Atmospheric models:

Why you should never ever ever trust them, why some people do, and why you will, too.

Evan Couzo ’05

MIT ESD
Massachusetts Institute of Technology
Engineering Systems Division

MIT JOINT PROGRAM
SCIENCE POLICY
GLOBAL CHANGE

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

These are not models

```
Last login: Mon Sep  3 14:03:13 on ttys001
Pythagoras:~ ecouzo
$
```

These are not models

```
Last login: Mon Sep  3 14:03:13 on ttys001
Pythagoras:~ ecouzo
$
```

```
import numpy as np
import pylab
from scipy import stats
import pdb

#################################################################
#####    Get the HCHO data     ##################################
#################################################################
HCHO_filepath = '/Users/ecouzo/Documents/research/EPA_AQS_data/datafiles_symlinks/'
HCHO_file     = file( ''.join((HCHO_filepath, 'HCHO_epaaqs_1hr_2003_2011.txt')), 'rU' )

HCHO_filelines = HCHO_file.readlines()
HCHO_datalines = HCHO_filelines[5:]
HCHO_datalines = [ l.rstrip('\n').split('|') for l in HCHO_datalines ] #split each file
HCHO_datalines = [ l[:13] for l in HCHO_datalines ] # remove empty fields

# Separate HCHO values by monitor
CLIN_HCHO = []
DRPK_HCHO = []

for line in HCHO_datalines:
    if line[4] == '1035': # CLIN is 1035
        date, time, value = line[10], line[11], line[12]
        date = '/'.join(( date[4:6], date[6:8], date[0:4] )) # convert to MM/DD/YYYY
        time = time[:2]
        if value != '':
            CLIN_HCHO.append( (date, time, value) )

    elif line[4] == '1039': # DRPK is 1039
        date, time, value = line[10], line[11], line[12]
        date = '/'.join(( date[4:6], date[6:8], date[0:4] )) # convert to MM/DD/YYYY
        time = time[:2]
        if value != '':
            DRPK_HCHO.append( (date, time, value) )

#################################################################
#####    Incorporate O3 data     ################################
#################################################################
NTOC_file = file('/Users/ecouzo/Documents/research/obs_NTOC_through_2011/data_text_file
CLIN_O3   = file('/Users/ecouzo/Documents/research/obs_NTOC_through_2011/one_hour_raw_0
DRPK_O3   = file('/Users/ecouzo/Documents/research/obs_NTOC_through_2011/one_hour_raw_0

# Prepare the O3 data for plotting
CLIN_O3_filelines = CLIN_O3.readlines()
CLIN_O3_datalines = CLIN_O3_filelines[6:]
CLIN_O3_datalines = [ l.strip('\n').split(',') for l in CLIN_O3_datalines ] #split each
DRPK_O3_filelines = DRPK_O3.readlines()
DRPK_O3_datalines = DRPK_O3_filelines[6:]
DRPK_O3_datalines = [ l.strip('\n').split(',') for l in DRPK_O3_datalines ] #split each

# Make a list of NTOC dates for each monitor
NTOC_filelines = NTOC_file.readlines()
NTOC_datalines = NTOC_filelines[2:-1]
NTOC_datalines = [ l.rstrip('\n').split(',') for l in NTOC_datalines ] #split each file
NTOC_datalines = [ [l[0],l[1].lstrip(' '),l[2].lstrip(' ')] for l in NTOC_datalines ] #

CLIN_NTOC_dates, DRPK_NTOC_dates, HCHV_NTOC_dates = [], [], []
for ntoc in NTOC_datalines:
    date = ntoc[0]
```

These are not models

```
Last login: Mon Sep  3 14:03:13 on ttys001
Pythagoras:~ ecouzo
$
```

```
import numpy as np
import pylab
from scipy import stats
import pdb

#########################################################
#####    Get the HCHO data      #########################
#########################################################
HCHO_filepath = '/Users/ecouzo/Documents/research/EP
HCHO_file     = file( ''.join((HCHO_filepath, 'HCHO_

HCHO_filelines = HCHO_file.readlines()
HCHO_datalines = HCHO_filelines[5:]
HCHO_datalines = [ l.rstrip('\n').split('|') for l i
HCHO_datalines = [ l[:13] for l in HCHO_datalines ]

# Separate HCHO values by monitor
CLIN_HCHO = []
DRPK_HCHO = []

for line in HCHO_datalines:

    if line[4] == '1035': # CLIN is 1035
        date, time, value = line[10], line[11], line
        date = '/'.join(( date[4:6], date[6:8], date
        time = time[:2]
        if value != '':
            CLIN_HCHO.append( (date, time, value) )

    elif line[4] == '1039': # DRPK is 1039
        date, time, value = line[10], line[11], line
        date = '/'.join(( date[4:6], date[6:8], date
        time = time[:2]
        if value != '':
            DRPK_HCHO.append( (date, time, value) )

#########################################################
#####    Incorporate O3 data     ########################
#########################################################
NTOC_file = file('/Users/ecouzo/Documents/research/obs_NTOC_through_2011/data_text_file
CLIN_O3   = file('/Users/ecouzo/Documents/research/obs_NTOC_through_2011/one_hour_raw_O
DRPK_O3   = file('/Users/ecouzo/Documents/research/obs_NTOC_through_2011/one_hour_raw_O

# Prepare the O3 data for plotting
CLIN_O3_filelines = CLIN_O3.readlines()
CLIN_O3_datalines = CLIN_O3_filelines[6:]
CLIN_O3_datalines = [ l.strip('\n').split(',') for l in CLIN_O3_datalines ] #split each
DRPK_O3_filelines = DRPK_O3.readlines()
DRPK_O3_datalines = DRPK_O3_filelines[6:]
DRPK_O3_datalines = [ l.strip('\n').split(',') for l in DRPK_O3_datalines ] #split each

# Make a list of NTOC dates for each monitor
NTOC_filelines = NTOC_file.readlines()
NTOC_datalines = NTOC_filelines[2:-1]
NTOC_datalines = [ l.rstrip('\n').split(',') for l in NTOC_datalines ] #split each file
NTOC_datalines = [ [l[0],l[1].lstrip(' '),l[2].lstrip(' ')] for l in NTOC_datalines ] #

CLIN_NTOC_dates, DRPK_NTOC_dates, HCHV_NTOC_dates = [], [], []
for ntoc in NTOC_datalines:
    date = ntoc[0]
```

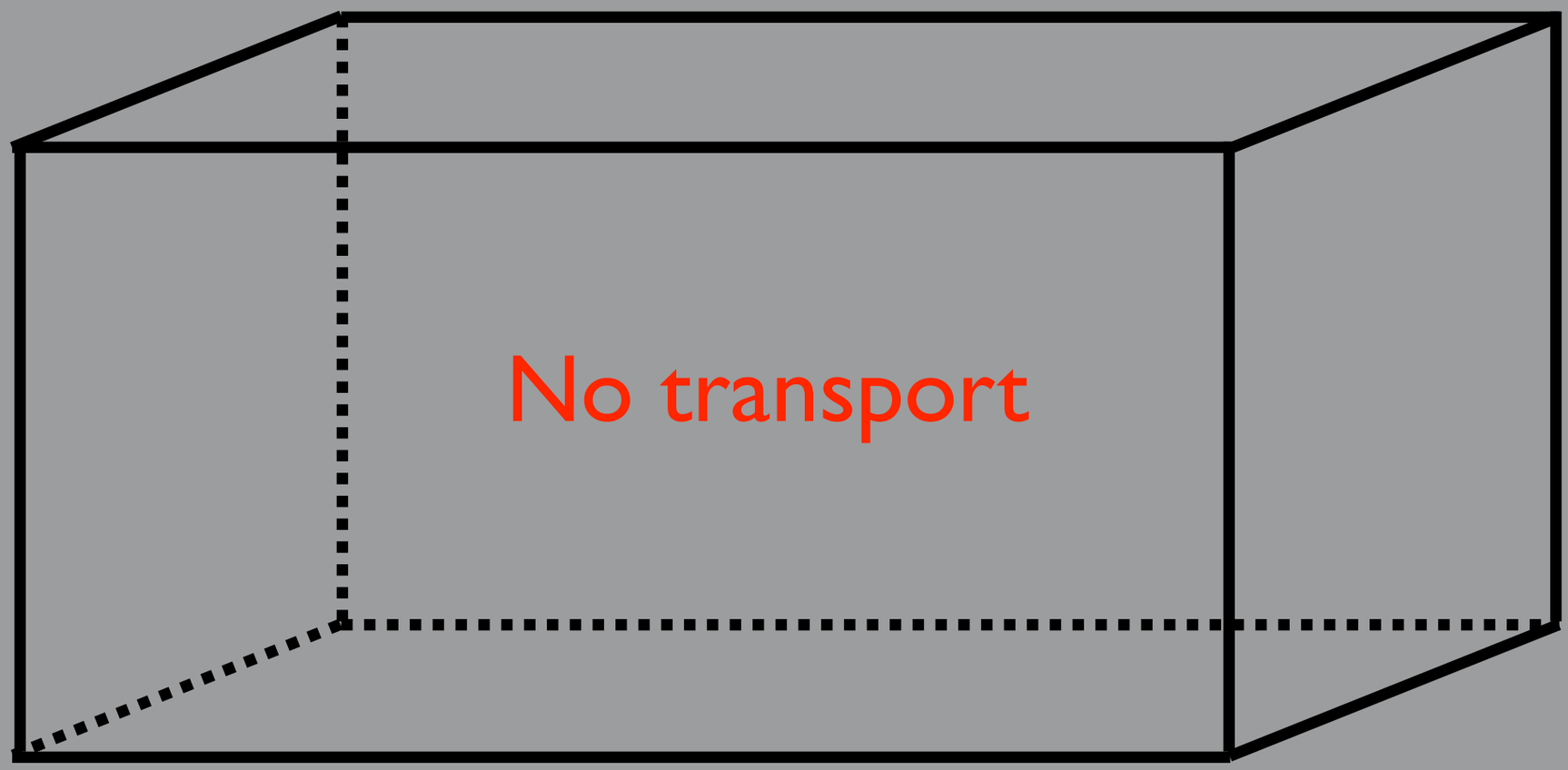
THIS is a model (a chemical transport model)

$$\frac{\partial C}{\partial t} = -\nabla \cdot (C\mathbf{v}) + \nabla \cdot (D\nabla C) + \left.\frac{\partial C}{\partial t}\right|_{\text{Emissions}} + \left.\frac{\partial C}{\partial t}\right|_{\text{Chemistry}} + \left.\frac{\partial C}{\partial t}\right|_{\text{Removal}}$$

Models describe how quantities (e.g. mass, momentum) are transported and transferred.

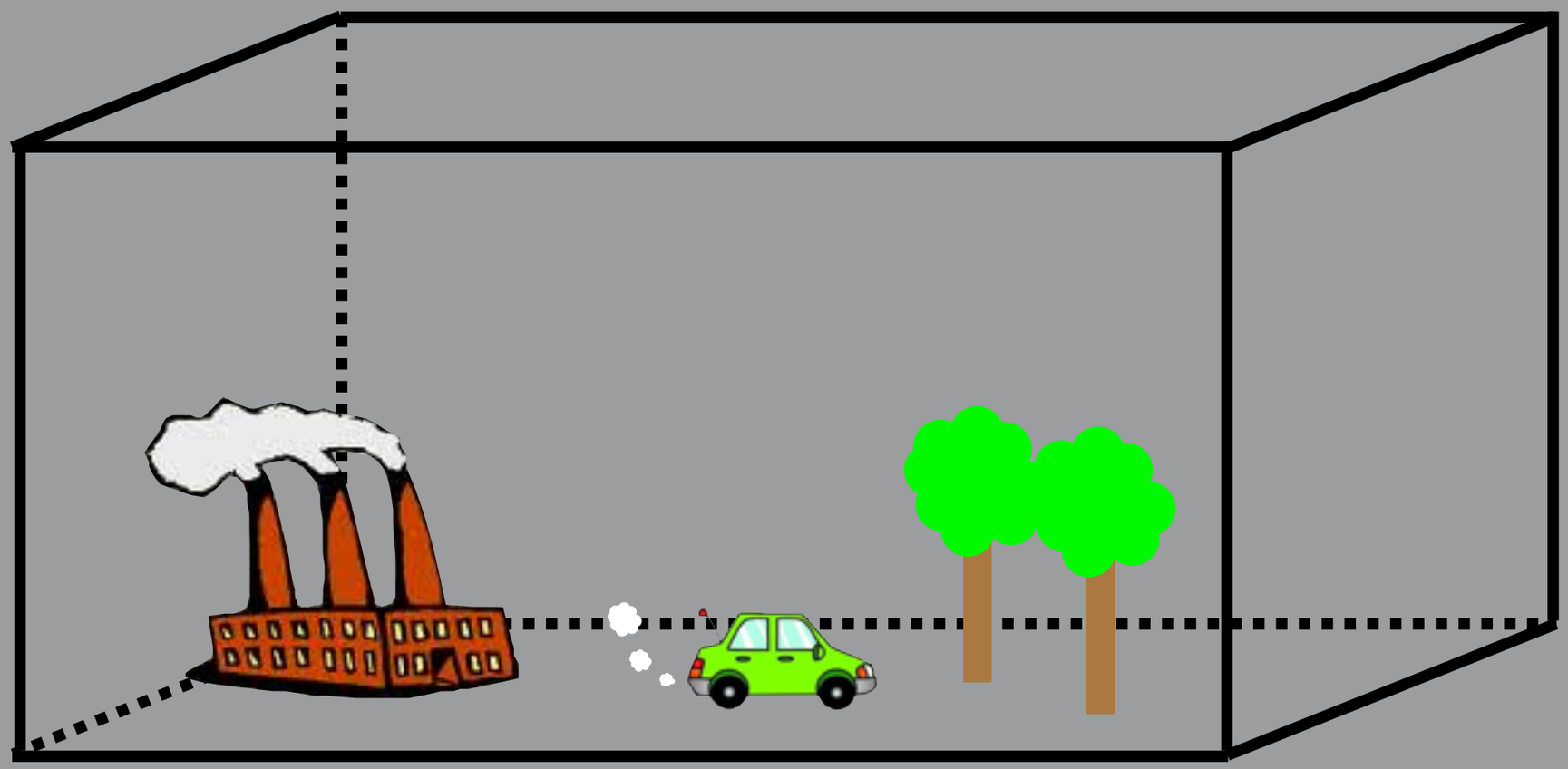
3-D box models

$$\frac{\partial C}{\partial t} = -\nabla \cdot (C\mathbf{v}) + \nabla \cdot (D\nabla C) + \left.\frac{\partial C}{\partial t}\right|_{\text{Emissions}} + \left.\frac{\partial C}{\partial t}\right|_{\text{Chemistry}} + \left.\frac{\partial C}{\partial t}\right|_{\text{Removal}}$$



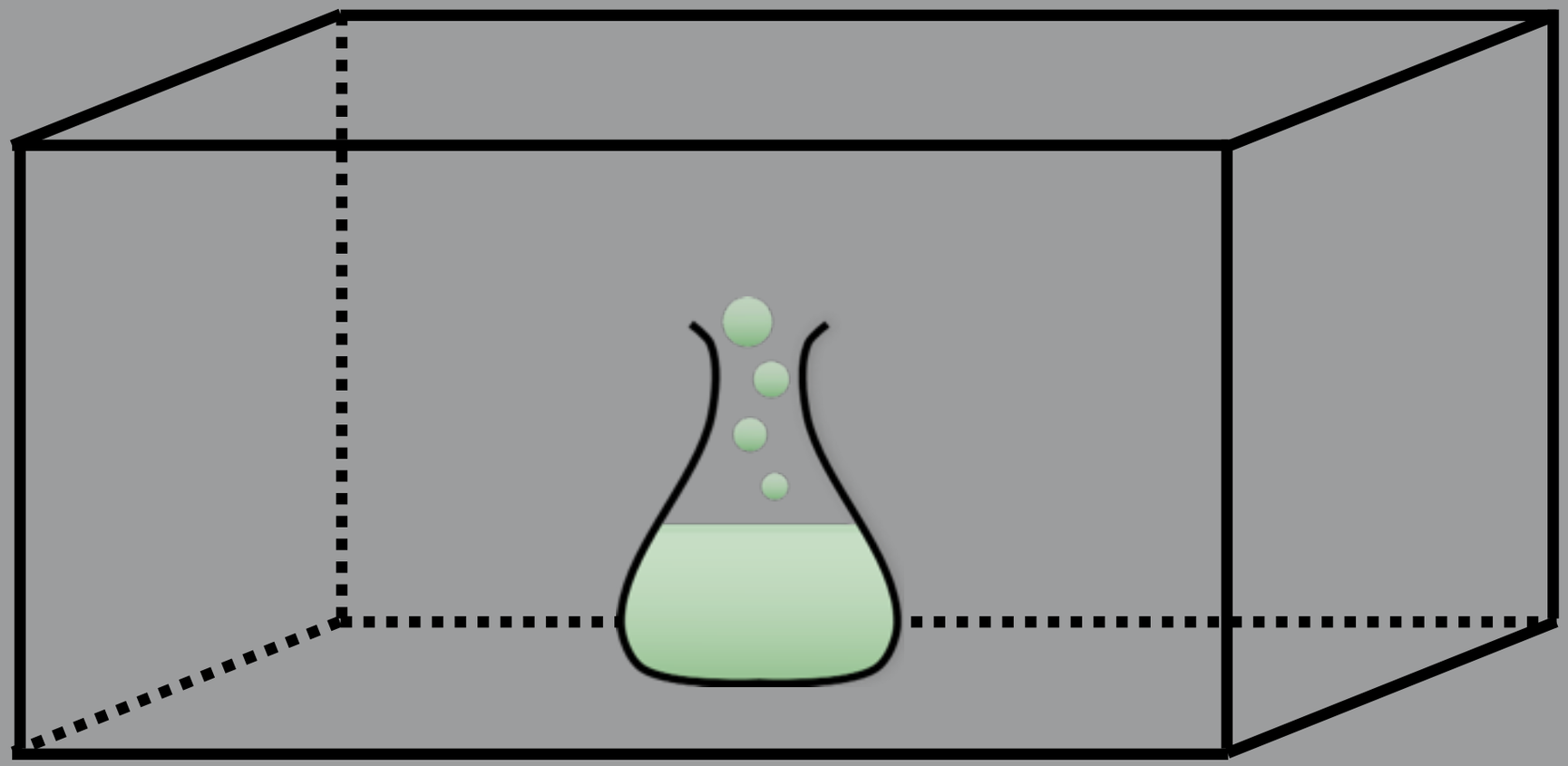
3-D box models

$$\frac{\partial C}{\partial t} = -\nabla \cdot (C\mathbf{v}) + \nabla \cdot (D\nabla C) + \left.\frac{\partial C}{\partial t}\right|_{\text{Emissions}} + \left.\frac{\partial C}{\partial t}\right|_{\text{Chemistry}} + \left.\frac{\partial C}{\partial t}\right|_{\text{Removal}}$$



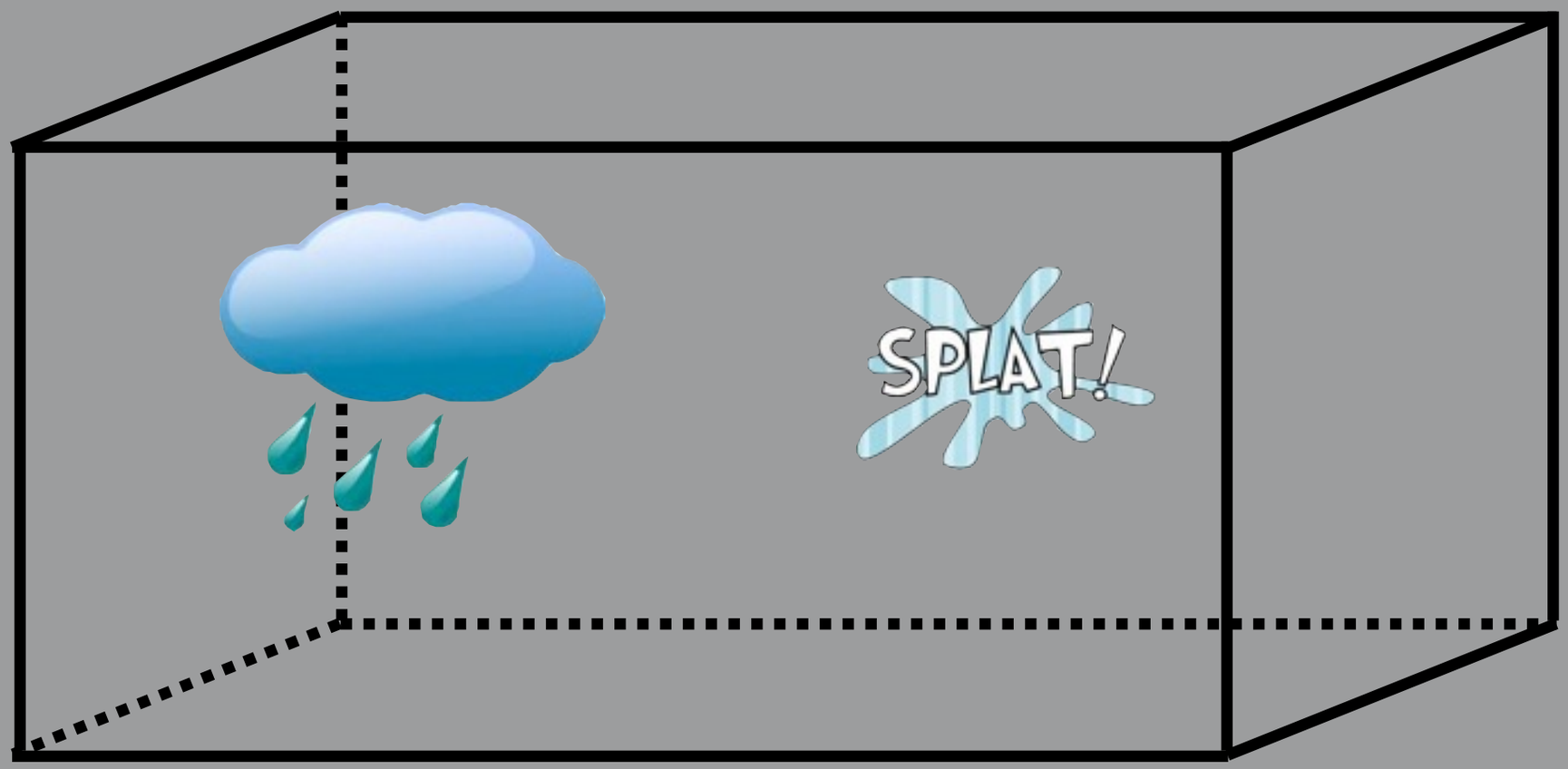
3-D box models

$$\frac{\partial C}{\partial t} = \nabla \cdot (D \nabla C) + \left.\frac{\partial C}{\partial t}\right|_{\text{Emissions}} + \left.\frac{\partial C}{\partial t}\right|_{\text{Chemistry}} + \left.\frac{\partial C}{\partial t}\right|_{\text{Removal}}$$

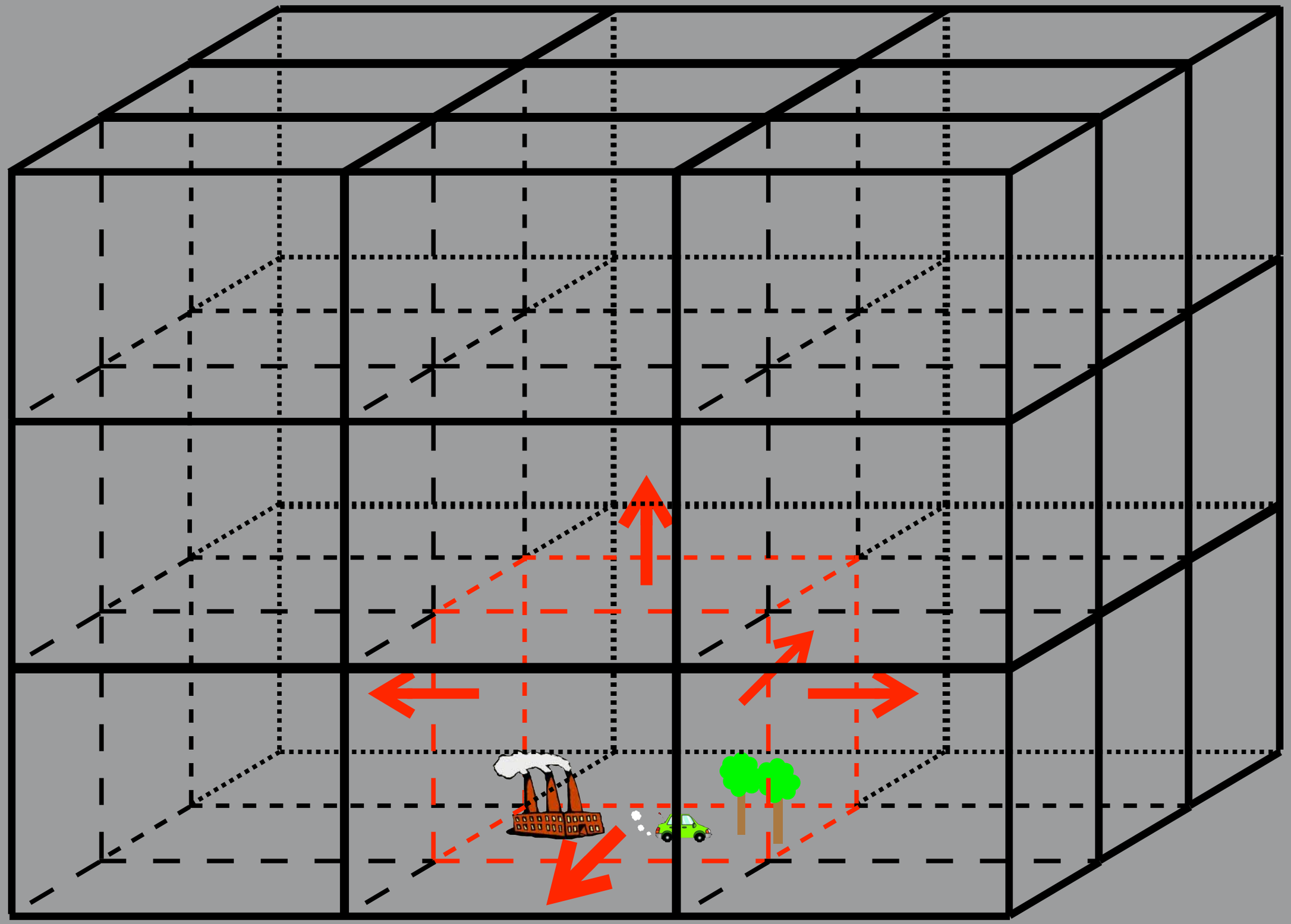


3-D box models

$$\frac{\partial C}{\partial t} = -\nabla \cdot (C\mathbf{v}) + \nabla \cdot (D\nabla C) + \left.\frac{\partial C}{\partial t}\right|_{\text{Emissions}} + \left.\frac{\partial C}{\partial t}\right|_{\text{Chemistry}} + \left.\frac{\partial C}{\partial t}\right|_{\text{Removal}}$$



3-D models simply connect lots of individual box models

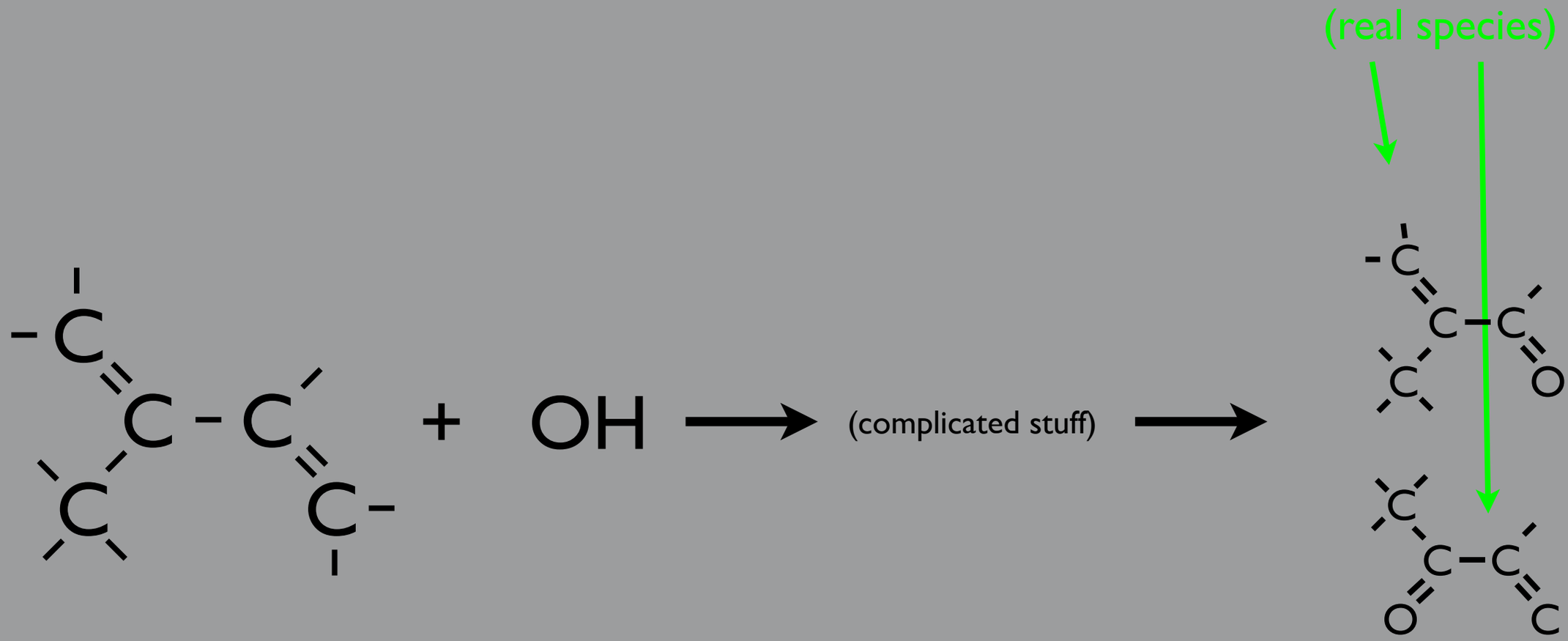


Model formulation introduces error

- No model is identical to an open environmental system.
- Models are always simplifications of the real world.
- Inputs are uncertain; parameters are subject to change.
- Even the best models can't be perfect.
- How does model formulation introduce error?

Treating different entities or processes as if they were the same for the sake of simplifying the description

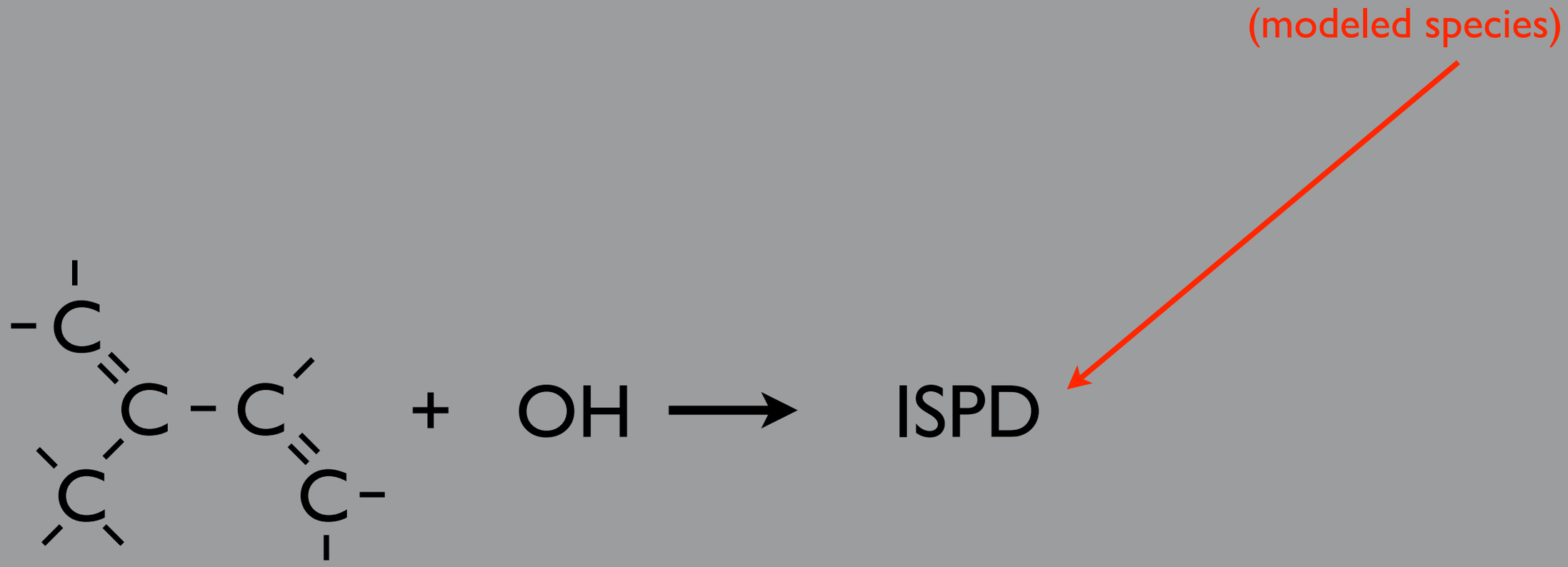
e.g. the concept of a "condensed" chemical mechanism



adapted from Harvey Jeffries, Current Status and Significant Issues in Chemical Mechanisms. International Conference on Chemical Mechanisms, 2006.

Treating different entities or processes as if they were the same for the sake of simplifying the description

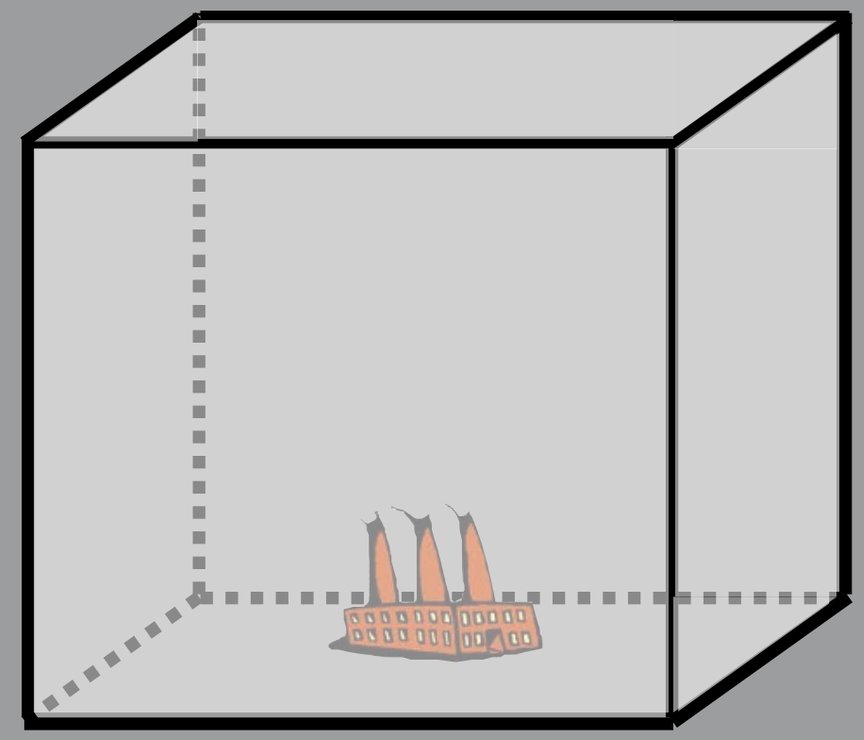
e.g. the concept of a "condensed" chemical mechanism



adapted from Harvey Jeffries, Current Status and Significant Issues in Chemical Mechanisms. International Conference on Chemical Mechanisms, 2006.

Changing the representation of entities or processes for the sake of simplifying the description

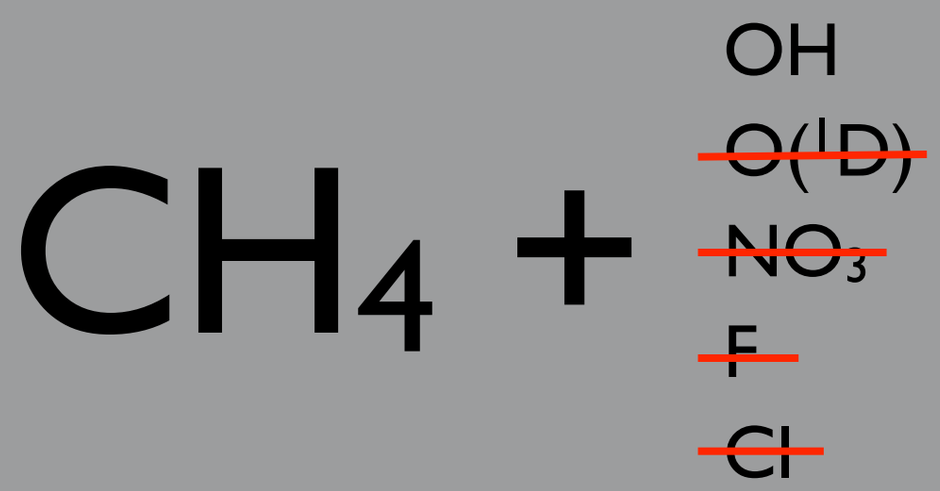
e.g. discretization of the atmosphere



adapted from Harvey Jeffries, Current Status and Significant Issues in Chemical Mechanisms. International Conference on Chemical Mechanisms, 2006.

Omitting entities or processes for the sake of simplifying the description

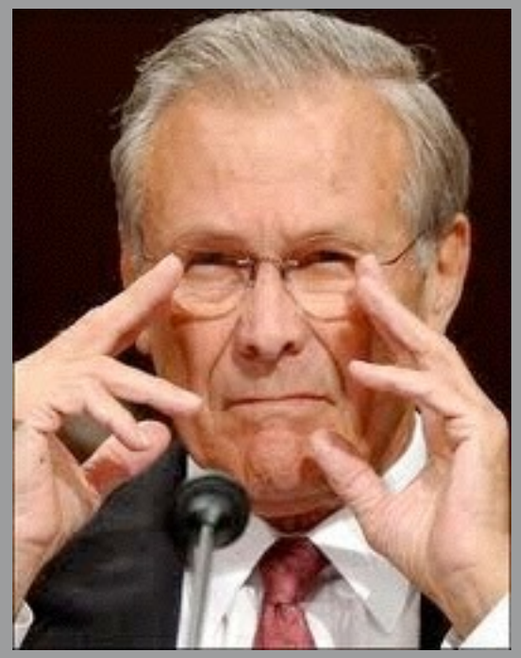
e.g. chemical reaction pathways



adapted from Harvey Jeffries, Current Status and Significant Issues in Chemical Mechanisms. International Conference on Chemical Mechanisms, 2006.

Unintentionally omitted entities or processes because of lack of knowledge

e.g. Isoprene oxidation can lead to secondary organic aerosol formation, but we used to not know that. Until Jason Surratt and others discovered the pathway, we didn't know we didn't know.



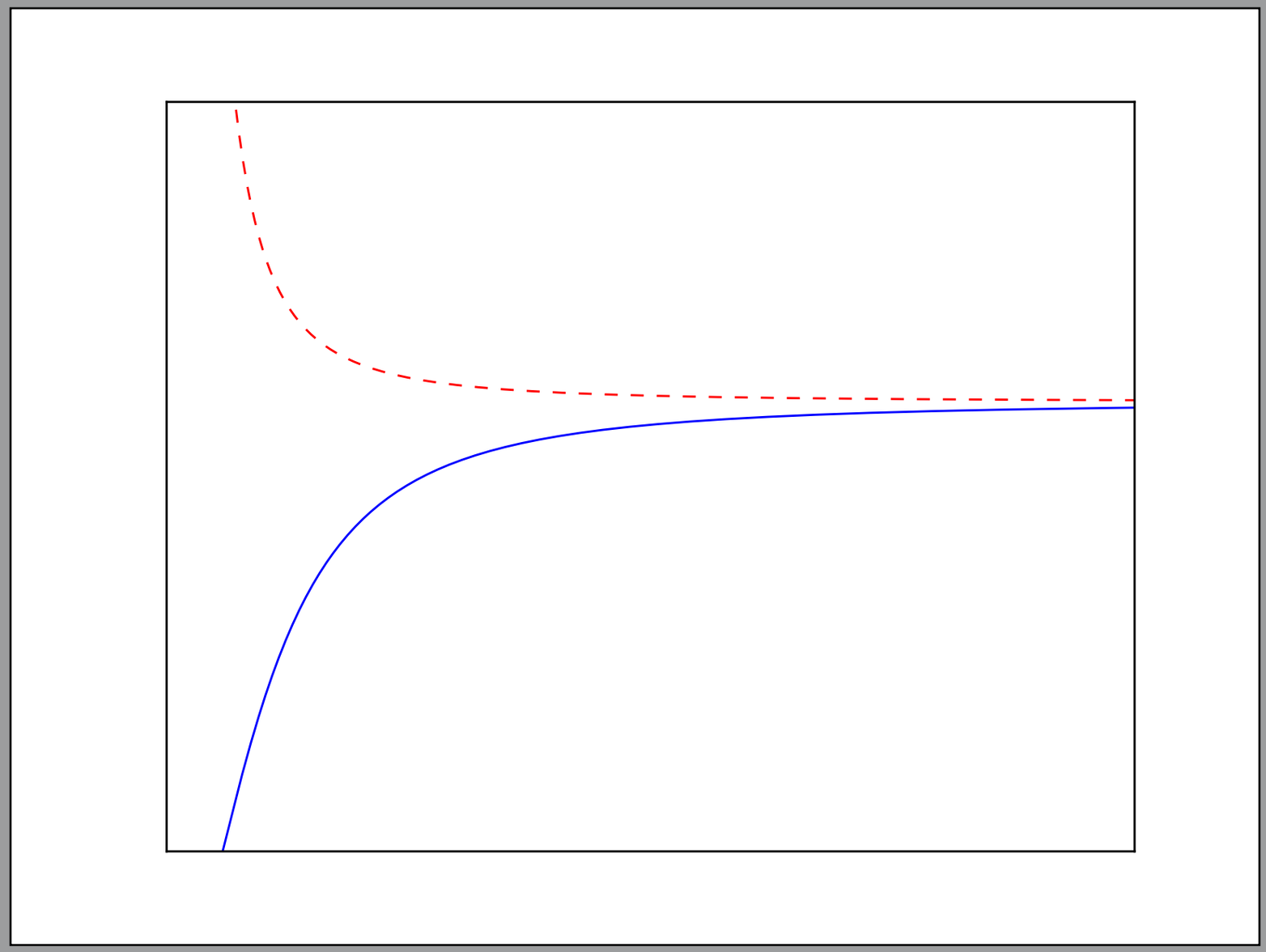
"But there are also unknown unknowns. There are things we do not know we don't know."

adapted from Harvey Jeffries, Current Status and Significant Issues in Chemical Mechanisms. International Conference on Chemical Mechanisms, 2006.

Semantic errors: verification and validation

- Models cannot be verified or validated.
- verify - to establish the truth of
  - Models say nothing about truth, and to say an environmental model has been verified is to claim that an inherently open system has been closed.
- validate - to establish the legitimacy of
  - Models that have no known flaws or inconsistencies might be valid, but this does not suggest the model is an accurate representation of reality.
- evaluate - to determine the value of
  - Model evaluation, usually by comparison to a set of observations, says nothing about the model. Rather, to evaluate a model is to test whether its predictions are acceptably similar to a historical observational record.
- But...we only *really* know something when the model fails.

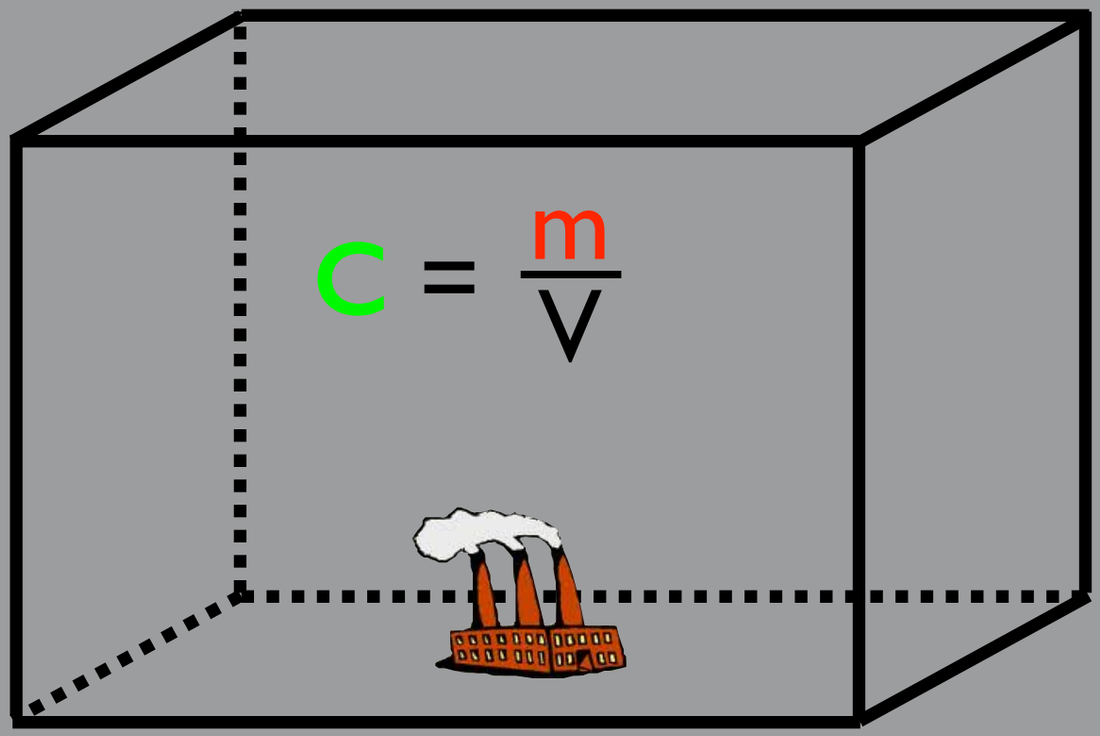
Equifinality - *equal ends*

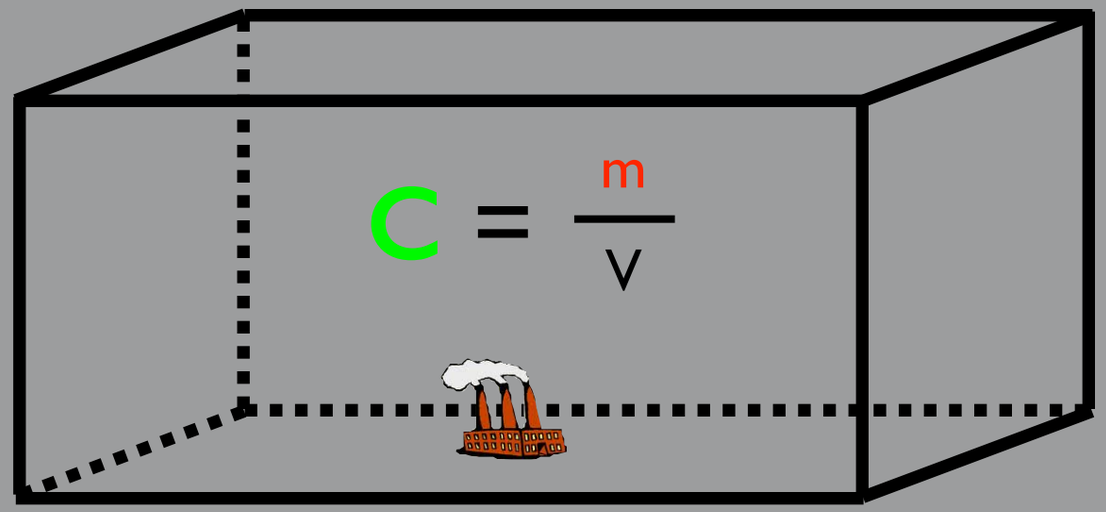


- Different model structure or parameters can lead to non-unique solutions.
- Fitting measurement data isn't good enough.
- "If two theories (or model realizations) are empirically equivalent, then there is no way to choose between them..."

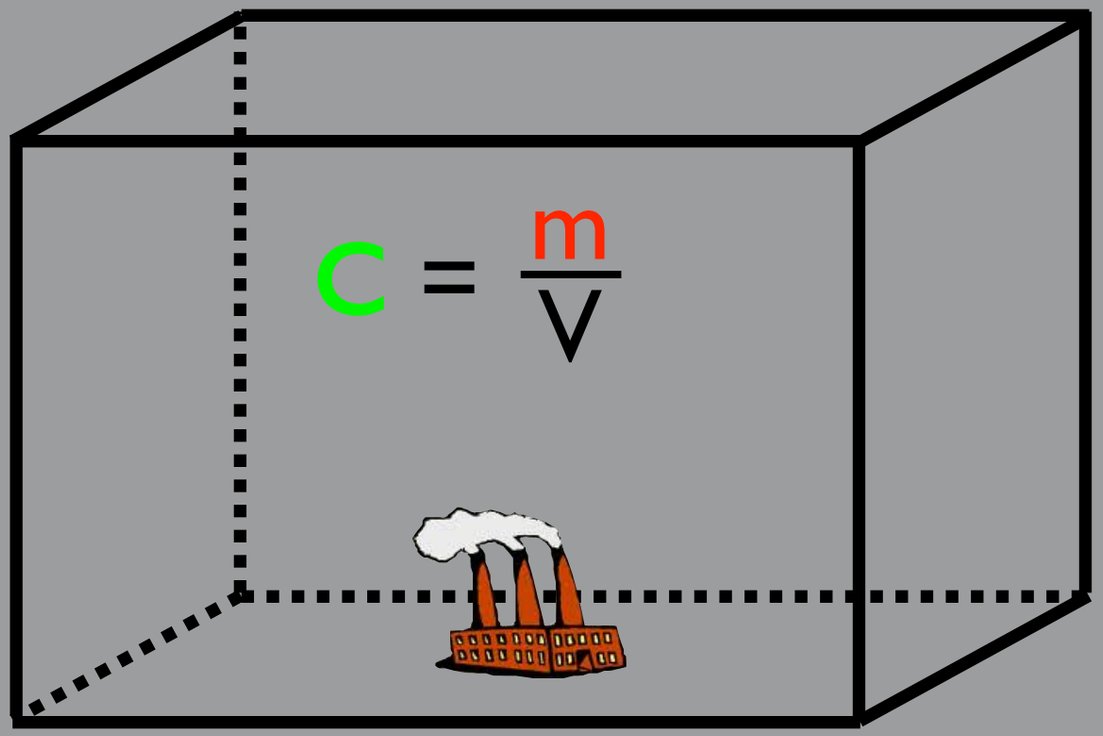
-Oreskes et al., Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science*, Vol. 263, Issue 5147, 641-646, 1994.

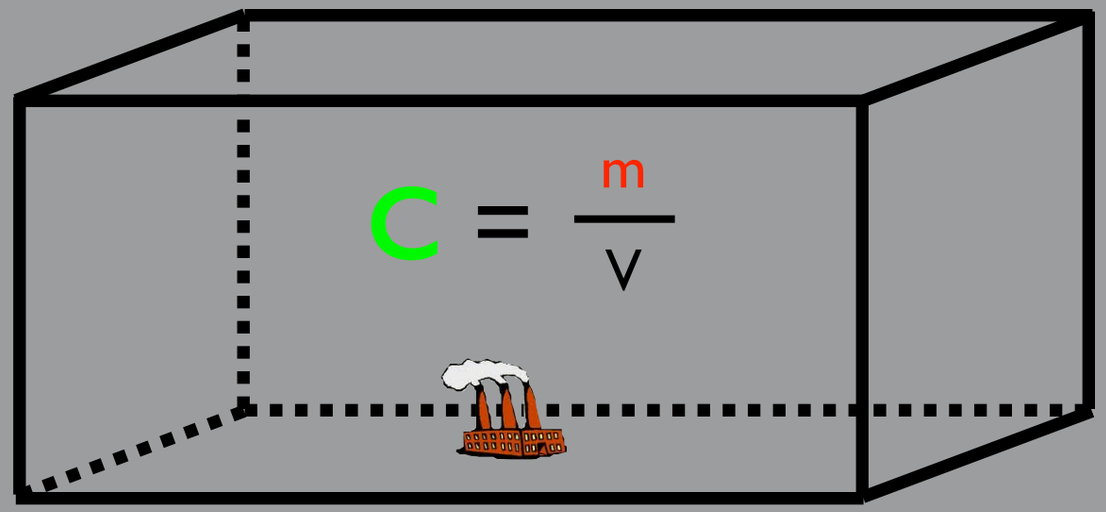
Compensating errors - truly terrifying for model evaluators

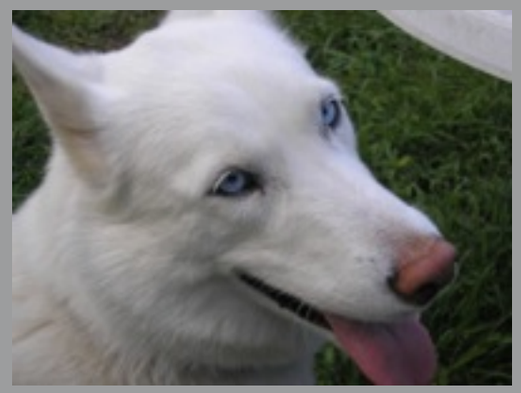




Compensating errors - truly terrifying for model evaluators







So what good are models, anyway?

archive:

a succinctly encoded archive of contemporary knowledge

- reaction kinetics
- physical processes

predict:

communicate:

explore:

from Beck, B. Model evaluation and performance. Encyclopedia of Environmetrics, Vol. 3, 1275-1279. Chichester: John Wiley & Sons, 2002.

So what good are models, anyway?

archive:

a succinctly encoded archive of contemporary knowledge

- reaction kinetics
- physical processes

predict:

an instrument of prediction in support of making a decision or formulating a policy

- pollution controls
- economic growth

communicate:

explore:

from Beck, B. Model evaluation and performance. Encyclopedia of Environmetrics, Vol. 3, 1275-1279. Chichester: John Wiley & Sons, 2002.

So what good are models, anyway?

archive:

a succinctly encoded archive of contemporary knowledge

- reaction kinetics
- physical processes

predict:

an instrument of prediction in support of making a decision or formulating a policy

- pollution controls
- economic growth

communicate:

a device for communicating scientific notions to a scientifically lay audience

- the extent of O_3 pollution
- possible solutions

explore:

from Beck, B. Model evaluation and performance. Encyclopedia of Environmetrics, Vol. 3, 1275-1279. Chichester: John Wiley & Sons, 2002.

So what good are models, anyway?

archive:

a succinctly encoded archive of contemporary knowledge

- reaction kinetics
- physical processes

predict:

an instrument of prediction in support of making a decision or formulating a policy

- pollution controls
- economic growth

communicate:

a device for communicating scientific notions to a scientifically lay audience

- the extent of O_3 pollution
- possible solutions

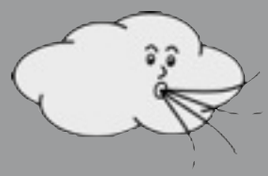
explore:

an exploratory vehicle for discovery of our ignorance

- unknown chemical species and reactions
- missing or underestimated emissions sources

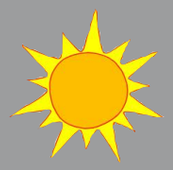
from Beck, B. Model evaluation and performance. Encyclopedia of Environmetrics, Vol. 3, 1275-1279. Chichester: John Wiley & Sons, 2002.

Models also help solve mind bogglingly complex problems.



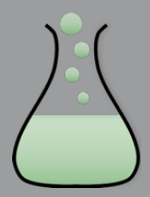
Meteorology

Geography



Photolysis

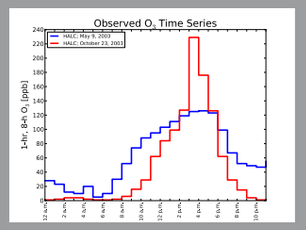
Emissions

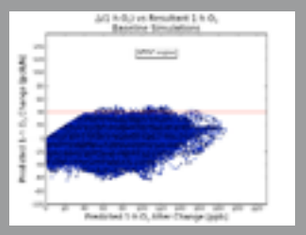


Chemistry

CAMx
("the" model)

Comprehensive Air Quality Model with Extensions (CAMx)



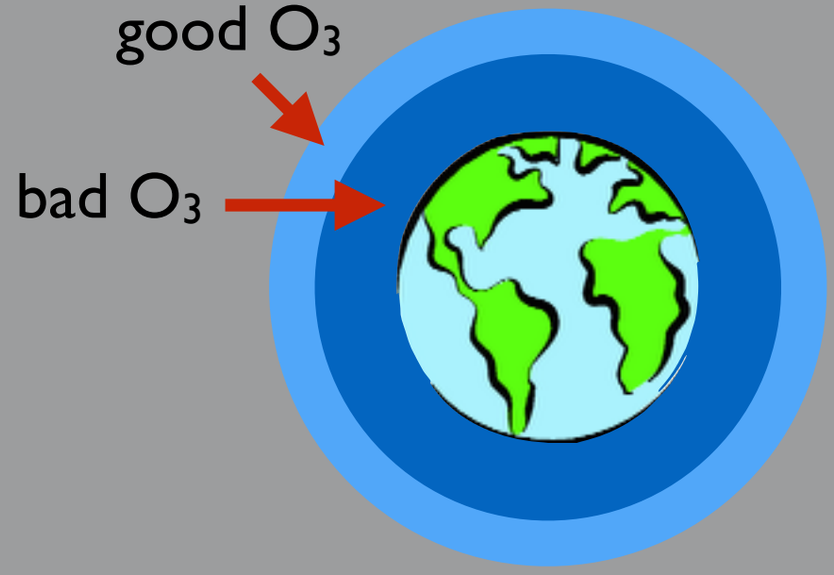


A model of a model of models.

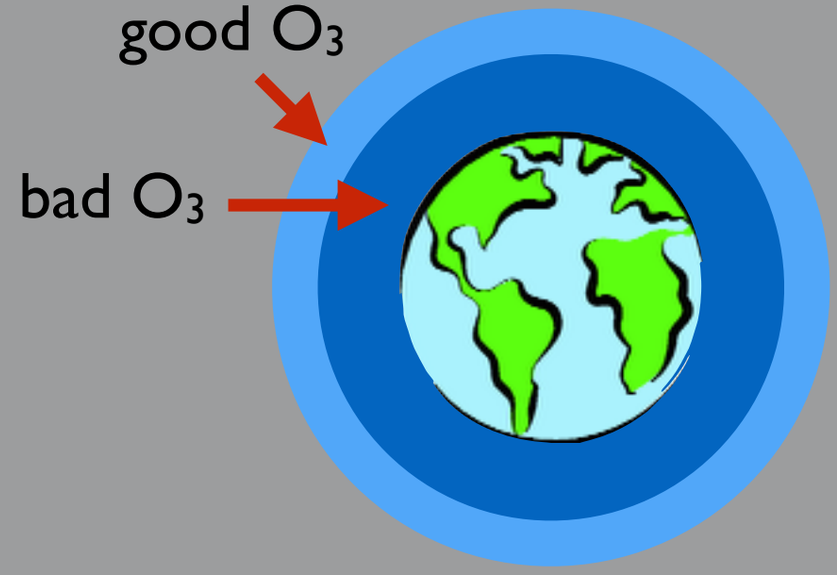
How to make ozone

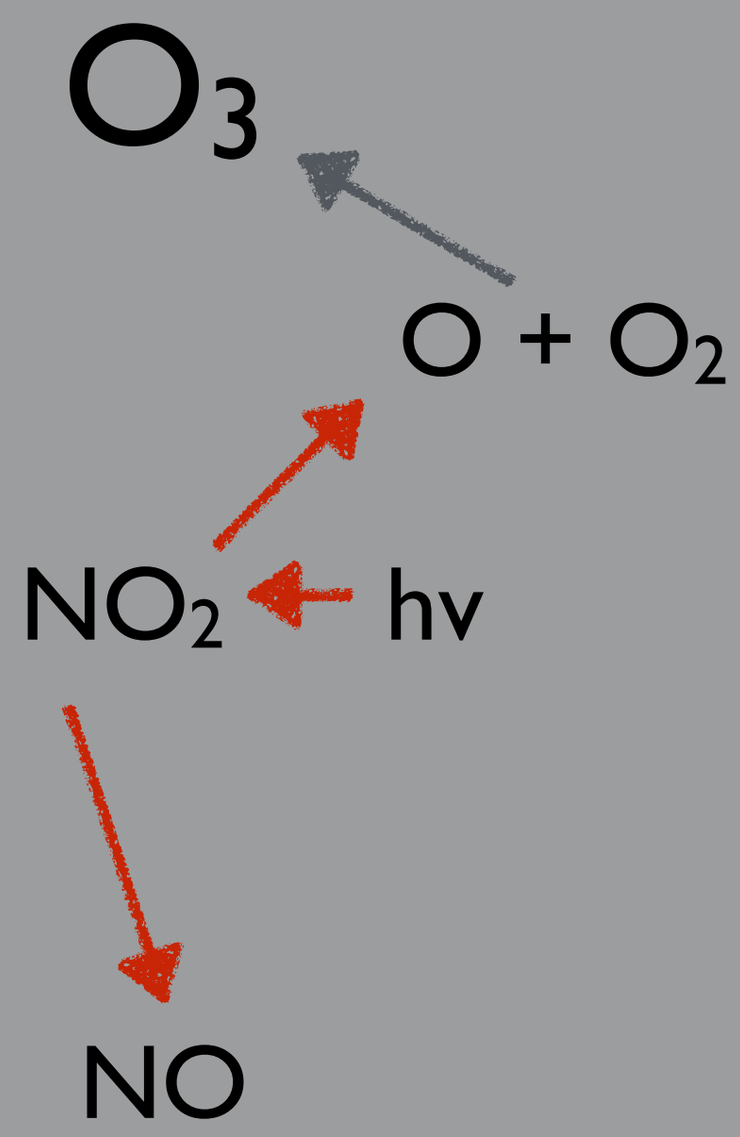
$$O_3$$

$$O + O_2$$



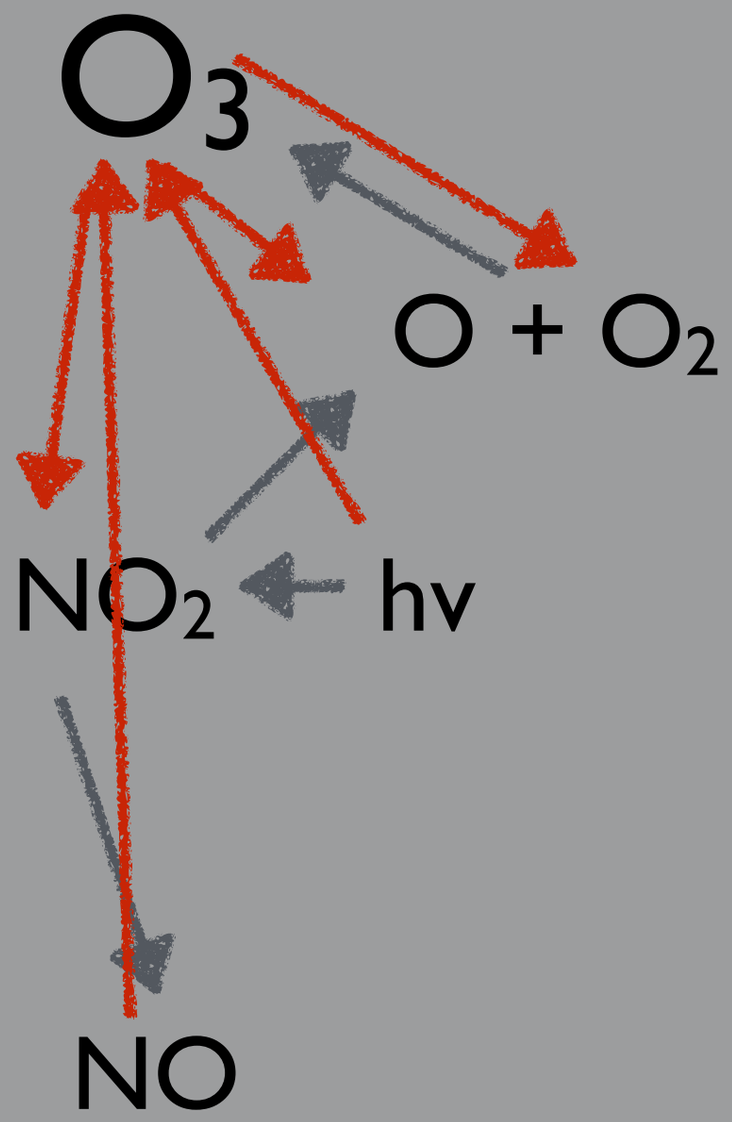
How to make ozone

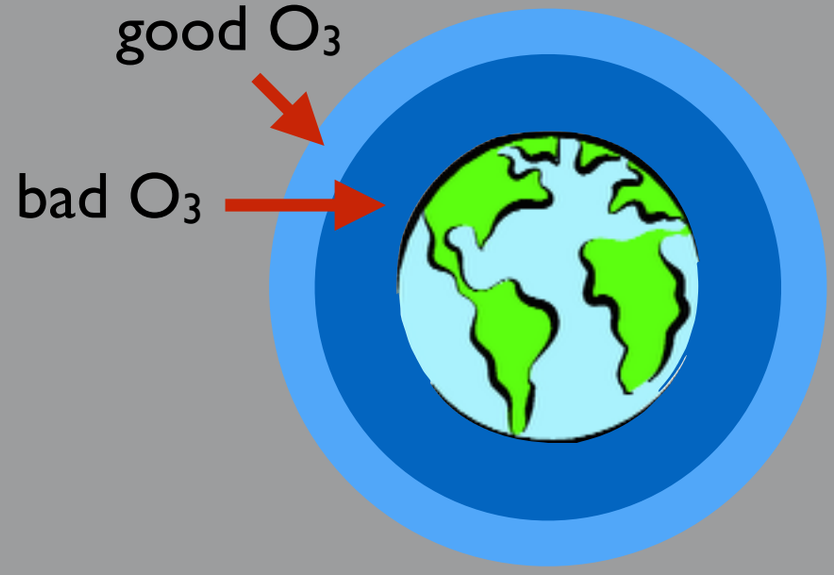




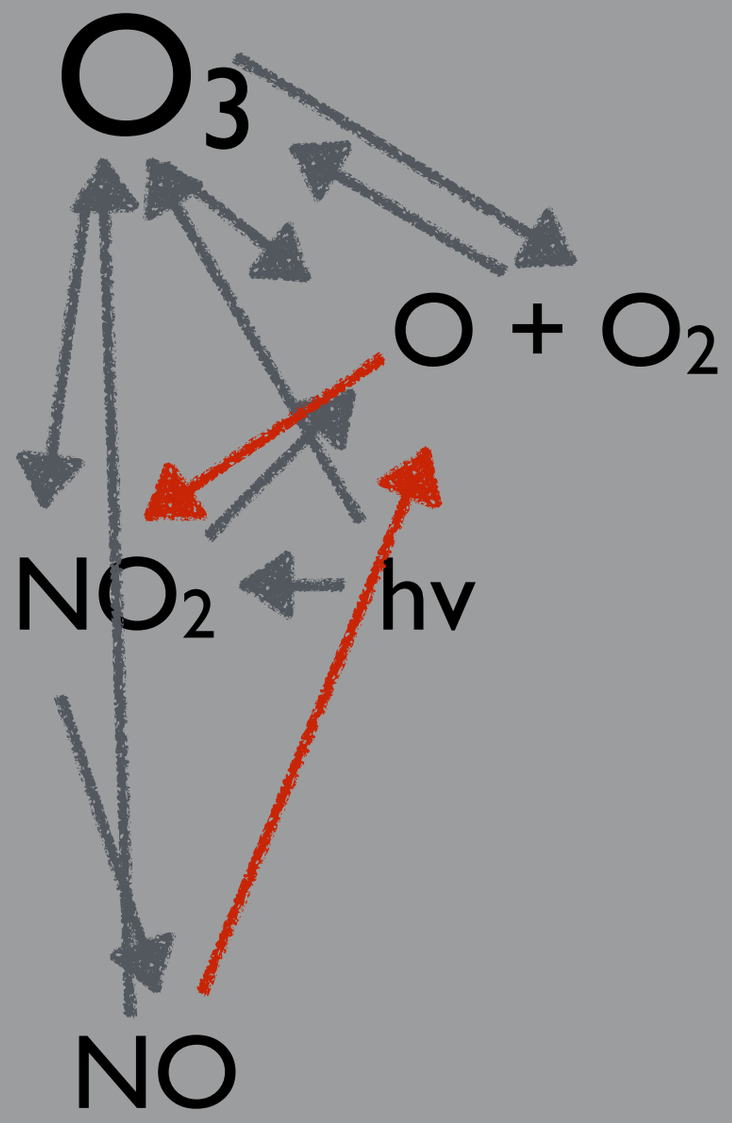
*NO and NO_2, collectively known as NOx, are emitted during combustion.

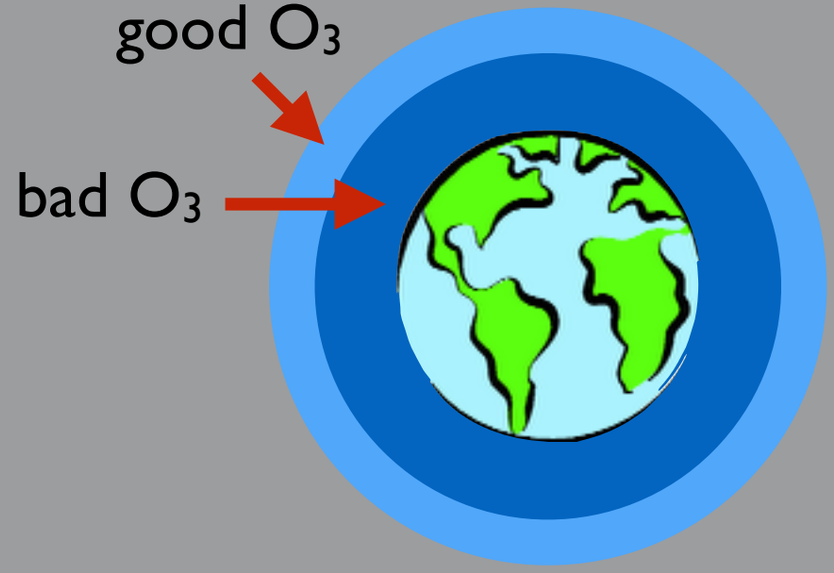
How to make ozone



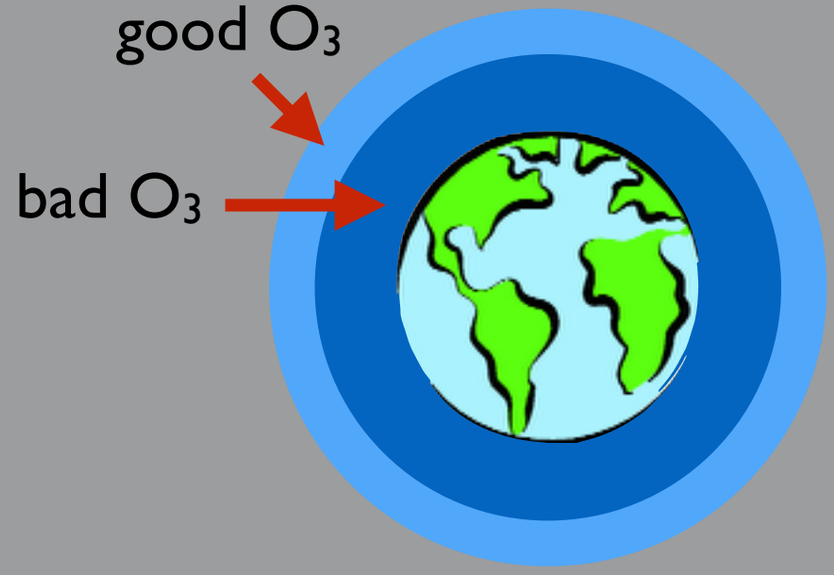


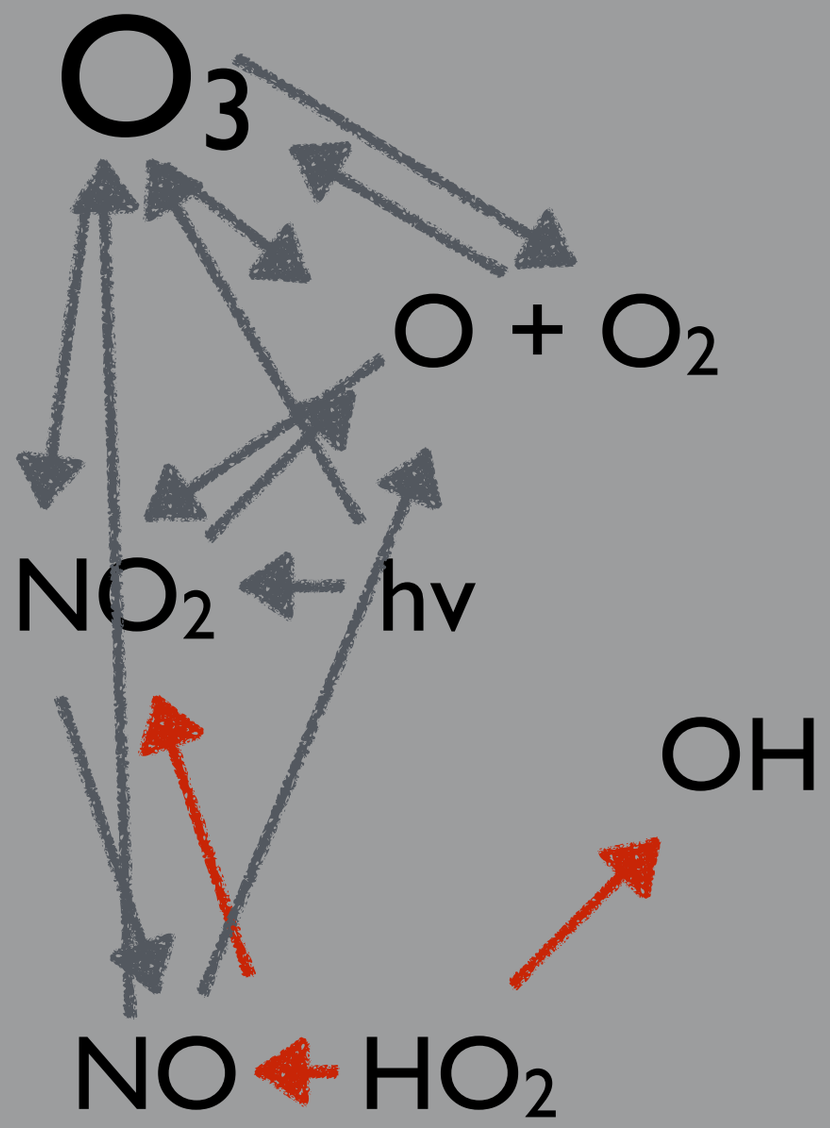
How to make ozone



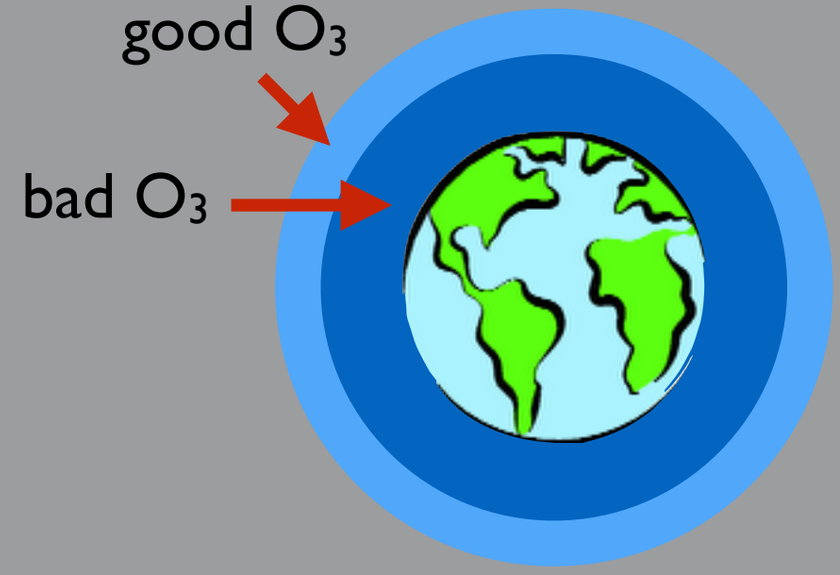


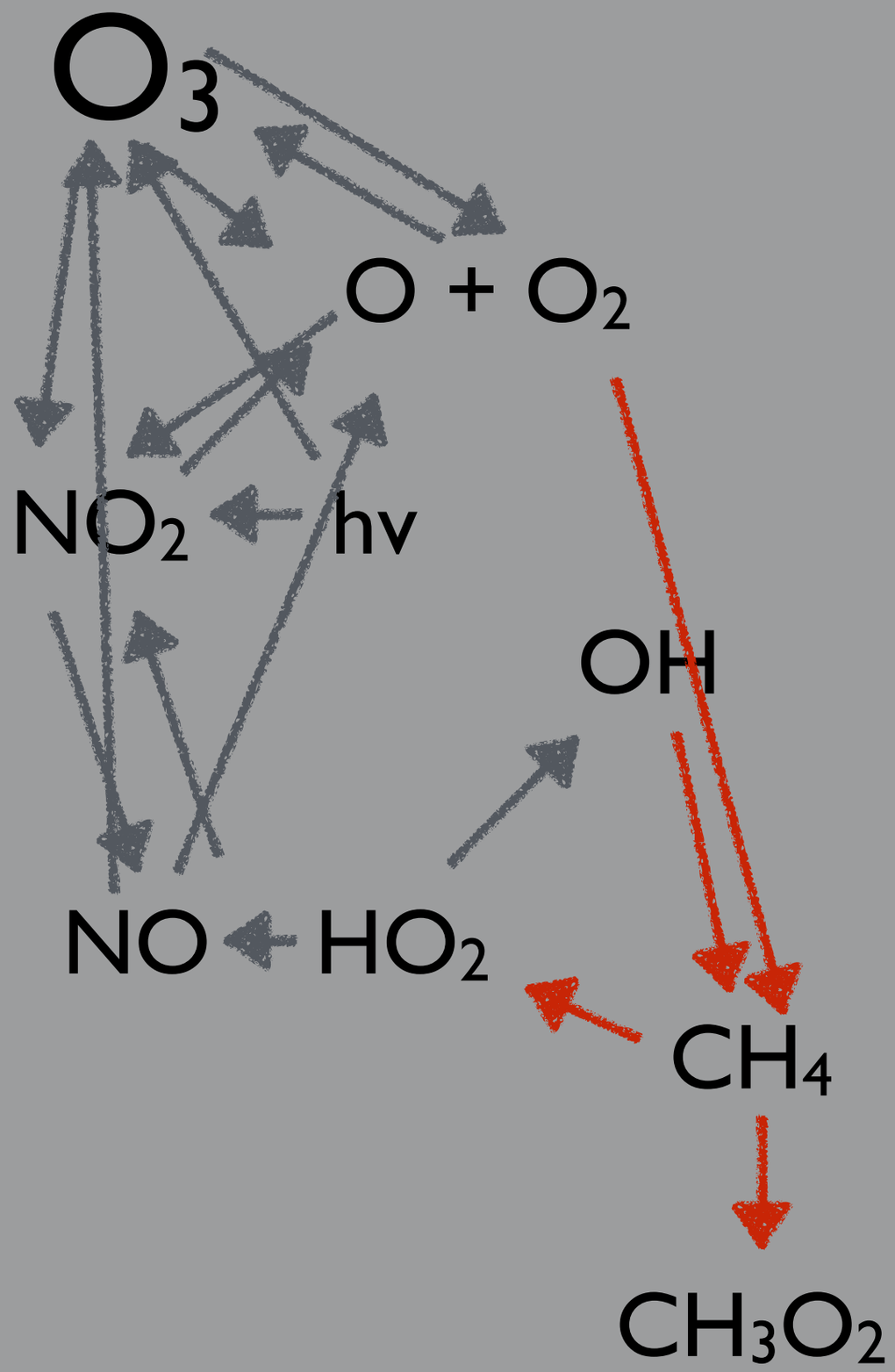
How to make ozone



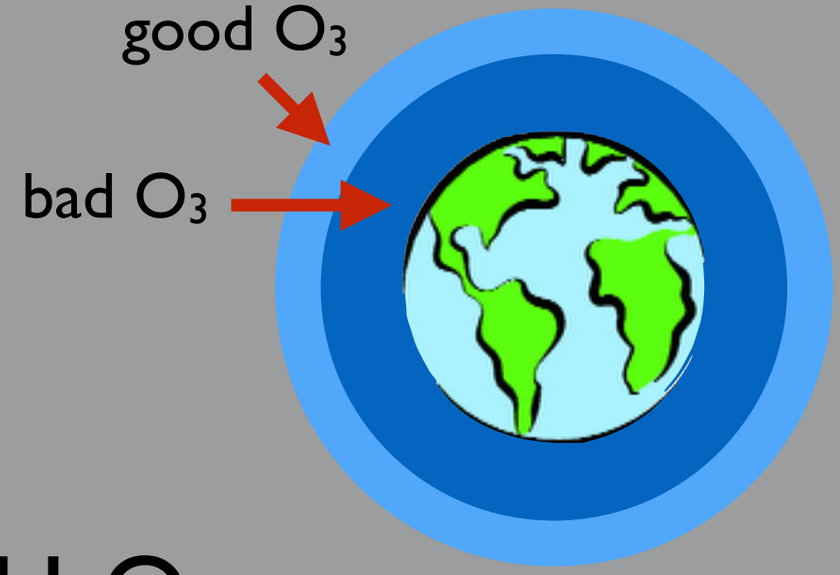


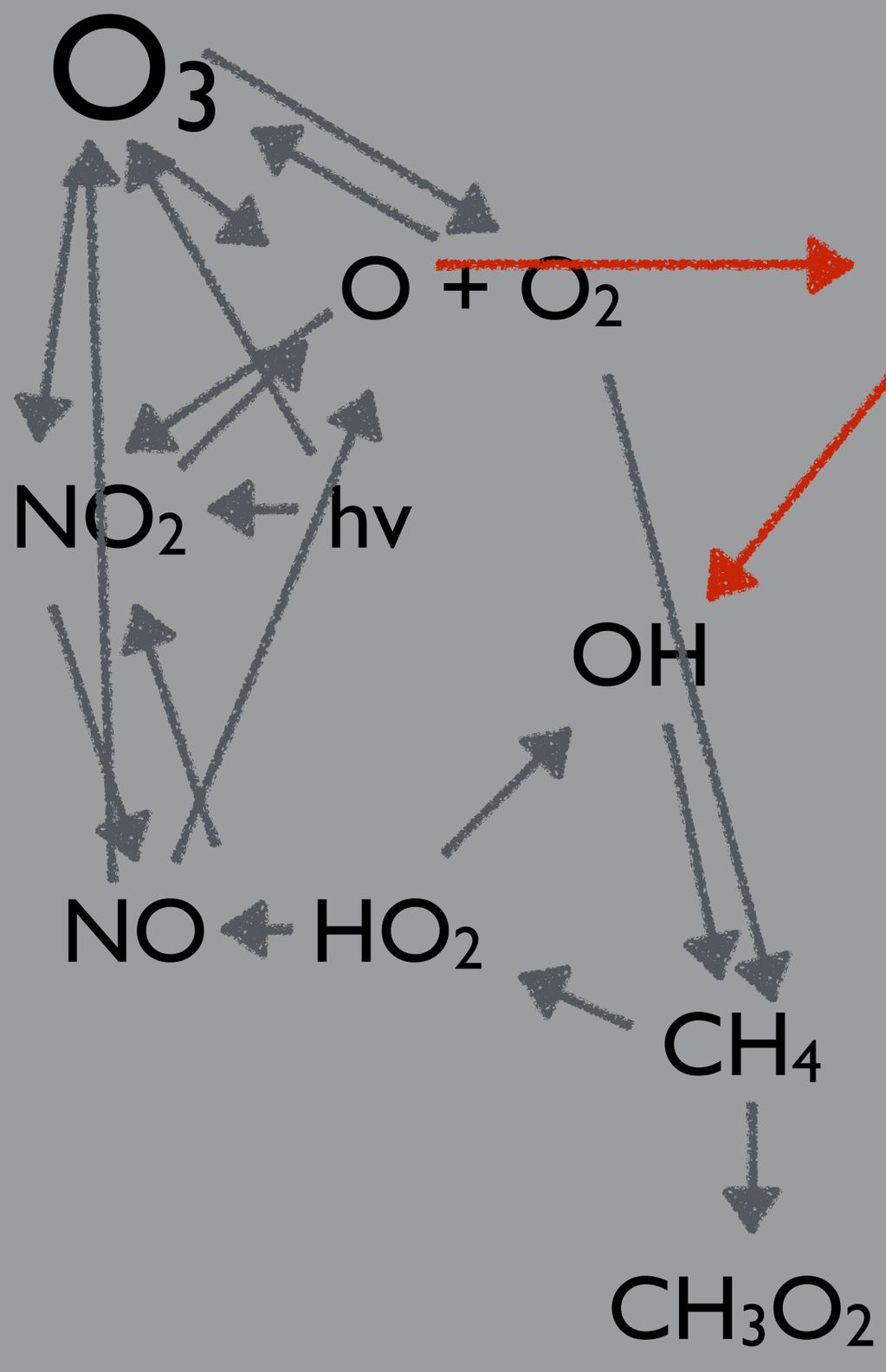
How to make ozone



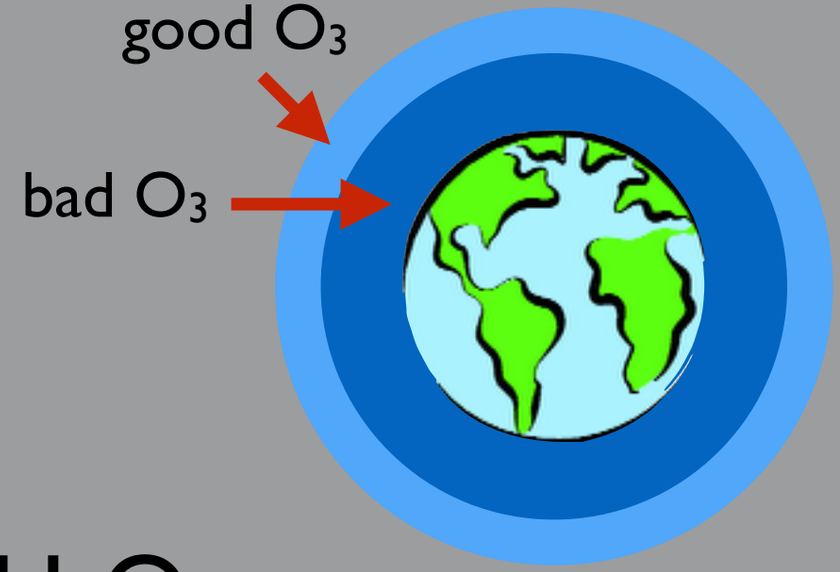


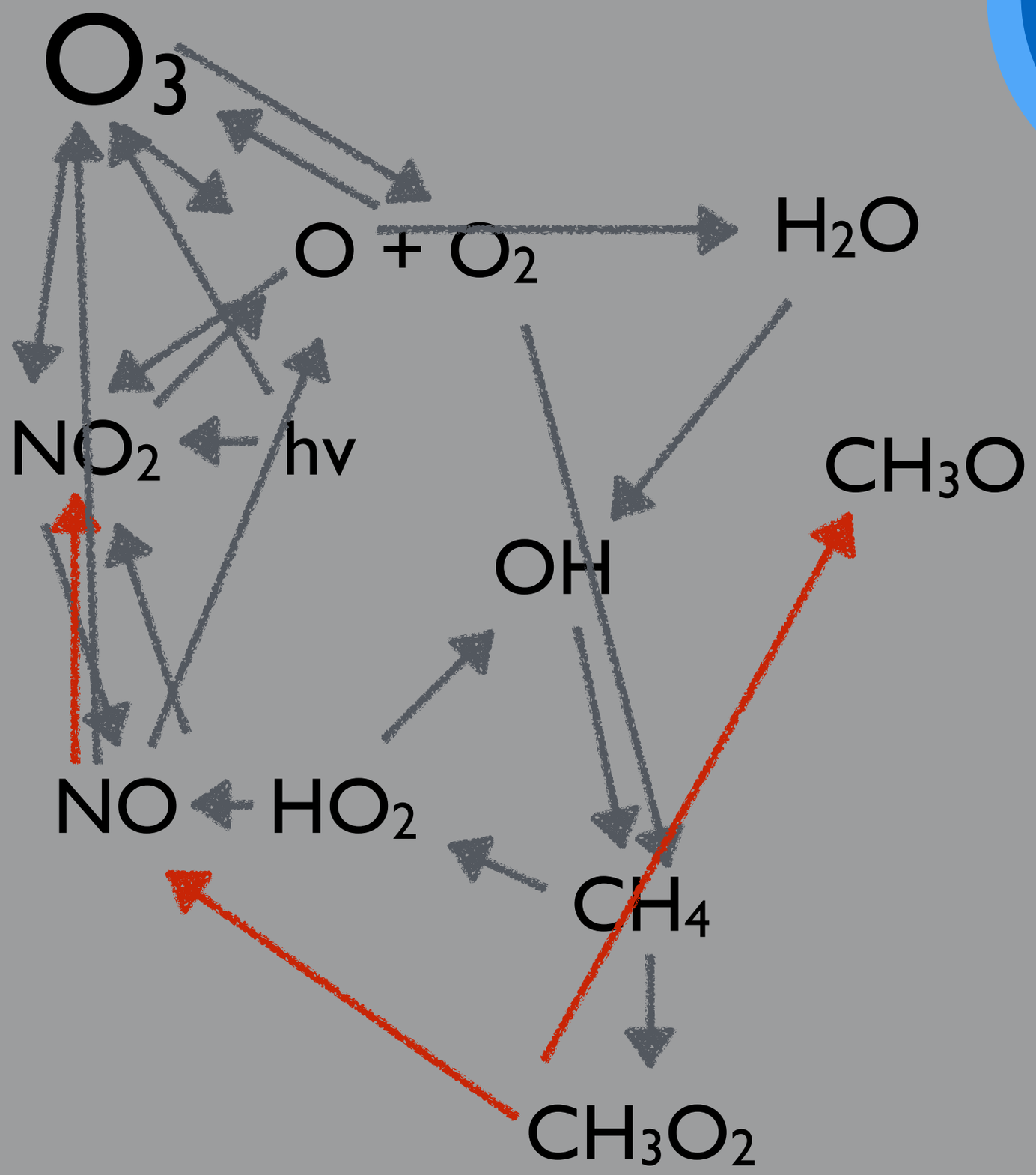
How to make ozone



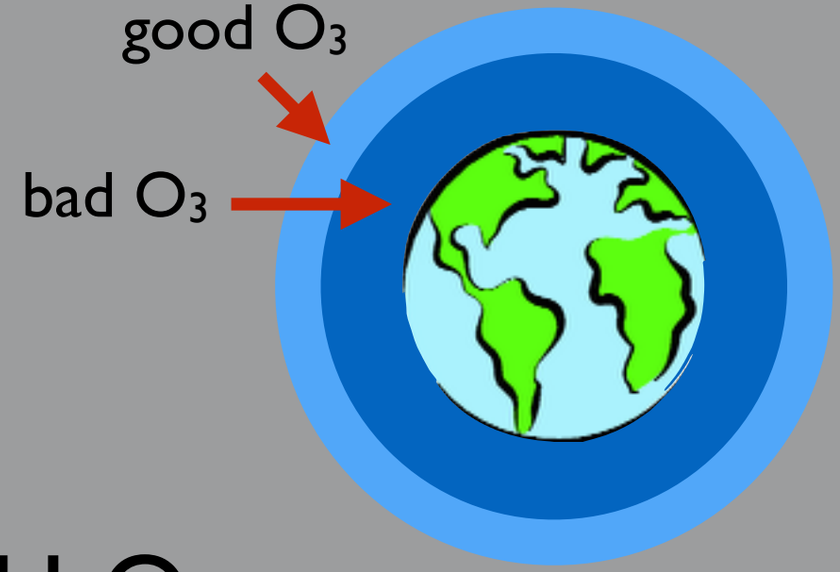


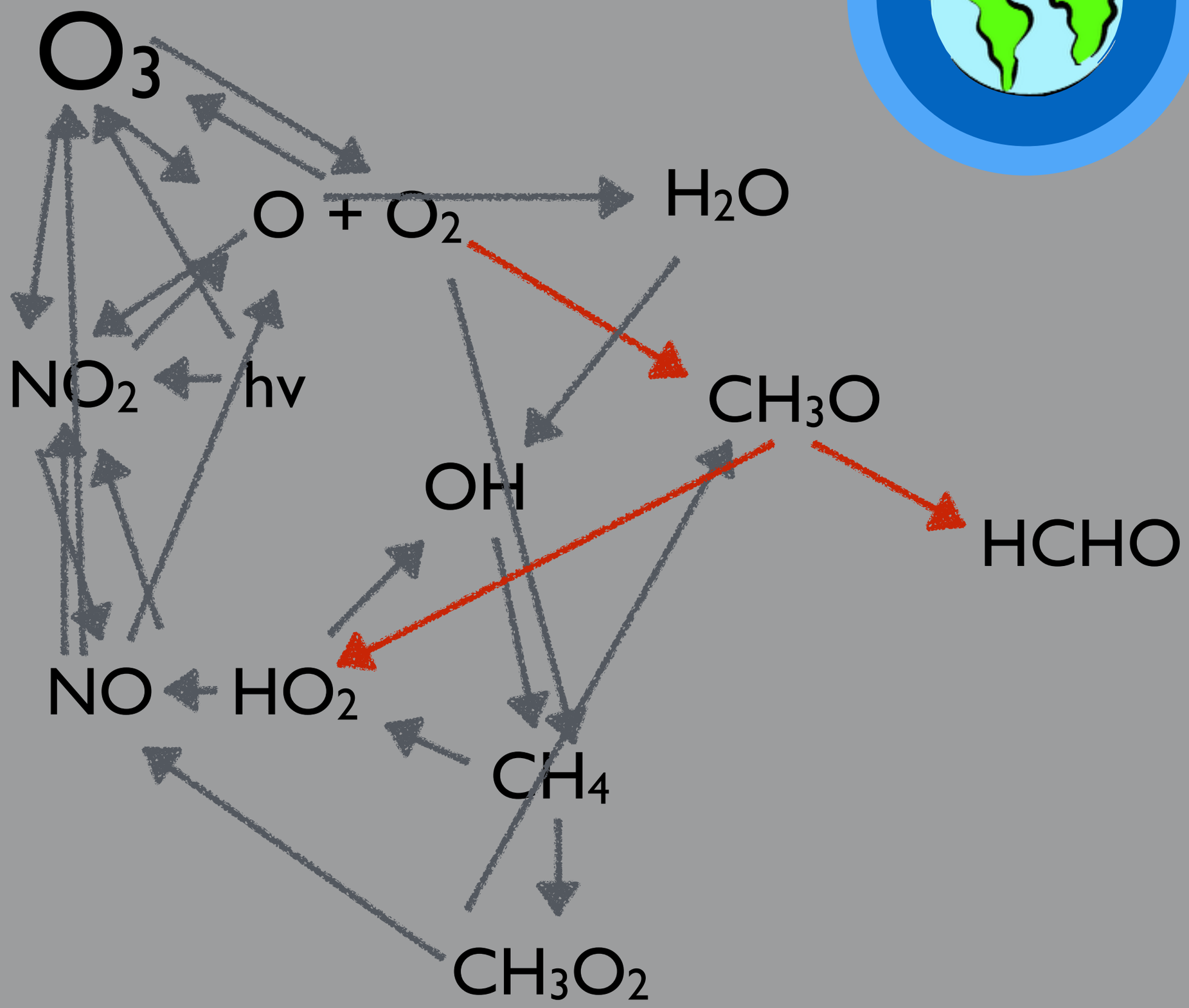
How to make ozone



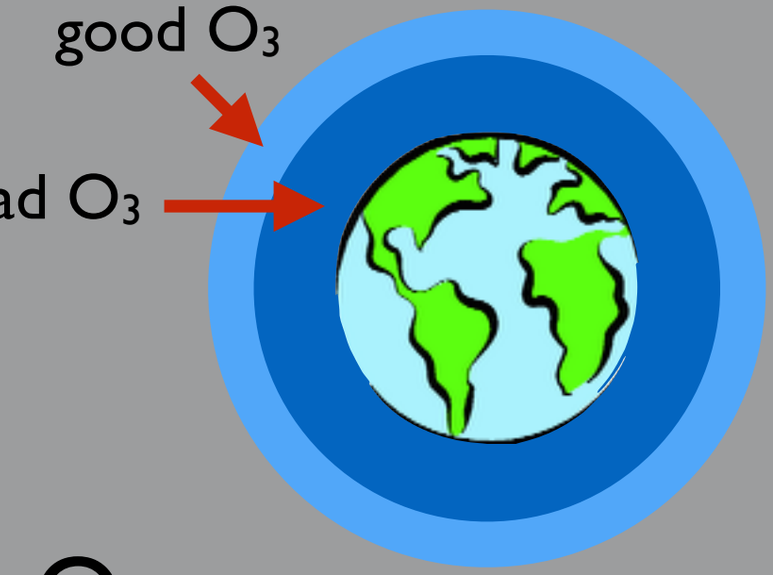


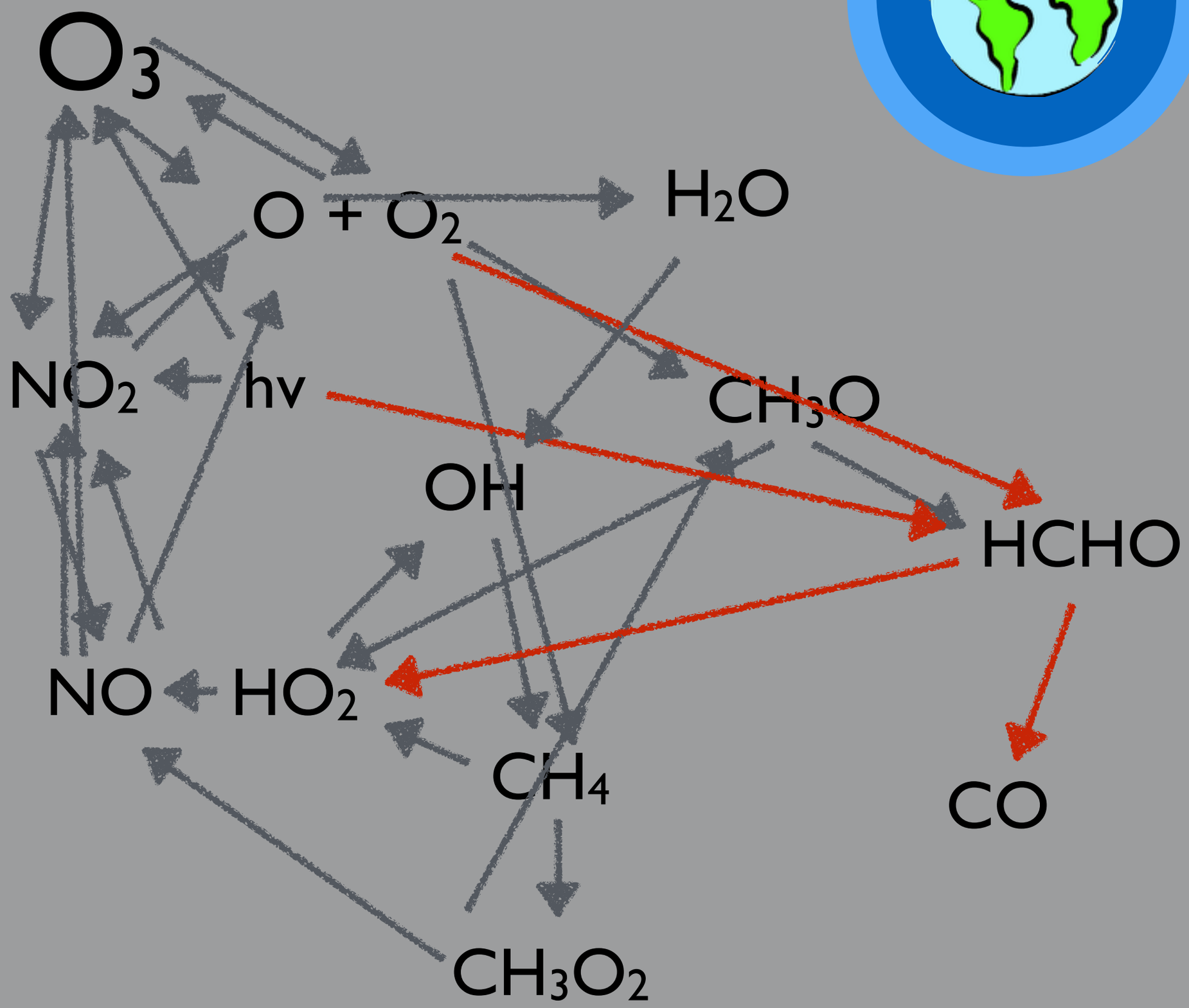
How to make ozone



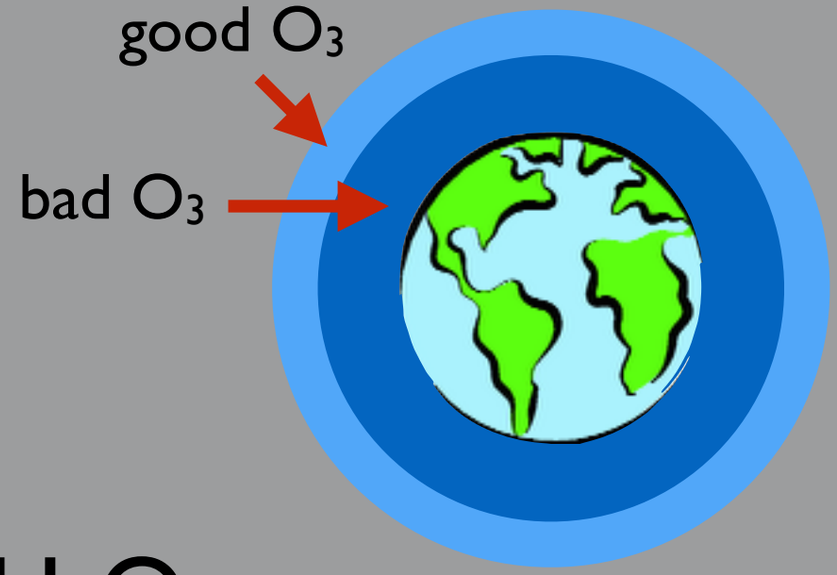


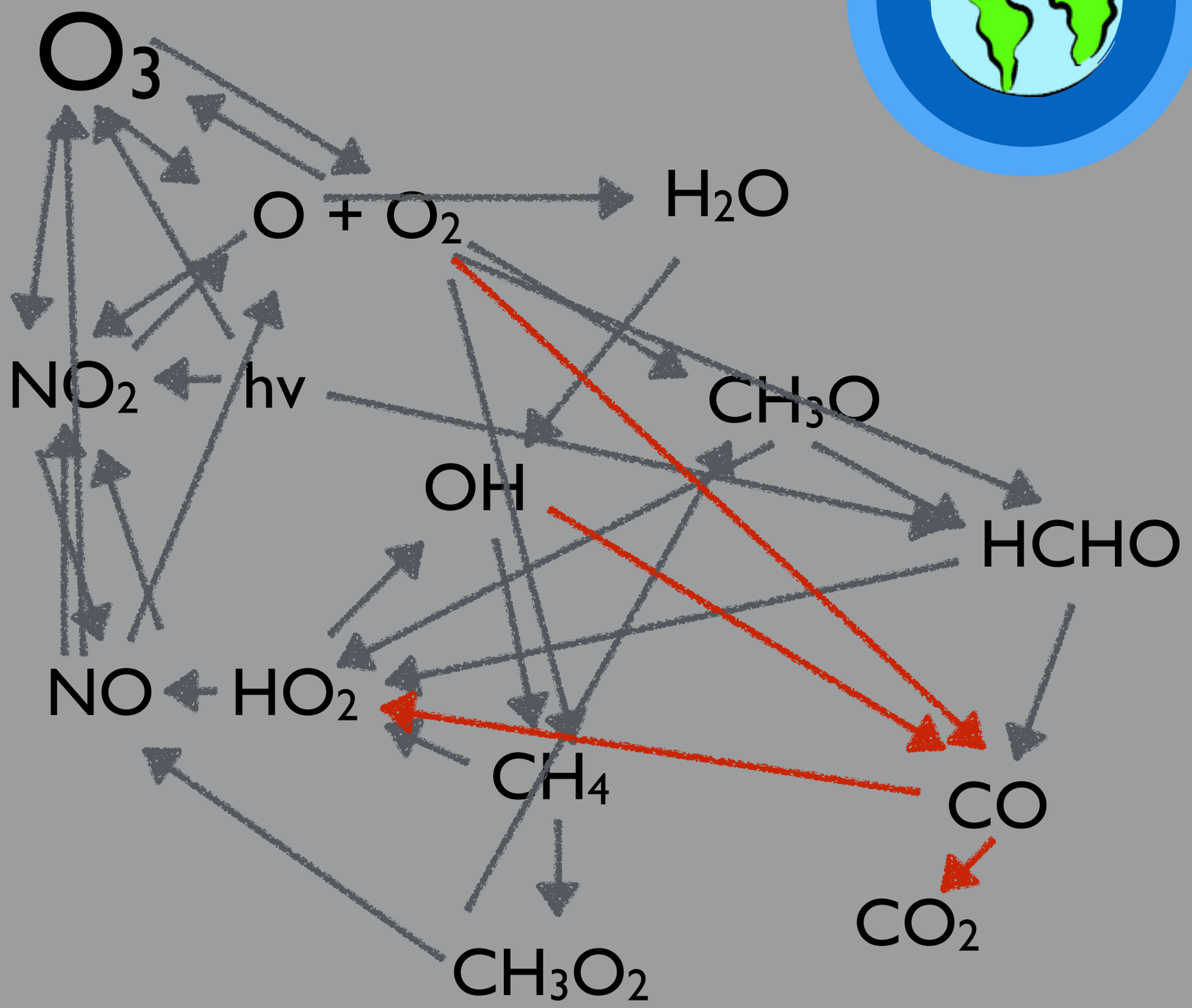
How to make ozone



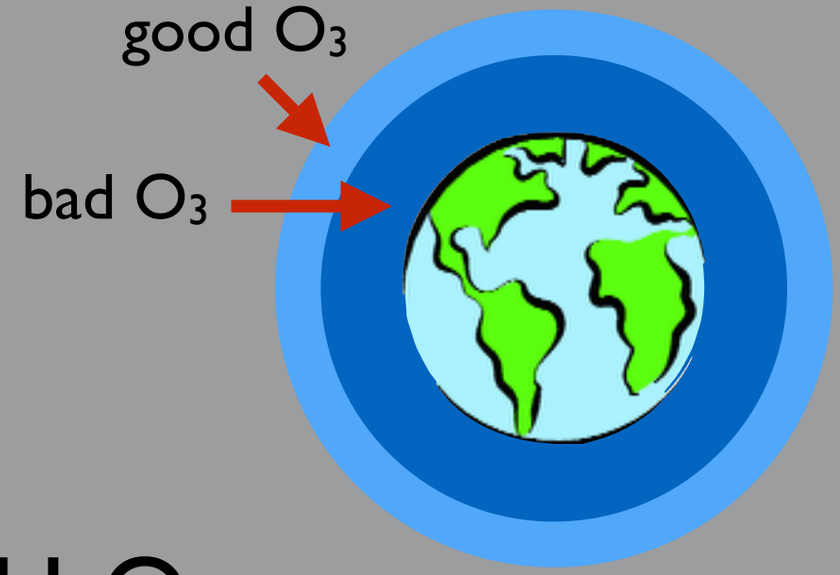


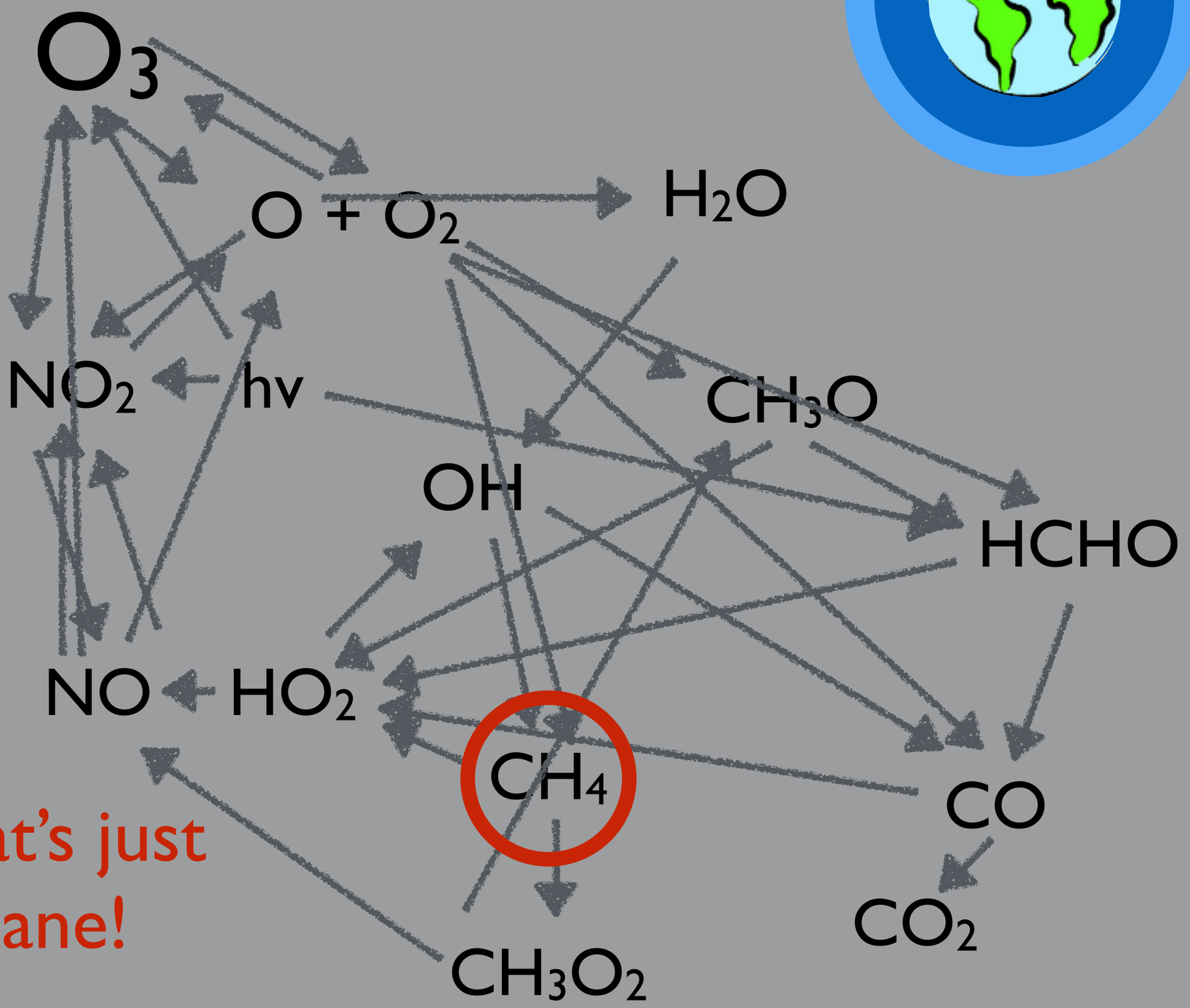
How to make ozone





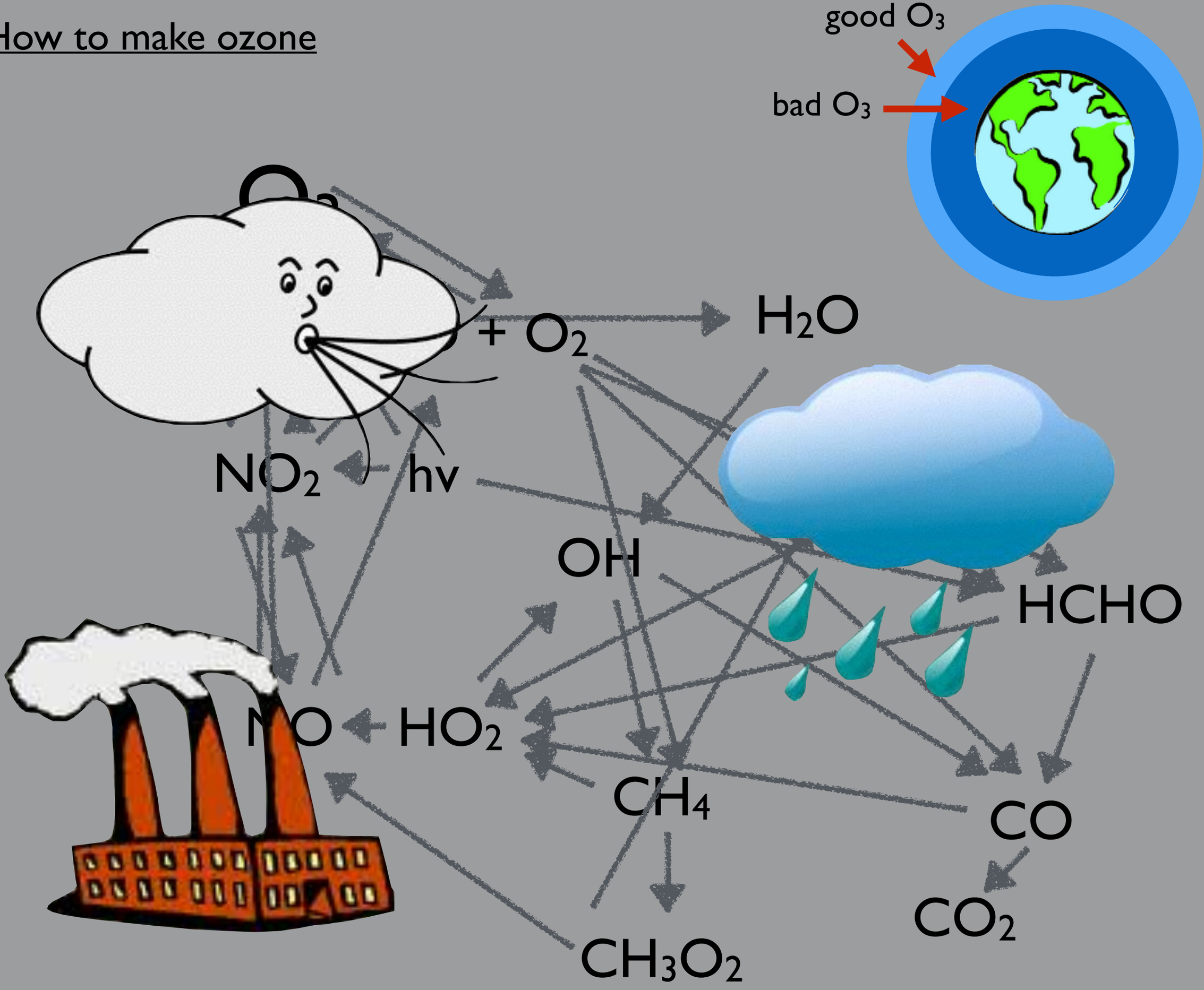
How to make ozone





And that's just methane!

How to make ozone



Whew.

Whew.

That's why we use models.

Impacts of heterogeneous HONO formation on radical sources and ozone chemistry in Houston, Texas

Evan Couzo[1,2], Barry Lefer[3], Jochen Stutz[4], Greg Yarwood[5], Prakash Karamchandani[5], Barron Henderson[6], and William Vizuete[2]

1 2 3 4 5 6

MIT ESD Massachusetts Institute of Technology Engineering Systems Division

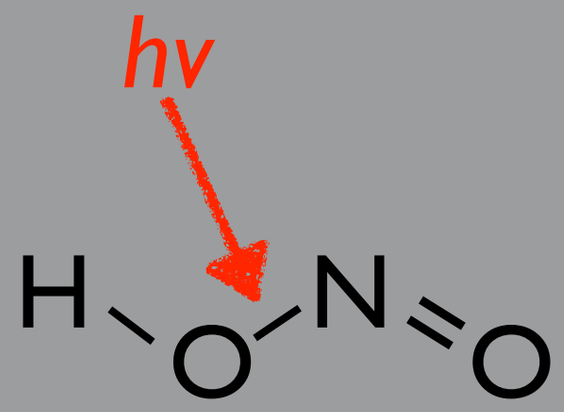
THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

UCLA

UF UNIVERSITY of FLORIDA

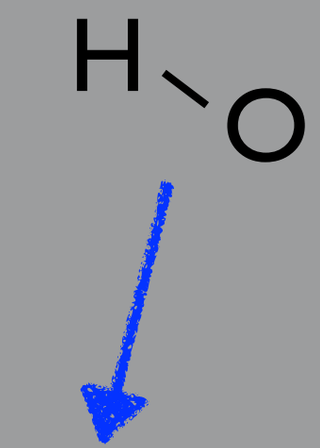
ENVIRON

Nitrous acid is an important radical source.

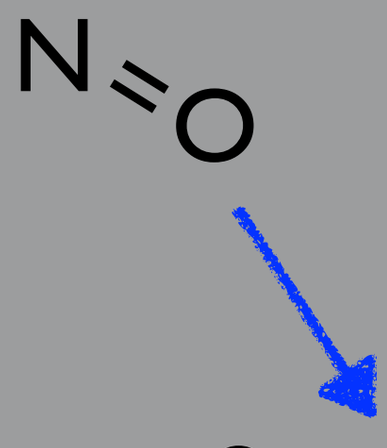


Early morning photolysis of HONO initiates radical formation before other radical sources (HCHO and O_3 photolysis) kick in.

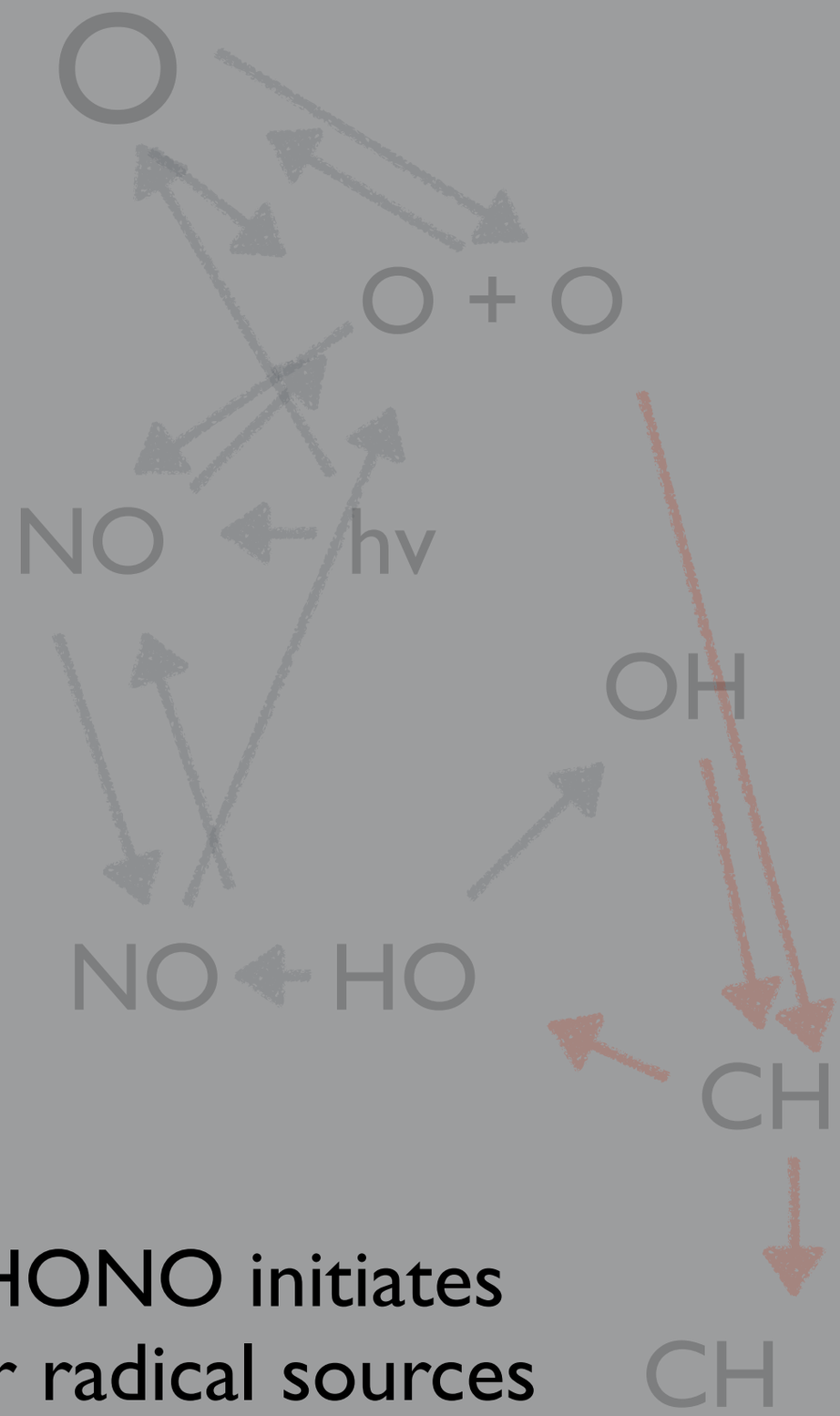
Nitrous acid is an important radical source.



radical

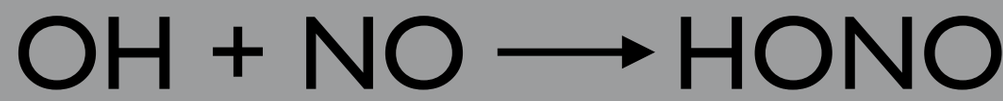


O_3 precursor



Early morning photolysis of HONO initiates radical formation before other radical sources (HCHO and O_3 photolysis) kick in.

Only one significant gas-phase HONO formation reaction.

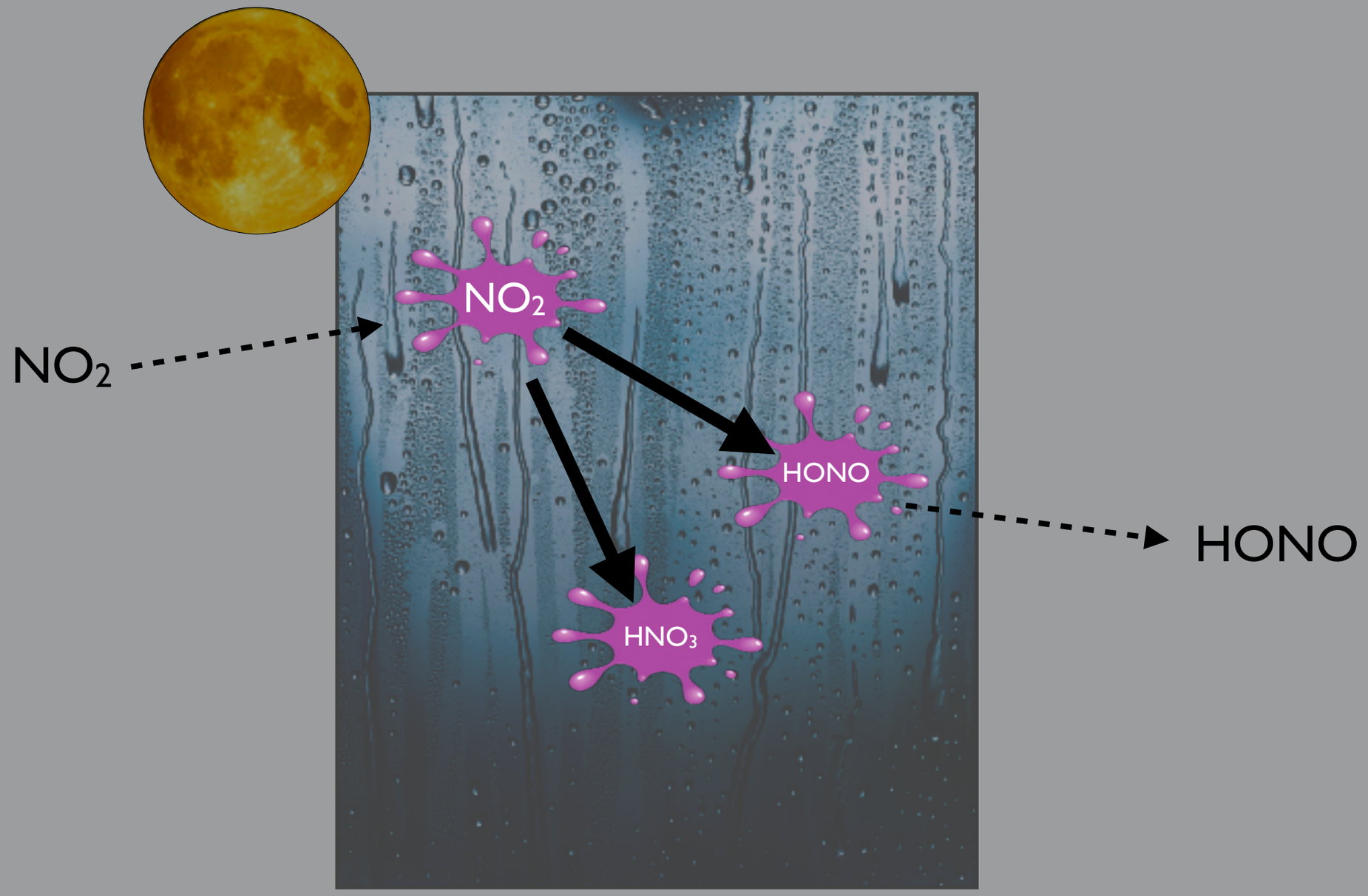


- Reverse of photolysis reaction.
- Included in air quality model chemical mechanisms.
- In air quality models, no other significant source of HONO exists.
- It is not an emitted species.

But, wait! There *are* other ways to make HONO.

$$2NO_{2,ads} + H_2O_{ads} \longrightarrow HONO_{ads} + HNO_{3,ads}$$

Lammel and Cape, 1996;
Finlayson-Pitts et al., 2003

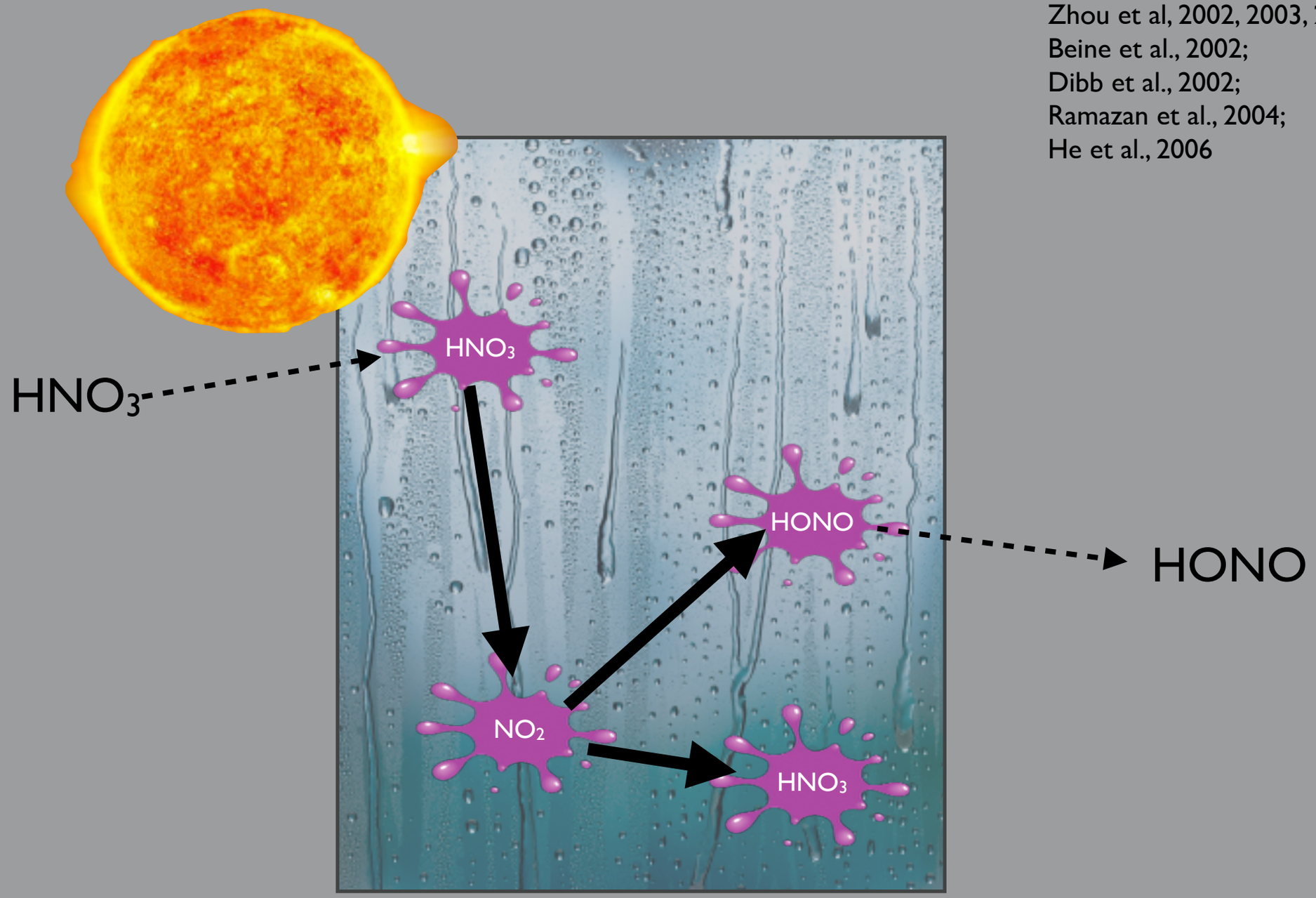


But, wait! There *are* other ways to make HONO.

$$HNO_{3,ads} + h\nu \longrightarrow NO_{2,ads} + OH$$

$$NO_{2,ads} + H_2O_{ads} \longrightarrow HONO_{ads} + HNO_{3,ads}$$

Zhou et al, 2002, 2003, 2007, and 2011;
Beine et al., 2002;
Dibb et al., 2002;
Ramazan et al., 2004;
He et al., 2006

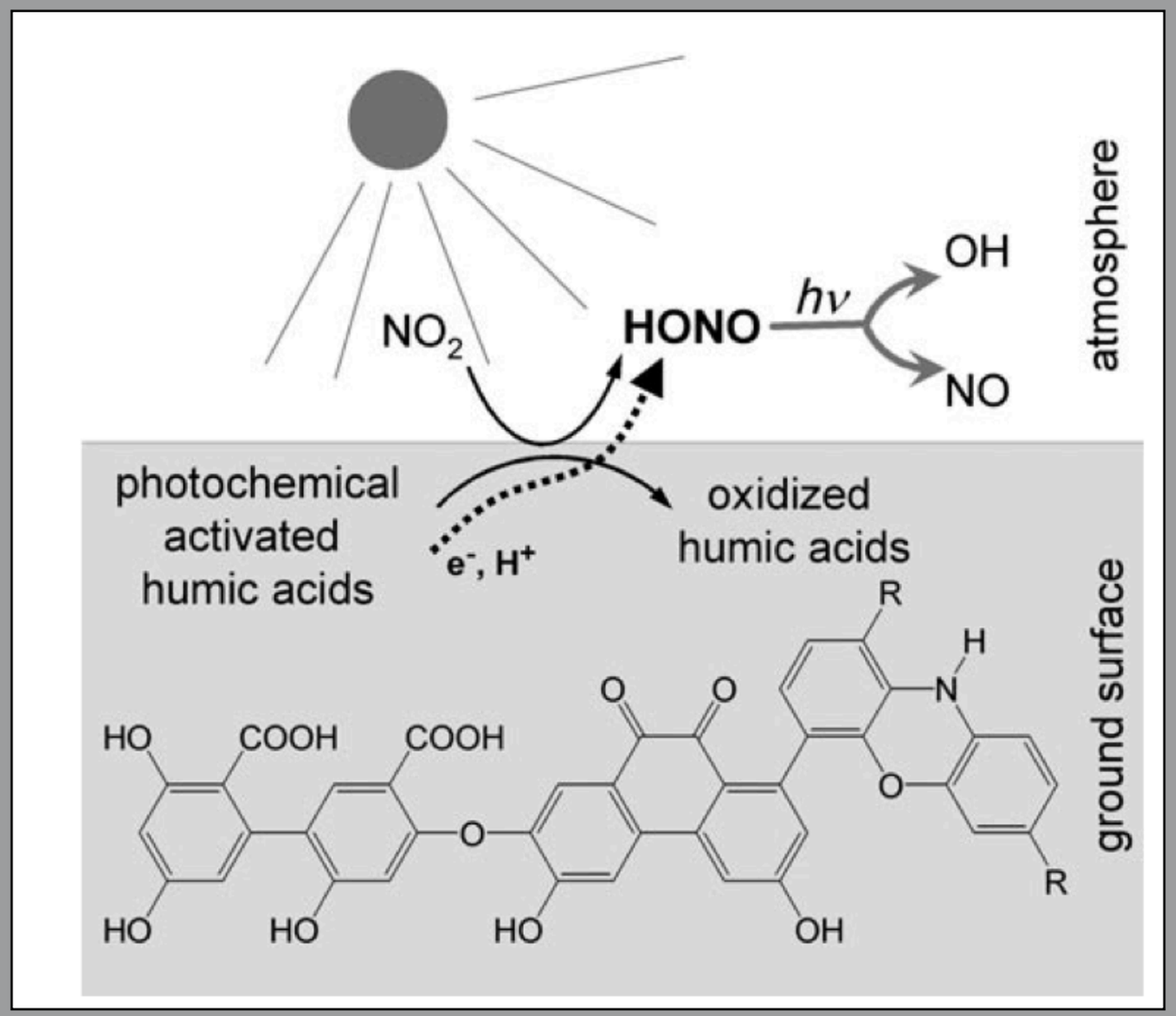


But, wait! There *are* other ways to make HONO.

$$HA + h\nu \longrightarrow A_{red}$$

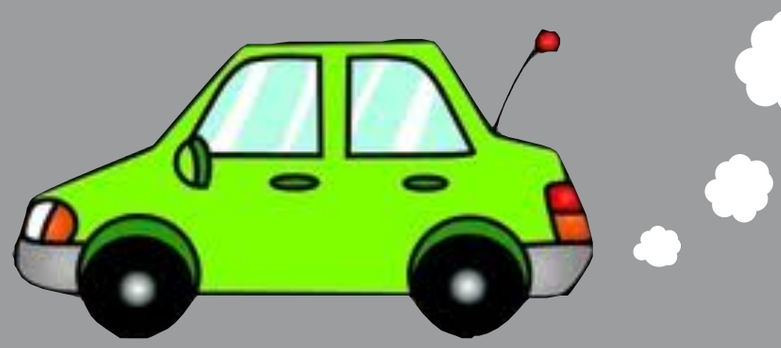
$$A_{red} + NO_2 \longrightarrow A' + HONO$$

George et al., 2005;
Stemmler et al., 2006 and 2007



from Kleffman et al., 2007

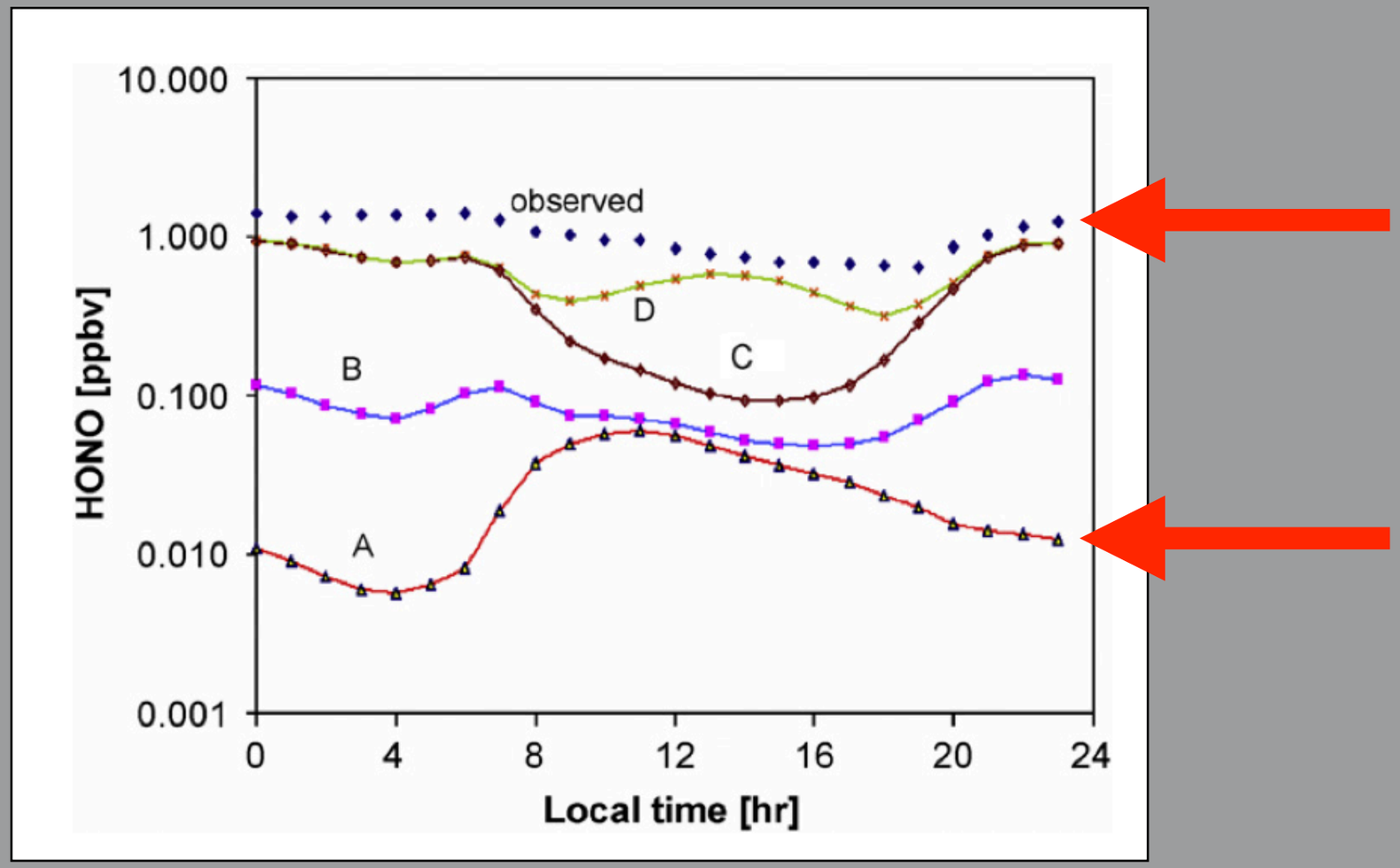
There is evidence that HONO is emitted during combustion, too.



$$\frac{\text{HONO}}{\text{NOx}} = 0.8\%$$

Pitts et al., 1984;
Kurtenbach et al., 2001

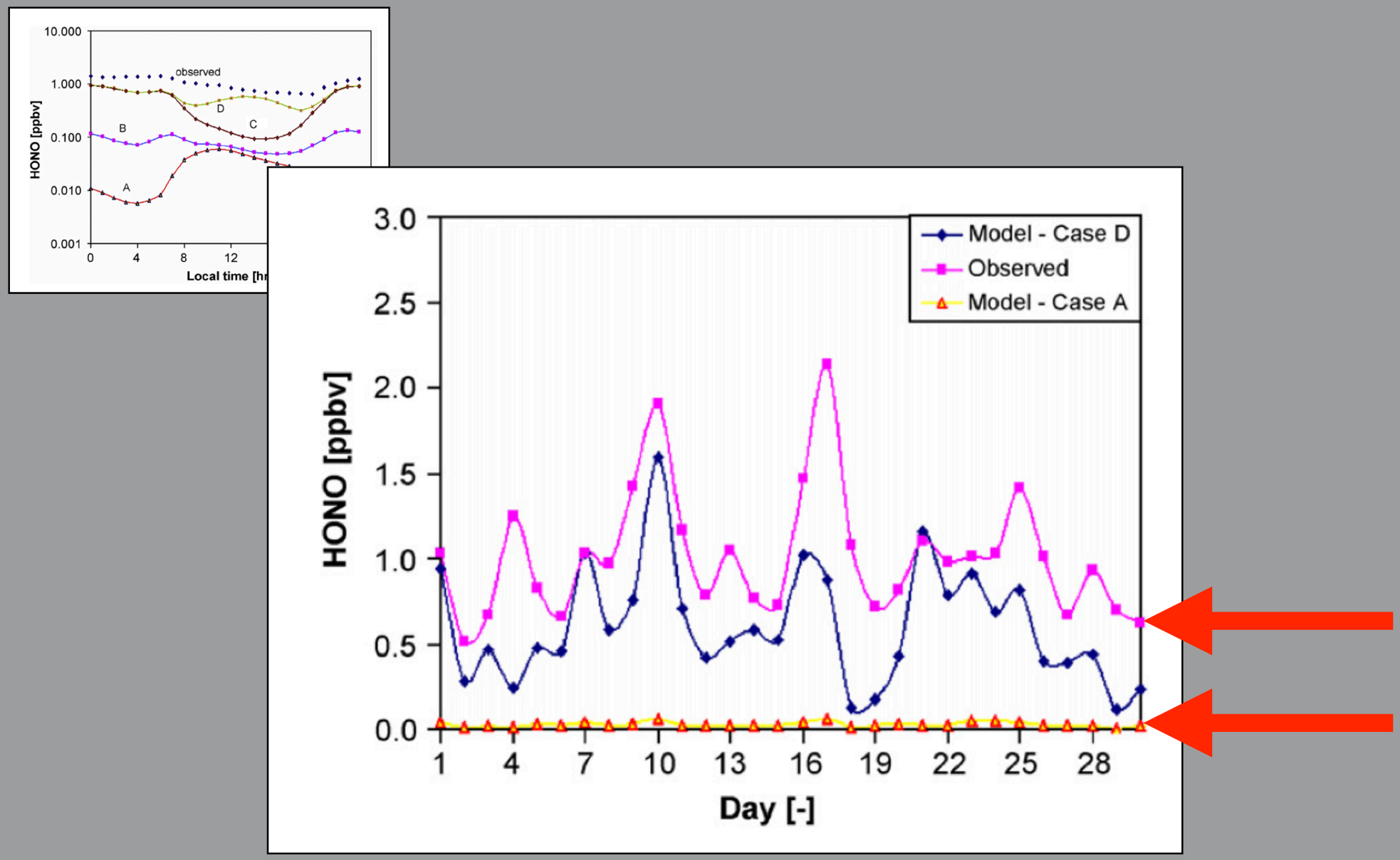
This adds up to quite a few missing HONO sources in air quality models.



from Sarwar et al., 2008

Other modeling studies found HONO under predictions compared to measurements.

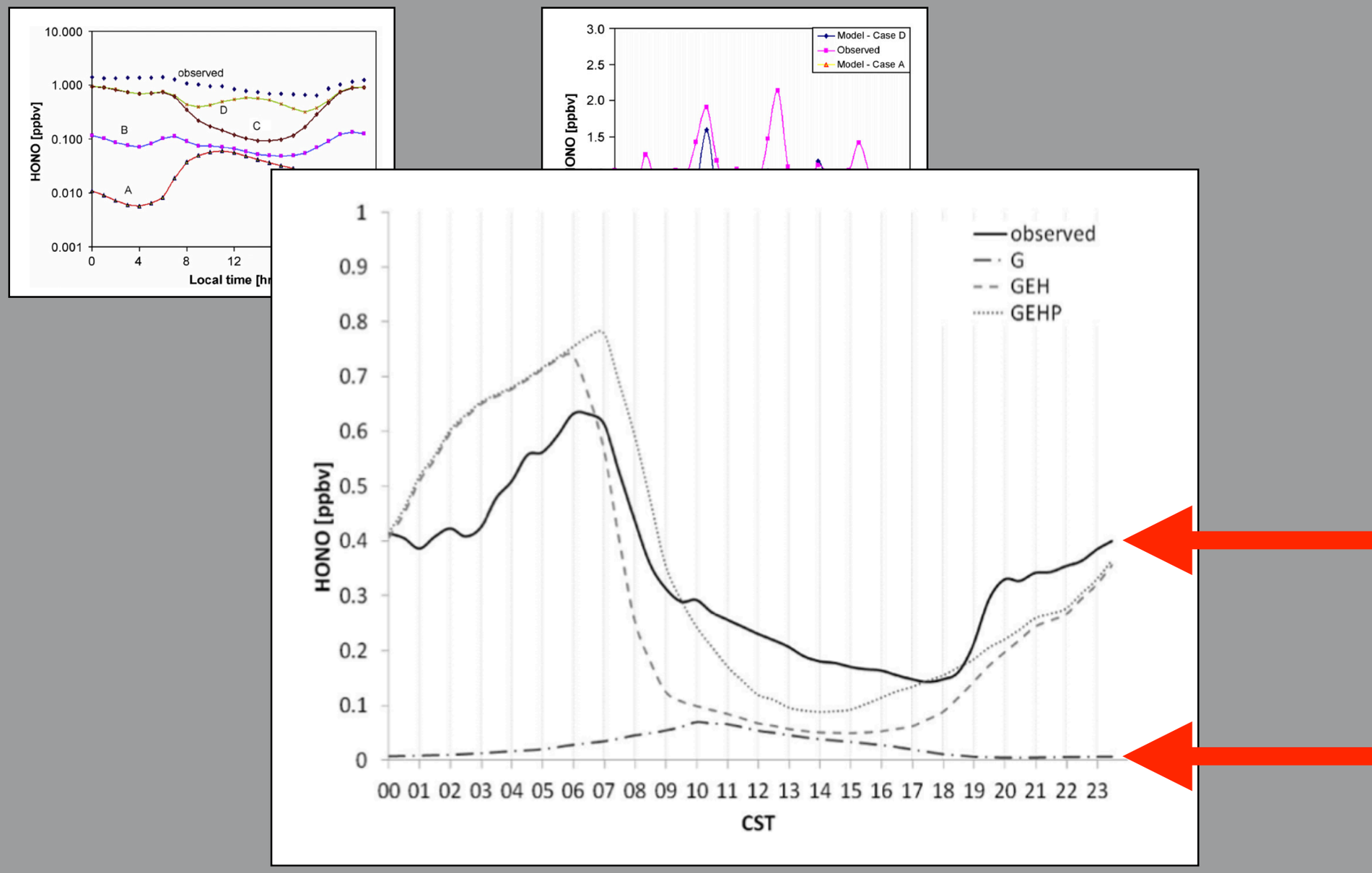
This adds up to quite a few missing HONO sources in air quality models.



from Sarwar et al., 2008

Other modeling studies found HONO under predictions compared to measurements.

This adds up to quite a few missing HONO sources in air quality models.



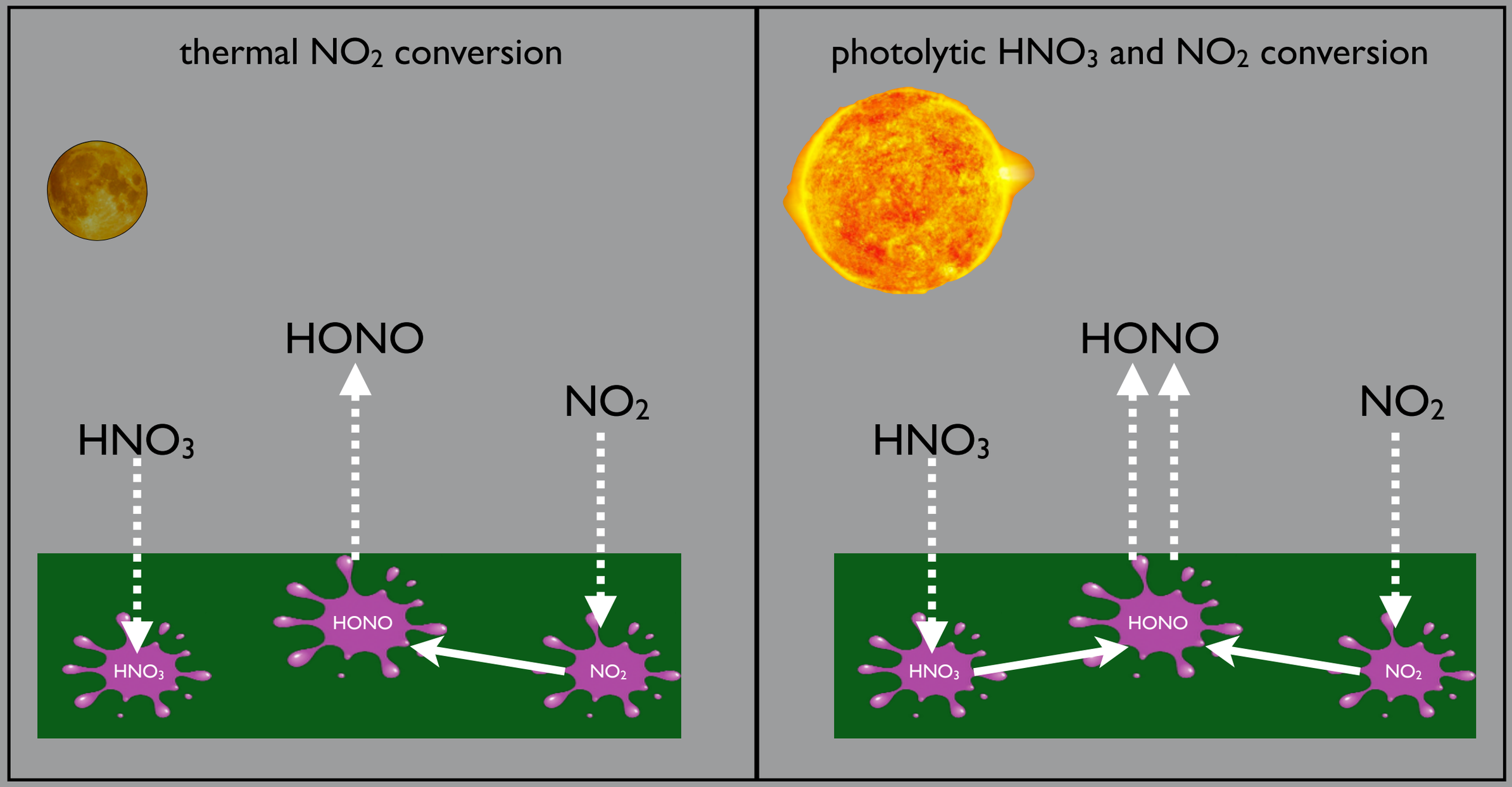
from Czader et al., 2012

Other modeling studies found HONO under predictions compared to measurements.

There have been attempts to model new HONO sources.

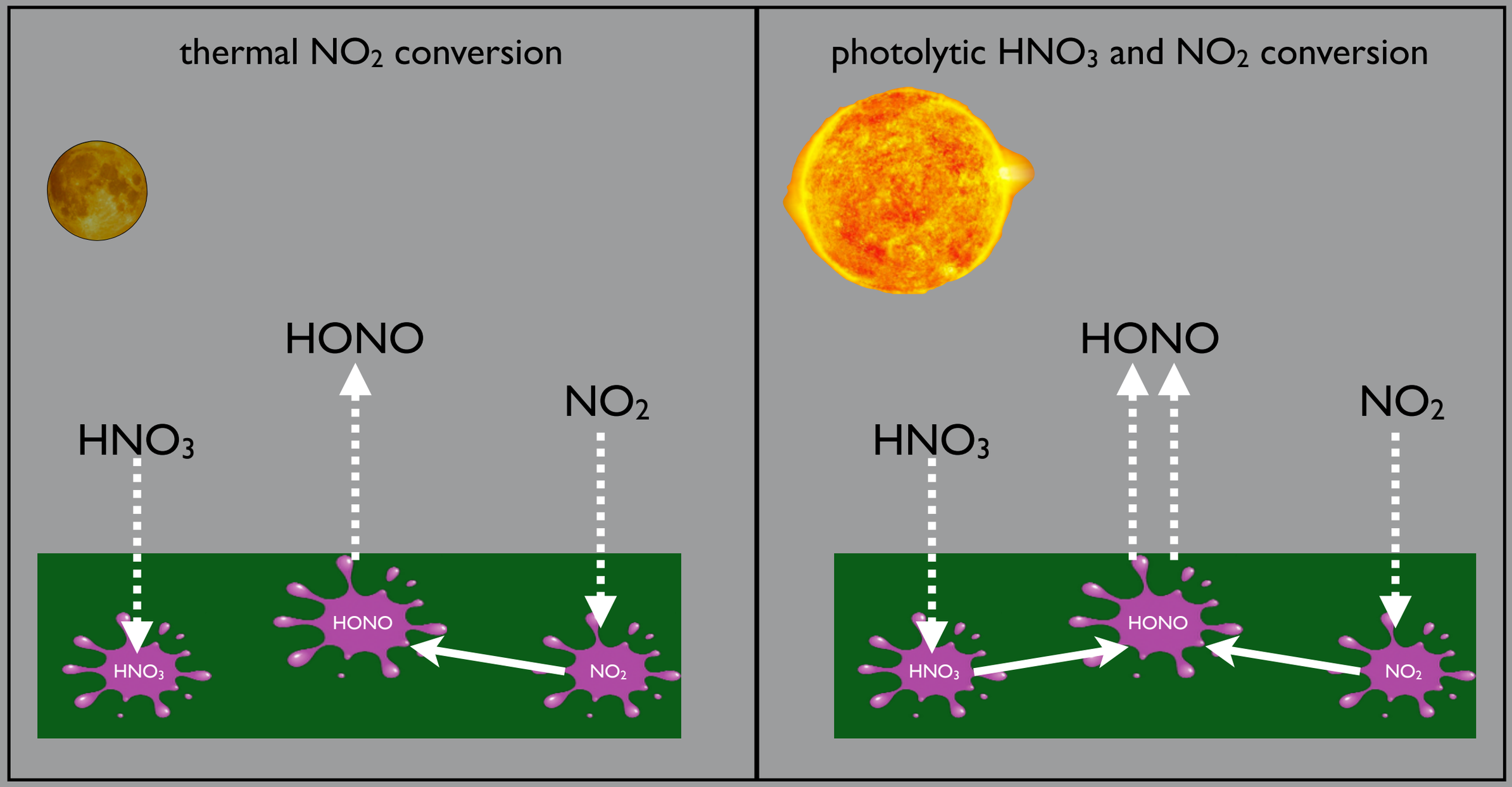
- Adding additional HONO sources improves model performance, especially at night.
- These attempts use surface area to volume parameterizations to approximate heterogeneous reactions.
- Though conceptually consistent, this approach is not physically realistic.

A new surface sub-model chemically processes HNO_3 and NO_2.



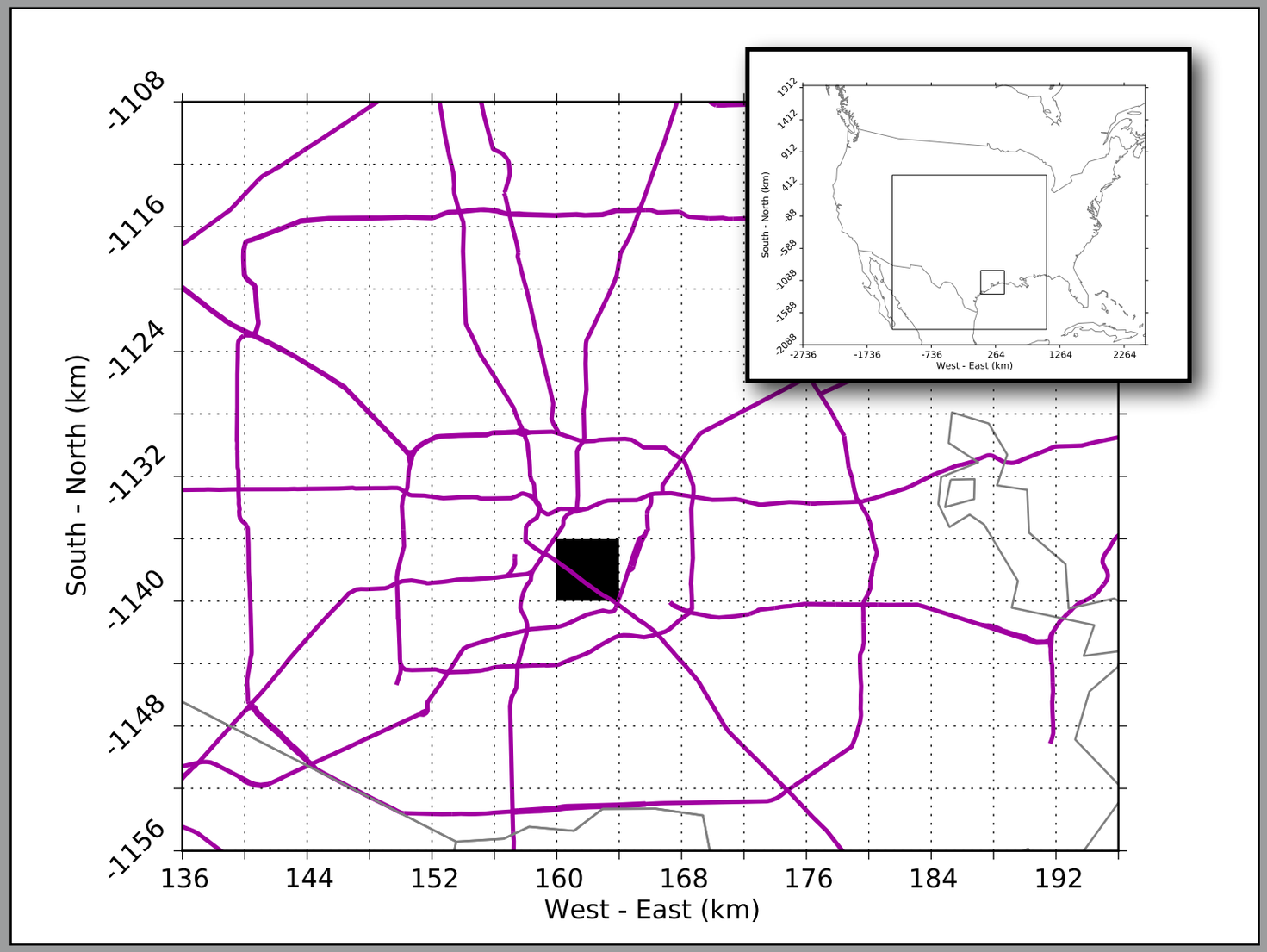
1. Dry deposition of HNO_3 and NO_2.
2. Nighttime NO_2 → HONO conversion; surface HNO_3 reservoir.
3. Daytime NO_2 → HONO and HNO_3 → HONO conversion.

A new surface sub-model chemically processes HNO_3 and NO_2.



Dry deposition is no longer a total loss process. This is a new, physically accurate way of modeling heterogeneous HONO chemistry in air quality models.

Modeling episode aligns with SHARP measurements.



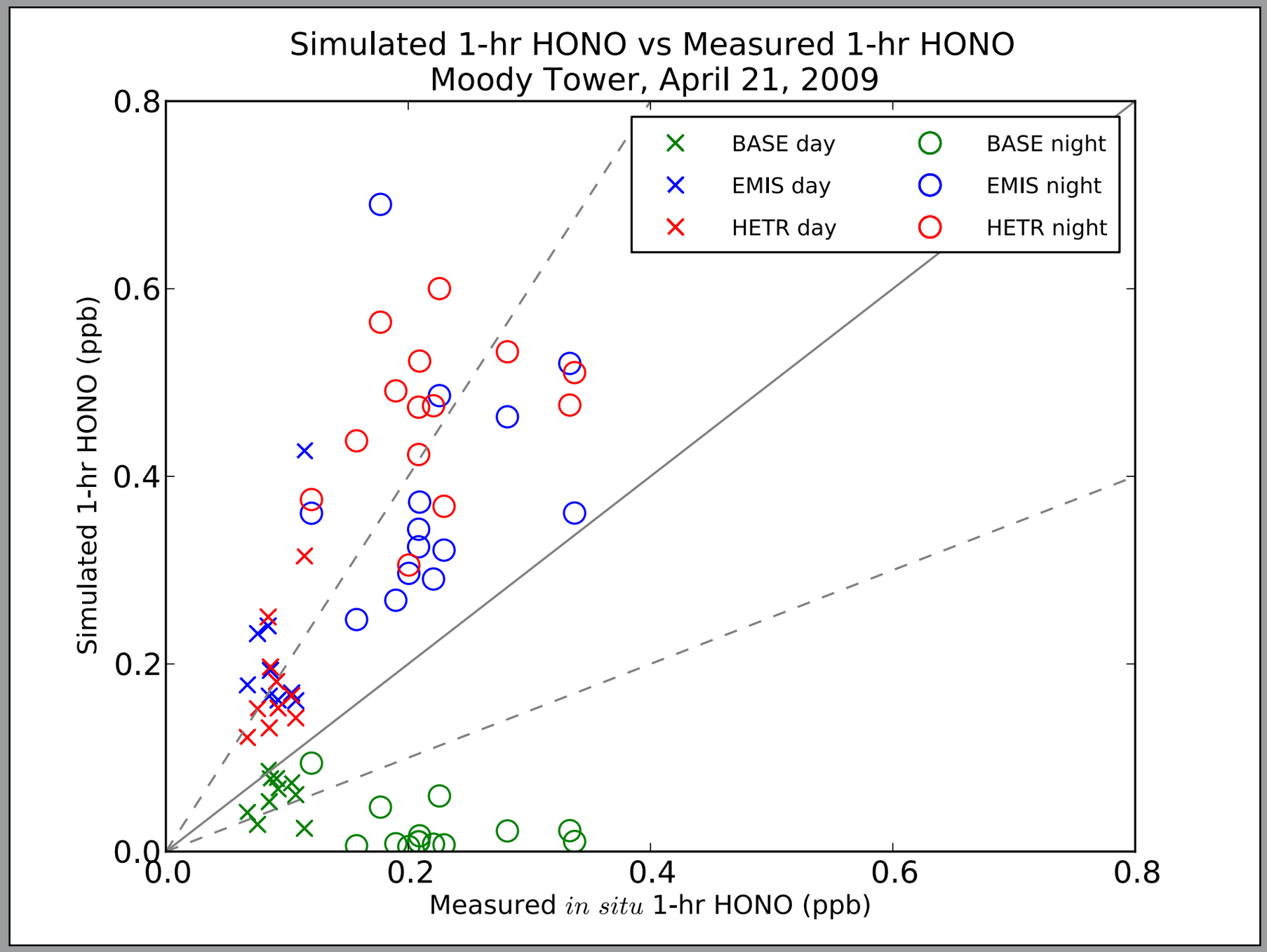
- Wide array of measurements taken at Moody Tower during SHARP.
- Our analysis focuses on Moody Tower grid cell on April 21, 2009.
- Greatest HONO and O_3 concentrations in April.
- Model data taken from 2nd vertical layer to match height of measurements.

Three different modeling scenarios.

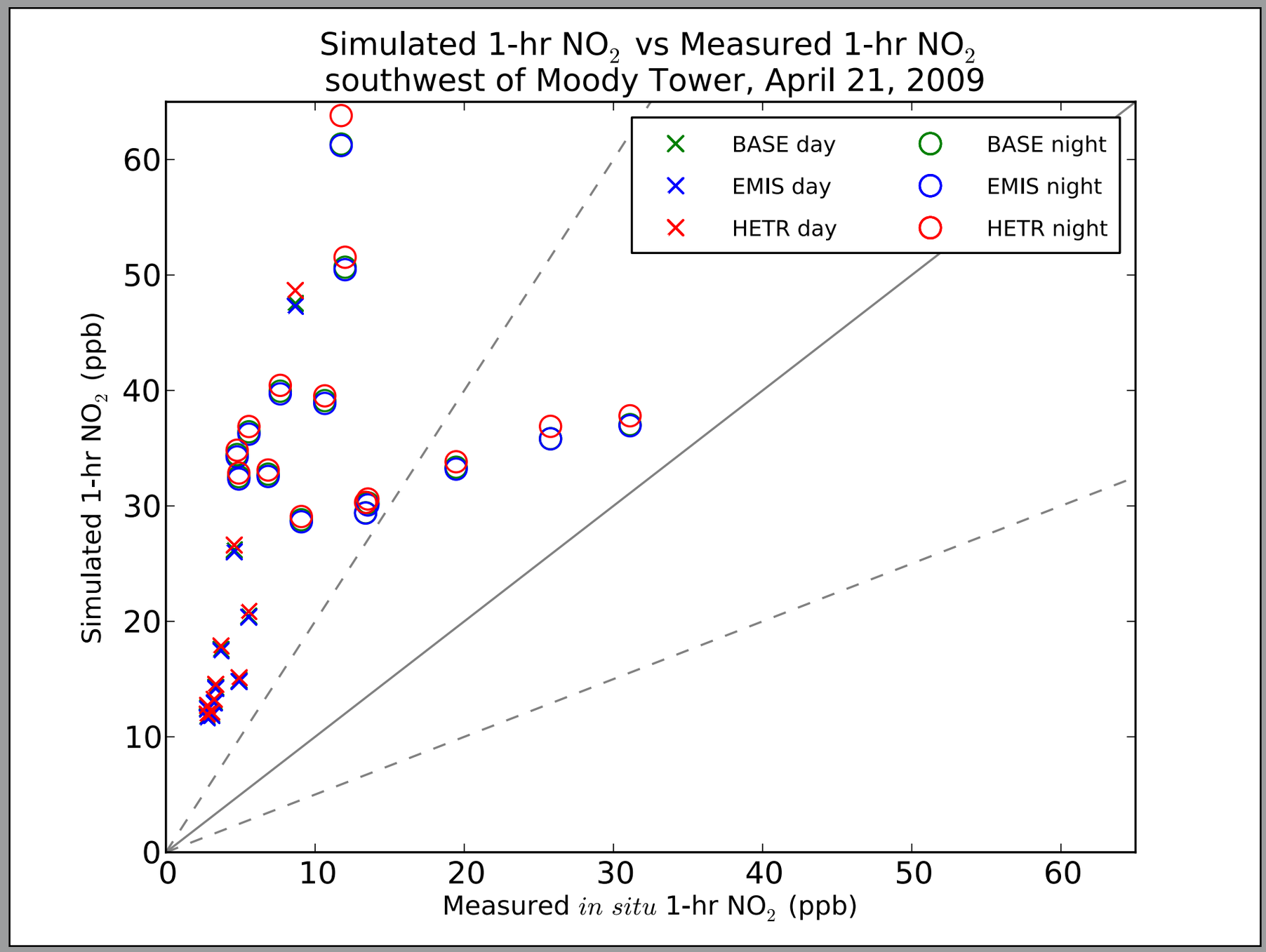
Scenario	Emission Inventory	Surface Model
BASE	base	no
EMIS	base + 0.8% HONO:NOx	no
HETR	base	yes

- Is it possible to model heterogeneous chemistry in a regional air quality model?
- Will additional HONO sources improve model performance?
- What is the effect on radical budgets and O_3 formation?

Severe HONO over predictions.

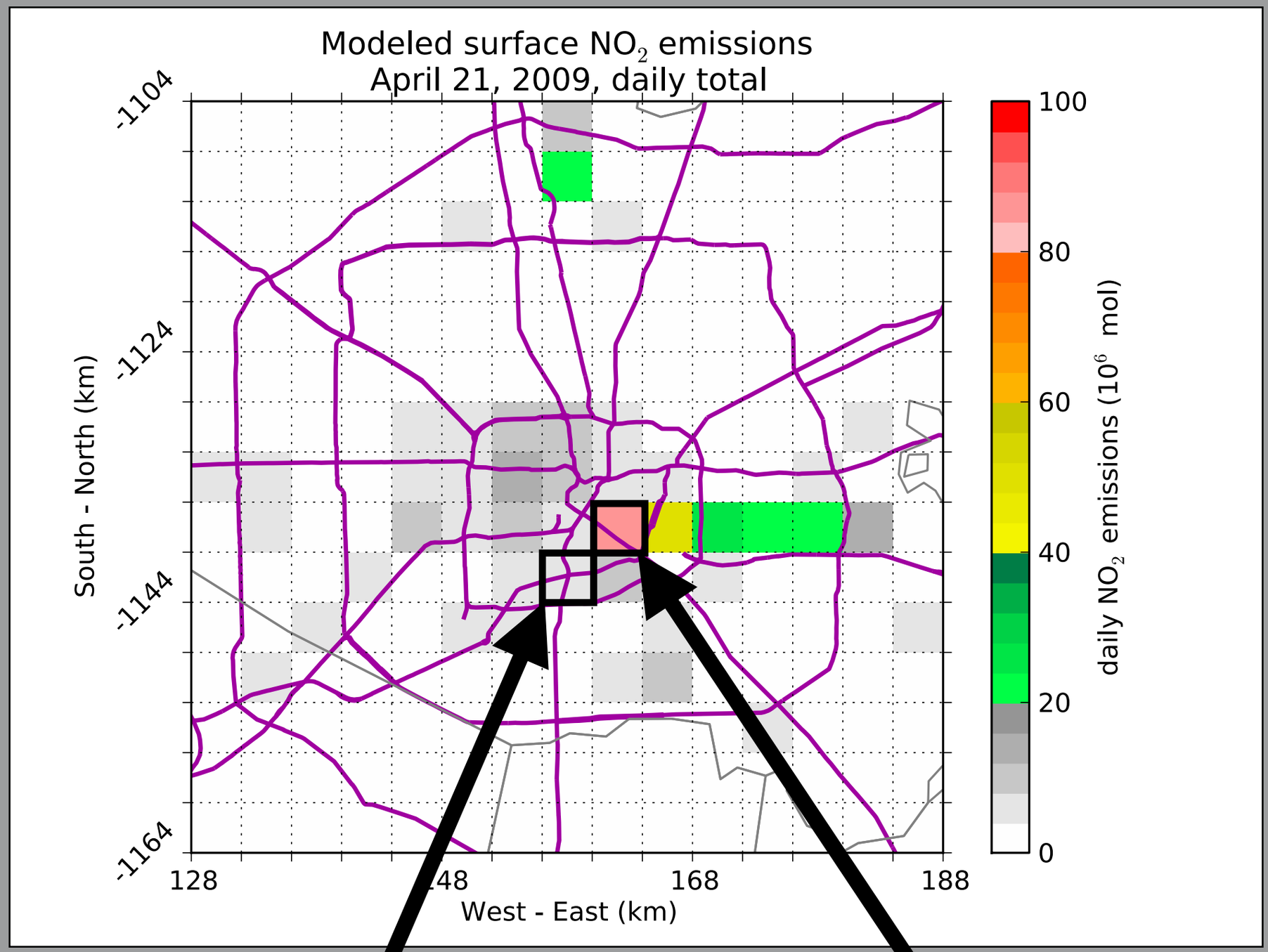


NO_2 performance is also bad.



(Why the NO_2 differences? In the HETR scenario, deposited nitrogen is recycled as HONO, and finds its way back to NO_2.)

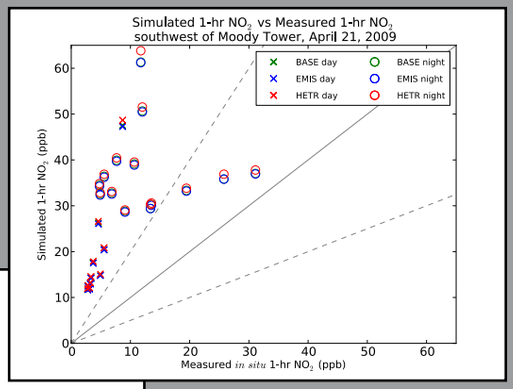
NO_2 emissions are very large in industrial region of Houston.



Moody Tower southwest

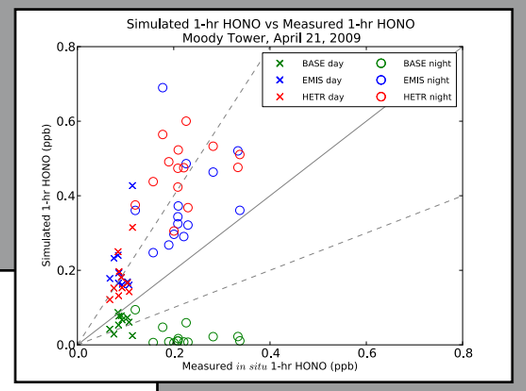
Moody Tower

Much better NO_2 performance.

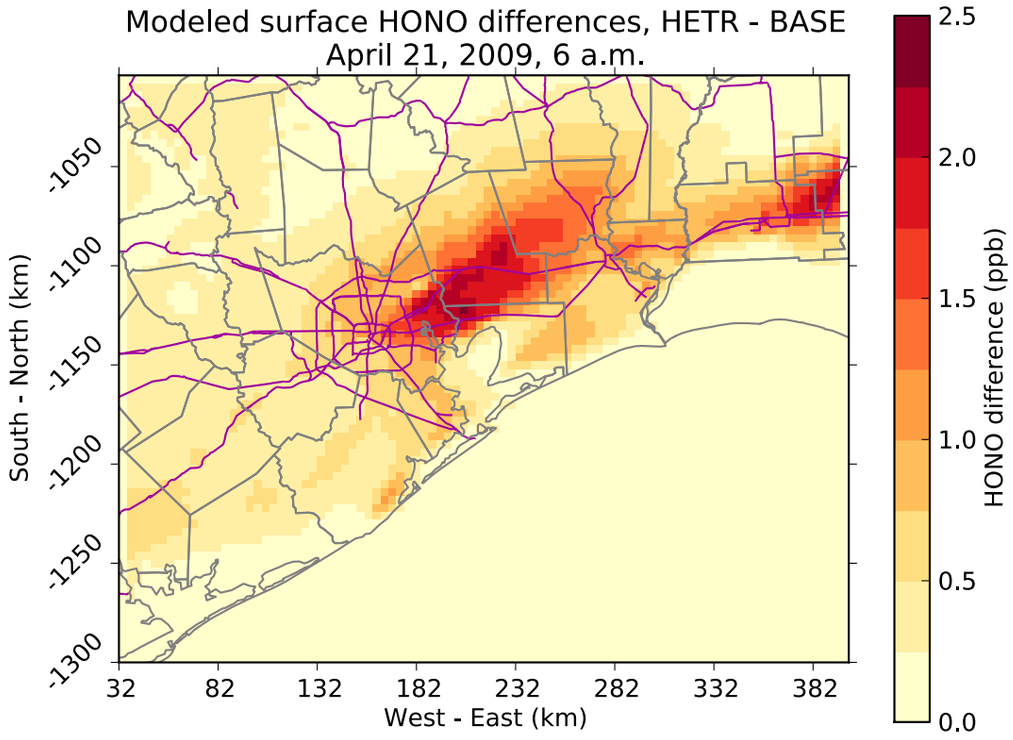


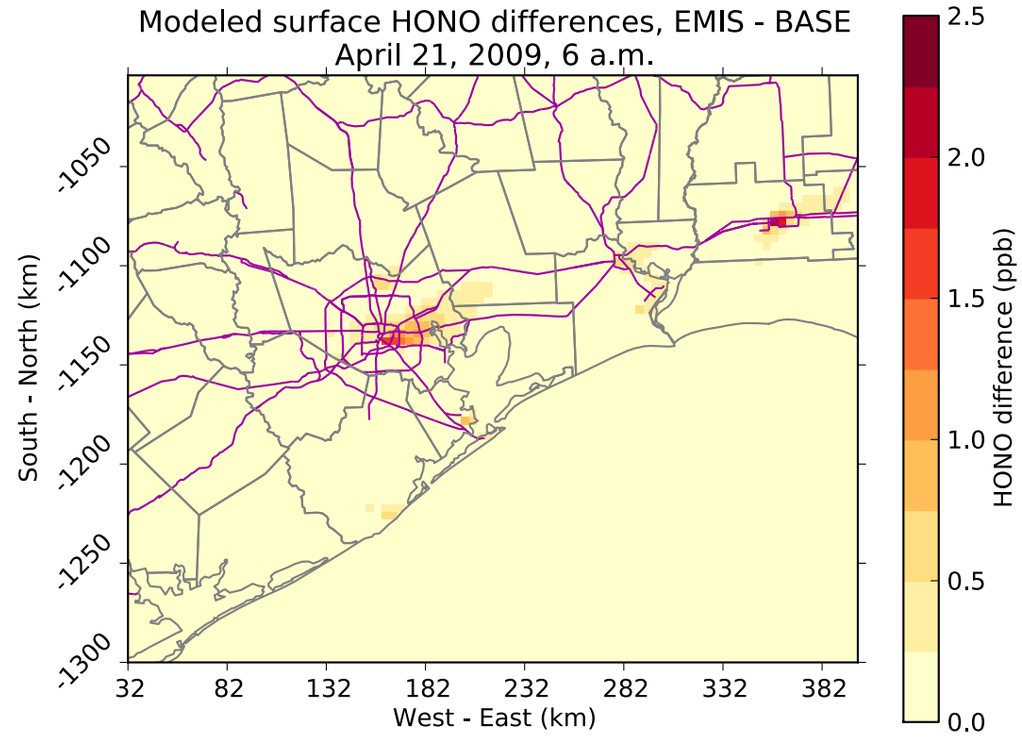
Simulated 1-hr NO_2 vs Measured 1-hr NO_2
southwest of Moody Tower, April 21, 2009
BASE day
EMIS day
HETR day
BASE night
EMIS night
HETR night
Simulated 1-hr NO_2 (ppb)
Measured *in situ* 1-hr NO_2 (ppb)

HONO also looks better.



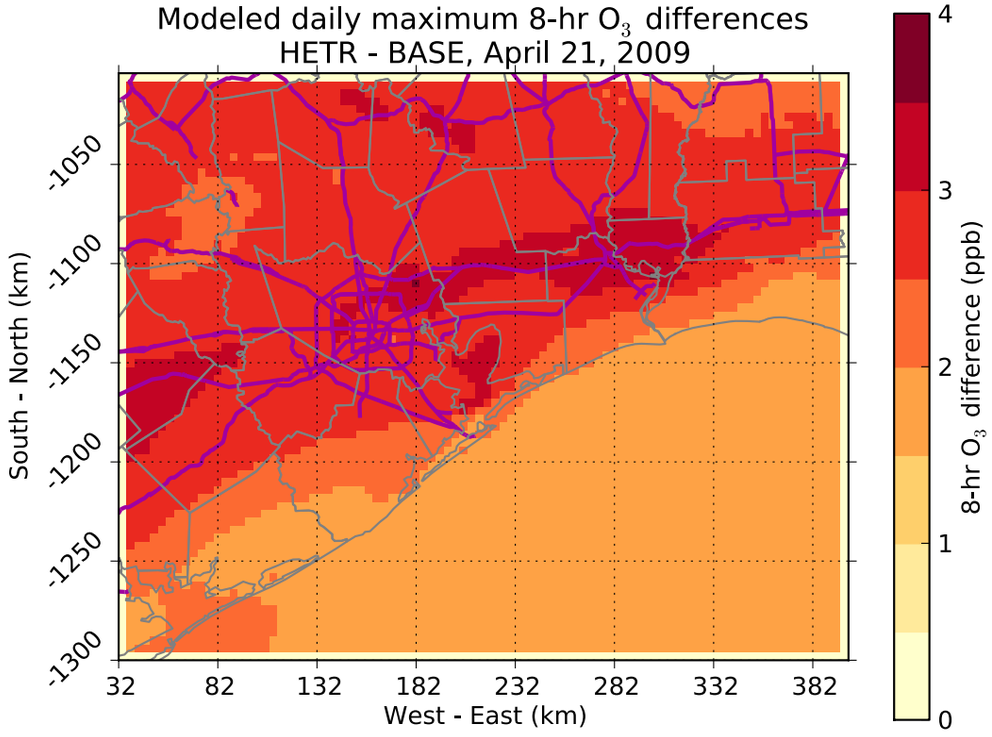
More HONO in more places with surface model.

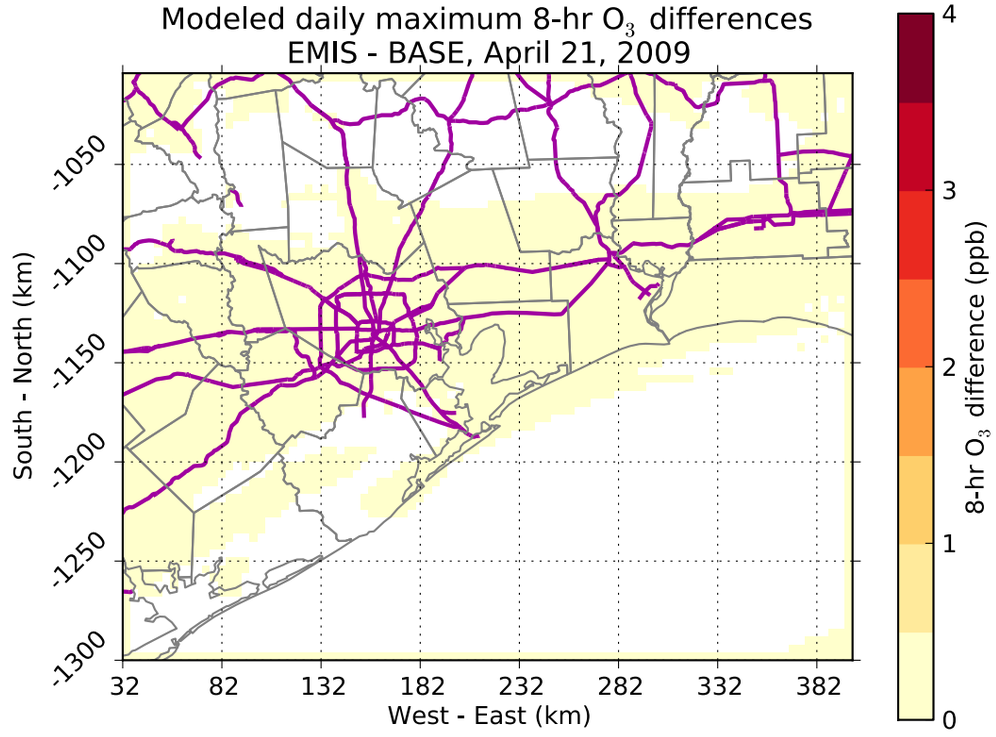




- HETR increase is regional - lots of soil and vegetation.
- EMIS increase is confined to local areas of high NOx emissions.
- In HETR scenario, there is a greater pool of radical precursors right at sunrise.

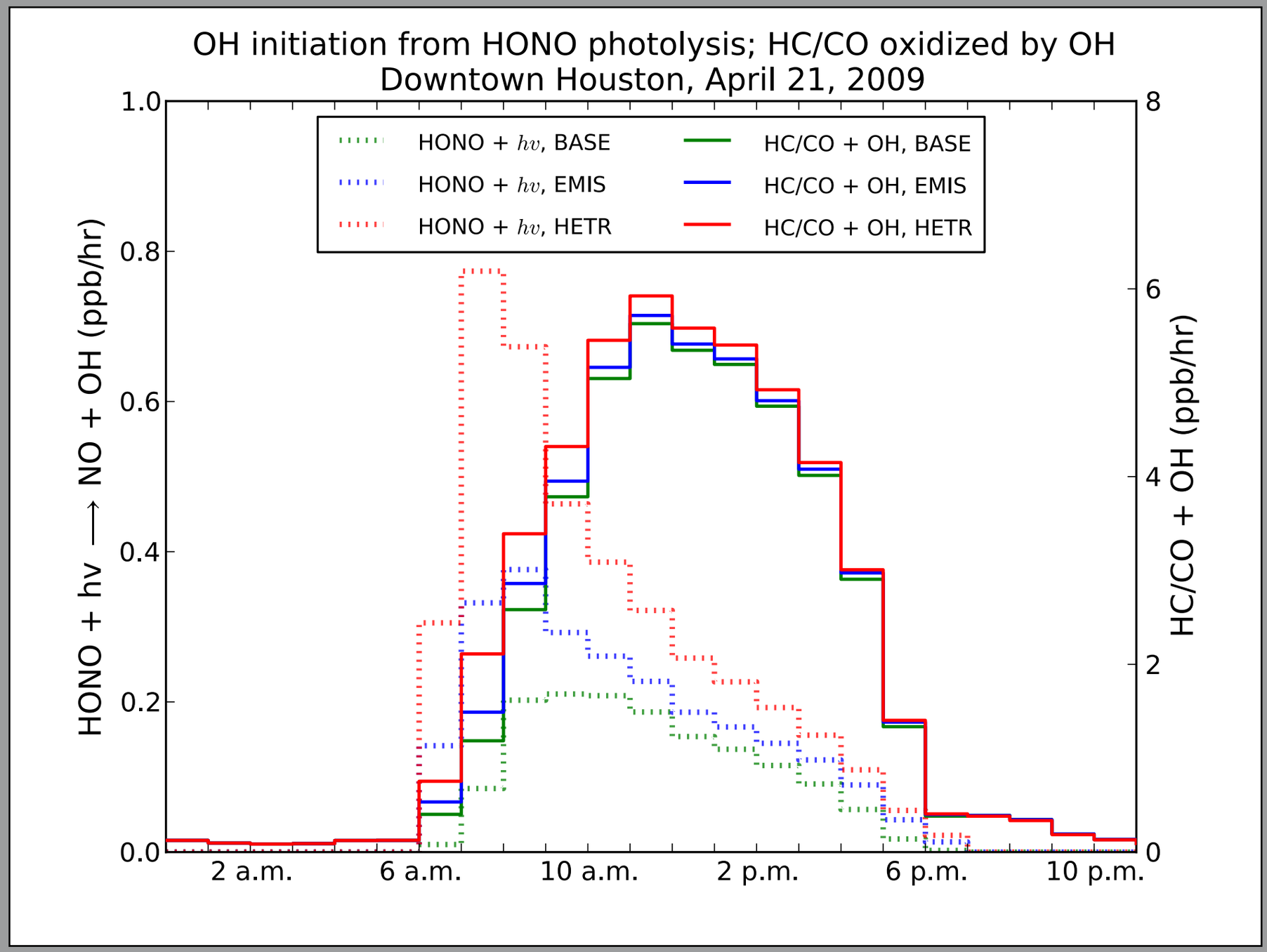
HETR makes more O_3 in more places.





- HETR daily maximum 8-hr O_3 is up to 3.5 ppb greater over BASE.
- Significant O_3 increases across modeling domain (urban and rural).
- Increases in EMIS are less than 0.5 ppb over BASE and are zero in rural areas.

Most of the action happens right after sunrise.



More HONO photolysis - because there is more HONO - produces more OH, which oxidizes more HC/CO.

Final thoughts and next steps.

- EMIS and HETR make HONO at night, which homogeneous chemistry alone does not do. Daytime performance is also improved.
- NO_x performance is important for HONO formation, especially EMIS.
- Quantified changes to the oxidative environment as a result of the additional HONO sources.
- Important proof-of-concept for a physically accurate surface sub-model.
- As a result, the surface model will be available in the latest release of CAMx.

Our model is still wrong.

- The modeled surface chemistry generalizes the heterogeneous reactions and ignores the intermediate chemical species.
- The ground is distorted by treating it as a flat surface covered fractionally by soil and vegetation.
- Known HONO formation mechanisms are deleted, such as formation on the sides of buildings and surfaces of aerosols.
- We could further improve the accuracy of our model and likely see modest performance gains as compared to measurements.
- But it's the hope of uncovering our own nescience that keeps us going.