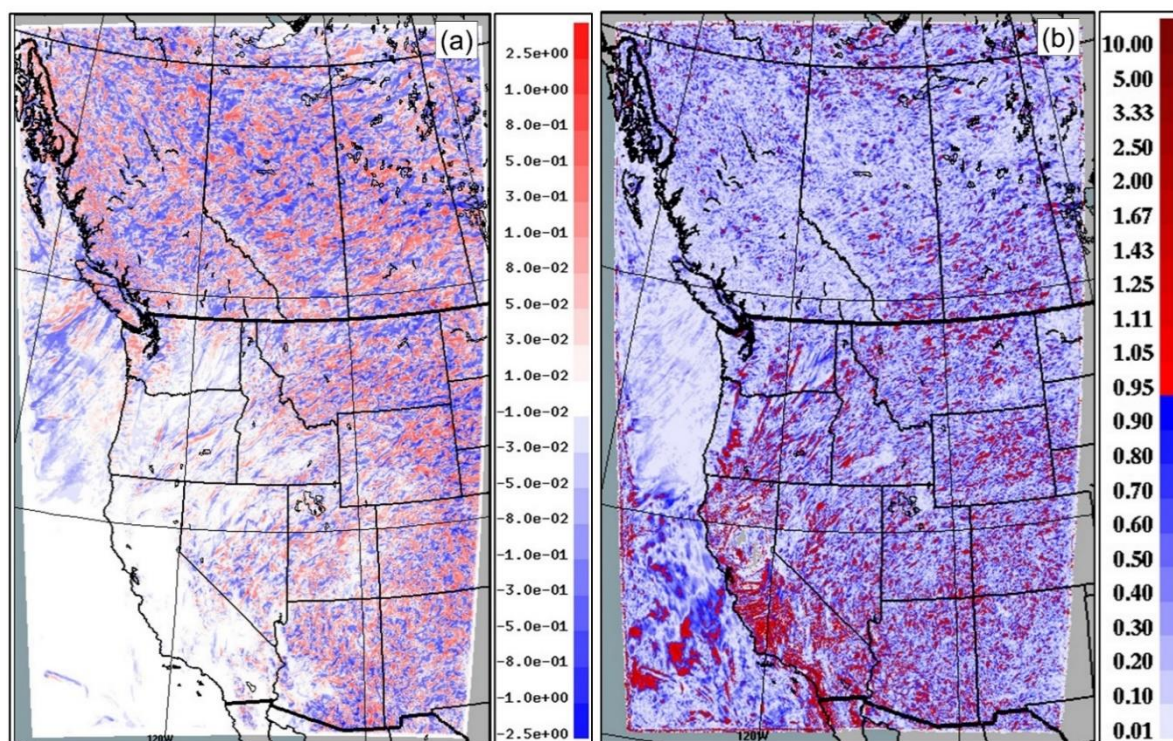


Figure 18: (a) Average (Feedback – No Feedback) total surface precipitation during the simulation period. (b) 90% confidence ratio – values greater than 1 indicate significantly different results at the 90% confidence level.

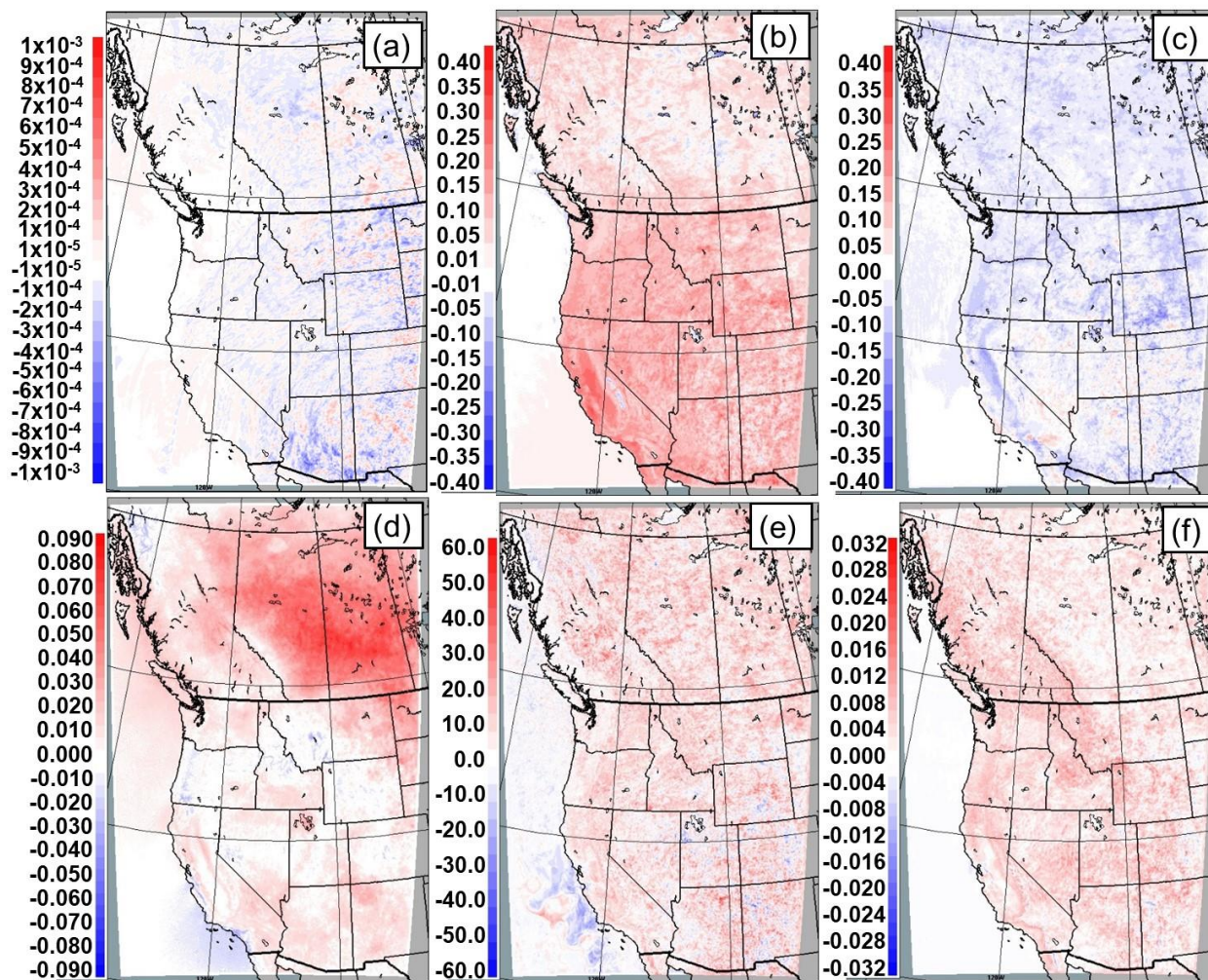


Figure 19: Differences in average meteorological fields (feedback – no-feedback; red values indicate more positive values in the feedback simulation than in the no-feedback simulation). Panels show average difference in: (a) specific humidity (g kg^{-1}); (b) air temperature (C), (c) dewpoint temperature (C), (d) surface pressure (mb), (e) planetary boundary layer height (m), (f) friction velocity (m s^{-1}).

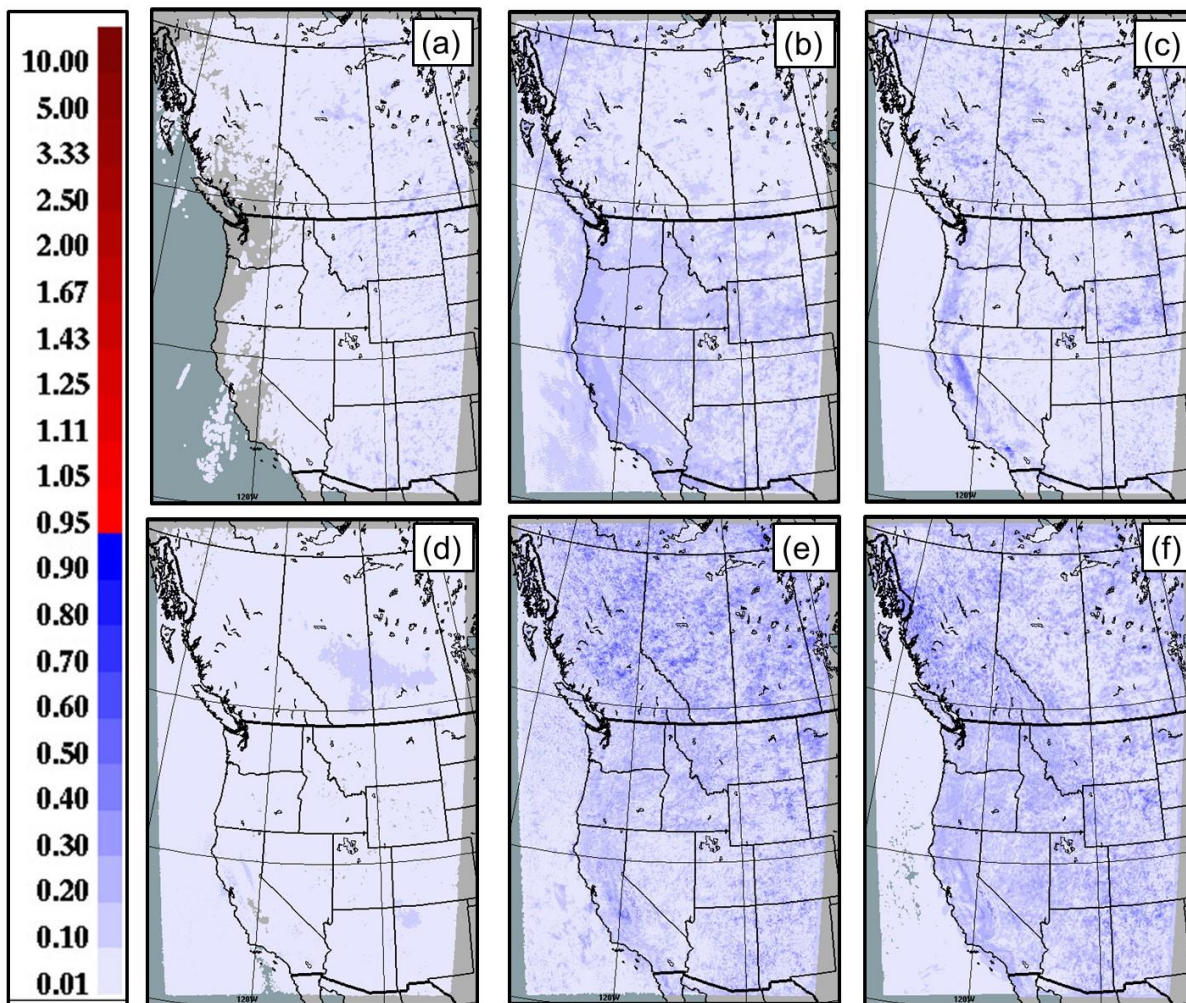


Figure 20: 90% confidence ratios, same fields as Figure 19. Values greater than 1 indicate significantly different results at or greater than the 90% confidence level.

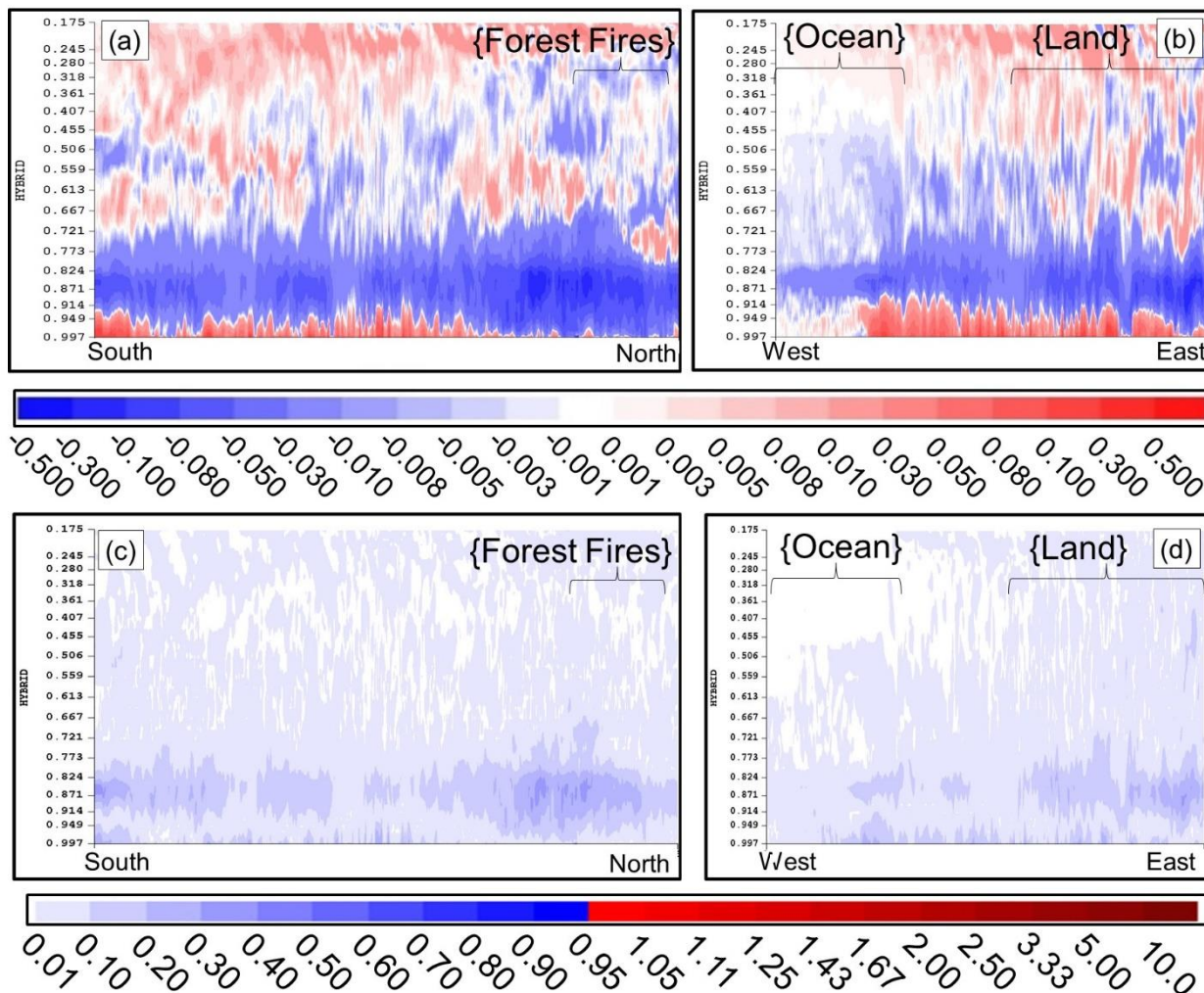


Figure 21: (a,b) Difference in mean (Feedback – No-Feedback) temperature simulations along south-to-north and east-to-west cross-sections through the middle of the model domain. (c,d) Corresponding confidence ratio of mean temperature differences– red areas indicate ratio values greater than unity, i.e., significance at or above the 90% confidence level. Note the reduction in temperatures between hybrid levels 0.90 to 0.70, similar to the findings of Saponaro *et al.*, (2017).

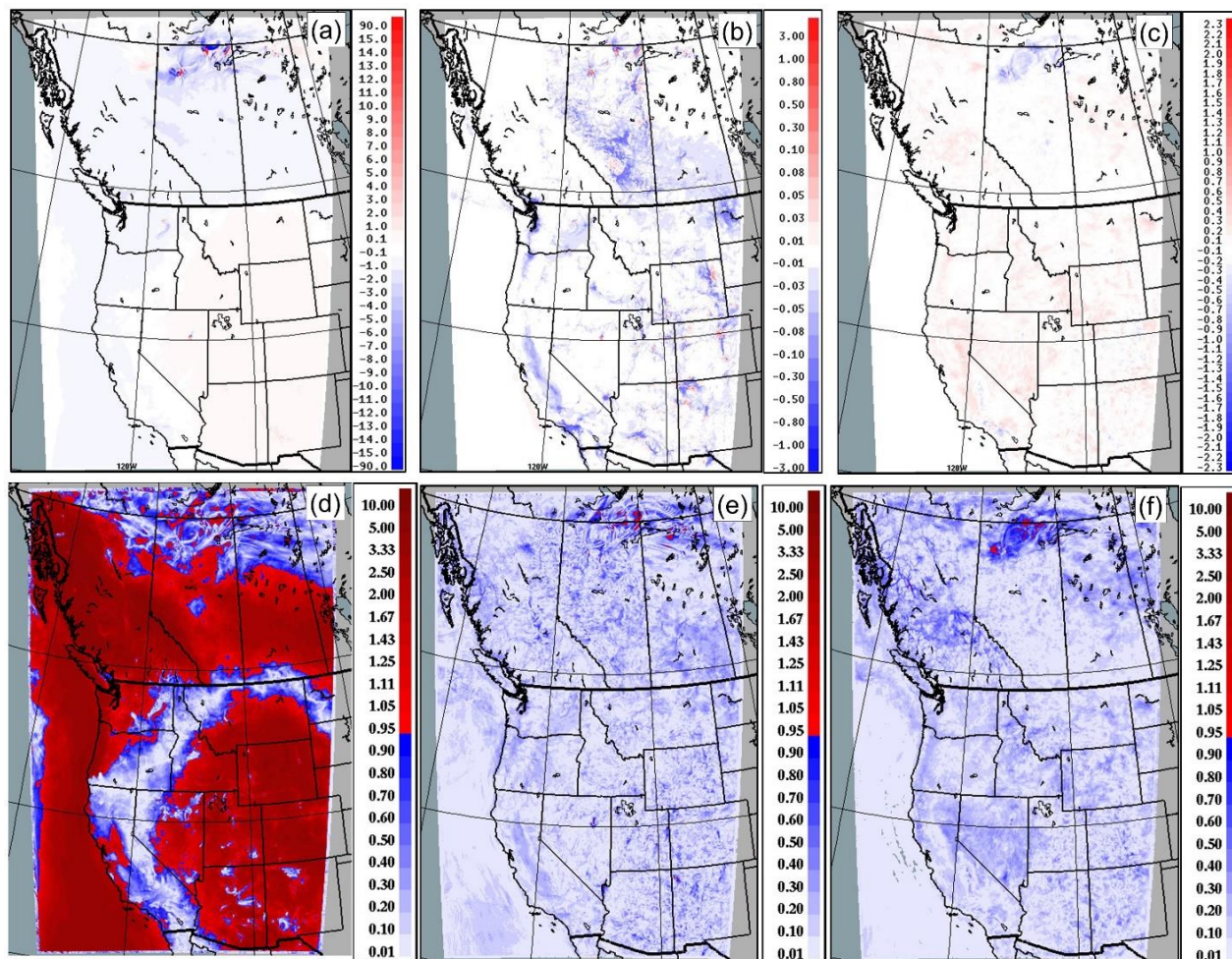


Figure 22: (a,b,c) Difference (Feedback – No-Feedback) in surface mean PM_{2.5} ($\mu\text{g m}^{-3}$), NO₂ (ppbv) and O₃ (ppbv), respectively. (d,e,f) Corresponding confidence ratio of mean differences – red areas indicate ratio values greater than unity, i.e., significance at or above the 90% confidence level.

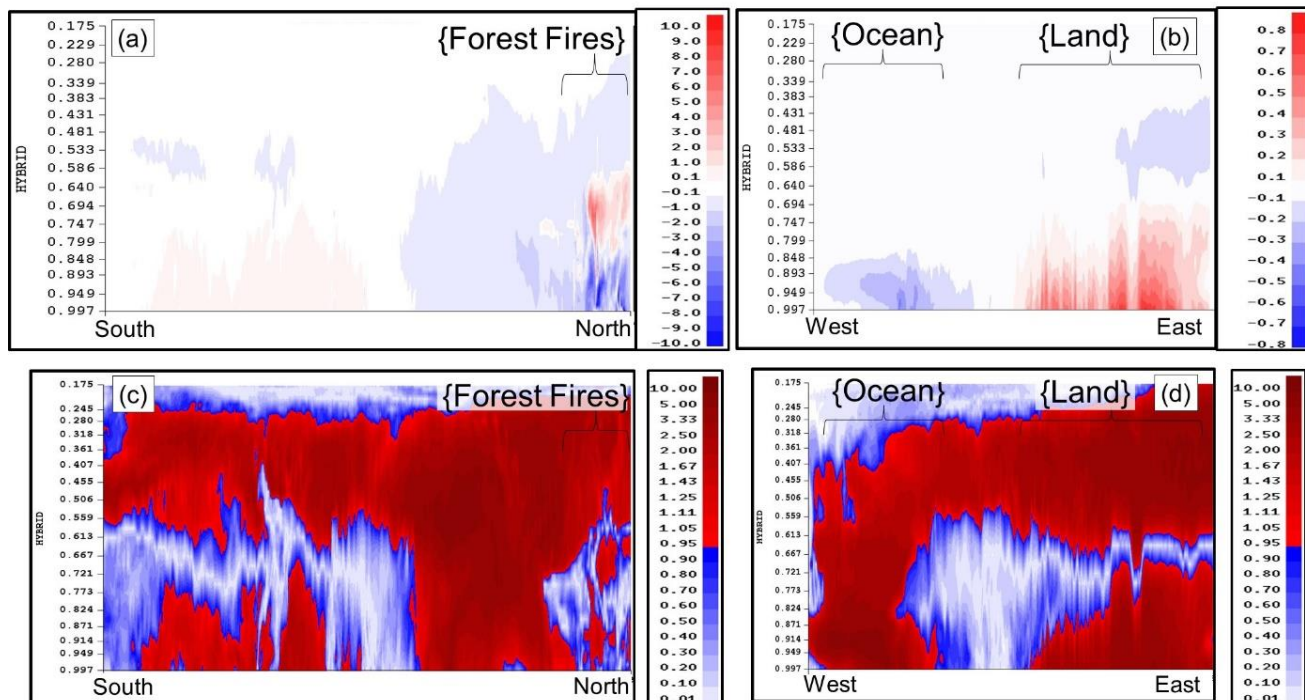


Figure 23: (a,b) Difference (Feedback – No-Feedback) in predicted mean PM_{2.5} (ug m⁻³), along domain-center South-North and West – East cross-sections. (c,d) Corresponding confidence ratio of mean differences – red areas indicate ratio values greater than unity, i.e., significance at or above the 90% confidence level. Note that colour bar scales differ between (a) and (b).

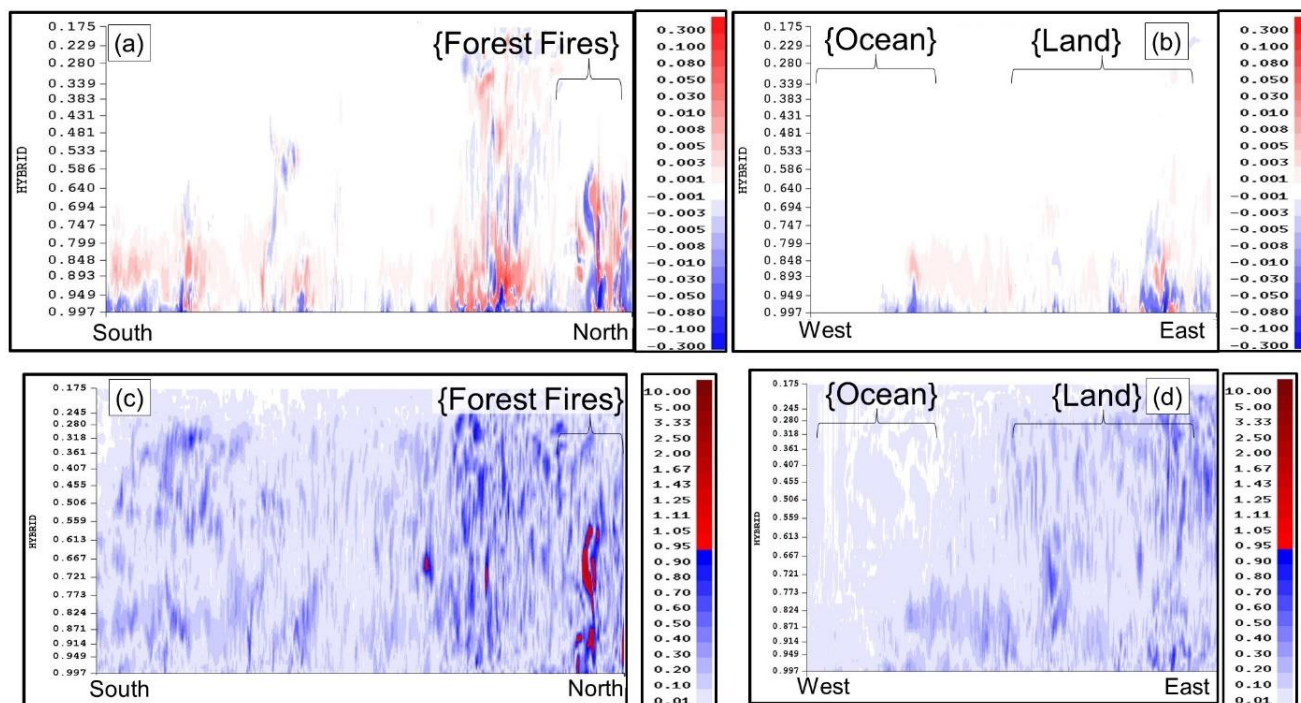


Figure 24: (a,b) Difference (Feedback – No-Feedback) in predicted mean NO₂ (ppbv), along domain-center South-North and West – East cross-sections. (c,d) Corresponding confidence ratio of mean differences – red areas indicate ratio values greater than unity, i.e., significance at or above the 90% confidence level.

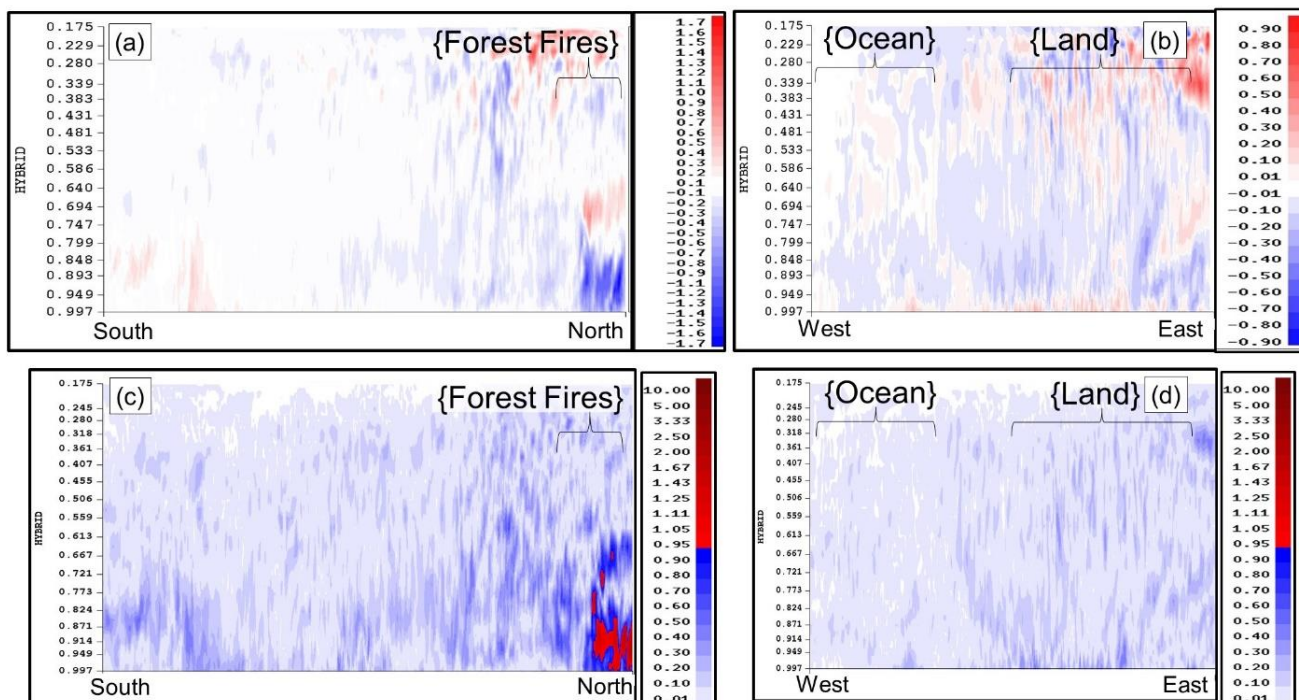


Figure 25: (a,b) Difference (Feedback – No-Feedback) in predicted mean O₃ (ppbv), along domain-center South-North and West – East cross-sections. (c,d) Corresponding confidence ratio of mean differences – red areas indicate ratio values greater than unity, i.e., significance at or above the 90% confidence level. Note that colour bar scales differ between (a) and (b).