Supplementary Material for:

The FireWork v2.0 air quality forecast system with biomass burning emissions from the Canadian Forest Fire Emissions Prediction System v2.03

Jack Chen[1], Kerry Anderson[2+], Radenko Pavlovic[3], Michael D. Moran[1], Peter Englefield[2], Dan K. Thompson[2], Rodrigo Munoz-Alpizar[3], and Hugo Landry[3]

[1] Air Quality Research Division, Environment and Climate Change Canada, Ontario, Canada

[2] Canadian Forest Service, Natural Resources Canada, Alberta, Canada

[3] Air Quality Modelling Applications Section, Environment and Climate Change Canada, Quebec, Canada

[+]Emeritus

Correspondence to: Jack Chen (jack.chen@canada.ca)

S1. Reference material

Table S1. List of acronyms.

AOD	Aerosol Optical Depth
APEI	Air Pollutant Emissions Inventory (Canada)
AQ	Air Quality
AQHI	Air Quality Health Index
AQS	Air Quality System measurement network (U.S.)
AVHRR	Advanced Very High Resolution Radiometer
CAM	Canadian Aerosol Module
CFB	Crown Fraction Burned
CFC	Crown Fuel Consumption
CFFDRS	Canadian Forest Fire Danger Rating System
CFFEPS	Canadian Forest Fire Emission Prediction System
CFS	Canadian Forest Service
CSI	Critical Success Index
CTM	Chemical Transport Model
CWFIS	Canadian Wildland Fire Information System
ECAN	Eastern Canada Region
ECCC	Environment Climate Change Canada
EPA	Environmental Protection Agency (U.S.)
EUSA	Eastern USA Region
FAR	False Alarm Rate
FBP	Canadian Forest Fire Behavior Prediction System
FEPS	Fire Emission Production Simulator (USFS)
FFMC	Fine Fuel Moisture Code

FRP	Fire Radiative Power
FWI	Canadian Forest Fire Weather Index System
GEM	Global Environmental Multiscale numerical weather prediction model
GEM-MACH	Global Environmental Multi-scale Modelling Air quality and Chemistry model
GOES	Geostationary Operational Environmental Satellite
GTAC	Geospatial Technology and Applications Center
HFI	Head Fire Intensity
LST	Local Standard Time
MB	Mean Bias
MODIS	Moderate Resolution Imaging Spectroradiometer
NAPS	National Air Pollution Surveillance network (Canada)
NASA	National Aeronautics and Space Administration (U.S.)
NEI	National Emissions Inventory (U.S.)
NMHC	Non-Methane Hydrocarbon
NOAA	National Oceanic and Atmospheric Administration (U.S.)
NRT	Near-Real-Time
NWP	Numerical Weather Prediction
NWS	National Weather Service (U.S.)
PBL	Planetary Boundary Layer
PM	Particulate Matter
POD	Probability of Detection
R	Pearson Correlation Coefficient
RAQDPS	Regional Air Quality Deterministic Prediction System (Canada)
RMSE	Root-mean-square Error
ROS	Rate of Spread
SFC	Surface Fuel Consumption
TFC	Total Fuel Consumption
USFS	United States Forest Service
UTC	Coordinated Universal Time
VCD	Vertical Column Density
VIIRS	Visible Infrared Imaging Radiometer Suite
VOC	Volatile organic compound
WCAN	Western Canada Region
WUSA	Western USA Region

Table S2. Details of NAPS measurement stations in the Northern-Canada Region (see Figure 1 for a plot of station locations).

Station ID	Name and Province/Territory	Latitude, Longitude
000129003	Yellowknife, NT	62.452084, -114.364031
000129601	Fort Smith, NT	60.004486, -111.893377
000129203	Inuvik, NT	68.35702, -133.7141
000080402	63 - 12th Street East, SK	53.201694, -105.7520
000080211	Saskatoon, SK	52.13613, -106.66293
000080110	Regina, SK	50.45017, -104.61722
000070203	Assiniboine College, MB	49.84225, -99.919
000070118	Winnipeg Pump Station, MB	49.93229, -97.11327
000070119	Ellen Street, MB	49.89795, -97.14665
000060808	Thunder Bay, ON	48.3794, -89.2902

S2. Additional model evaluation material for the 2017 fire season

Table S3. Model performance statistics for daily maximum surface O₃ volume mixing ratio (ppbv) for the Aug. 1 to Sept. 18, 2017 period for measurement stations in the AB+BC, ID+MT+OR+WA, and Northern-Canada regions.

O_3	AB+BC (64 stations)			ID+MT+OR+WA (25 stations)			Northern-Canada (10 stations)		
Model	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS
count	3131			1220			389		
$\bar{O}$	41			57			33		
$\bar{M}$	48	71	51	60	90	64	37	41	37
MB	7	30	9	4	33	8	4	9	5
R	0.55	0.37	0.55	0.70	0.36	0.67	0.73	0.61	0.71
RMSE	19	68	22	18	78	21	9	15	9

Table S4. O₃ categorical scores based on a threshold of 65 ppbv for the Aug. 1 to Sept. 18, 2017 period for measurement stations in the AB+BC, ID+MT+OR+WA, and Northern-Canada regions.

O_3	AB+BC			ID+MT+OR+WA			Northern-Canada		
Model	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS
POD	54%	84%	67%	53%	83%	65%	Inf.	Inf.	Inf.
FAR	85%	94%	88%	52%	74%	58%	100%	100%	100%
CSI	13%	6%	11%	34%	25%	34%	0%	0%	0%

Table S5. Model performance statistics for daily maximum surface NO₂ volume mixing ratio (ppbv) for the Aug. 1 to Sept. 18, 2017 period for measurement stations in the AB+BC, ID+MT+OR+WA, and Northern-Canada regions. Note that there is only one measurement station (in Oregon) for the ID+MT+OR+WA region.

NO_2	AB+BC (71 stations)			ID+MT+OR+WA (1 station in OR)			Northern-Canada (10 stations)		
Model	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS
count	3477			46			342		
$\bar{O}$	14			17			10		
$\bar{M}$	20	22	20	41	47	42	15	15	15
MB	7	8	7	24	29	24	5	5	5
R	0.59	0.58	0.59	0.53	0.57	0.54	0.44	0.44	0.45
RMSE	15	18	15	30	37	31	11	11	11

Table S6. NO₂ categorical scores based on a threshold of 30 ppbv for the Aug. 1 to Sept. 18, 2017 period for measurement stations in the AB+BC, ID+MT+OR+WA, and Northern-Canada regions. Note that there is only one measurement station (in Oregon) for the ID+MT+OR+WA region.

NO_2	AB+BC			ID+MT+OR+WA (1 station in OR)			Northern-Canada		
Model	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS
POD	47%	46%	47%	89%	89%	89%	7%	7%	12%
FAR	91%	92%	91%	69%	73%	70%	98%	98%	97%
CSI	8%	7%	8%	30%	26%	29%	1%	1%	3%

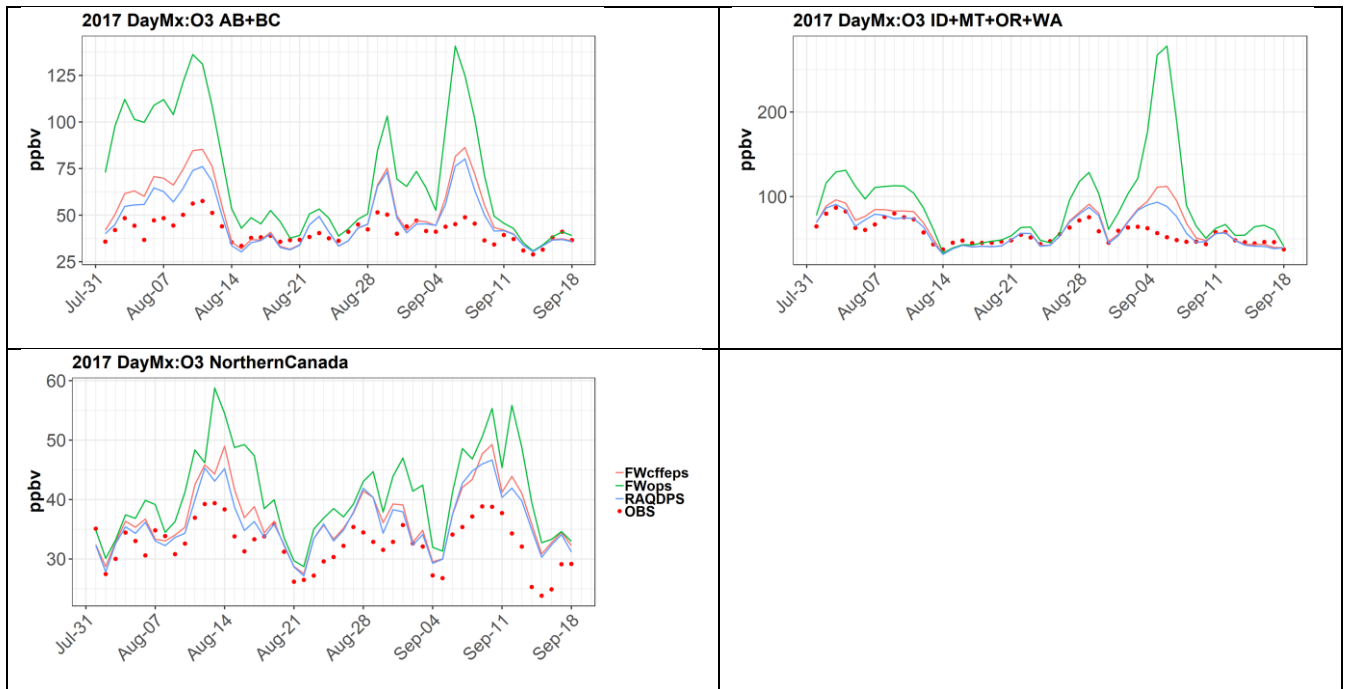
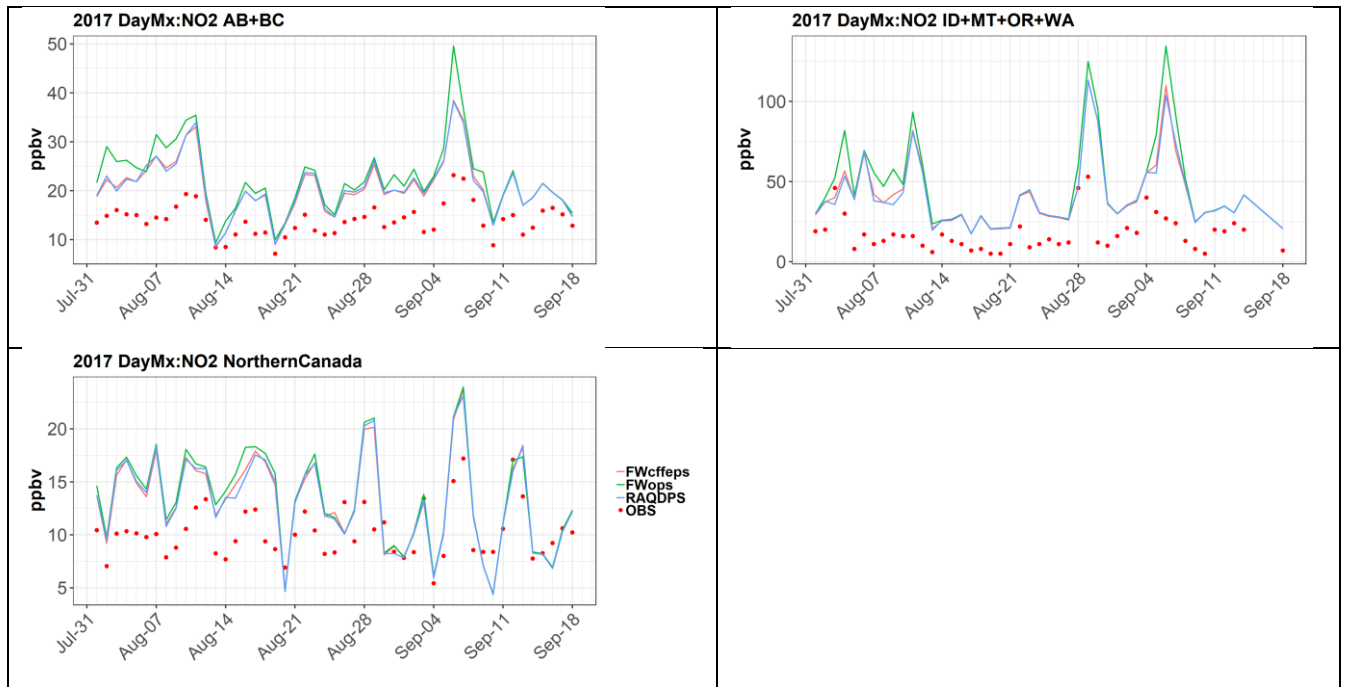



Figure S1. Time series of mean daily maximum O₃ volume mixing ratio (ppbv) from Aug. 1 to Sept. 18, 2017 for the three forecast models and surface measurements (OBS) averaged across measurement stations in the AB+BC, ID+MT+OR+WA, and Northern-Canada regions.

5   **Figure S2. Time series of mean daily maximum NO₂ volume mixing ratio (ppbv) from Aug. 1 to Sept. 18, 2017 the three forecast models and surface measurements (OBS) averaged across measurement stations in AB+BC, ID+MT+OR+WA and Northern-Canada regions**

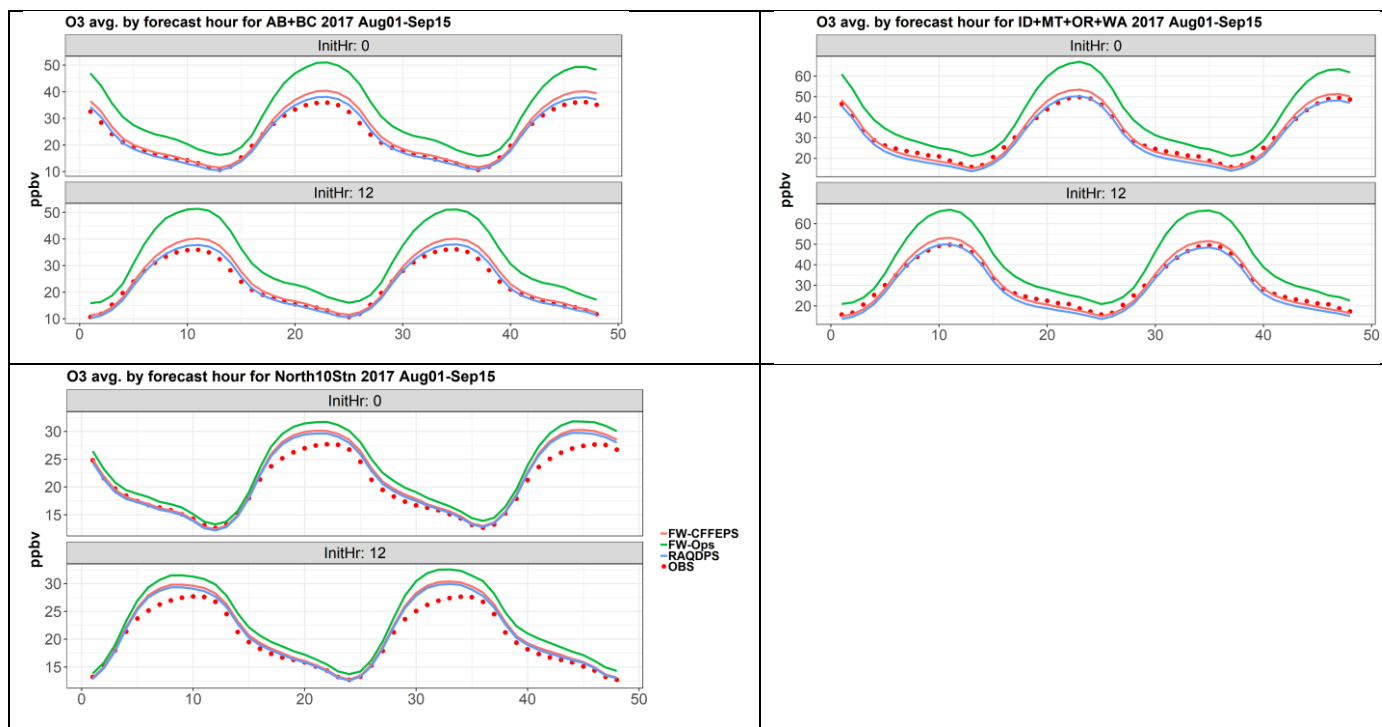
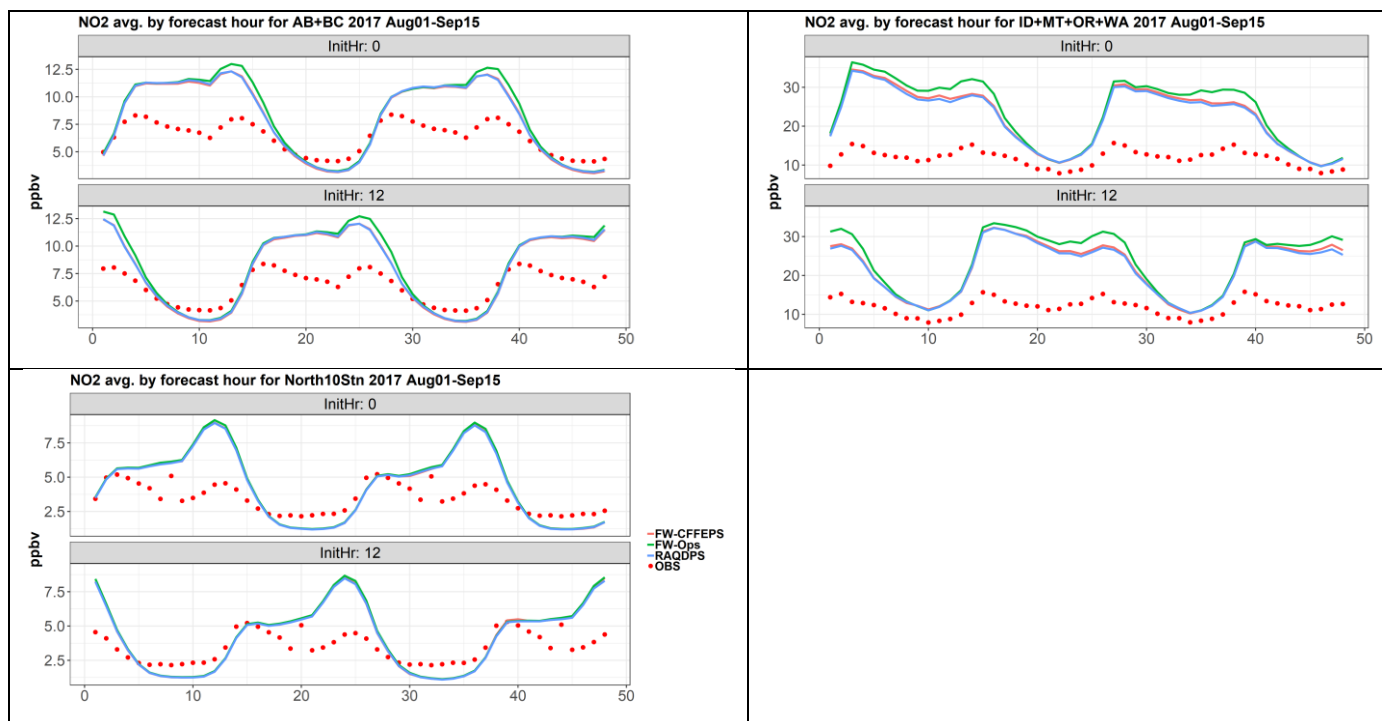



Figure S3. Mean O₃ volume mixing ratio (ppbv) by forecast hour (f00-f48) for the period from Aug. 1 to Sept. 16, 2017 for the three forecast models and surface measurements (OBS) for measurement stations in the AB+BC, ID+MT+OR+WA, and Northern-Canada regions.

5   **Figure S4. Mean NO₂ volume mxing ratio (ppbv) by forecast hour (f00-f48) from Aug. 1 to Sept. 16, 2017 for the three forecast models and surface measurements (OBS) for measurement stations in the AB+BC, ID+MT+OR+WA and Northern-Canada regions.**

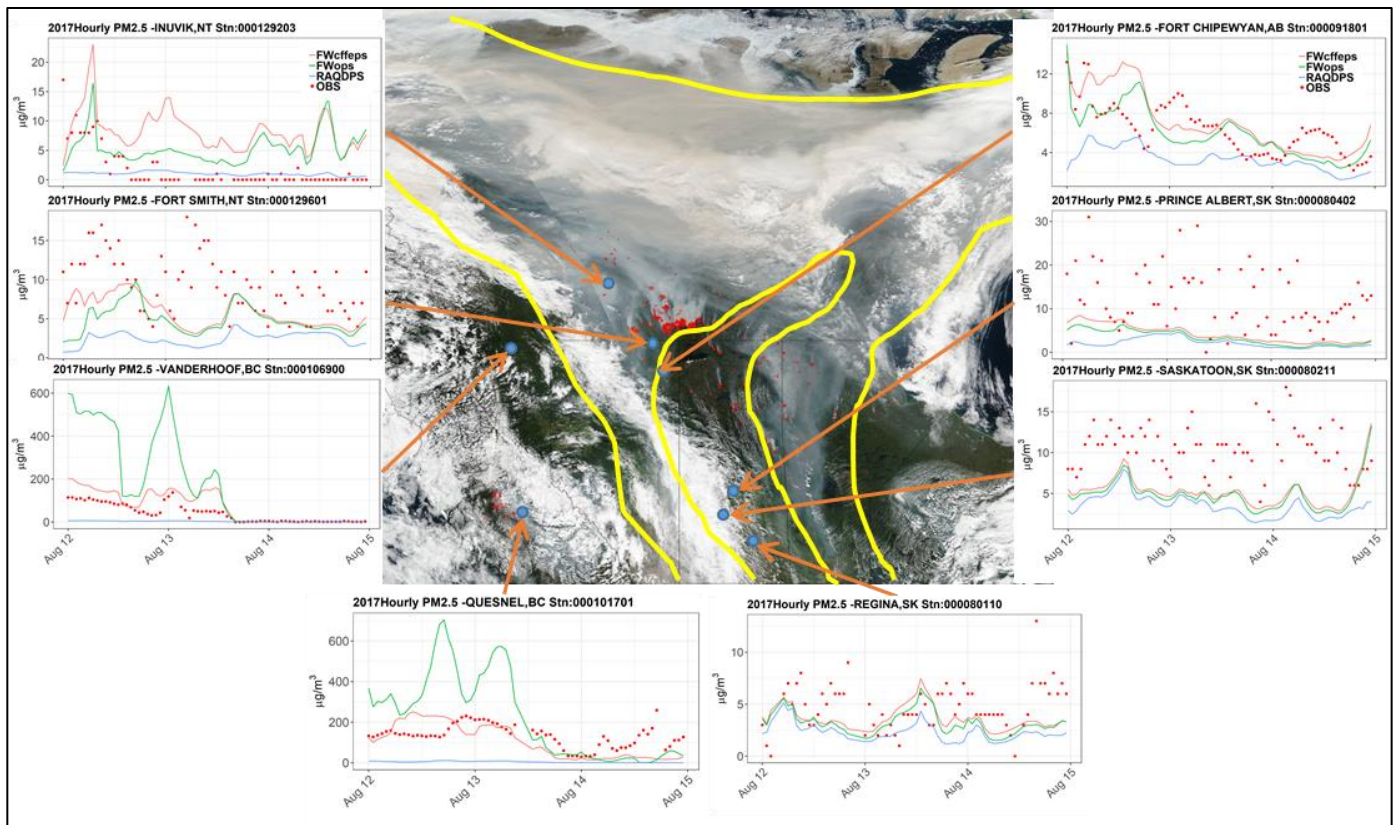



Figure S5. Time series of hourly surface PM₂.₅ concentations (μg m⁻³) from Aug. 12–15, 2017 for three model simulations and measurements (OBS) for selected stations with station locations identified by red arrows, and (centre) VIIRS true colour satellite image for Aug. 13, 2017 with yellow lines superimposed to aid comparison. Fire hotspots are represented in red dots in VIIRS image. VIIRS image source: NASA https://www.nasa.gov/image-feature/goddard/2017/smoke-and-clouds-obscure-skies-in-northern-canada

S3. Model Forecast Evaluation for the 2018 Fire Season

Similar to the model forecast evaluation for the 2017 fire season presented in Section 3.1, model outputs from FireWork-CFFEPS (FireWork v2.0) and FireWork-Ops (FireWork v1.0) were analysed and benchmarked against RAQDPS for the 2018 season. FireWork-CFFEPS was run in hindcast mode with the same forecast configuration as FireWork-Ops for three months (July to August) in 2018 that covered periods of high wildfire activity in Canada.

2018 was an extreme year for fire activity in Canada. Most fires occurred in BC in August but there was some activity in eastern Canada earlier in the summer. Figure S6 shows the Canada-wide weekly burned area starting in May 1, 2018 compared to the 10-year average. The last three weeks of August 2018 stand out for higher-than-average fire activity.

The $PM_{2.5}$ contribution from fire emissions (fire-$PM_{2.5}$), both primary and secondary, can be derived from the difference in model output between RAQDPS and FireWork during the same time period. (Figure S7) shows monthly mean forecast surface fire-$PM_{2.5}$ concentrations from FireWork-CFFEPS and FireWork-Ops for the 2018 fire season, with the contribution from fire-$PM_{2.5}$ dominating near areas of fire activity. Similar to the analysis for the 2017 fire season (Figure 7), there are generally lower surface concentrations near areas of hotspot sources but overall larger spatial distributions of lower level $PM_{2.5}$ concentrations for FireWork-CFFEPS compared to FireWork-Ops.

Daily maximum hourly model forecast surface concentrations for $PM_{2.5}$, O_3, and NO_2 were analysed against available hourly surface measurements from the Canadian NAPS network and the U.S. AQS network. The number of available measurement stations and their distribution across the model domain are indicated in Figure S8. Model results are paired by forecast time and by the grid cell enclosing the station location. Only stations with greater than >75% measurement completeness were included. Most of the measurement stations are located in urban and suburban areas with high population density. Also, most of the NAPS network stations have co-located measurements for all three pollutants whereas there are comparatively fewer stations in the U.S. AQS network that measure NO_2.

Model forecast performance statistics and categorical scores were calculated for stations across the four large geographic regions (WCAN, ECAN, WUSA, EUSA) shown in Figure 1 and separately for stations within the two westernmost Canadian provinces (BC, AB) and two western U.S. states (WA, MT) that were strongly impacted by fire emissions during the fire season. Table S7 provides a comparison of forecast statistics for FireWork-CFFEPS and FireWork-Ops with the RAQDPS. Table S8 provides a similar comparison of categorical scores for the same concentration threshold levels that were used in the 2017 evaluation analysis: $PM_{2.5}$ threshold of 30 µg m^{-3}, O_3 threshold of 65 ppbv, and NO_2 threshold of 30 ppbv. Similar to the results shown in Section 3.1 for 2017, FireWork also outperforms the RAQDPS for $PM_{2.5}$ and FireWork-CFFEPS outperforms FireWork-Ops in 2018. Figure S9 and Figure S10 show time series of mean daily maximum $PM_{2.5}$, O_3, and NO_2 abundances over the three months averaged over stations within two western Canadian provinces (AB, BC) and two western U.S. states (MT, WA). The FireWork-CFFEPS time series generally reduce pollutant over-predictions associated with FireWork-Ops while at the same time capturing fire contributions to $PM_{2.5}$ that are missed by the RAQDPS. The number of stations used to calculate the averages as well as selected model evaluation statistics over the three-month period are also listed within the two figures.

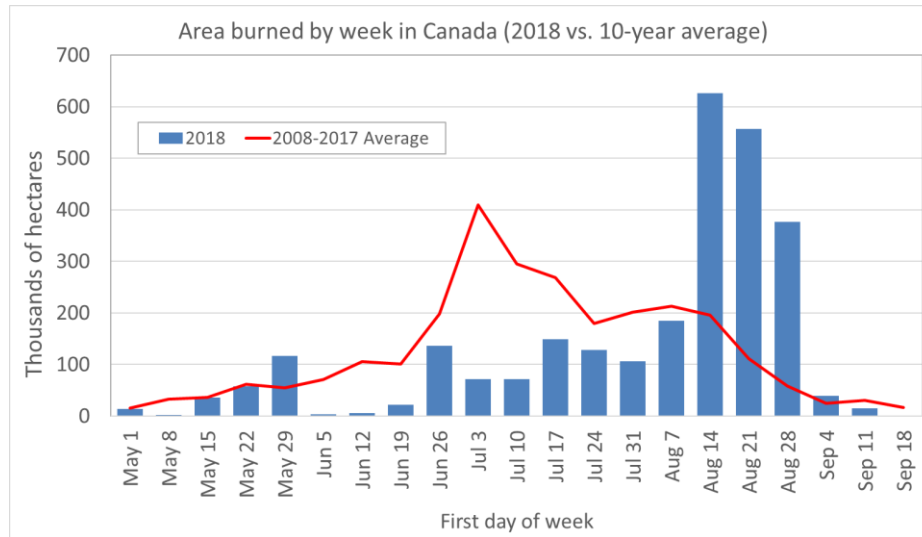
Model performance by forecast hour in terms of diurnal variability predicted by the three forecast systems for the stations in the two regions (AB+BC and MT+WA) are presented in Figure S11. Similar to Figure 9, Figure S3 and Figure S4, modelled and measured surface $PM_{2.5}$, O_3 and NO_2 were paired and averaged by forecast hour (f00-f48) over all stations within the region during period of high fire activity from Aug. 13 to Aug. 26 2018. As with the evaluation for 2017, there are improvements to O_3 hourly

variability with FireWork-CFFEPS compared to FireWork-Ops, and minimal changes to NO_2 concentrations. However, unlike the 2017 evaluation, the hourly forecast performance for $PM_{2.5}$ is under-predicted with FireWork-CFFEPS, and the concentration ranges were better represented with FireWork-Ops. Upon further analysis, this is due to effects of averaging over periods of under-prediction in both FireWork systems (Aug. 13-20 2018), together with periods of large over-predictions in the FireWork-Ops

5   system (Aug. 20-26 2017). The compensation of forecast $PM_{2.5}$ errors in the FireWork-Ops system resulted in the overall better diurnal average compared to FireWork-CFFEPS. This can be observed from examples of station specific hourly $PM_{2.5}$ concentrations time series comparisons from AB and BC in Figure S12.

Similar to Section 3.1.5, fire-$PM_{2.5}$ VCD fields (g m^{-2}) from the FireWork-CFFEPS and FireWork-Ops forecasts were subjectively compared against NASA's VIIR-true-colour image for a case study. Figure S13 shows the 24-hour forecast fire-$PM_{2.5}$ VCD field

10   from the FireWork-Ops and FireWork-CFFEPS simulations from the Aug. 22 12 UTC simulation, valid at Aug. 23 12 UTC. These fields are compared against a VIIRS true colour satellite image for Aug. 23, 2018. A visual comparison reveals lower $PM_{2.5}$ VCD near fire hotspots, but in general a more widespread spatial distribution for FireWork-CFFEPS compared to FireWork-Ops. Figure S14 shows a quantitative comparison of hourly surface $PM_{2.5}$ concentration time series over the same period (Aug. 22-25) from selected stations across Canada. Forecast concentrations were well represented from FireWork-CFFEPS for stations in the west,

15   near sources of fire hotspots. However, surface $PM_{2.5}$ concentrations were generally under-predicted for stations further downwind of the fire hotspots in Ontario and Quebec, though with FireWork-CFFEPS showing slightly higher forecast concentrations than FireWork-Ops.

20   **Figure S6. National fire burn area in Canada by week for 2018 fire season (blue vertical bars) and previous 10-year average (red line).**

FireWork-CFFEPS	FireWork-Ops

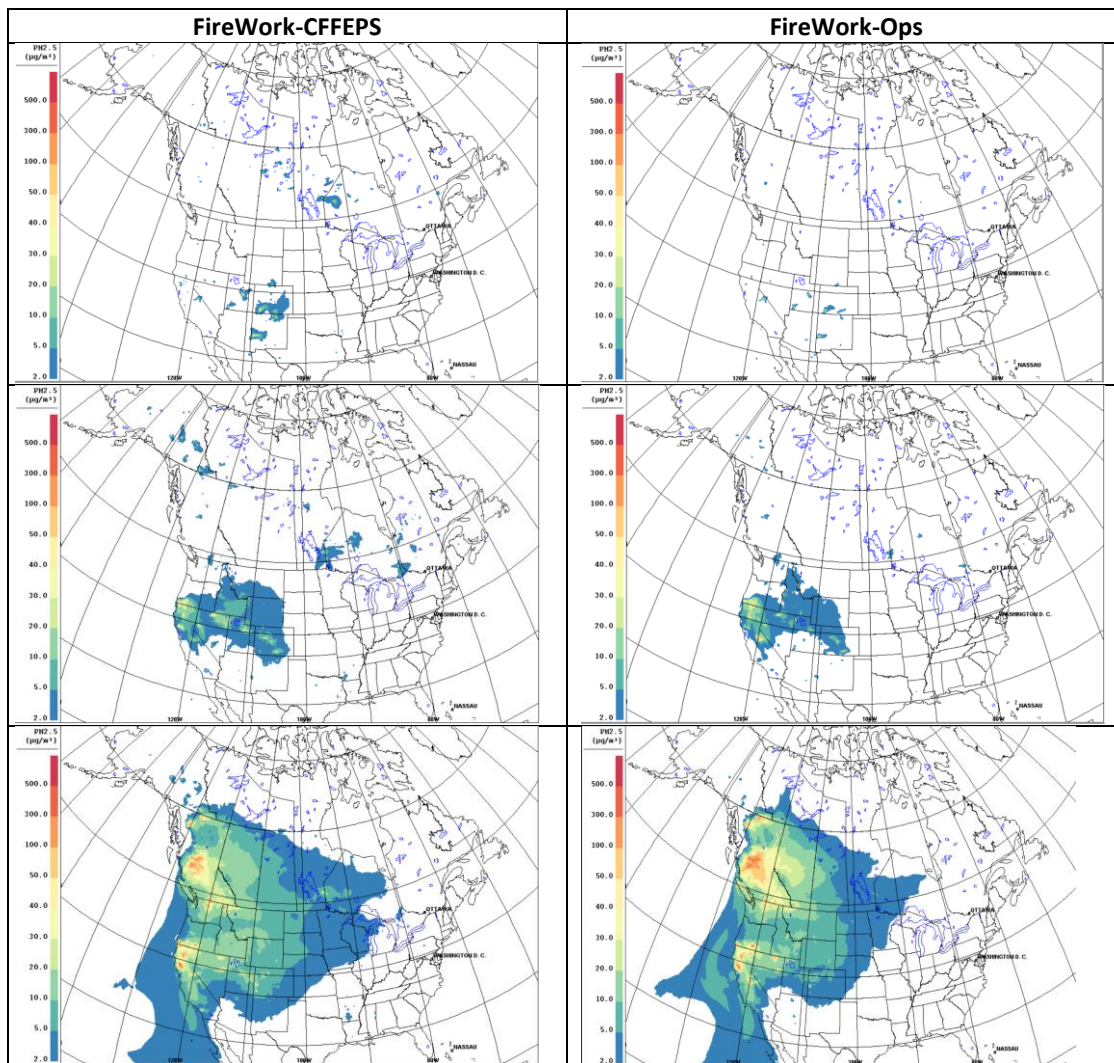



Figure S7. Mean monthly surface fire-PM$_{2.5}$ concentrations (µg m^{-3}) from FireWork-CFFEPS (left) and FireWork-Ops (right) for June (top row), July (middle row), and August (bottom row) 2018.

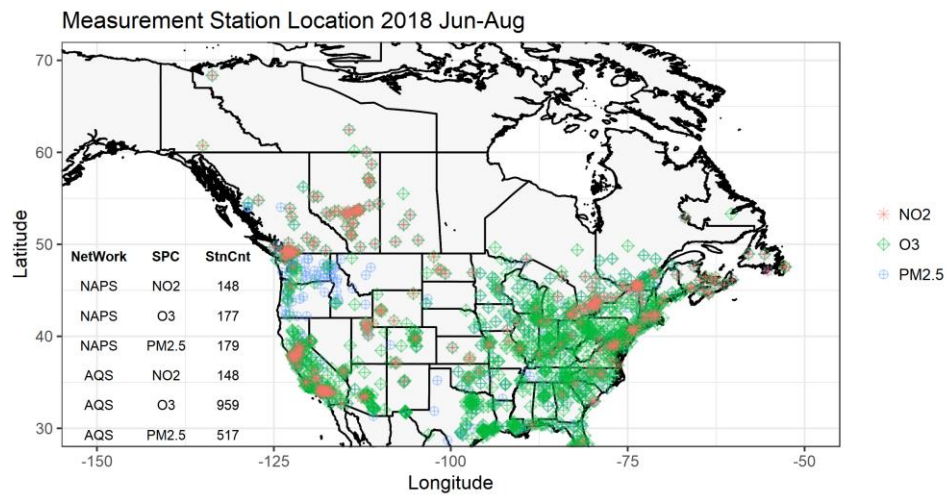



Figure S8. Locations of Canadian NAPS and U.S. AQS stations used for 2018 model evaluation with 75% measurement completeness criterion. Some stations measure more than one species.

Table S7. Model performance statistics for daily maximum surface (a) PM₂.₅, (b) O₃, and (c) NO₂ concentrations for June-August 2018 for three modelling systems and four geographic regions (shown in Figure 1). Units of Ō, M̄, MB, and RMSE are µg m⁻³ for PM₂.₅ and ppbv for O₃ and NO₂.

(a) PM$_{2.5}$	WCAN (81 stations, n=7369)			ECAN (98 stations, n=8870)			WUSA (198 stations, n=17852)			EUSA (319 stations, n=28664)		
Model	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS
$\bar{O}$	27			14			22			17		
$\bar{M}$	10	30	21	13	14	15	12	26	25	16	16	17
MB	-16.9	2.9	-5.9	0.0	0.2	1.3	-10.1	3.5	2.4	-0.8	-1.0	0.2
R	0.16	0.53	0.64	0.19	0.27	0.32	0.10	0.49	0.50	0.19	0.20	0.20
RMSE	45	78	35	13	13	13	30	56	47	13	13	13

(b) O$_3$	WCAN (76 stations, n=6919)			ECAN (101 stations, n=9181)			WUSA (270 stations, n=24545)			EUSA (689 stations, n=62559)		
Model	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS
$\bar{O}$	41			42			61			48		
$\bar{M}$	42	50	43	50	50	50	72	78	73	69	69	70
MB	1	9	3	8	8	8	10	17	12	21	21	22
R	0.69	0.45	0.68	0.73	0.73	0.73	0.65	0.53	0.65	0.65	0.66	0.66
RMSE	13	29	14	15	16	16	22	34	23	27	27	27

(c) NO$_2$	WCAN (77 stations, n=7035)			ECAN (71 stations, n=6467)			WUSA (74 stations, n=6743)			EUSA (74 stations, n=6652)		
Model	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS
$\bar{O}$	11			10			14			12		
$\bar{M}$	16	17	16	16	16	16	29	29	29	26	26	26
MB	5	6	5	6	6	6	14	15	15	14	14	14
R	0.59	0.59	0.60	0.62	0.63	0.62	0.61	0.60	0.61	0.66	0.66	0.66
RMSE	12	13	12	13	13	13	22	23	22	21	21	21

Table S8. Categorical scores for June-August 2018 by geographic region for (a) PM$_{2.5}$ based on threshold of 30 µg m^{-3}, (b) O$_3$ based on threshold of 65 ppbv, (c) NO$_2$ based on threshold of 30 ppbv

(a) PM$_{2.5}$	WCAN			ECAN			WUSA			EUSA		
Model	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS
POD	2%	46%	42%	1%	2%	4%	3%	41%	44%	3%	2%	3%
FAR	76%	30%	26%	100%	99%	98%	85%	50%	49%	96%	96%	96%
CSI	1%	38%	37%	0%	1%	1%	2%	29%	31%	1%	1%	2%

(a) O$_3$	WCAN			ECAN			WUSA			EUSA		
Model	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS
POD	31%	48%	39%	72%	72%	73%	54%	62%	58%	81%	81%	81%
FAR	82%	91%	82%	87%	87%	87%	51%	56%	52%	87%	87%	87%
CSI	13%	8%	14%	12%	12%	12%	35%	35%	35%	13%	13%	13%

(a) NO$_2$	WCAN			ECAN			WUSA			EUSA		
Model	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS	RAQDPS	FW-Ops	FW-CFFEPS
POD	41%	41%	42%	24%	25%	24%	76%	76%	76%	53%	54%	54%
FAR	95%	95%	95%	98%	98%	98%	93%	93%	93%	95%	95%	95%
CSI	5%	5%	5%	2%	2%	2%	6%	6%	6%	5%	5%	5%

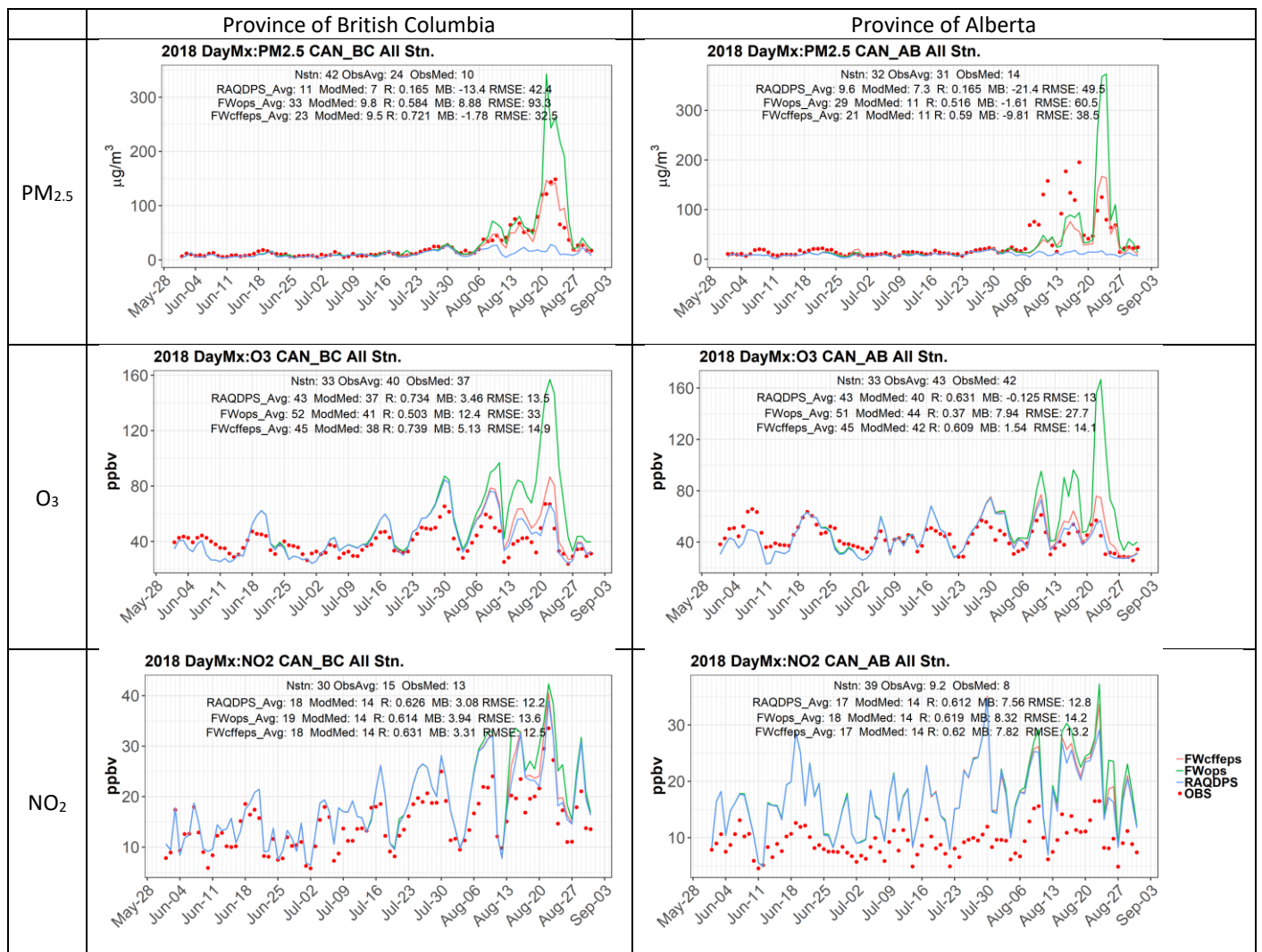



Figure S9. Time series of mean daily maximum PM2.5, O₃, and NO₂ abundances for BC and Alberta in Canada from June–Aug. 2018 for forecasts by the RAQDPS (blue lines), FireWork-Ops (green lines), and FireWork-CFFEPS (red lines). The number of stations within each province and values of some model evaluation statistics for the three-month period are also listed within each panel, where Nstn is the number of measurement stations used in the calculation, ObsAvg and ObsMed are measurement average and median concentrations, respectively, RAQDPS_Avg, FWops_Avg, and FWcffeps_Avg and the ModMed on the same line are average and median forecast concentrations from the RAQDPS, FireWork-Ops, and FireWork-CFFEPS systems, respectively, MB is mean bias, and RMSE is root mean square error.

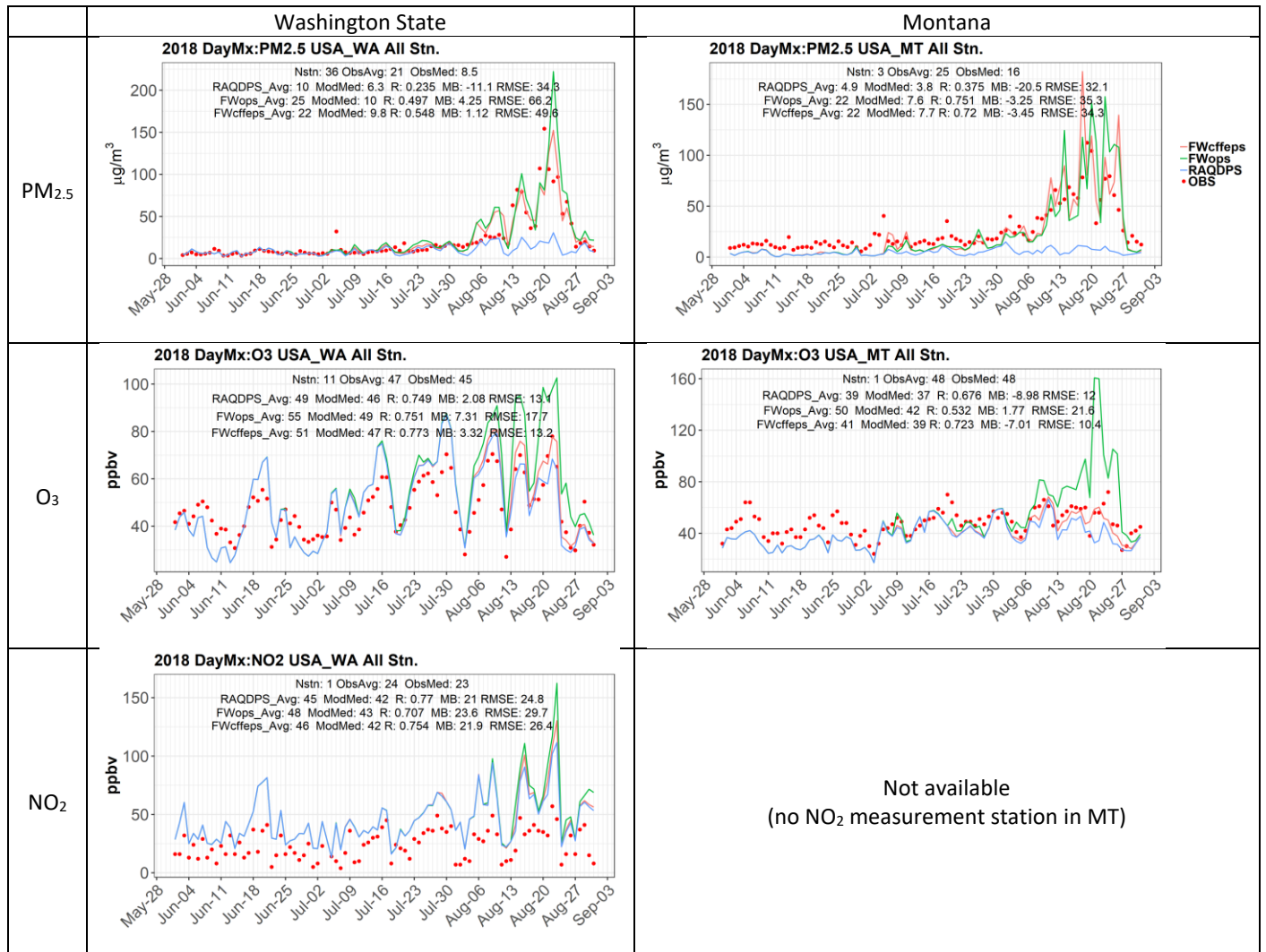



Figure S10. Same as Figure S9 but for WA and MT states in the U.S.

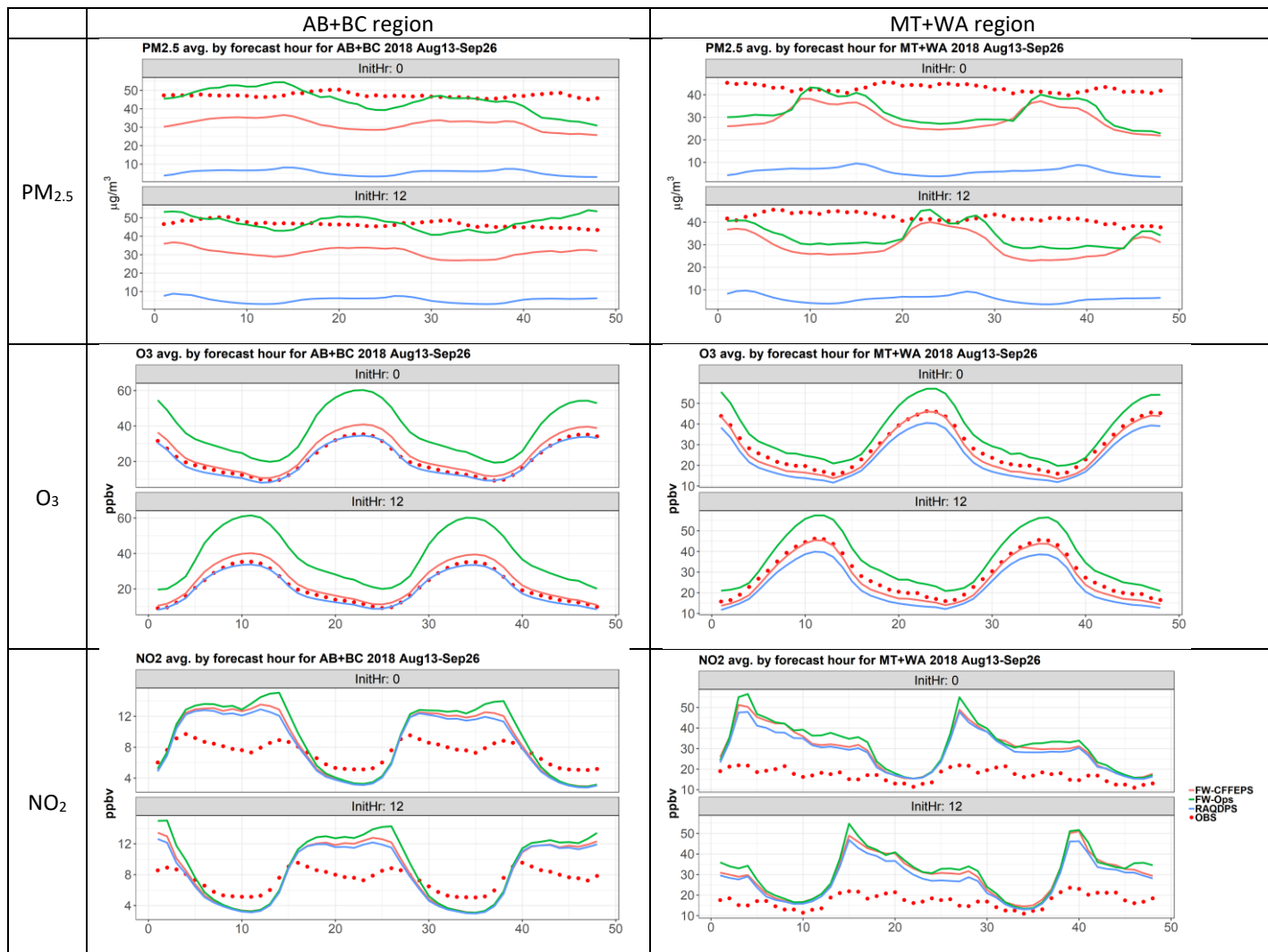



Figure S11. Mean PM₂.₅ concentration (top), O₃ mixing ratio (middle) and NO₂ mixing ratio (bottom) by forecast hour (f00-f48) for the period from Aug. 13 to Aug. 26, 2018 for the three forecast models and surface measurements (OBS) for measurement stations in the AB+BC, and MT+WA regions. Note there is only one NO₂ measurement station in WA for the MT+WA region.

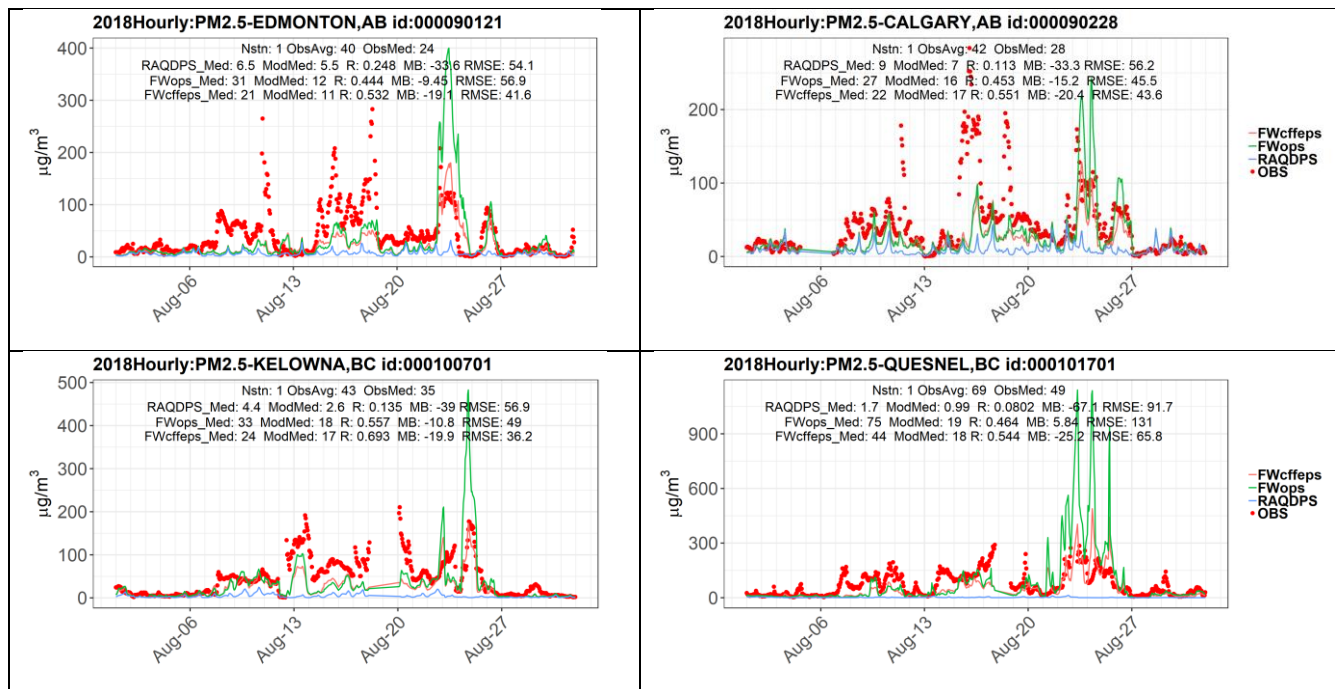



Figure S12. Hourly time series of PM2.5 concentrations for Edmonton (lat./lon.: 53.54823/-113.3681) , Calgary (lat./lon.: 51.04701/- 114.0756) in AB, and Kelowna (lat./lon.: 49.86234/ -119.4774), Quesnel (lat./lon.: 52.98169 / -122.4932) in BC for Aug. 1-31 2018 for forecasts by the RAQDPS (blue lines), FireWork-Ops (green lines), and FireWork-CFFEPS (red lines). Model evaluation statistics for the period are listed within each panel similar to Figure S9.

5   .

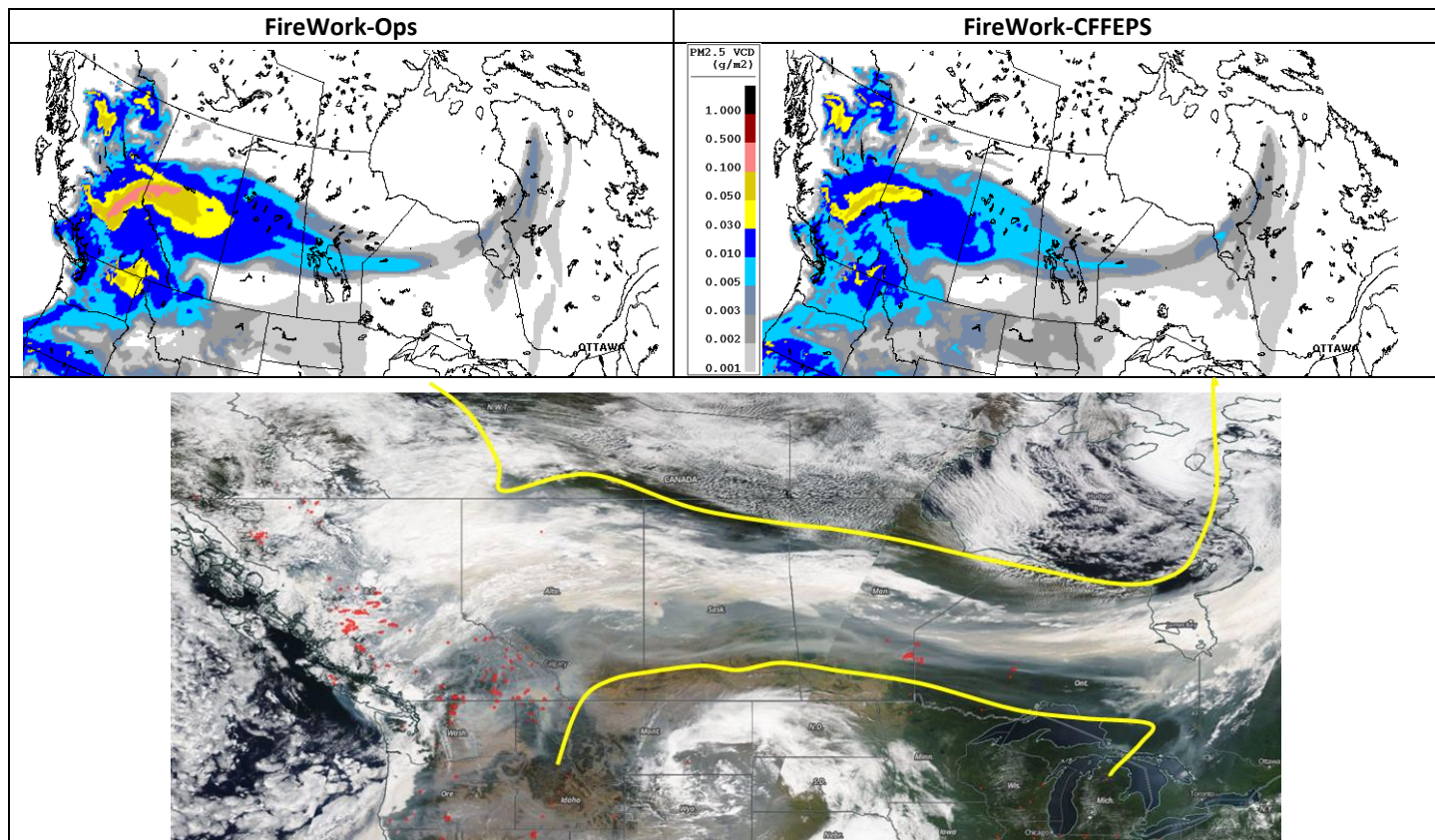



Figure S13. Fire-PM$_{2.5}$ VCD fields (g m^{-2}) forecast by FireWork-Ops (top left) and FireWork-CFFEPS (top right) from the Aug. 22, 2018 12 UTC forecast run valid for August 23, 2018 12 UTC, and (bottom) VIIRS true colour satellite image for August 23, 2018 with yellow lines superimposed to aid comparison. Fire hotspots are represented as red dots. VIIRS image source: NASA https://worldview.earthdata.nasa.gov

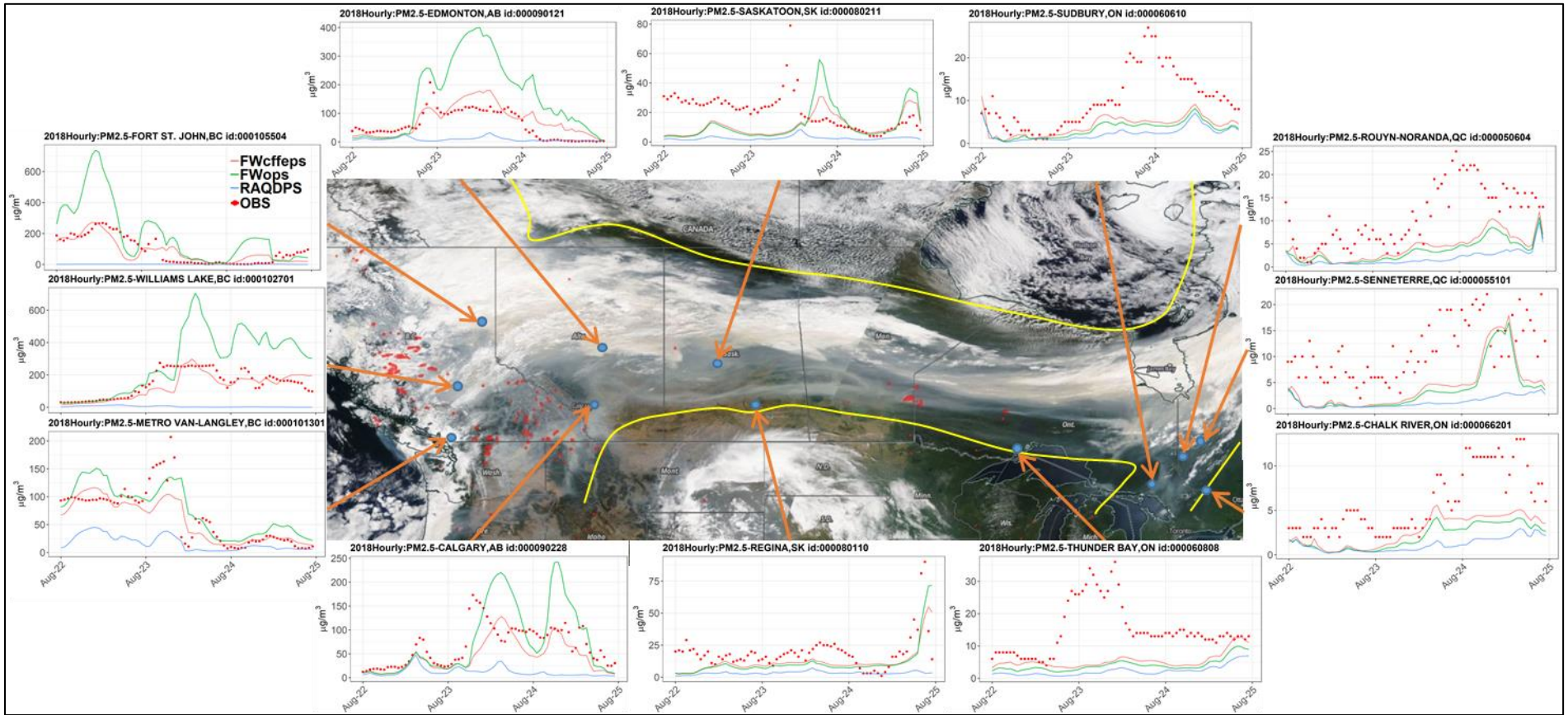



Figure S14. Time series of 2018 hourly surface PM$_{2.5}$ concentration from three model simulations and measurements (OBS) for selected stations for Aug. 22-25, 2018, and (centre) VIIRS true colour satellite image for August 23, 2018 with yellow lines superimposed to aid comparison. RAQDPS forecasts are shown with blue lines, FireWork-Ops forecasts with green lines, and FireWork-CFFEPS forecasts with red lines in the inset time series plots. Station locations are represented by blue dots and fire hotspots are represented in red dots superimposed on the satellite image. VIIRS image source: NASA https://worldview.earthdata.nasa.gov.