Regional Air Quality Impact of Northern South America Biomass Burning Emissions

Juan Felipe Mendez Espinosa, Ricardo Morales Betancourt, Luis Carlos Belalcazar Cerón

Barranquilla, 2019
Introduction
Seasonality of Particulate Matter in Bogotá

Monthly average in Bogotá 2006-2016

$\text{PM}_{10}$ 24.8 $\mu g m^{-3}$

$\text{PM}_{2.5}$ 14.8 $\mu g m^{-3}$
Introduction

Impact of meteorological conditions on the annual cycle of pollutants

Dry season, December, January and February

Wet season of June, July August

Outgoing Longwave Radiation (OLR) anomalies (contours)
Surface winds climatology (arrows)

Source: NCAR/NCEP Reanalysis
Introduction
Previous studies

(Gaitan et al., 2007)
Introduction
Previous studies

(Chacón and Belalcazar, 2016)

EVENTS
Bogotá: $r$: 0.4
Bucaramanga: $r$: 0.42
Introduction

Previous studies

*Mixing Height, $H_{mix}$ Modification of Holzworth method, (1967)
Biomass burning emissions have a substantial impact on:

- **Regional air quality** (Forster et al., 2001)
- **Climate** (Thornhill et al., 2017)
- **Health** (Youssouf et al., 2014)
- **Economy** (Fann et al., 2018)

Schematic showing the injection of biomass burning emissions into atmospheric layers and their subsequent long-range transport in the upper atmosphere (Keywood, 2013).
Introduction

Biomass burning – Regional issue

Colombia

BB can episodically add up to 100 $\mu g m^{-3}$

to daily-average baseline $PM_{10}$ concentration (35 $\mu g m^{-3}$) in Arauca and Yopal (Hernandez et al., 2019)

(a.) True colour

(b.) Short-wave infrared (SWIR) light

Towering plume of smoke from a single fire in Orinoco basin from Sentinel 2
400 thousand fires occurred during 2006-2016 with ≈ 23 million MW of emitted energy.

During the dry season (November to March), massive biomass burning occur (natural and anthropogenic).

**Satellites:** Aqua – Terra, NASA  
**Sensor:** MODIS C6  
**Spatial Resolution:** 1km at nadir  
**Confidence ≥75%**  
**Regional impact fires ~120 ha**

2006 to 2016
- Number of fires per day (Nf)
- Fire Radiative Power (FRP) [MW]

Spearman correlation coefficient 0.99 (p-value < 1%)
More species were included
- Most recent analysis (2006 - 2016)
- Temporal representativity (≥75%)

| PM_{10}, | Better Biomass burning tracer |
| PM_{2.5}, |
| CO |

A more detailed spatial-analysis in a systematic way
- Selected those active fires in the vicinity of the air masses arriving at each city.
- Coupling pollutant concentrations and back-trajectories of air masses

Contribution of local variables to variations in air pollution levels to those associated with regional biomass burning emissions were taken into account.
Our Goal

• Establish the degree of association between air pollutants indicators (PM$_{10}$, PM$_{2.5}$, and CO) in Colombian cities and biomass burning events in the North of South America domain using remote sensing data for number of fires and back-trajectory modeling.
Methodology
Air pollution data

Temporal representativity (≥75%)

**BOGOTA**
- PM$_{10}$ (2006-2016)
- PM$_{2.5}$ (2006-2016)
- CO (2009-2016)

**MEDELLIN**
- PM$_{10}$ (2008-2016)

**BUCARAMANGA**
- PM$_{10}$ (2008-2016)
Methodology

Spatio-temporal distribution of fires

How do we count active fires in the vicinity of the air masses arriving at each city?

[Map showing fire distribution in January 2016]
Methodology

Spatio-temporal distribution of fires

Selecting fires based on a back-trajectory analysis

Related fires with air quality, $N_f$, inside buffer radius of 50, 100 and 150 km
Methodology
Spatio-temporal distribution of fires

Selecting fires based on a back-trajectory analysis

Counting method includes:

- Location of active fires in the vicinity of the air masses arriving at each city
- Lag between a fire and its possible impact over observed concentrations
Methodology

Back-trajecotories of air masses

HYSPLIT back-trajecotories run **SYSTEMATICALLY**

8 arrival times daily
Total run time per trajectory = 96 hours
Height = 50 to 3000 meters AGL
Meteorology = Reanalysis, GDAS1
Period = 2006 - 2016

CITIES ANALYZED
- Bogotá
- Medellín
- Bucaramanga
Results

Time series

Monthly average of number of active fires, and PM from 2006 to 2016.
Results

Association analysis

A seven-day centered moving average was applied to remove the concentration variations between workdays and weekends.

Comparative Spearman's rank correlation between $\overline{N_f}$ and $\overline{PM_{10}}$, $\overline{PM_{2.5}}$ for the 2006-2016 period. All p-values are <1%

<table>
<thead>
<tr>
<th>Receptor height (m)</th>
<th>Buffer radius (km)</th>
<th>Bogotá $N_f$-PM$_{2.5}$</th>
<th>Bogotá $N_f$-PM$_{10}$</th>
<th>Medellín $N_f$-PM$_{10}$</th>
<th>Bucaramanga $N_f$-PM$_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>50</td>
<td>0.22</td>
<td>0.31</td>
<td>0.27</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.20</td>
<td>0.29</td>
<td>0.26</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.22</td>
<td>0.30</td>
<td>0.25</td>
<td>0.60</td>
</tr>
<tr>
<td>1500</td>
<td>50</td>
<td>0.35</td>
<td>0.43</td>
<td>0.29</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.35</td>
<td>0.41</td>
<td>0.30</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.35</td>
<td>0.42</td>
<td>0.30</td>
<td>0.50</td>
</tr>
<tr>
<td>3000</td>
<td>50</td>
<td>0.50</td>
<td>0.55</td>
<td>0.36</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.49</td>
<td>0.54</td>
<td>0.37</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.50</td>
<td>0.55</td>
<td>0.37</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Association increases
Results

Association analysis

- The highest association is obtained for Bucaramanga
- Bucaramanga is the closest place to biomass burning emissions
- Long-range transport is not as important as in the Bogotá case - short receptor height
Results

Association analysis

- Long-range transport is important — Highest association for high receptor height (Medellin, Bogotá)
- the trajectories at greater height are longer
We do not know...

Results
Association analysis

Biomass burning emissions

Mixing height

BOGOTÁ
Results – spatial analysis

IT IS NOT JUST A TIME SERIES TOPIC...

(b.) July

HYSLIPT back trajectories density and MODIS fires (2006 -2016)
All Januaries vs. All Julies
Spatial attribution of source regions

Impact of receptor height over Back-trajectories - $\overline{PM}_{2.5}$

GDAS1 meteorology
Results

Comparison between meteorological database

**GDAS1 (1°)**  **Reanalysis (2.5°)**

Impact of meteorological data set over Back-trajectories - $PM_{10}$

(c.) MEDELLIN
Results

Contribution over seasonal cycle local meteorological vs. BB emissions

Multiple linear regression analysis using MONTHLY MEANS (2006 - 2016)

\[ \langle \cdot \rangle = \beta_0 + \beta_{N_f} \langle N_f \rangle + \beta_{H_{mix}} \langle H_{mix} \rangle + \varepsilon \]

- Uni-variate regressions
  - \( N_f \) and \( H_{mix} \) showed the highest correlation with air pollutant concentrations

- Multiple linear regressions (2, 3, 4 independent variables)

- Statistical assumptions
  - Independence of error – It was corrected using Praise Winter method
  - Linearity
  - Homoscedasticity
  - Normality
  - Multicollinearity
  - Unusual observations

Akaike Information Criterion (AIC)
Results

Contribution over seasonal cycle
local meteorological vs. BB emissions

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\beta_0$ (s.e.)</th>
<th>$\beta_N$ (s.e.)</th>
<th>$\beta_{H_{mix}}$ (s.e.)</th>
<th>p-value</th>
<th>p-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle PM_{10} \rangle$</td>
<td>72.4 (5.9)</td>
<td>+0.23 (0.06)</td>
<td>-16.6 (4.5)</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td></td>
<td>(60.6, 84.2)</td>
<td>(0.98, 0.37)</td>
<td>(-25.6, -7.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\langle PM_{2.5} \rangle$</td>
<td>41.1 (3.7)</td>
<td>+0.15 (0.04)</td>
<td>-9.5 (2.9)</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td></td>
<td>(33.6, 48.7)</td>
<td>(0.06, 0.24)</td>
<td>(-15.3, -3.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\langle CO \rangle$</td>
<td>1180 (109)</td>
<td>-0.76 (1.27)</td>
<td>-225 (83)</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td></td>
<td>(962, 1397)</td>
<td>(-3.3, 1.7)</td>
<td>(-392, -58.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$R^2 = 0.32$ \quad \varepsilon = 7

$R^2 = 0.31$ \quad \varepsilon = 4

$R^2 = 0.08$ \quad \varepsilon = 119

$\varepsilon$ means Residual standard error.
## Results

Contribution over seasonal cycle
local meteorological vs. BB emissions

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\beta_0$ (s.e.)</th>
<th>C.I. $\beta_0$</th>
<th>p-value</th>
<th>$\beta_{N_f}$ (s.e.)</th>
<th>C.I. $\beta_{N_f}$</th>
<th>p-value</th>
<th>$\beta_{H_{mix}}$ (s.e.)</th>
<th>C.I. $\beta_{H_{mix}}$</th>
<th>p-value</th>
<th>$R^2$</th>
<th>$\varepsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle PM_{10} \rangle$</td>
<td>72.4 (5.9)</td>
<td>(60.6, 84.2)</td>
<td>&lt; 1%</td>
<td>+0.23 (0.06)</td>
<td>(0.98, 0.37)</td>
<td>&lt; 1%</td>
<td>-16.6 (4.5)</td>
<td>(-25.6, -7.5)</td>
<td>&lt; 1%</td>
<td>0.32</td>
<td>7</td>
</tr>
<tr>
<td>$\langle PM_{2.5} \rangle$</td>
<td>41.1 (3.7)</td>
<td>(33.6, 48.7)</td>
<td>&lt; 1%</td>
<td>+0.15 (0.04)</td>
<td>(0.06, 0.24)</td>
<td>&lt; 1%</td>
<td>-9.5 (2.9)</td>
<td>(-15.3, -3.7)</td>
<td>&lt; 1%</td>
<td>0.31</td>
<td>4</td>
</tr>
<tr>
<td>$\langle CO \rangle$</td>
<td>1180 (109)</td>
<td>(962, 1397)</td>
<td>&lt; 1%</td>
<td>-0.76 (1.27)</td>
<td>(3.3, 1.7)</td>
<td>&lt; 1%</td>
<td>-225 (83)</td>
<td>(-392, -58.9)</td>
<td>&lt; 1%</td>
<td>0.08</td>
<td>119</td>
</tr>
</tbody>
</table>

$\varepsilon$ means Residual standard error.
## Results

Contribution over seasonal cycle
local meteorological vs. BB emissions

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\beta_0$ (s.e.) C.I. $\beta_0$ p-value</th>
<th>$\beta_{Nf}$ (s.e.) C.I. $\beta_{Nf}$ p-value</th>
<th>$\beta_{H_{mix}}$ (s.e.) C.I. $\beta_{H_{mix}}$ p-value</th>
<th>$R^2$</th>
<th>$\varepsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle PM_{10} \rangle$</td>
<td>72.4 (5.9) (60.6, 84.2) &lt; 1%</td>
<td>+0.23 (0.06) (0.98, 0.37) &lt; 1%</td>
<td>-16.6 (4.5) (-25.6, -7.5) &lt; 1%</td>
<td>0.32</td>
<td>7</td>
</tr>
<tr>
<td>$\langle PM_{2.5} \rangle$</td>
<td>41.1 (3.7) (33.6, 48.7) &lt; 1%</td>
<td>+0.15 (0.04) (0.06, 0.24) &lt; 1%</td>
<td>-9.5 (2.9) (-15.3, -3.7) &lt; 1%</td>
<td>0.31</td>
<td>4</td>
</tr>
<tr>
<td>$\langle CO \rangle$</td>
<td>1180 (109) (962, 1397) &lt; 1%</td>
<td>-0.76 (1.27) (-3.3, 1.7) &lt; 1%</td>
<td>-225 (83) (-392, -58.9) &lt; 1%</td>
<td>0.08</td>
<td>119</td>
</tr>
</tbody>
</table>

Weaker seasonality

$\varepsilon$ means Residual standard error.
Results

Contribution over seasonal cycle local meteorological vs. BB emissions

\[
\langle \cdot \rangle = \beta_0 + \beta_{N_f} \langle N_f \rangle + \beta_{H_{mix}} \langle H_{mix} \rangle + \varepsilon
\]

\[
\Delta \langle H_{mix} \rangle \approx 290 \text{m} \quad \Rightarrow \quad 2.7 \pm 0.8 \mu g m^{-3} \quad 23\% \pm 6.8\%
\]

\[
\Delta \langle N_f \rangle \approx 33 \quad \Rightarrow \quad 4.9 \pm 1.3 \mu g m^{-3} \quad 41\% \pm 11\%
\]

\[
\Delta \langle H_{mix} \rangle \approx 290 \text{m} \quad \Rightarrow \quad 4.8 \pm 1.3 \mu g m^{-3} \quad 25\% \pm 6.8\%
\]

\[
\Delta \langle N_f \rangle \approx 33 \quad \Rightarrow \quad 8 \pm 2 \mu g m^{-3} \quad 39\% \pm 10\%
\]
Conclusion

Having in mind:

- **the degree of association** between *air pollutants* indicators (PM$_{10}$, PM$_{2.5}$) in Colombian cities and *biomass burning* events in the NSA domain, $\rho > 0.5$ (p-value <1%)

- Spatial analysis agree with temporal association

- **Fires** can contribute to seasonal cycle of $PM_{2.5}$ with $4\mu g m^{-3}$

Our findings *support the possibility* that *fires* in the Orinoco river basin *deteriorate air quality* in highly populated urban centers. Indeed, having *regional impact*
ACKNOWLEDGEMENT

Co-Authors

Proyecto Colciencias:
"Cuantificación del Impacto de Incendios Forestales Regionales Sobre la Calidad del Aire de la Ciudad de Bogotá“. Contrato No. FP44842-050-2017

Ph.D. Ricardo Morales, Director
Centro de Investigación en Ingeniería Ambiental, Universidad de los Andes
GRACIAS

Ricardo Morales Betancourt r.moralesb@uniandes.edu.co
Luis Carlos Belalcázar Ceron lcbelalcazarc@unal.edu.co
Juan Felipe Mendez Espinosa jf.mendez@uniandes.edu.co

Centro de Investigación en Ingeniería Ambiental
CIIA
UNIVERSIDAD DE LOS ANDES
Air Quality Group

Ricardo Morales, Assistant Professor

Karen Ballesteros (Ph.D. Student)

Juan Manuel Rincón (estudiante M.Sc.)
María Alejandra Rincón (estudiante M.Sc.)
Juan Felipe Méndez (estudiante M.Sc.)
Daniela Méndez Molano (estudiante M.Sc.)
Sebatian Espitia Cano (estudiante M.Sc.)
Catalia Pinto (estudiante M.Sc)
Alejandra Montejo (estudiante M.Sc)

Egresados:
María Paula Perez (M.Sc.)
Juan Pablo Ayala (M.Sc.)
Yadert Contreras (M.Sc.)
Miguel Quirama (M.Sc.)
Supplementary material
Regional air quality impact of northern South America biomass burning emissions

J.F. Mendez-Espinosa\textsuperscript{a}, L.C. Belalcazar\textsuperscript{b}, R. Morales Betancourt\textsuperscript{a,∗}

\textsuperscript{a} Department of Civil and Environmental Engineering, Universidad de Los Andes, Bogota, Colombia
\textsuperscript{b} Department of Chemical and Environmental Engineering, Universidad Nacional de Colombia, Bogota, Colombia

Transboundary transport of biomass burning aerosols and photochemical pollution in the Orinoco River Basin

Andrea J. Hernandez\textsuperscript{a}, Luis A. Morales-Rincon\textsuperscript{a}, Dien Wu\textsuperscript{b}, Derek Mallia\textsuperscript{b}, John C. Lin\textsuperscript{b}, Rodrigo Jimenez\textsuperscript{a,∗}

\textsuperscript{a} Universidad Nacional de Colombia – Bogota, Department of Chemical and Environmental Engineering, Air Quality Research Group, Bogota, DC, 111321, Colombia
\textsuperscript{b} Land-Atmosphere Interactions Research Group, Department of Atmospheric Sciences, University of Utah, Salt Lake City, UT, 84112, USA
More local study, April–May 2015

- BB can episodically add up to 100 ug m\(^{-3}\) to daily-average baseline PM10 concentration (35 ug m\(^{-3}\)).

- 16,376 tons of TPM associated to BB was emitted during May, 2015 according to Global Fire Emission Database (GFED4s) in the Orinoco basin.

- DAILY “PM10 concentrations” in both cities display rather reasonable correlations with CFER (\(r^2=0.34\) for Yopal and \(r^2=0.12\) for Arauca)

**CFER:** Indicator of the contribution at the receptor of particles from fire emissions, using a footprint from the **lagrangian model** STILT and the MODIS-FRP.
Overview:

Previous and Subsequent studies

Effect of wildfires on the air quality of Northern South America (2003-2013) (Belalcazar et al., 2015) - 108th A&WMA Annual Conference & Exhibition

Pearson correlation coefficient

PM$_{10}$ concentrations

Bogotá: 0.4
Bucaramanga: 0.42

Transboundary transport of biomass burning aerosols and photochemical pollution in the Orinoco River Basin (Hernandez et al., 2019)

More local study (Llanos region), April–May 2015

BB can episodically add up to 100 ug m$^{-3}$ to daily-average baseline PM10 concentration (35 ug m$^{-3}$ ) in Arauca and Yopal
Percent difference in occurrence for high $\overline{CO}$ concentrations (>70th percentile) over the 2014 - 2016 period for Bogotá. Calculations are performed for GDAS1.
Comparative Spearman's rank correlation between $N_f$ and $CO$, $PM_{10}$, $PM_{2.5}$ for the 2006-2016 period. All p-values are <1%

Highest correlation for the city closest to the fires region

and

From those obtained by Belalcazar et al. (2015)

<table>
<thead>
<tr>
<th>Receptor height (m)</th>
<th>Buffer radius (km)</th>
<th>Bogotá $N_f$-PM$_{2.5}$</th>
<th>Bogotá $N_f$-PM$_{10}$</th>
<th>Medellín $N_f$-PM$_{10}$</th>
<th>Bucaramanga $N_f$-PM$_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>50</td>
<td>0.22</td>
<td>0.31</td>
<td>0.26</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.20</td>
<td>0.29</td>
<td>0.25</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.22</td>
<td>0.30</td>
<td>0.26</td>
<td>0.56</td>
</tr>
<tr>
<td>1500</td>
<td>50</td>
<td>0.35</td>
<td>0.43</td>
<td>0.29</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.35</td>
<td>0.41</td>
<td>0.30</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.35</td>
<td>0.42</td>
<td>0.30</td>
<td>0.50</td>
</tr>
<tr>
<td>3000</td>
<td>50</td>
<td>0.50</td>
<td>0.55</td>
<td>0.36</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.49</td>
<td>0.54</td>
<td>0.37</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.50</td>
<td>0.55</td>
<td>0.37</td>
<td>0.50</td>
</tr>
</tbody>
</table>
**Methodology – Chapter 1**

Impact of meteorological conditions on the annual cycle of pollutants

- **Dry season**
- **Wet season**

**OLR anomalies** (contours)

**Surface winds** climatology (arrows)

*Mixing Height, $H_{mix}$*  
Modification of Holzworth method, (1967)

**Bogotá**

*Monthly means of PM$_{10}$, PM$_{2.5}$*
Mathematical concepts

\[ \overrightarrow{R}_{tp(tp-nh,z)} = \{Pn_{x,y,tp-nh}\}_n \]  
\hspace{1cm} (1.1)

In which, \( \overrightarrow{R}_{tp} \) represents the set of points that determine the trajectory L which is followed by an air parcel that returns n hours in time. \( tp \) in Equation (1.1) represents the departure time of a back-trajectory from a receptor, while \( z \) is the receptor height. \( Pn_{x,y,tp-nh} \) means a point with latitude and longitude which represents the location of an air parcel for \( tp - nh \). Where \(-nh\) is the number of returned hours in time (e.g. 0 to 96 hours). \( Ni_{R_{tp,d}} \) is mathematically represented by

\[ Ni_{R_{tp-day}} = \sum_{n=-1}^{n=each-24} \left\{ f_{x',y',t'/} (X_t' - Pn_{x,tp-nh})^2 + (Y_t' - Pn_{y,tp-nh})^2 \leq R_b^2 \right\} \]  
\hspace{1cm} (1.2)

In which, \( Ni_{R_{tp,d}} \) represents the counting of fires per trajectory day returned (i.e. Counting fires per window of 24 hours until complete 96 hours. 4 days). \( f_{x',y',t'} \) means a fire with latitude and longitude inside NSA for a specific time. \( R_b^2 \) is a circumference radius (km) where the central point is \( Pn_{x,y,tp-nh} \). The number of fires not repeated per day, \( N_{f_d} \), is defined by

\[ N_{f_d} = \sum_{n=0-day}^{n=4-th-day} Ni_{R_{tp-day}} \times e^{i/T} \]  
\hspace{1cm} (1.3)

In which, \( e^{i/T} \) is a temporal dilution factor. More distant fires receive a higher dilution.
Results – Chapter 1

Monthly average of number of active fires, PM and Accumulated precipitation (NSA) from 2006 to 2016.

Spearman correlation coefficient (Nf and precipitation) = -0.6 (p-value < 2.2e-16 bilateral).
Linear regressions where the dependent variable is PM$_{10}$ ($\mu g \ m^{-3}$), and there is only one independent variable. Coefficients, $\beta$, standard errors, (s.e), and 95% confidence intervals, C.I. Mixing height, $H_{mix}$ is in units of km. Precipitation, $prec$ is in units of mm. Precipitable Water, $PWAT$ is in units of mm. $Nf$ is number of fires within a 150 km buffer zone of the NCEP/NCAR Reanalysis back-trajectories. Residual standard error, $\epsilon$, and Akaike Information Criterion (AIC) to compare models. Models with smaller AIC values-indicating adequate fit with fewer parameters were preferred.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\beta_0$ (s.e)</th>
<th>$\beta_{Nf}$ (s.e)</th>
<th>$\beta_{prec}$ (s.e)</th>
<th>$\beta_{H_{mix}}$ (s.e)</th>
<th>$\beta_{PWAT}$ (s.e)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C.I.</td>
<td>C.I.</td>
<td>C.I.</td>
<td>C.I.</td>
<td>C.I.</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>$\beta_{Nf}$</td>
<td>$\beta_{prec}$</td>
<td>$\beta_{H_{mix}}$</td>
<td>$\beta_{PWAT}$</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>p-value</td>
<td>p-value</td>
<td>p-value</td>
<td>p-value</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nf</td>
<td>50.4 (0.9)</td>
<td>0.4 (0.1)</td>
<td>1 (0.6)</td>
<td>$R^2 = 0.35$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(48.6, 52.3)</td>
<td>(0.3, 0.5)</td>
<td>&lt;0.001</td>
<td>$\epsilon = 7.2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;0.001</td>
<td>&lt;0.01</td>
<td></td>
<td>$AIC = 580$</td>
<td></td>
</tr>
<tr>
<td>Prec</td>
<td>51.4 (1.7)</td>
<td>-25.6 (4.3)</td>
<td>$R^2 = 0.3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(47.9, 54.7)</td>
<td>(-34, -17)</td>
<td>$\epsilon = 7.4$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;0.001</td>
<td>&lt;0.01</td>
<td></td>
<td>$AIC = 586$</td>
<td></td>
</tr>
<tr>
<td>Hmix</td>
<td>87.6 (8.3)</td>
<td>-1.7 (0.4)</td>
<td>$R^2 = 0.16$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(71.0, 104.2)</td>
<td>(-2.5, -0.8)</td>
<td>$\epsilon = 8.1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;0.001</td>
<td>&lt;0.01</td>
<td></td>
<td>$AIC = 602$</td>
<td></td>
</tr>
</tbody>
</table>
statistical assumptions

• INDEPENDENCE OF ERRORS Durbin-Watson-Test

The significant p-value (p < 5%) suggest autocorrelation. 
*If the data are ordered in some way, you'll get a significant DW test.*

• LINEARITY

Taking into account the line that models the residuals of the predictor against the dependent variable.
Statistical assumptions

- **HOMOSCEDASTICITY** (constant error variance)

  nvcTest - The non-significant p-value (p = 0.8) suggest homoscedasticity.

- **NORMALITY**
statistical considerations

• MULTICOLLINEARITY

It can be detected using a statistic called the variance inflation factor (VIF)

\[ \sqrt{VIF} \] the degree to which the confidence interval for that variable's regression parameter is expanded relative to a model with uncorrelated predictors

\[ \sqrt{VIF} > 2 \] multicollinearity problem

THERE IS NOT MULTICOLLINEARITY

• UNUSUAL OBSERVATIONS

Bonferroni test
Cook distance