Assimilation of lightning data in NWP?

Philippe Lopez

ECMWF (European Centre for Medium-Range Weather Forecasts), Reading, UK

Outline

- Purpose and ingredients of data assimilation (DA) in NWP.
- Data assimilation methods.
- Issues with the assimilation of lightning observations.
- Summary and prospects.

Joint MTG LI Mission Advisory Group & GOES-R GLM Science Team workshop, Rome, 27-29 May 2015

ECMWF

The purpose and ingredients of Data Assimilation in NWP

* The purpose of DA in NWP is to merge information coming from observations with *a priori* information coming from a forecast model to obtain an optimal 3D representation of the atmospheric state at a given time (= the "analysis").

This 3D analysis can then provide initial conditions to the numerical forecast model.

* The main ingredients of DA are:

- a set of **observations** available over a period of typically a few hours.
- a previous short-range forecast from the NWP model ("**background**" information).
- some statistical description of the **errors** of both observations and model background.
- a data assimilation **method** (e.g. nudging, variational DA, EnKF,…).

Assimilation methods for lightning data (1)

* **Nudging:** Newtonian relaxation of the model towards a proxy for lightning, x_{obs}:

$$\left(\frac{dx}{dt}\right)_{nudg} = \alpha\,(x_{obs} - x)$$

where x denotes the profile of latent heating rate or relative humidity or temperature, and α is the relaxation coefficient.

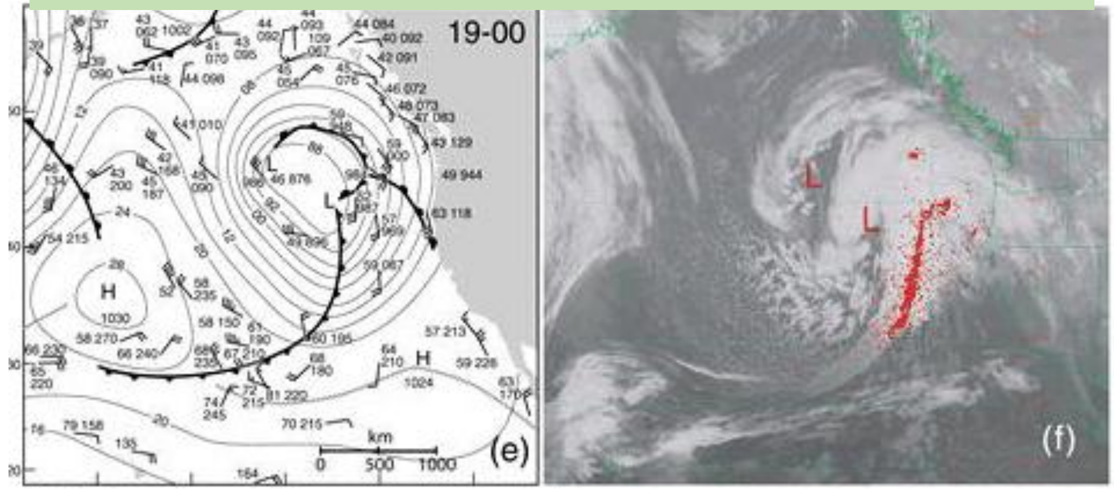
☺ Simple to apply.
☺ Possible to initiate convection from an initially non-convective state.

☹ Rather empirical and usually applicable to a small number of observation types.
☹ Relaxation factor α cannot be too large to avoid the generation of acoustic waves in the model.
☹ Spurious noise from very large increments can appear when convection does not pre-exist.
☹ Nothing ensures the dynamical consistency of increments in x.

Examples: Papadopoulos *et al.* (2005), Mansell *et al.* (2007), Pessi and Businger (2009), Fierro *et al.* (2012), Marchand and Fuelberg (2014).

Assimilation methods for lightning data (2)

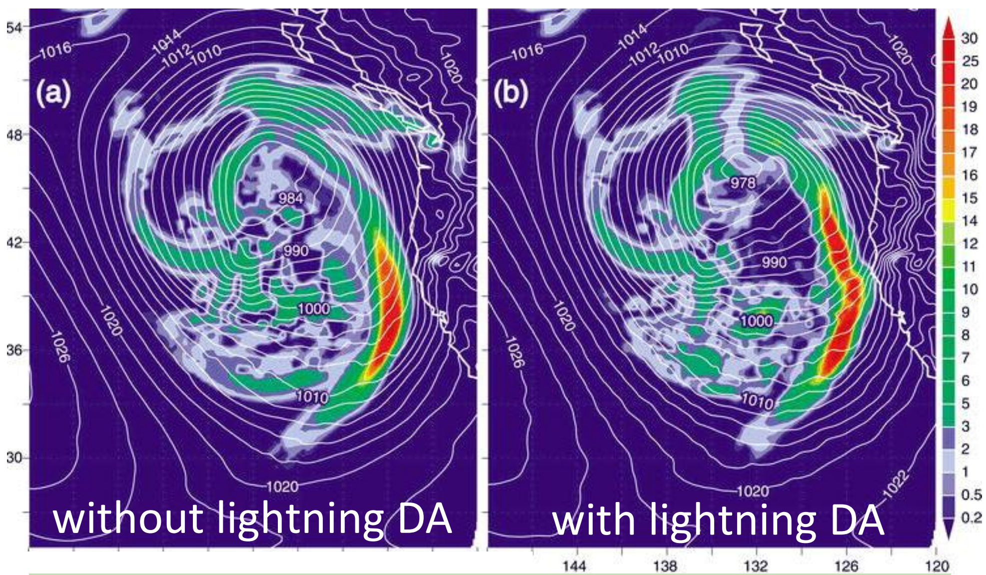
Situation at 00Z 19 December 2002 over Northeast Pacific



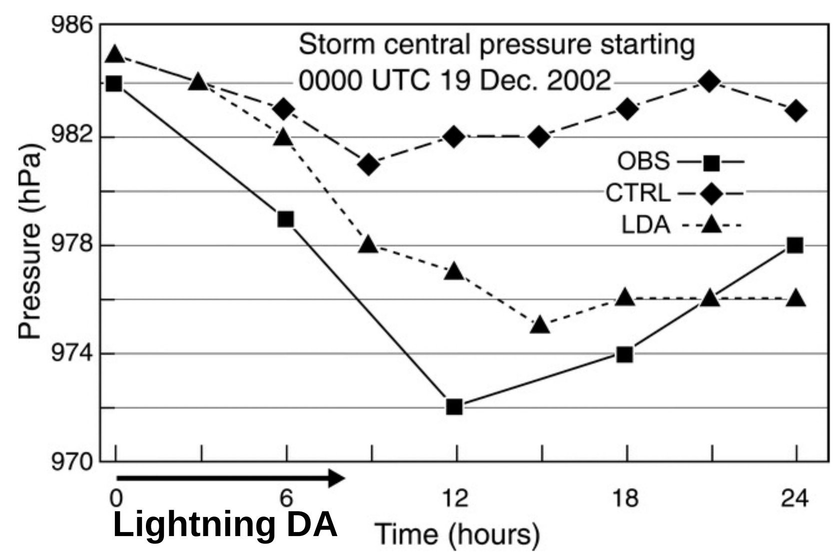
Assimilation of PacNet lightning flash rates using **Latent Heat Nudging** (rain as proxy).

MM5 forecast started at 00Z 19 Dec 2002.

from Pessi and Businger (2009)



Impact of LDA on 3h accum. surface precipitation forecast valid at 12Z 19 December 2002



Assimilation methods for lightning data (3)

* **Variational data assimilation** (3D-Var, 4D-Var): Optimal model state, $\mathbf{x}$, obtained by minimizing the following cost function:

$$J = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \mathrm{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \sum_{n=1}^{nsteps} \frac{1}{2}(H_n(\mathbf{x})-\mathbf{y}_n)^T \mathrm{R}^{-1}(H_n(\mathbf{x})-\mathbf{y}_n) \quad \text{(4D−Var case)}$$

where $\mathbf{B}$ and $\mathbf{R}$ are error covariance matrices for model background and observations, resp. $\mathbf{x}_b$ is the model background state, $\mathbf{y}_n$ the observations and H_n the observation operator.

☺ Multiple types of observations can be handled simultaneously, on the global scale.
☺ 4D-Var increments are dynamically well-balanced (via operator H_n).

☹ Variational DA relies on the linearity assumption during the minimization.
☹ It also requires the coding of tangent-linear and adjoint versions of operator H_n.

Examples: Fierro *et al.* (2014; 3D-Var), Stefanescu *et al.* (*submitted* 2013; 1D+4D-Var).

ECMWF

Assimilation methods for lightning data (4)

* **Ensemble Kalman Filter** (EnKF): Recursive filter based on an ensemble approach to propagate the model state ($\mathbf{x}$) and to estimate its error covariance matrix ($\mathbf{B}_n$).

$$\bar{\mathbf{x}}_n^b = M_n(\bar{\mathbf{x}}_{n-1}^a) \qquad \text{(forecast: time step n–1} \rightarrow \text{time step n)}$$

$\mathbf{B}_n$ is computed over the ensemble members.

$$\bar{\mathbf{x}}_n^a = \bar{\mathbf{x}}_n^b + \mathbf{K}_n(\mathbf{y}_n - H\bar{\mathbf{x}}_n^b) \quad \text{with} \quad \mathbf{K}_n = \mathbf{B}_n\mathbf{H}^T(\mathbf{R} + \mathbf{H}\mathbf{B}_n\mathbf{H}^T)^{-1} \quad \text{(analysis step)}$$

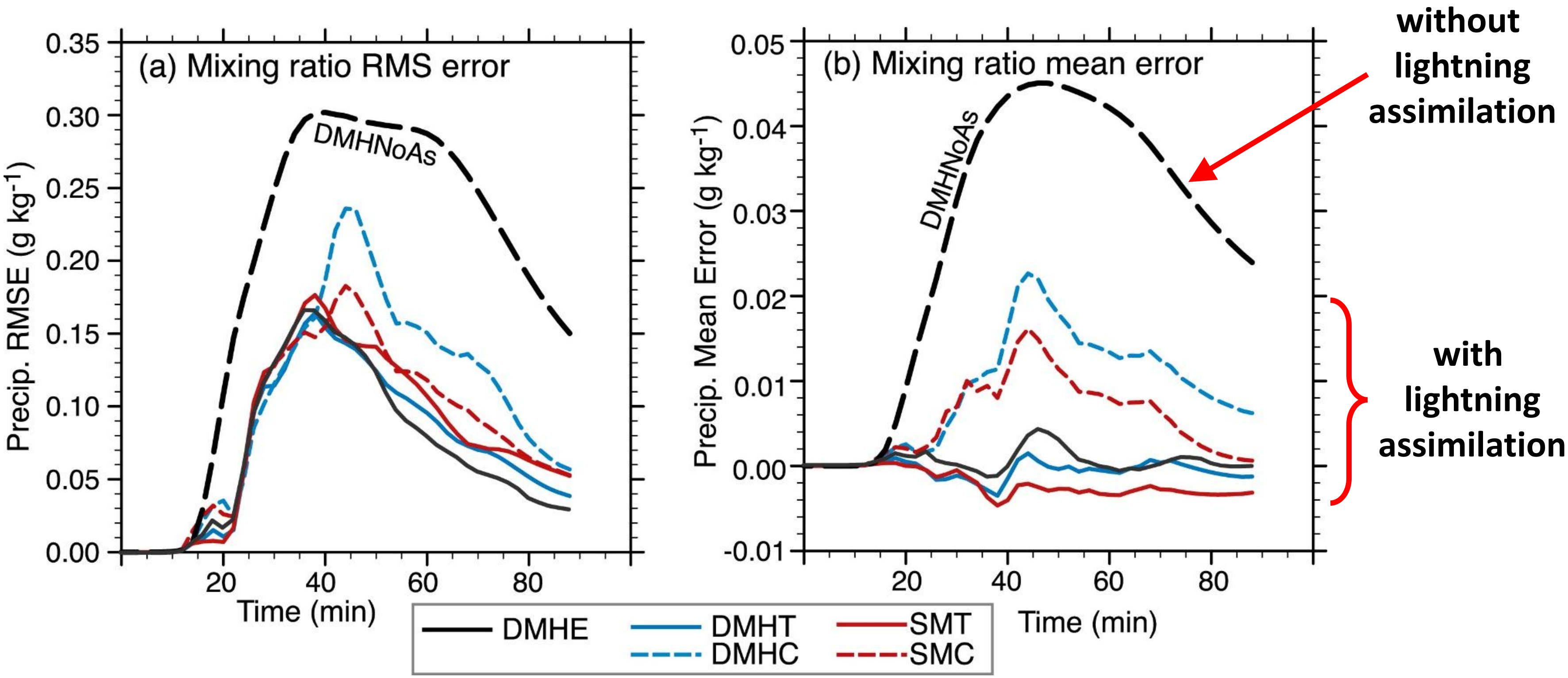
☺ No need for linearized versions of the forecast model (unlike 4D-Var).
☺ Operators may be non-linear (unlike 4D-Var).

☹ Still some underlying linearity assumption.
☹ Running enough members in the ensemble can be expensive.
☹ Some localization needs to be applied to reduce computational cost.
☹ No dynamical consistency of analysis increments (unlike 4D-Var).
☹ The model error covariances from the ensemble are often underestimated (too low spread), which can lead to a drift of the EnKF solution (unless some inflation factor is used).

Examples: Allen and Mansell (2013), Mansell *et al.* (2014).

Assimilation methods for lightning data (5)

Assimilation of GOES-GLM proxy lightning flash rates using an Ensemble Kalman Filter.
Single idealized Observing System Simulation Experiment from Mansell (2014).



Time evolution of precipitation mixing ratio RMS and mean errors with and without lightning data assimilation, and using either a "perfect" (DMH*) or "imperfect" model (SM*).

DMHE included an explicit parameterization of electrification processes. DMHT/C and SMT/C used two statistical fits between lightning flash rate and graupel volume.

Assimilation of lightning observations: potential issues (1)

* In 4D-Var, the linearity assumption might not be very valid for lightning processes.

→ Better to assimilate information averaged over a few hours rather than instantaneous observations?

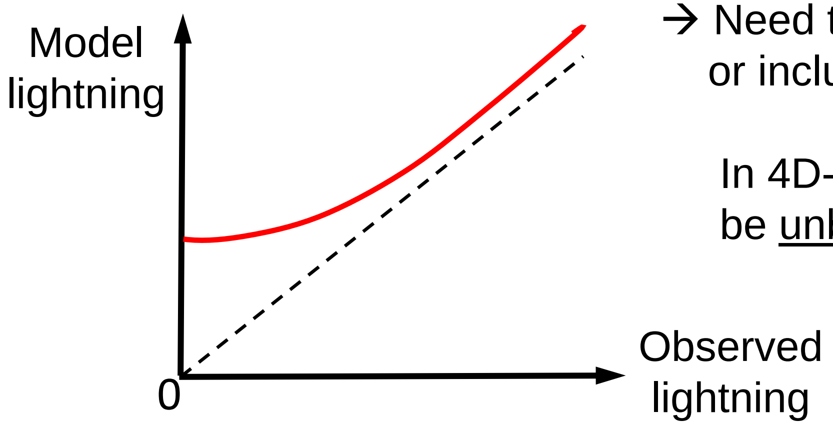
* In 4D-Var and EnKF, the model background and observation errors should be Gaussian.

→ Some transform might need to be applied to lightning data before their assimilation.

* The possible misplacement of convective cells in the model will make the assimilation of lightning data problematic (difficult to move cloud systems in the analysis).

→ Time-averaging might help.

* Limited lightning detection efficiency can bias the analysis.



→ Need to apply a bias-correction procedure to observations or include detection efficiency in model's lightning simulator.

In 4D-Var, model and observations errors are assumed to be unbiased.

Assimilation of lightning observations: potential issues (2)

* **“0-obs” case:**

The ambiguity between “undetected” and true “no-lightning” observation could be dangerous in DA.

→ However, rejecting those cases only is likely to bias the analysis.

* **“0-model” case:**

If the model background state has no lightning (4D-Var) or if none of the ensemble members have lightning (EnKF), the local gradient of lightning with respect to the model’s input variables is zero and the assimilation can do nothing.

On the other hand, the nudging approach may lead to excessive adjustments in order to trigger convection from scratch.

→ Should both “0-obs” and “0-model” cases be discarded from the assimilation?
If so, however, the impact of lightning data in the analyses might be rather limited.

* Is it better to include prognostic variables for graupel and hail in the DA process?

Summary and prospects

* There has been a growing interest in assessing the feasibility and potential benefits of assimilating lightning observations in NWP systems over the last decade.

* So far, all studies have focused on the assimilation of lightning data in high-resolution limited-area NWP models for individual convective cases, with some success.

* With the advent of GOES-GLM and MTG-LI, lightning DA should also be investigated with global coarser-resolution models (including its impact on medium-range forecasts, typ. over a few days).

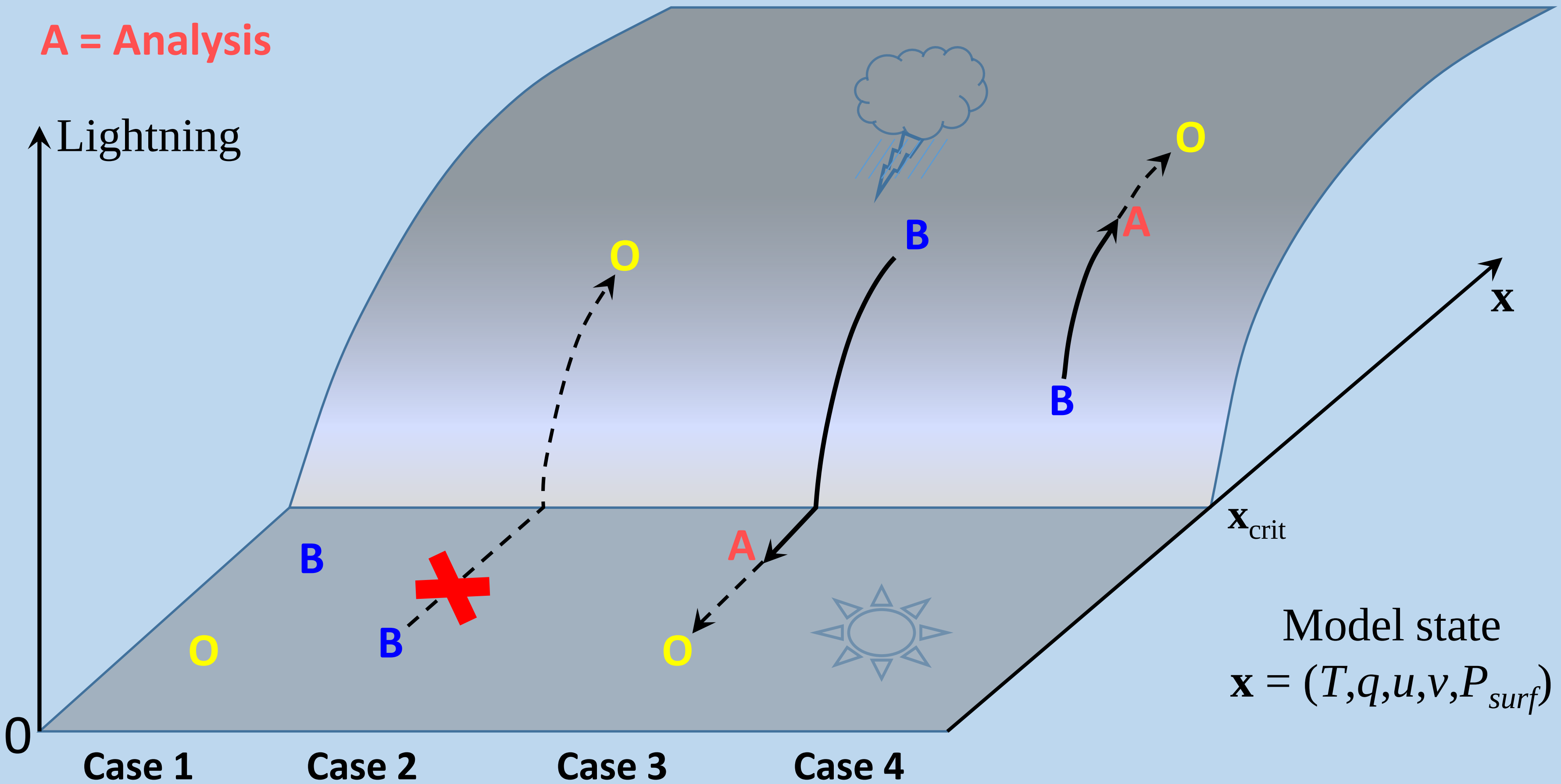
* Important issues remain which are related to:
 - the choice of the most appropriate DA method (nudging, *3*D-Var, 4D-Var, EnKF).
 - the selection of model variables to be adjusted.
 - the great uncertainty in the relationships between lightning and other meteorological variables.
 - the inclusion of information about lightning observation errors and detection efficiency.
 - "no-lightning" cases in obs (ambiguous) or in model (no sensitivity or too strong adjustment).

* Lightning DA might also improve the analysis of NOx concentrations in the atmosphere, provided better chemistry parameterizations and better estimations of the vertical distribution of lightning energy release become available.

Thank you!

The “0-lightning” issue (in obs or NWP model)

B = Background
O = Observation
A = Analysis



$$\frac{\partial \mathrm{Lightning}}{\partial \mathbf{x}} = 0$$

Case 1 is irrelevant (model = obs = 0).

Case 2 cannot work (no sensitivity in NWP model).

Case 3 is ambiguous (undetected vs no-lightning).

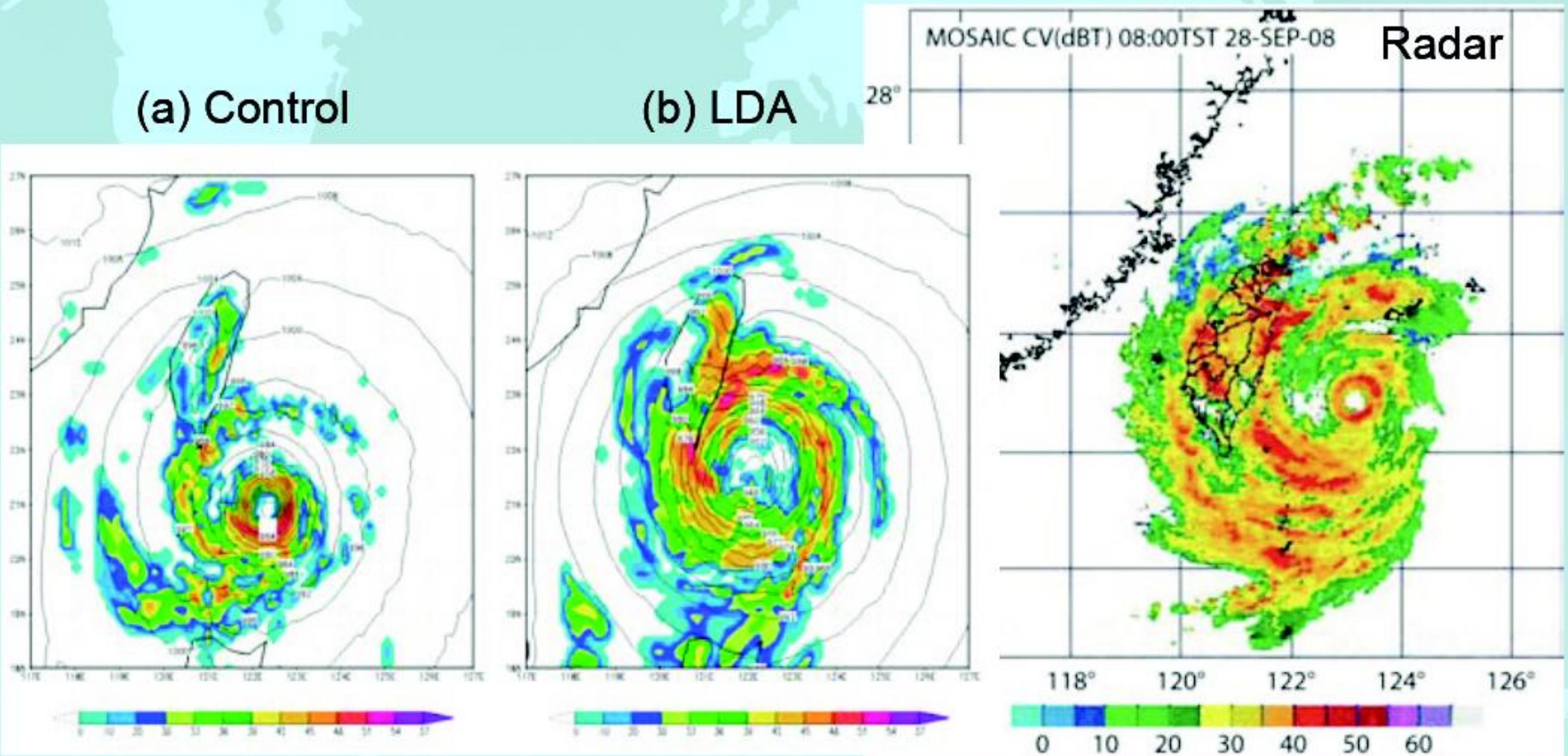
Case 4 is the only safe case.

Lightning data assimilation using latent heat nudging method

(Courtesy of Steven Businger 2010)

Typhoon Jangmi (near Taiwan)

Impact on MSL pressure and reflectivity in 36h forecast (WRF model)



Sea-level pressure and simulated radar reflectivity at 0000 UTC 28 September. (a) Control run, (b) LDA run. The model was initialized with a bogus vortex. The central pressures of the control and LDA runs were 934 and 945 hPa. Observed value was 926 hPa.

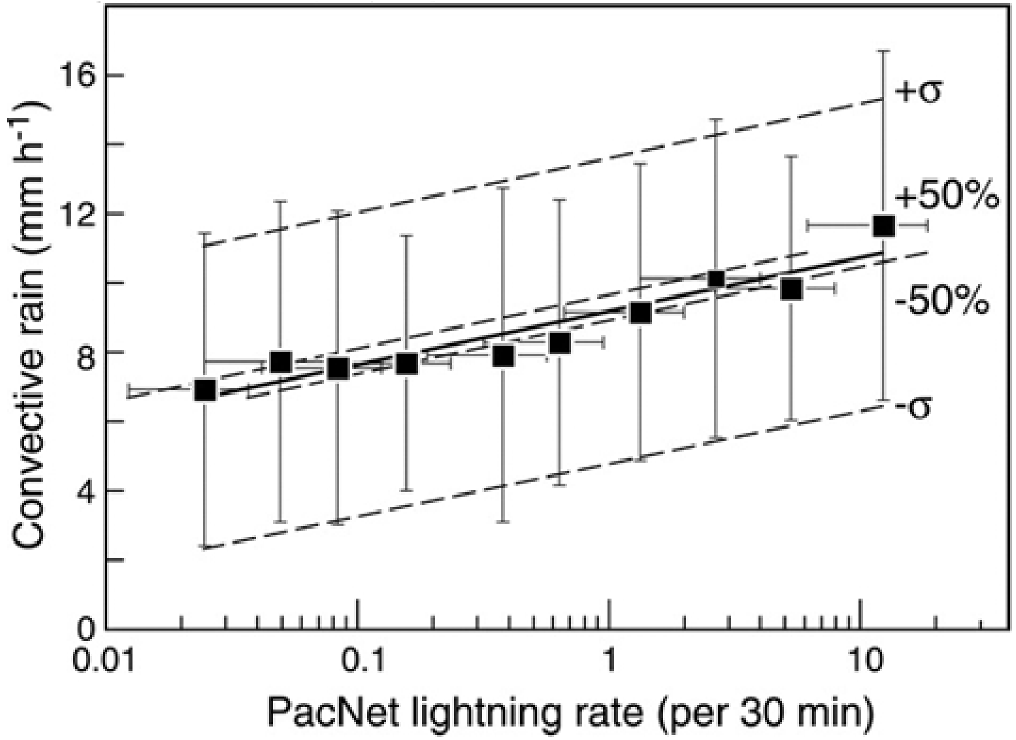
Approaches to compare model with lightning observations (1)

Two possible opposite approaches:

☹ The **statistical regression** is not universally valid (region/regime dependent) and thus difficult to apply in a global NWP model.

Approaches to compare model with lightning observations (2)

Example of approach #1: Regression of precipitation on lightning flash rate.



Regression of TRMM convective rain on PacNet lightning flash rate in the North Pacific (from Pessi and Businger 2009).

Approaches to compare model with lightning observations (3)

Two possible opposite approaches:

1 Lightning obs $\xrightarrow{\text{Statistical regression}}$ Precipitation (proxy) $\xleftrightarrow{\text{Comparison}}$ NWP model precipitation

☹ The **statistical regression** is not universally valid (region/regime dependent) and thus difficult to apply in a global NWP model.

2 NWP Model outputs $\xrightarrow{\text{Observation operator}}$ Simulated lightning $\xleftrightarrow{\text{Comparison}}$ Lightning obs

The **observation operator** can be:

* a simple parameterization of lightning (for global models) : Flash_Rate = f(predictors).

Predictors can be convective precipitation, cloud top height or depth, CAPE, updraught vertical velocity, graupel or cloud ice concentrations (e.g. Price and Rind 1994; Kurz and Grewe 2002, McCaul *et al.* 2009, Dahl *et al.* 2011).

* a more complex lightning simulator describing cloud electrification (suitable for high-resolution cloud-resolving models; e.g. Mansell *et al.* 2005; Barthe and Pinty 2007) .