ECOLE POLYTECHNIQUE
PROMOTION X2011
Tom BEUCLER

RESEARCH INTERNSHIP REPORT

Self-Aggregation Phenomenon in Cyclogenesis

NON CONFIDENTIAL

Option	: MEC/PHY 596
Internship Coordinator	: Pr Hervé LE TREUT
Internship Supervisor	: Pr Kerry EMANUEL
Internship Dates	: 31/03/2014 - 18/07/2014
Host Laboratory	: Dept. of Earth, Atmospheric and Planetary Sciences Massachusetts Institute of Technology 77 Massachusetts Avenue Cambridge, MA 02139-4307

Contents

Acknowledgments

Working on this project has an exciting and intense experience. Although I was working around the clock, I ultimately learned that stopping a bit and taking a step back are necessary conditions to stay creative and keep a good understanding of the underlying physics of the studied phenomena.

To begin with, I would like to thank my advisor Dr. Kerry Emanuel. I am always amazed by his insight on the physics of Cyclogenesis and his ability to improve my own understanding of the subject. His advice combined an in-depth knowledge of fluid dynamics and thermodynamic equations with an impressive meteorological sense of the different storms I was working on. I hope that the next five years of my PhD at MIT will help me better refine these skills. He also taught me to be very independent in my research and to constantly question and improve upon my project, which are, in my opinion, two of the most essential qualities for a successful researcher.

I would also like to thank my French professors: Hervé Le Treut, Thomas Dubos, Albert Hertzhog, and Riwal Plougonven, who introduced me to the fantastic field of atmospheric and oceanic science. They were the ones that made this internship possible, for which I will always be grateful.

I would also like to thank Dr. Alison Wing. Her impressive PhD work on self-aggregation opened a lot of doors, and made my internship a reality. Even when she was swamped in work, she always took time to answer all my questions, even the most elementary ones, in a very supportive way. Her postdoc work at Columbia University will further research about Cyclogenesis, and will hopefully allow us to continue this project.

My internship would have been impossible without the help and the advice of the many students and post docs of this department, and I would like to thank them all. I would especially like to thank the people from my group who have always been welcoming and friendly: Dr. Emanuel Vincent, Morgan O'Neil, Timothy Cronin, Vincent Agard, Sandy Shedd, Diamilez Peret-Betancourt. I would also like to thank the administrative assistants Jennifer Sell, Jacqueline Taylor, and Robyn Jepson, who helped me get oriented in my office, better understand the department, and countless other things that made my research possible.

Finally, I would like to thank my office mates, without whom I would never have survived through my internship: Zuchang Zhang and Ruyi Wei. I don't think I would have worked half as much if they hadn't been so present and supportive. A special thank goes to David Hibbert, who has been there for me during this whole internship, and who helped me adjust to living in America.

Definitions and Notations

b	Physical quantity
c_p	Specific Heat at constant pressure
c_{p_d}	Specific Heat of dry air at constant pressure
c_{p_v}	Specific Heat of water vapor at constant pressure
$\frac{D}{Dt}$	Material Derivative with respect to time
$\frac{\partial}{\partial t}$	Partial Derivative with respect to time
DIA	Diabatic term in MSE Variance budget
DYN	Dynamical term in MSE Variance budget
e_c	Kinetic energy
e_{lat}	Latent Heat
$ECMWF$	European Center for Medium-Range Weather Forecasts
g	Gravity constant
H	Integral Moist Static Energy of the Moist Column
h	Moist Static Energy
L_{vap}	Latent Heat of Vaporization
LHF	Latent Heat Flux
IFS	Integrated Forecast System
M_{air}	Molar mass of dry air
M_{H_2O}	Molar mass of water vapor
MCS	Mesoscale Convective System
MSE	Moist Static Energy
$NetLW$	Net Long Wave Radiation
$NetSW$	Net Short Wave Radiation
p	Pressure
p_b	Pressure at Bottom Level
p_t	Pressure at Top Level
PV	Potential Vorticity
$\dot{Q}$	Diabatic heating per time unit per mass unit
q	Specific humidity
R	Mean Earth radius
r^2	Correlation coefficient
r_d	Specific gas constant of dry air
r_v	Specific gas constant of water vapor
RCE	Radiative-Convective Equilibrium
$RESID$	Residual of the MSE variance budget
(S)	Surface of the Moist Column
SAM	System for Atmospheric Modeling
SHF	Sensible Heat Flux
T	Temperature
t	Time
TOA	Top of the Atmosphere

u	Zonal component of the velocity
$\underline{u}$	Velocity vector
$\underline{u_2}$	Horizontal velocity vector
$\underline{u_{abs}}$	Absolute velocity vector
$\underline{u_{box}}$	Box velocity vector
v	Meridional component of the velocity
w	Vertical component of the velocity
x	Zonal horizontal coordinate
$\underline{x}$	Position vector
y	Meridional horizontal coordinate
$YOTC$	Year of Tropical Convection
z	Vertical coordinate
z_b	Altitude at Bottom Level
z_t	Altitude at Top Level
Δ	Variation operator
ζ	Vertical component of the vorticity
Θ	Virtual potential temperature
θ	Latitude
λ	Longitude
Π	Potential vorticity
ρ	Density
ρ_d	Density of dry air
ρ_v	Density of water vapor
τ	E-folding time of Moist Static Energy variance
ϕ	Geopotential
ω	Vertical velocity in pressure coordinates
$\underline{\omega_a}$	Absolute vorticity vector
$\underline{\Omega}$	Earth's angular velocity vector
$\underline{\nabla}$	Nabla operator
$\underline{\nabla_2}$	Horizontal Nabla operator

We introduce the following operators for a physical quantity b:

- Difference:
$$[b]_{z_b}^{z_t} = b(z = z_t) - b(z = z_b)$$
- Density-weighted Vertical integral (under hydrostatic approximation):
$$\hat{b} = \int_{z_b}^{z_t} \rho b dz = \int_{p_t}^{p_b} b \frac{dp}{g}$$
- Horizontal average:
$$< b >= \frac{1}{S_{domain}} \iint_{(Domain)} b dx dy$$
- Horizontal perturbation:
$$b' = b - < b >$$

- Vertical average:

$$\{b\} = \frac{g}{(p_b - p_t)}\hat{b}$$

- Vertical perturbation:

$$b'' = b - \{b\}$$

We will use the classical geophysical system of coordinate by using a Cartesian tangent plan at latitude θ, so that we will use the coordinates (x, y, z) associated to the spherical coordinates in pressure coordinates (λ, θ, p):

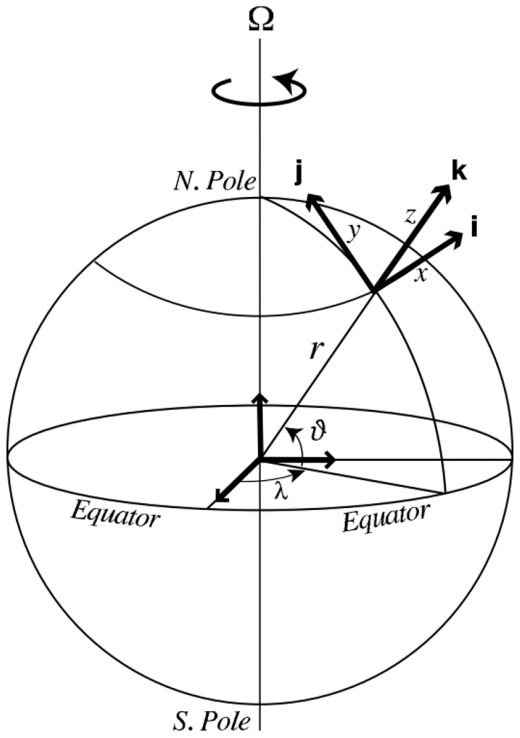

Figure 1: The spherical coordinates system. The orthogonal unit vectors $\underline{i}$, $\underline{j}$ and $\underline{k}$ point in the direction of increasing longitude λ, latitude θ, and altitude z. Locally, one may apply a Cartesian system with variables x, y and z measuring distances along $\underline{i}$, $\underline{j}$ and $\underline{k}$.

Introduction

Tropical cyclones, also referred to as typhoons in the Western North Pacific and hurricanes if their intensity is sufficient, are, with floods, the most lethal geophysical hazards. They can generate strong winds, storm surges, wind-driven waves, heavy rainfalls, flooding, and tornadoes, resulting in important direct damage (infrastructure loss, crops destruction, coastal erosion) as well as indirect ones (water contamination, oil price increases). For instance, tropical cyclone Bhola killed nearly half a million of people in Bangladesh in 1970, and in 2005 Hurricane Katrina damages were estimated at $108 billion. With such societal and economic impacts, forecasts of tropical cyclones, and more generally of severe weather events, are vital to warn authorities and the public, allowing appropriate action to be taken. Not only is a precise knowledge of their strength and track required, but the event also has to be predicted as soon as possible for logistic reasons. Currently, a good order of magnitude of such predictions is 3-4 days, leaving much room for improvement.

Unfortunately, the intriguing physics of tropical cyclogenesis are only partially understood, and there is not even a universally accepted definition of what cyclogenesis is. From a very general point of view, one can apply the term to any system which undergoes a transition from a disturbance not driven by surface fluxes, to a more symmetric, warm-core storm with a low pressure center at the surface. Over the past fifty years, positive factors (warm sea surface temperature, high tropospheric relative humidity, large enough values of the Coriolis parameter, positive wind-surface flux feedbacks) and inimical factors (mostly wind shear) to tropical cyclogenesis have been established, and numerically and empirically verified. According to Kerry Emanuel's work in the 90s, a necessary condition for cyclogenesis to occur is the establishment of an order 100km-wide column of nearly saturated air, to prevent disruptive low entropy downdrafts to occur. The question of how such a moist region is produced is thus of capital importance. Recently, the "Marsupial Pouch" paradigm has emerged: The genesis tends to occur near the critical layer of a tropical wave. This layer is a region of closed circulation which protects the growing disturbance from its surrounding environment, by preventing the entrainment of dry air and allowing moistening by convection, resulting in such a moist column.

In parallel, in the past twenty years, numerous studies of convection have shown that under certain conditions, a phenomenon called self-aggregation can occur. Usually, when cloud-resolving simulations of convection in radiative-convective equilibrium (RCE, ie a state where radiative cooling is balanced by convective heating) are run, convection is nearly random in space and time. However, when certain conditions (mainly depending on the sea surface temperature and the domain size) are met, the convection self-organizes into a single convecting moist cluster, surrounded by a broad region of dry subsiding air. Such a typical behavior of the tropical atmosphere has been confirmed by numerous numerical experiments using different models. From the observational point of view, order 100km-wide cloud clusters have been long known in mid latitudes as mesoscale convective systems (MCS). However, squall lines ($\approx$ 10km), tropical cyclones ($\approx$ 1000km) and perhaps the Madden Julian Oscillation ($\approx 10^4$km) can also be regarded as self-aggregation phenomena.

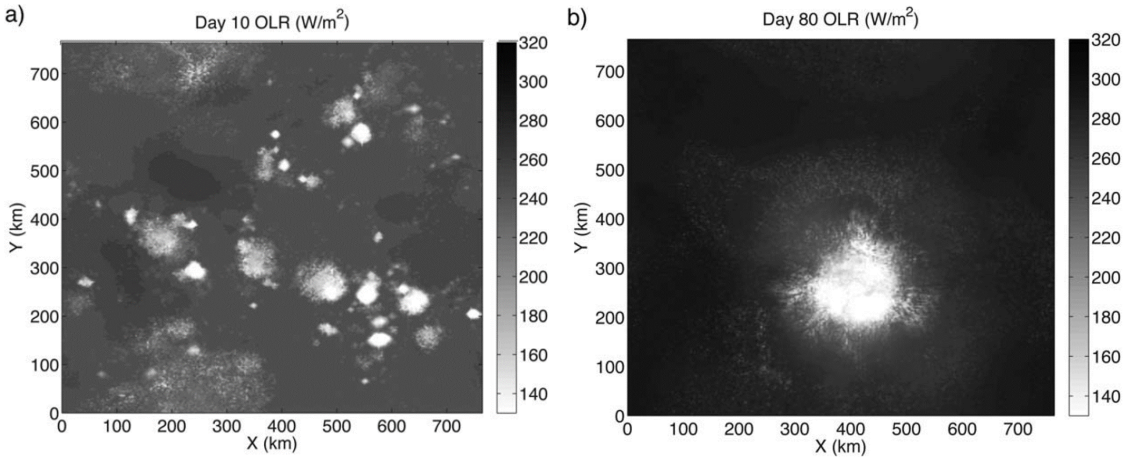

Figure 2: Snapshot of Outgoing Long Wave Radiation (OLR) at day 10 and 80 of an RCE simulation. Credits go to Alison Wing.

Consequently, it is not unreasonable to think that the physics of cyclogenesis are similar to those governing self-aggregation. In this work, we are address the following questions:

1. How does cyclogenesis compare to idealized self-aggregation phenomena?

2. How important are radiative, advective, and surface feedbacks during cyclogenesis?

During her PhD, Dr Alison Wing developed a new analysis technique to quantify the various radiative and surface feedbacks essential to self-aggregation in RCE numerical simulations. She worked with a 3D cloud resolving model well adapted to model the tropical atmosphere, called System for Atmospheric Modeling (SAM). This work adapts her technique to 10-days forecasts, produced by the European Center for Medium-Range Weather Forecasts (ECMWF), to get more insight into tropical cyclogenesis.

We will first establish an adapted framework to analyze self-aggregation on a moving domain, and then use it study the cyclogenesis of multiple storms.

1 Self-aggregation on a moving domain

Wing's analysis technique allows one to quantify the radiative and surface feedbacks on a fixed domain with periodic boundary conditions. In this part, we adapt this framework to a moving domain following the cyclone in the ECMWF model.

1.1 Derivation of the Moist Static Energy variance budget

1.1.1 Definition of the Moist Static Energy

We are interested in the Moist Static Energy (MSE), ie the total enthalpy of the air mass [1] plus its latent heat, assimilated to an ideal homogeneous mix of air and water vapor, and which is conserved during moist adiabatic motion:

$$h = \phi + e_t + e_{lat}$$

with:

- $\phi = gz$ the geopotential, with $g = 9.80665 m.s^{-2}$ the Earth's mean gravity constant and z the geopotential height.
- $e_t = c_p T$ the internal energy of the moist air, involving the specific heat at constant pressure of a homogeneous mix of dry air and water vapor:
  $$c_p = c_{p_d} + (c_{p_v} - c_{p_d})q$$
  with the specific heat at constant pressure of dry air c_{p_d}, the specific heat at constant pressure of water vapor c_{p_v}, and the specific humidity:
  $$q = \frac{\rho_v}{\rho_d + \rho_v}$$
  Knowing that in standard atmospheric conditions: $c_{p_d} \approx 1 kJ.kg^{-1}.K^{-1}$, $c_{p_v} \approx 4 kJ.kg^{-1}.K^{-1}$ and $q < 3\%$, we can find an upper bound and a lower bound for the ratio:
  $$1 < \frac{c_p}{c_{p_d}} < 1.1$$
  so that within a 10% error we can assume:
  $$c_p \approx c_{p_d}$$
  Knowing a posteriori that variations in internal energy are more than one order of magnitude smaller than the variations in MSE, we make use of this approximation in the computation of MSE.
- $e_{lat} = L_{vap} q$ the latent heat, with $L_{vap} = 2.5.10^6 J.kg^{-1}$ the Latent heat of vaporization of water, assumed to be constant.

The local thermodynamic equation for MSE can be immediately derived from the first principle of thermodynamics, assuming hydrostatic balance and neglecting the temporal pressure fluctuations, as well as the kinetic energy contribution [2]:

$$\frac{Dh}{Dt} = \frac{\partial h}{\partial t} + \underline{\nabla h} \cdot \underline{u} = \dot{Q}$$

[1] The enthalpy is by definition the sum of the internal energy and the opposite of the work of the pressure forces. For an ideal gas air column in hydrostatic balance, it can be proven that the total opposite of the work of the pressure forces equals the total gravitational potential energy of the column.

[2] This is justified as we assume the air column to be motionless when we derive the vertically integrated MSE budget.

with $\dot{Q}$ the diabatic heating per mass unit and per time unit, which does not include phase transitions between liquid and vapor.

1.1.2 Vertically integrated moist static energy budget

We are interested in the MSE of an air column, delimited by the ocean or the ground at the bottom and the top of the atmosphere (TOA) at the top. The integral MSE budget for the air column is simply the volume integral of the local budget:

$$\frac{DH}{Dt} = \frac{D}{Dt} \iint_{(S)} \int_{(z)} \rho h dz d^2S = \iint_{(S)} \int_{(z)} \dot{Q} dz d^2S = \iint_{(S)} (DIA) d^2S$$

The total diabatic heating of this column can be modeled by 4 types of flux (DIA):

1. The Sensible Heat Flux from the ocean, which is the sum of all direct sources of diabatic heating:

   $$SHF$$

   where SHF is expressed in $W.m^{-2}$ (surface and radiative fluxes are accumulated energy fields in the ECMWF model).

2. The Latent Heat Flux from the ocean, which is the energy due to the latent heat of the vaporized water at the surface:

   $$LHF$$

   where LHF is expressed in $W.m^{-2}$.

3. The Long Wave (LW) radiative fluxes, which come from the terrestrial or thermal part of the spectrum (mostly Infrared). At the top of the atmosphere, the Net LW flux is defined as the difference between the downward and the upward radiative flux (the ECMWF convention is to count downward radiative fluxes as positive):

   $$(NetLW)(p = p_t) = (NetLW \downarrow)(p = p_t) - (NetLW \uparrow)(p = p_t)$$

   where p_t is the pressure at the top of the atmosphere, defined here as the highest level in the ECMWF model. We define similarly the Net LW flux at the surface:

   $$(NetLW)(p = p_b) = (NetLW \downarrow)(p = p_b) - (NetLW \uparrow)(p = p_b)$$

   where p_b is the pressure at the surface, defined here as the lowest level in the ECMWF model.

4. The Short Wave (SW) radiative fluxes, which come from the solar part of the spectrum (mostly visible); with the same conventions:

   $$\begin{cases} (NetSW)(p = p_t) = (NetSW \downarrow)(p = p_t) - (NetSW \uparrow)(p = p_t) \\ (NetSW)(p = p_b) = (NetSW \downarrow)(p = p_b) - (NetSW \uparrow)(p = p_b) \end{cases}$$

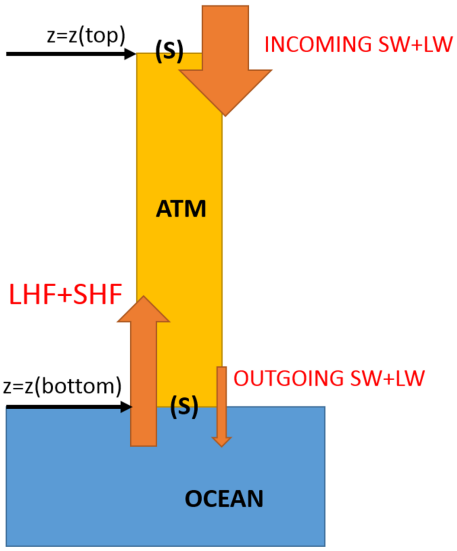

Figure 3: Diabatic fluxes for an air column delimited by the TOA and the ocean

Applying the integral budget on a motionless air column with surface (S), delimited by the ocean $(z = z_b) \leftrightarrow (p = p_b)$ and the top of the atmosphere $(z = z_t) \longleftrightarrow (p = p_t)$:

$$\frac{DH}{Dt} = \frac{D}{Dt} \iint_{(S)} \int_{z_b}^{z_t} \rho h dz d^2S = \iint_{(S)} (SHF + LHF + [NetLW]_{z_b}^{z_t} + [NetSW]_{z_b}^{z_t}) d^2S = \iint_{(S)} (DIA) d^2S$$

where we have used the notation:

$$[b]_{z_b}^{z_t} = b(z = z_t) - b(z = z_b)$$

and defined the total diabatic term:

$$DIA = SHF + LHF + [NetLW]_{z_b}^{z_t} + [NetSW]_{z_b}^{z_t}$$

Grouping all the terms under the surface integral, we can obtain a MSE budget per surface unit, which is equivalent to a vertically integrated MSE budget:

$$\frac{D\hat{h}}{Dt} = DIA$$

where we have defined the vertical density-weighted integral of a quantity b and assumed a hydrostatic atmosphere so that:

$$\widehat{b} = \int_{z_b}^{z_t} \rho b dz = \int_{p_t}^{p_b} b \frac{dp}{g}$$

with ρ the density of the atmosphere, linked to the pressure by the perfect gas law for a homogeneous mix of dry air and water vapor:

$$p = \rho r_d T_v = \rho [r_d + q(r_v - r_d)] T$$

with $r_d = 287.058 J.kg^{-1}.K^{-1}$ the gas constant of dry air, $r_v = 461.523 J.kg^{-1}.K^{-1}$ the gas constant of water vapor, and T_v the virtual temperature.

1.1.3 Moist Static Energy variance budget

The self-aggregation phenomenon can be seen as an increase in the horizontal variance of MSE: From a starting point where convection is nearly random (corresponding to zero spatial variance), self-aggregation concentrates all the moisture, and thus all the MSE (mostly in form of latent heat) in a single cluster, drastically increasing the horizontal variance of MSE. We thus define the horizontal 2D average of a physical quantity b :

$$< b >= \frac{1}{S_{domain}} \iint_{(Domain)} b dx dy$$

and the deviation to that average:

$$b' = b - < b >$$

We then subtract the averaged vertically integrated MSE budget to its average to obtain:

$$(\frac{D\hat{h}}{Dt})' = \frac{\partial \hat{h}'}{\partial t} + (\widehat{\underline{\nabla h} \cdot \underline{u}})' = DIA'$$

and multiply it by the MSE vertically integrated perturbation $\hat{h}'$ to obtain a budget for the MSE variance:

$$\frac{\partial}{\partial t} \frac{(\hat{h}')^2}{2} + \hat{h}'(\widehat{\underline{\nabla h} \cdot \underline{u}})' = \hat{h}' DIA'$$

As self-aggregation occurs, the variance of MSE increases, fed by 5 types of terms:

- The four radiative and surface feedback terms $\hat{h}' DIA'$. By definition, these terms represent horizontal correlation: if the MSE perturbation is positively correlated with one of the surface or radiative fluxes, the total term will be a "positive feedback" for the variance of MSE, enhancing the moist cluster formation. Thus, each physical feedback (LHF,SHF,NetLW,NetSW) can be important. Wing has shown that in idealized RCE numerical simulations, these terms are of equal order of magnitude, and they play different roles depending on the stage of self-aggregation.
- The advection correlation $-\hat{h}'(\widehat{\underline{\nabla h} \cdot \underline{u}})'$, which dominates the budget before the self-aggregation phenomenon. Once the cluster is formed and a steady state is reached, the budget for the MSE variance shows that this term will balance the diabatic feedback term.

1.2 Moist Static Energy variance budget on a moving domain

In the case of domain moving at speed $\underline{u_{box}}$, the budget still applies if we consider the relative speed:

$$\underline{u} = \underline{u_{abs}} - \underline{u_{box}}$$

where $\underline{u_{abs}}$ is the absolute velocity of the air in the domain, that we can commonly refer to as wind speed. Consequently, both the increase of the MSE variance and the advection correlation term strongly depend on $\underline{u_{box}}$, preventing us from directly studying the self-aggregation phenomenon. We thus look for a Galilean-invariant version of the budget, ie a budget that does not vary if a constant vector is added to $\underline{u}$, and so which does not depend on the domain translation speed.

1.2.1 Pressure coordinates

The obtention of such a budget is easier if we work in pressure coordinates, and working in these coordinates seems natural as the ECMWF model vertical levels are pressure levels. We thus introduce the vertical velocity in pressure coordinates ω in $Pa.s^{-1}$:

$$\omega = \frac{Dp}{Dt}$$

Assuming a hydrostatic field, a mass budget on an elementary volume allows us to see that the velocity in pressure coordinates is divergent free:

$$\underline{\nabla} \cdot \underline{u} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial \omega}{\partial p} = 0$$

We can also use the approximate relation to link the vertical velocity in full Cartesian coordinates to the vertical velocity in pressure coordinates:

$$\omega = \frac{Dp}{Dt} = \frac{Dp}{Dz}\frac{Dz}{Dt} \approx -\rho g w$$

if the pressure field is assumed to be steady and independent of the horizontal coordinates (ie $\frac{Dp}{Dz} = \frac{\partial p}{\partial z}$).

1.2.2 Derivation of the Galilean invariant MSE budget

We assume that rigid boundary condition are satisfied at top and bottom levels:

$$\omega(p = p_b) = \omega(p = p_t) = 0$$

This leads us to rewriting the advection term as:

$$\underline{\nabla h} \cdot \underline{u} = \underline{\nabla} \cdot (h\underline{u}) = \underline{\nabla_2} \cdot (h\underline{u_2}) + \frac{\partial (h\omega)}{\partial p}$$

where we use the horizontal gradient operator:

$$\underline{\nabla_2} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix}$$

Applying the vertical density-weighted integral:

$$\widehat{\underline{\nabla h} \cdot \underline{u}} = \underline{\nabla_2} \cdot \widehat{(h\underline{u_2})} + [\frac{h\omega}{g}]_{p_t}^{p_b} = \underline{\nabla_2} \cdot \widehat{(h\underline{u_2})}$$

We develop the integrand, looking for a horizontal and a vertical advection terms:

$$\underline{\nabla_2} \cdot (h\underline{u_2}) = \underline{\nabla_2 h} \cdot \underline{u_2} + h(\underline{\nabla_2} \cdot \underline{u_2}) = \underline{\nabla_2 h} \cdot \underline{u_2} - h\frac{\partial \omega}{\partial p}$$

We finally decompose the horizontal speed into its vertical average and a vertical perturbation:

$$\begin{cases} \underline{u_2} = \{\underline{u_2}\} + \underline{u_2}'' \\ \{\underline{u_2}\} = \frac{g}{(p_b - p_t)}\widehat{\underline{u_2}} \end{cases}$$

This leads to the following vertical density-weight integral advection term:

$$\underline{\widehat{\nabla_2 h}} \cdot \underline{u_2} = \underline{\widehat{\nabla_2 h}} \cdot \{\underline{u_2}\} + \underline{\nabla_2 h} \cdot \underline{u_2}'' - h\widehat{\frac{\partial \omega}{\partial p}} = \underline{\widehat{\nabla_2 h}} \cdot \{\underline{u_2}\} + \underline{\nabla_2 h} \cdot \underline{u_2}'' + \omega\widehat{\frac{\partial h}{\partial p}}$$

Noticing that:

$$\underline{\nabla_2 \widehat{h}} = \underline{\nabla_2 \widehat{h}'}$$

We can write our initial budget with three Galilean invariant terms:

$$\widehat{\frac{D}{Dt}}(\frac{\widehat{h}'^2}{2}) = \widehat{h}' DYN' + \widehat{h}' DIA'$$

with:

- $\widehat{\frac{D}{Dt}} = \frac{\partial}{\partial t} + \widehat{\{\underline{u_2}\}} \cdot \underline{\nabla_2}$ a horizontal, vertically integrated equivalent of the material derivative, which is Galilean invariant because it is the sum of the two following Galilean invariant terms:
- $DYN = -\underline{u_2}'' \cdot \widehat{\underline{\nabla_2 h}} - \omega\widehat{\frac{\partial h}{\partial p}}$ a dynamical term, which is Galilean invariant as we can see by adding a constant vector to $\underline{u_2}$.
- $DIA = SHF + LHF + [NetLW]_{z_b}^{z_t} + [NetSW]_{z_b}^{z_t}$ the total diabatic term.

1.3 Adaptation to the ECMWF data

The ECMWF forecast model gives outputs on a Gaussian regular grid with resolution (0.5^olon,0.5^olat), and on 25 pressure levels, which leads us to work in a locally Cartesian frame in pressure-spherical coordinates.

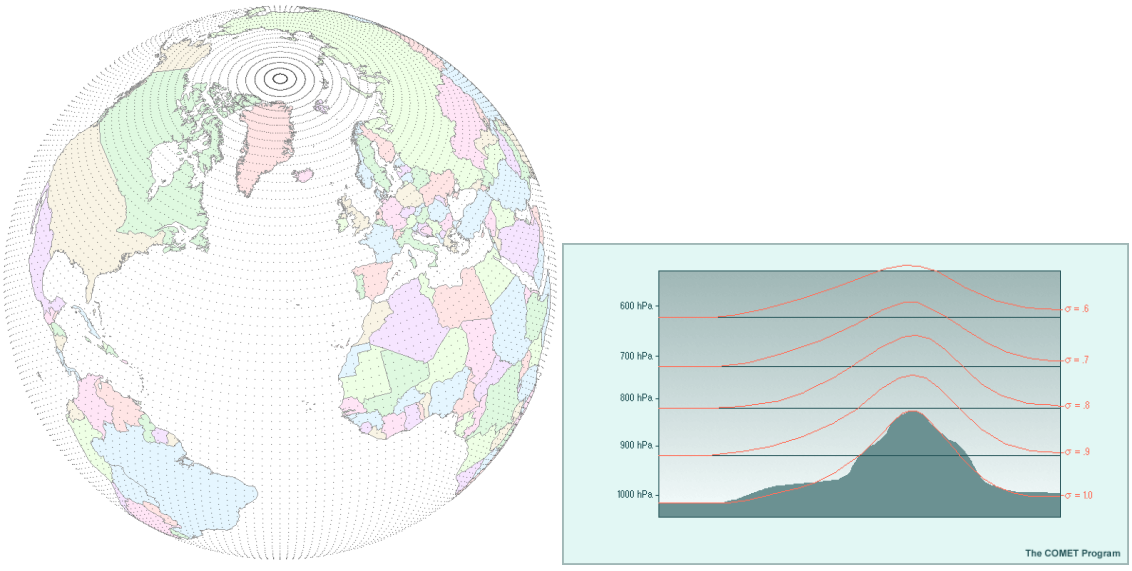

Figure 4: A Gaussian regular grid (left) and a typical view of sigma-coordinates, based on pressure levels (right)

In this subsection, we describe the different steps to apply the previously described framework, assuming we are already in possession of all the necessary physical quantities in the MSE budget. A description of the ECMWF model and its outputs is available in Appendix A, and the numerical discretization schemes are available in Appendix B and Appendix C, at the end of the report.

1.3.1 Definition of the moving domain

The first step to obtain the moving domain is to define its center, which should be close to the "center of the cyclone". For that purpose, we use the potential vorticity at level $p = 850hPa$ (PV850):

- Potential Vorticity is conserved during frictionless moist adiabatic motion, and high values of the potential vorticity is a good indicator of a strong cyclonic flow.
- This level is typically above the boundary layer and thus not subject to diurnal variation, and is classically used to follow air masses (eg cold and warm fronts), it is thus well-adapted to track cyclones.

As cyclones are very easy to track once they are formed, we start at the final time step, and choose visually the center of the cyclone:

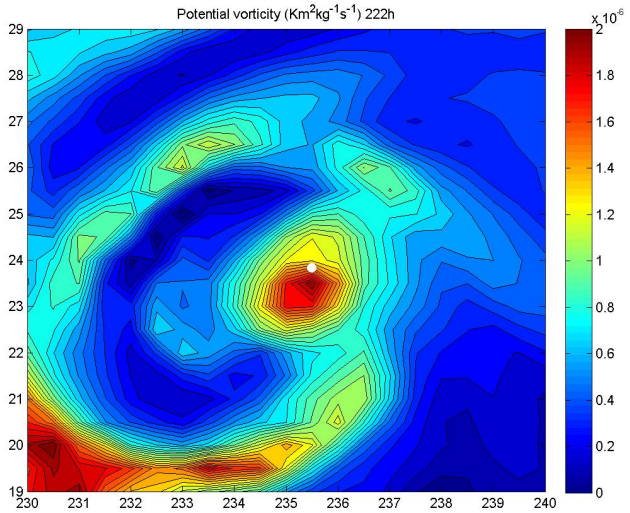

Figure 5: Potential vorticity at the final time step of the analysis (222h). The tracker is represented by the white dot at the center of the $(16^{o}\text{lon}, 16^{o}\text{lat})$ domain

Then going backwards in time, at each time step, we define the "center of the cyclone" as the barycenter of PV850 in a small domain (usually in the order of $(2.5^{o}\text{lon}, 2.5^{o}\text{lat})$ near the previous location of the "center"). We obtain tracks that are similar to the cyclone's Best Tracks, which can for example be found for the Atlantic case in the National Hurricane Center archives:

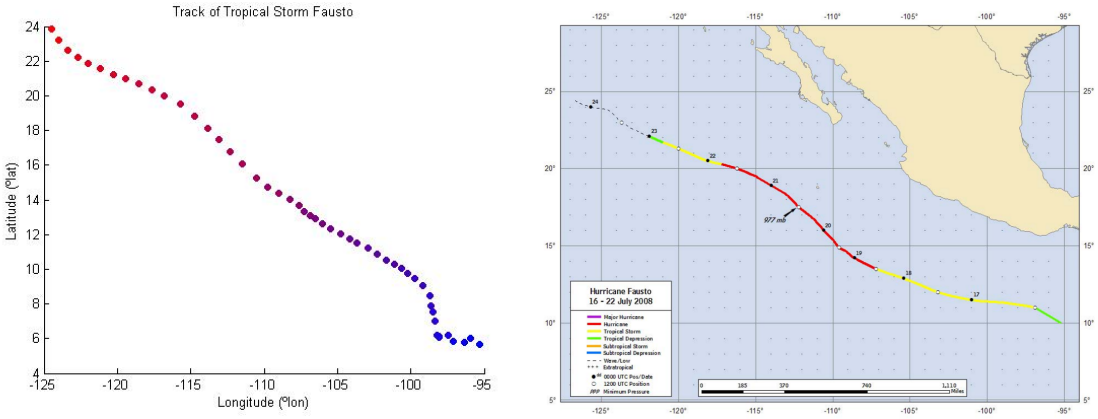

Figure 6: Trajectory of the PV tracker (left) vs Best track of Tropical Hurricane Fausto from the National Hurricane center (right)

We then define the domain as a rectangular box of grid points, which center is the previously defined "center of the cyclone". The domain is of constant size for a given analysis.

1.3.2 Computation of the MSE budget

- Horizontally, note that the "center of the cyclone" is not constrained to adopt integer values, which implies that the data has to be bilinearly interpolated, so that we can use a grid that corresponds to our domain. This simple numerical scheme is described in Appendix C.

Furthermore, the average $< b >$ is computed by taking an average over all grid points.

- Vertically, we will only be interested in the phenomena that happen under the tropopause, since stratospheric waves that can strongly perturb the geopotential ϕ and thus the MSE h are irrelevant if we want to focus on the self-aggregation phenomenon. Obtaining the following quasi-constant temperature profile for points at different places of the domain and different time steps:

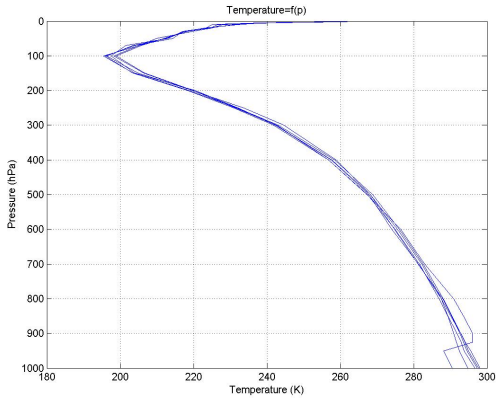

Figure 7: Temperature profile at different time steps and 4 distinct locations in the domain

We are led to consider:

$$\begin{cases} p_b = 1000hPa & Level\ i = 25/25 \\ p_t \approx 100hPa & Level\ i = 11/25 \end{cases}$$

and use the following discretization scheme:

$$\widehat{b} = \sum_{i=11}^{24} b(i) \cdot \frac{p(i+1) - p(i-1)}{2} + b(25) \cdot \frac{p(25) - p(24)}{2}$$

- This allows us to compute each quantity of the budget at each grid point and at each time step. We then horizontally average the budget on every grid point to obtain the domain budget:

$$< \widehat{\frac{D}{Dt}}(\frac{\widehat{h}'^2}{2}) >=< \widehat{h}' DYN' > + < \widehat{h}' DIA' >$$

- We generally apply a daily moving-average on all the physical variables to eliminate the diurnal cycle (it also removes some of the temporal noise that we do not wish to study).

2 Application to cyclogenesis

We apply the previously described framework to several cyclone cases that are listed in Appendix D. For all the studied cases, there is a fast MSE variance increase as the cyclone forms during the forecast. As expected, the variance is primarily due to the development of a moist and a dry regions, which have respectively higher and lower MSE. The differences in diabatic heating between those two regions then allows us to quantify the correlations between this heating and the MSE, resulting in the diabatic terms on the right-hand side of the variance budget.

2.1 Investigation of MSE variance increase

We investigate the MSE variance increase by taking the example of Tropical Hurricane Fausto. The storm was reasonably intense and it aggregated in one single phase, making it a good case study: it can be insightfully compared to the self-aggregation phenomenon in idealized RCE studies. We choose the domain size to be (20°lon,20°lat); the domain size sensitivity will be explored later in the report.

2.1.1 Magnitude and timescale of aggregation

We remember that the MSE has three components:

$$h = \phi + e_t + e_{lat}$$

with:

$$\begin{cases} \phi = gz \\ e = c_p T \\ e_{lat} = L_{vap} q \end{cases}$$

Observing the horizontally-averaged vertically integrated MSE variance, we can see on figure 8 that it monotonically increases from $< (\hat{h}')^2_{min} >$ to $< (\hat{h}')^2_{max} >$ with time during the self-aggregation period:

$$\begin{cases} < (\hat{h}')^2_{min} > = < (\hat{h}')^2 > (t = 90h) = 2.0.10^{14} J^2.m^{-4} \\ < (\hat{h}')^2_{max} > = < (\hat{h}')^2 > (t = 174h) = 1.3.10^{15} J^2.m^{-4} \end{cases}$$

giving typical variation rate and time scale:

$$\begin{cases} \frac{\Delta < (\hat{h}')^2 >}{\Delta t} \approx 10^{13} J^2.m^{-4}.h^{-1} \\ \Delta t \approx 100h \approx 4d \end{cases}$$

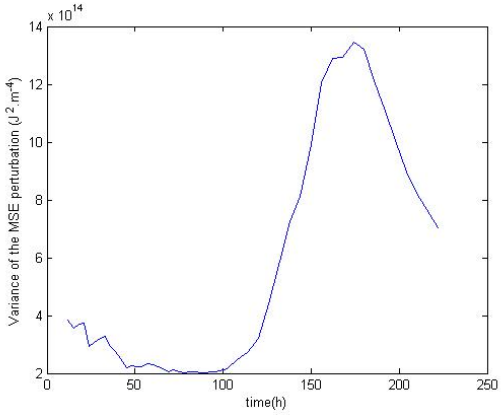

Figure 8: Horizontally averaged vertically integrated MSE variance vs time

We can compare our results to those obtained by Wing in SAM simulations. We plot the MSE variance in log-lin scale, and use a linear fit to determine a typical e-folding time.

Considering a fit of type:

$$< (\hat{h}')^2 >=< (\hat{h}')_0^2 > .e^{\frac{t}{\tau}}$$

we obtain the following coefficients in both cases:

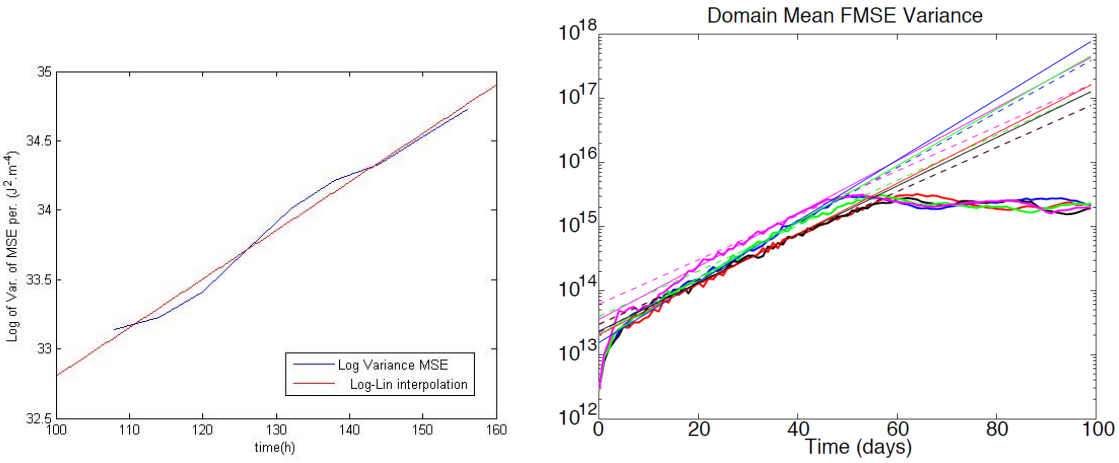

Figure 9: MSE variance in log-lin scale, linearly interpolated (left: Tropical storm Fausto; right: RCE numerical simulations at 305K)

	ECMWF Forecast	SAM Model at 305K
Start/End of the increase of variance(days)	3.75/7.25	5/55
Magnitude of MSE variance $< (\hat{h}')_0^2 > (J^2 m^{-4})$	$5.32.10^{12}$	$2.96.10^{13}$
E-folding time τ (days)	1.19	12.57
Correlation coefficient r^2	0.988	0.991

Several observations can be made:

- In the Cyclogenesis case, the self-aggregation is considerably faster (an order of magnitude faster, considering the e-folding times). This suggests that the mechanism is considerably different. We can assume that the main difference is the presence of the external environment in the ECMWF, which helps trigger the Cyclogenesis through the dynamical term $< \widehat{h}' DYN' >$, and emphasizes the latent heat flux feedback $< \widehat{h}' LHF' >$, as we will see later in the report.

- A very interesting point is that the magnitude of MSE variance $< (\hat{h}')_0^2 >$ is approximately 6 times smaller in the ECMWF case. Knowing that the domain size are:

$$\begin{cases} (20^{\text{o}}\text{lon}, 20^{\text{o}}\text{lat}) \approx (2200km, 2200km) & in\ ECMWF \\ (768km, 768km) & in\ SAM \end{cases}$$

  the total areas have an approximate ratio:

$$\frac{S(SAM)}{S(ECMWF)} \approx 8$$

  meaning that the self-aggregation is considerably more advanced in the RCE case, reaching a state that is not stable in the case of Cyclogenesis (due to external factors such as the extra-tropical transition). However, the orders of magnitude of the maximum MSE variance are similar, suggesting that the beginning of self-aggregation in RCE can still be compared to Cyclogenesis.

2.1.2 Decomposition of MSE variance

We can obtain a first fundamental information in the horizontally-averaged vertically integrated MSE variance increase by studying the variance of the three components of MSE:

$$< (\hat{e_{lat}}')^2 >, < (\hat{e_t}')^2 >, < (\hat{\phi}')^2 >$$

Of course, the $< (\hat{'})^2 >$ operator is strongly nonlinear so that $< (\hat{h}')^2 >$ is not the sum of the previous terms. However, their separate order of magnitude can help us understand the nature of the MSE variance increase:

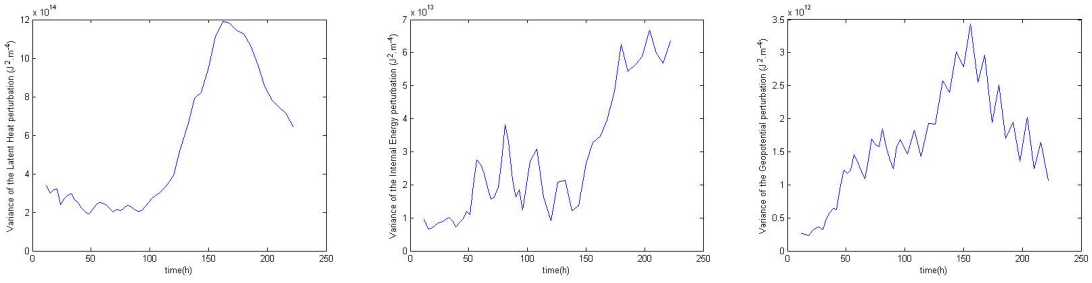

Figure 10: Horizontally averaged Vertically integrated Variance of Latent Heat (left), Internal Energy (center) and Geopotential (right)

- First we note that $< (\hat{e_{lat}}')^2 >$ and $< (\hat{h}')^2 >$ are extremely close, which means that increase of the MSE variance is mainly due to increase of latent heat variance, thus confirming the fact that self-aggregation of

moisture is the main energetic variance source. The maximum in latent heat variance being slightly smaller, we can deduce that the contribution from the two other terms have a positive effect on MSE variance.

- Indeed, the thermic term $< (\hat{e_t}')^2 >$ shows that the variance in internal energy also increases, confirming the formation of a warm-core center in the moist column if we look at $\hat{e_t}'$. However, it is considerably smaller than the previous term:

$$< (\hat{e_t}')^2 >_{max} = 6.10^{13} J^2.m^{-4} \approx \frac{1}{20} < (\hat{e_{lat}}')^2 >_{max}$$

- The variance of geopotential $< (\hat{\phi}')^2 >$ is even smaller, and is due to the decrease in altitude of isobar surfaces in the center of the storm. This term would be much bigger and considerably affect the MSE variance if we included the highest pressure levels in our vertical integral operator $\hat{}$. It justifies a posteriori to focus on the troposphere, assuming that stratospheric waves are not fundamental in the physics of self-aggregation.

2.2 Dry and moist regions

2.2.1 Development of dry and moist regions

In SAM RCE simulations, the radiative-convective instability, leading to a moist cluster surrounded by a dry area, begins by the extension of a dry patch, fed by short wave radiations, which ends up concentrating the moisture in a single cluster. In our case, observations of the latent heat versus time gives the same impression of a single dry region amplifying and encircling the moisture:

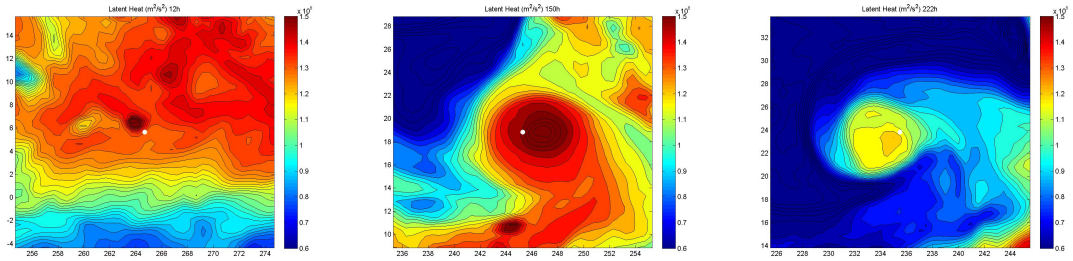

Figure 11: Spatial repartition of the vertically integrated latent heat ($J.m^{-2}$) at 12h (left), 150h (middle) and 222h (right)

A clear identification of those two regions is the key to understanding the diabatic feedback terms. We begin by defining the 10% moistest and the 10% driest regions, based on the vertically integrated latent heat $\hat{e_{lat}}$. At each time step, we can average a physical variable in those two regions to get a feeling for the correlation between the moisture and the variable. A good starting point is to apply this method to the latent heat:

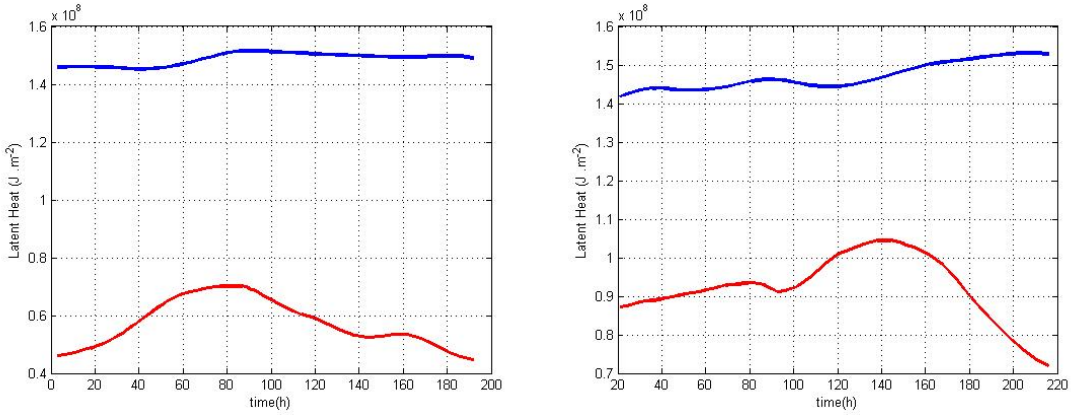

Figure 12: Latent heat vs time averaged in the 10% moistest columns (blue) and in the 10% driest columns (red) for cyclones Marie (left) and Genevieve (right)

It is a direct illustration of the radiative-convective instability process: During the self-aggregation process, moist regions get (slightly) moister and dry regions become significantly dryer.

2.2.2 Diabatic terms in both regions

We now study the difference between the diabatic terms in both regions, to see how the formation of the cluster correlates the energy fluxes from the environment:

- We can begin with the latent heat flux, which is the biggest source of energy for the hurricane once it is formed. The results allow us to distinguish two phases:

1. At first (before the self-aggregation begins) the driest columns, corresponding to a negative anomaly of MSE ($\hat{h}' < 0$), have the strongest latent heat flux ($LHF' > 0$), leading to a negative correlation/feedback ($< \hat{h}' LHF' >< 0$).

2. The tendency is then inverted, and the moistest columns ($\hat{h}' > 0$) receive the strongest latent heat flux ($LHF' > 0$), leading to a positive correlation/feedback ($< \hat{h}' LHF' >> 0$). As the cyclone forms, this feedback increases, to rapidly become the predominant one, as expected.

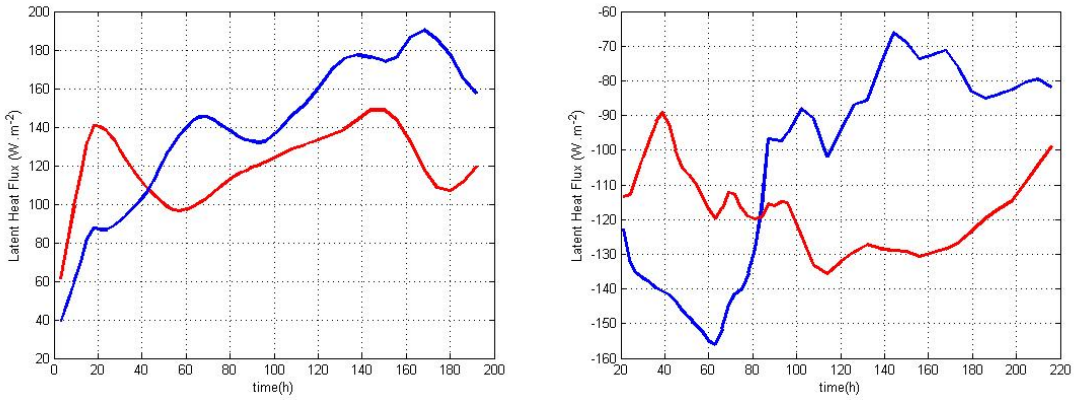

Figure 13: Latent heat flux vs time averaged in the 10% moistest columns (blue) and in the 10% driest columns (red) for cyclones Marie (left) and Genevieve (right)

- The other most important feedback is the net long wave radiative flux, and is always positive ($< \hat{h}' NetLW' >> 0$). By considering the clear sky longwave feedback, we see that the longwave feedback is almost entirely due to the clouds: the clouds, formed above the moist region of positive MSE anomaly ($\hat{h}' > 0$), absorb the terrestrial radiation (mostly IR) much more efficiently than dry air, leading to a positive longwave anomaly ($NetLW' > 0$) and thus a positive feedback.

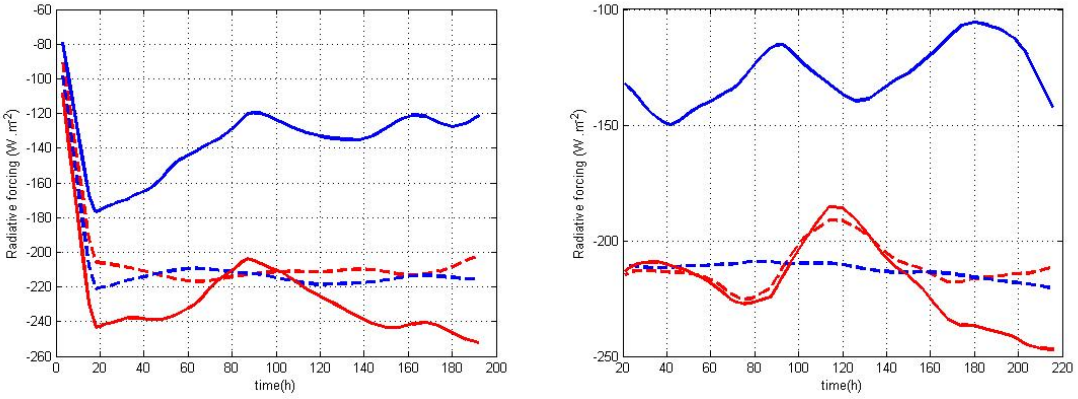

Figure 14: Net longwave flux vs time averaged in the 10% moistest columns (blue) and in the 10% driest columns (red) for cyclones Marie (left) and Genevieve (right) in the normal (full line) and clear sky (dashed line) cases

- The shortwave feedback is much smaller and more subtle, but we can see that it varies from positive to negative. By isolating the clear sky component, we can see that this shift is mostly due to the clouds, that end up reflecting the solar radiation ($NetSW' < 0$) in the moistest regions ($\hat{h}' > 0$) once the cluster is formed.

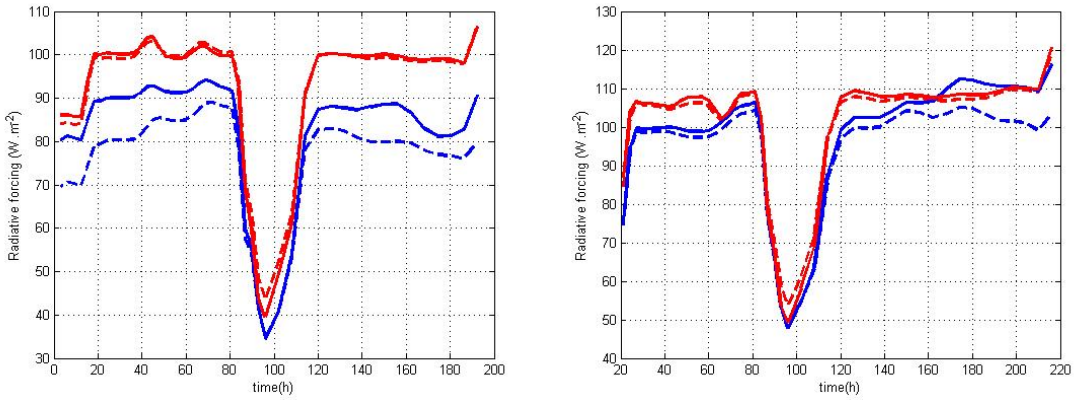

Figure 15: Net longwave flux vs time averaged in the 10% moistest columns (blue) and in the 10% driest columns (red) for cyclones Marie (left) and Genevieve (right) in the normal (full line) and clear sky (dashed line) cases

Thus, the identification and study of the two extremal behaviors in our system, governed by the moisture, allows us to physically explain the feedback terms. However, we cannot obtain exact results by this method, as we have only studied 20% of the regions of the system (the 10% moistest and the 10% driest), and an exact computation of the feedback terms is necessary to confirm the previous physical interpretations.

2.3 Feedback analysis

In this part, we directly compute the feedback terms from the Galilean invariant MSE variance budget.

2.3.1 General results

This budget can be written:

$$< \widehat{\frac{D}{Dt}}(\frac{\widehat{h}'^2}{2}) >=< \widehat{h}' DYN' > + < \widehat{h}' DIA' >$$

We compute the MSE variance $< \widehat{\frac{D}{Dt}}(\frac{\widehat{h}'^2}{2}) >$ based on ECMWF data, with a high incertitude, as explained in Appendix B. We then compute the diabatic correlation term based on ECMWF surface and radiative fields, and compute the dynamical term as the residual of the budget:

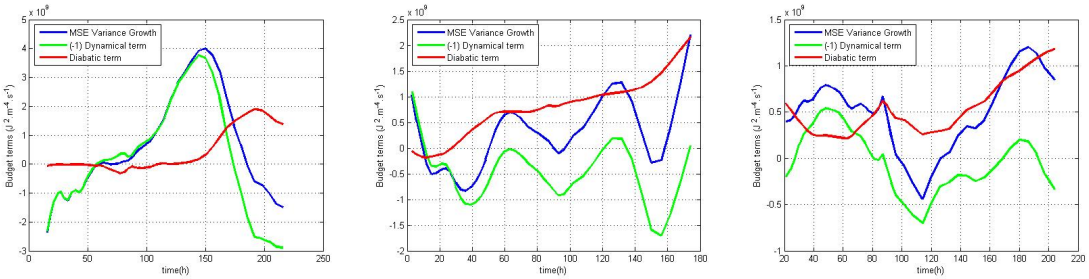

Figure 16: Horizontally averaged MSE variance budget for cyclone Fausto (left), Marie (middle) and Genevieve (right). The MSE variance growth is blue, the dynamical term is green and the diabatic term is red.

We notice that the diabatic term is close to zero before the Cyclogenesis happens, resulting in a balance between MSE variance and its transport. We can roughly distinguish two phases for the total diabatic term:

- The beginning of the self-aggregation (for $t \in [60; 110]h$ in the case of Tropical storm Fausto), where the diabatic correlation is negative, and a transport surplus helps the Cyclogenesis begin.
- The end of the self-aggregation and after (for $t \geq 140h$ in the case of Tropical storm Fausto), where the cluster is formed and the diabatic correlation is positive and at first increasing in time, helping the cluster to sustain.

2.3.2 Diabatic correlations

Decomposing the diabatic term in its different components helps understanding the role of different surface and radiative feedbacks:

$$DIA = SHF + LHF + [NetLW]_{z_b}^{z_t} + [NetSW]_{z_b}^{z_t}$$

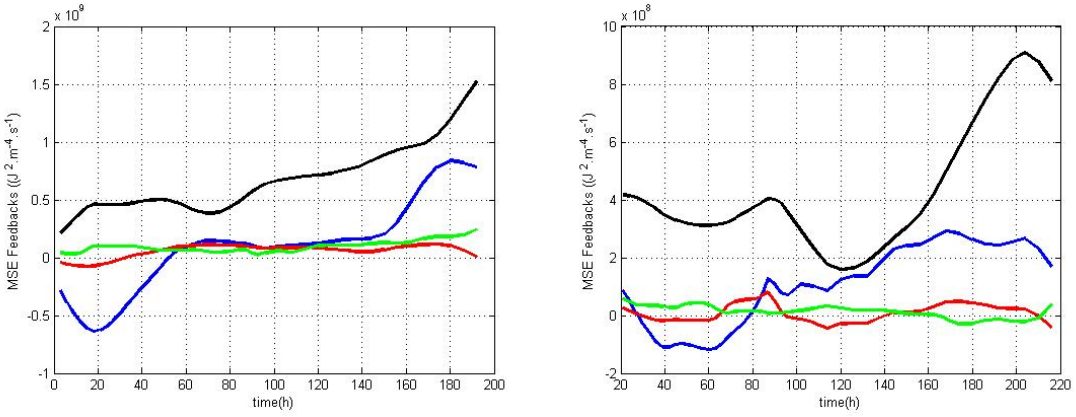

Figure 17: Diabatic correlation terms ($J^2 m^{-4} s^{-1}$) for cyclones Marie (left) and Genevieve (right). The latent heat flux term is in blue, the net longwave term is in black, the net shortwave term is in green and the sensible heat flux term is in red.

We recover the results obtained when studying the 10% driest and the 10% moistest air columns. Indeed:

- The two dominant feedbacks are clearly the latent heat flux and the net longwave radiation fields.
- The two dominant feedbacks increase and become strongly positive once the cluster is formed.
- The latent heat feedback is negative at the beginning of the self-aggregation, and increases until it is strong and positive once the cluster is formed. This term is closely linked to the increase of MSE variance, so that both terms become positive approximately at the same time.
- The longwave feedback is almost always positive, helping the self-aggregation to occur.

The Radiative feedbacks can be further studied by comparing them to their "clear sky" values, ie their values with the ECMWF model is run without any clouds:

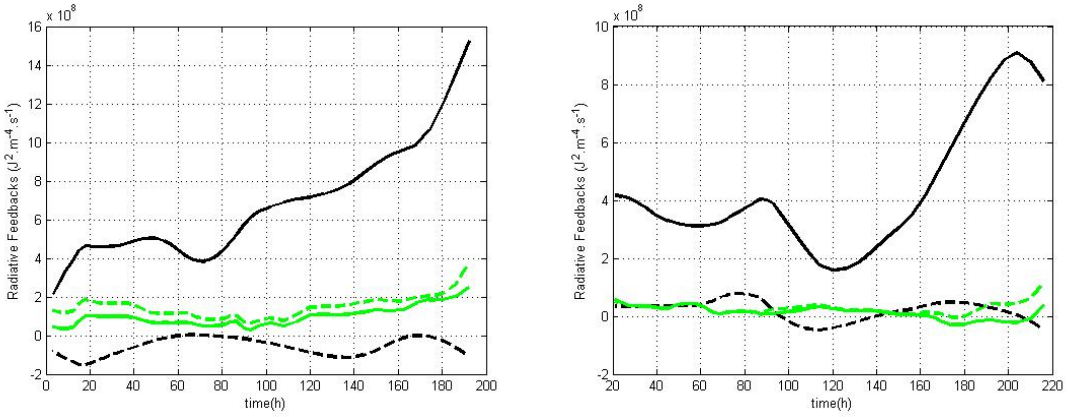

Figure 18: Longwave (black) and shortwave (green) radiative feedbacks for cyclones Marie (left) and Genevieve (right) in the normal case (full line) or clear sky case (dashed line) ($J^2m^{-4}s^{-1}$)

- In the longwave case, it can be seen again that the positive feedback is entirely due to the clouds: Clouds of the storm absorb terrestrial radiation, increasing the MSE and thus the moisture in the cluster, resulting in a positive feedback.
- In the shortwave case, the clouds have a slight negative effect, which can be explained by the reflection of the solar radiation at the top of the clouds once the cluster is formed.

2.3.3 Estimate of the precision

An estimate of the precision can be done by computing the residual of the budget:

$$RESID = < \widehat{\frac{D}{Dt}}(\frac{\widehat{h}'^2}{2}) > - < \widehat{h}'DYN' > - < \widehat{h}'DIA' >$$

Knowing that among the outputs of the ECMWF model, the biggest uncertainty is on the vertical velocity, we visualize the residual by computing the vertical dynamical feedback in three different ways:

1. We compute it as a residual from the MSE variance budget:

$$\widehat{h'\omega\frac{\partial h}{\partial p}}' = \widehat{h'DIA'} - \widehat{\frac{D}{Dt}(\frac{h'^2}{2})} - \widehat{h'\underline{u_2}''\cdot\underline{\nabla_2 h}}$$

2. We compute it directly from the ECMWF vertical velocity data:

$$\widehat{h'\omega\frac{\partial h}{\partial p}}' = -\widehat{h'}[\frac{\widehat{pwg}}{(r_d+q(r_v-r_d))T}\frac{\partial h}{\partial p}]'$$

3. Using the incompressibility in pressure coordinates, we compute it from the horizontal velocity data, which is more precise than the previous method:

$$\widehat{h'\omega\frac{\partial h}{\partial p}}' = -\widehat{h'h\frac{\partial \omega}{\partial p}}' = \widehat{h'(h\frac{\partial u}{\partial x}+h\frac{\partial v}{\partial y})'}$$

If the budget were perfectly balanced ($RESID = 0$), the three methods would give the exact same results. However, the high uncertainty in the vertical velocity data makes method 2 very unprecise (typically ω is much harder to sample than the other variables), and we compare the terms obtained by method 1 and 3 to estimate the precision of the feedback analysis.

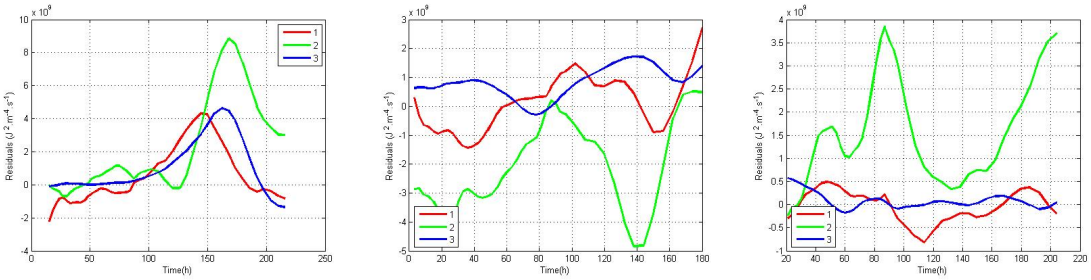

Figure 19: Vertical velocity computed using the three methods for cyclones Fausto (left) Marie (middle) and Genevieve (right)

The results are acceptable because of the high precision of the finite difference scheme, but not acceptable because we did not have access to data in the spherical harmonics space. As expected, the vertical dynamical term obtained by method 2 differs significantly from those which have been obtained by methods 1 and 3. In conclusion, we must keep in mind that the global terms in the MSE variance budget have the right order of magnitude, but we cannot know the precise value of the dynamical term, computed as a residual. However, the precision of the feedback terms is much higher as ECMWF radiative outputs are accumulated fields, making the finite difference method much more appropriate; thus the previous interpretations are not invalidated by this estimate.

2.4 Sensitivity tests

2.4.1 Dependence on the domain size

The choice of the domain size must correspond to the scale of the cyclone. If we take the example of tropical storm Fausto, we can see that at a given time, the MSE variance on a (16°,16°) domain and the MSE variance on

a (30°,30°) are fundamentally different, mainly because of Mesoscale convective systems (MCS) generated by the model around the cyclone.

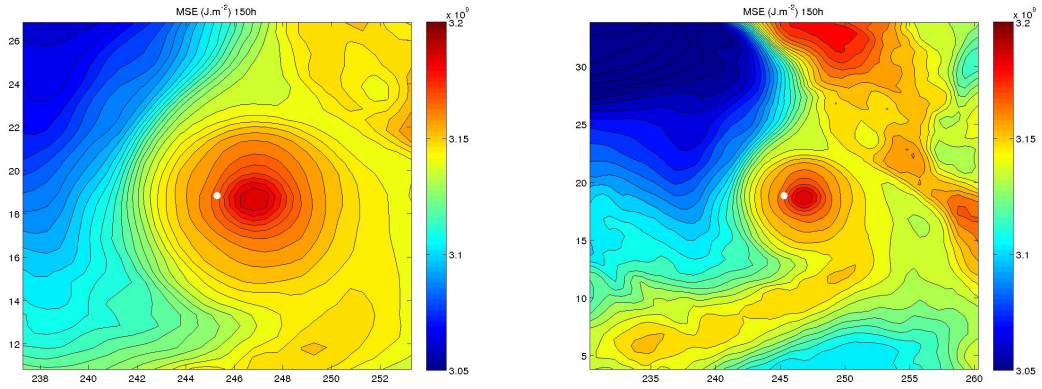

Figure 20: Spatial plot of vertically-integrated MSE at a given time (150h) on a (16°,16°) domain (left) and on a (30°,30°) domain (right)

Because of these MCS and because of the natural MSE gradient in the ocean, the MSE variance always increases when the domain size is increased. For instance, the main features of hurricane Fausto can be captured on a (16°,16°) domain, but the formation of the storm can still clearly be seen on a (30°,30°) domain.

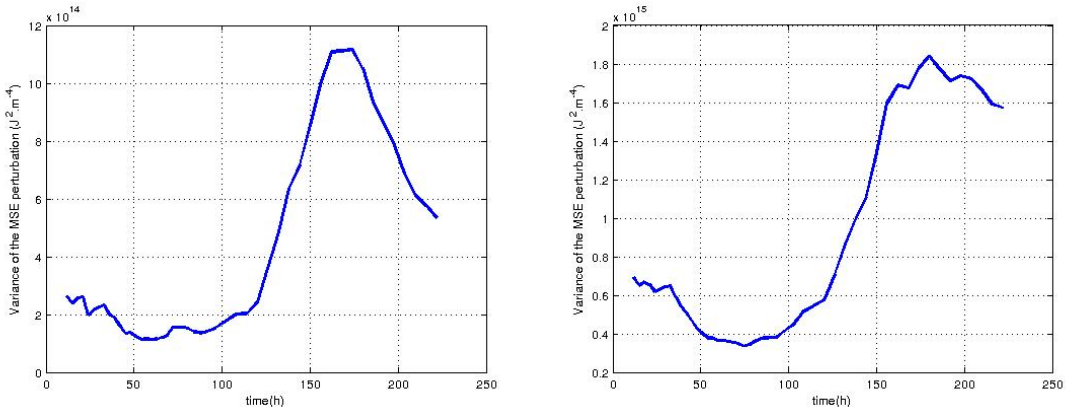

Figure 21: MSE variance vs time on a (16°,16°) domain (left) and on a (30°,30°) domain (right)

A more quantitative way of understanding the influence of the domain size is to fit again an exponential to the MSE variance increase:

$$< (\hat{h}')^2 >=< (\hat{h}')_0^2 > .e^{\frac{t}{\tau}}$$

We obtain the following results:

Domain size (°lon,°lat)	(16,16)	(20,20)	(25,25)	(30,30)
Start/End of the increase of variance(days)	2.5/7.25	3.75/7.25	3.25/7.5	3/7.5
Magnitude of MSE variance $< (\hat{h}')_0^2 > (J^2 m^{-4})$	$3.25.10^{12}$	$5.32.10^{12}$	$1.54.10^{13}$	$3.25.10^{13}$
E-folding time τ (days)	1.12	1.19	1.45	1.68
Correlation coefficient r^2	0.973	0.988	0.988	0.985

We can draw several conclusions:

- The start/end day of the increase of variance are pretty similar, which means the typical time of self-aggregation is scale free for a given cyclone.
- The magnitude of the MSE variance indeed increases with the domain size.
- The e-folding time decreases with the domain size: The relative variations of MSE due to cyclogenesis are bigger on a smaller scale, which means the cyclone has a bigger impact near his center. As we "zoom out", the MSE variations due to the cyclogenesis slowly merge into the natural MSE gradient.

The influence of the domain size on the diabatic feedbacks is more complex:

- The latent heat flux, which is much stronger in the cyclone, is thus relatively larger on small domain sizes. In the case of hurricane Fausto, a very interesting feature can be observed: The latent heat flux feedback becomes negative at some point, eventually stopping the cyclogenesis (the variance of MSE falls down rapidly at the end of the simulation). It confirms the idea that the latent heat flux feedback is a critical term for cyclogenesis.
- The net longwave term increases with the domain size, as clouds are spread everywhere in the domain. Consequently, this feedback is the most important one if we study the self-aggregation phenomenon on a very large scale.

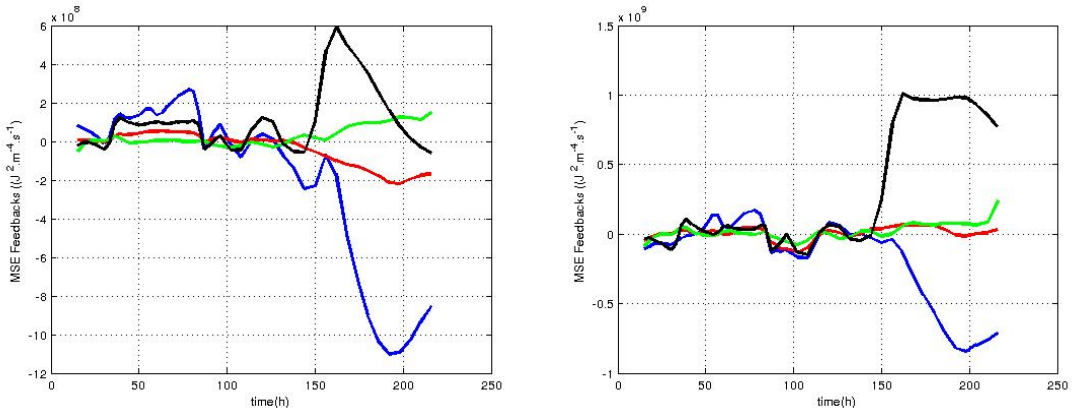

Figure 22: Diabatic correlation terms ($J^2 m^{-4} s^{-1}$) for a (16°,16°) domain (left) and for a (30°,30°) (right). The latent heat flux term is in blue, the Net longwave term is in black, the Net shortwave term is in green and the sensible heat flux term is in red.

Finally, we can explore the influence of the domain size on the residuals from the budget: is there an optimal domain size for our feedback analysis? An obvious result is that the finite difference scheme only works for a grid which has enough points, preventing us from working on very small domains (typically, smaller than (15°lon,15°lat)). The minimal residual is reached around a (20°,20°) in most cases, which is a good estimate of the scale of the cyclone, and justifies a posteriori our initial domaine size choice.

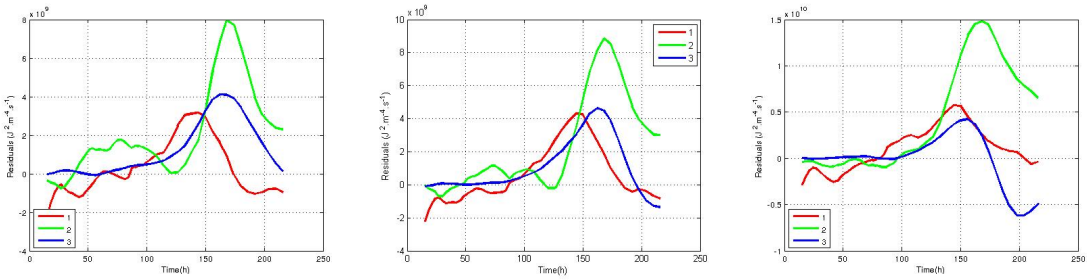

Figure 23: Vertical velocity computed using the three methods for a (16°,16°) domain (left), a (20°,20°) domain (middle) and a (30°,30°) domain (right)

2.4.2 Dependence on the cyclone type

The three presented cases (Fausto, Marie, and Genevieve) are Eastern Pacific Categorie 1 hurricanes. We have also studied bigger cyclones (for instance Hernan, a categorie 3 Eastern Pacific hurricane) and cyclones at different locations, such as typhoons Vongfong and Nuri, formed in the Western Pacific (see Appendix D for a list of cyclone cases). All hurricanes budgets have similar behaviors if the feedback analysis method applies (ie if the residual of the budget is sufficiently small).

For instance, we can consider the category 3 typhoon Nuri, and repeat the feedback analysis on a (20°,20°) domain.

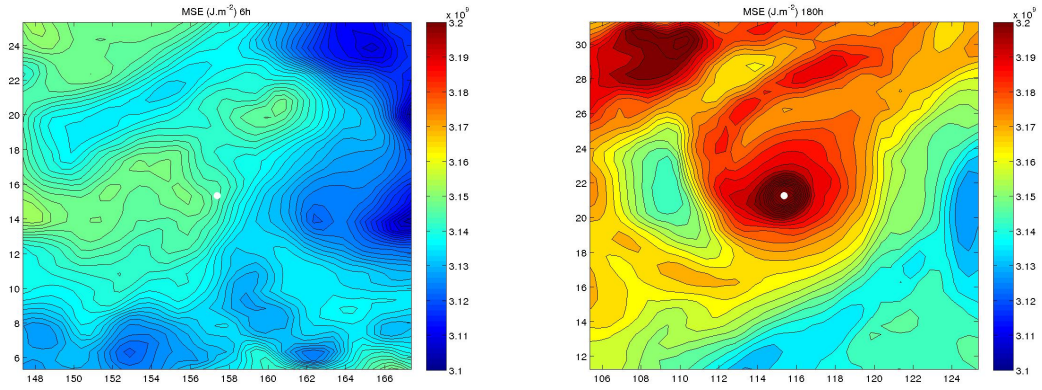

Figure 24: Spatial plot of vertically integrated MSE at the beginning (6h) and at the end (180h) of the study

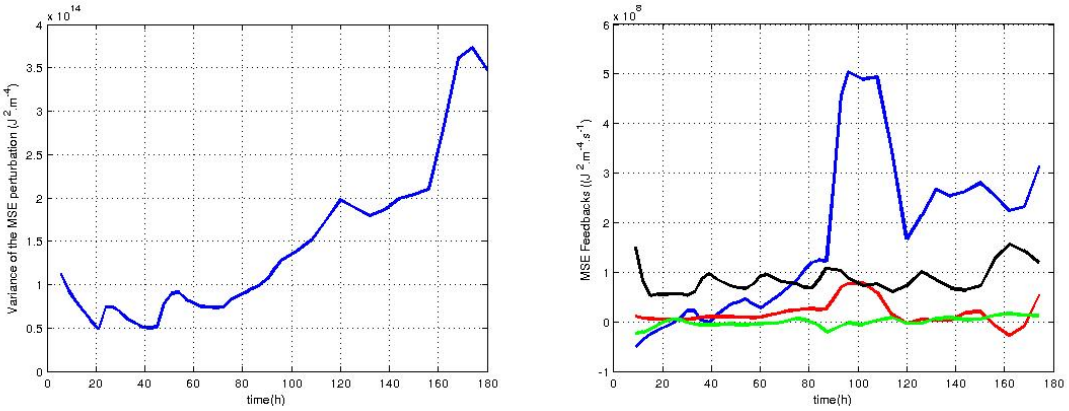

Figure 25: MSE variance vs time (left) and Diabatic correlation terms (right). The latent heat flux term is in blue, the Net longwave term is in black, the Net shortwave term is in green and the sensible heat flux term is in red.

All the diabatic feedbacks behave as expected; once again the net longwave feedback is initially the largest contributor to the MSE variance increase, but it is when the latent heat flux feedback becomes positive that this increase is actually triggered. In this case, the cyclone is especially intense, which means that the latent heat flux feedback ends up dominating all the other diabatic feedbacks.

Conclusion

During this project:

- We have developed a theoretical framework to analyze self-aggregation in a moving domain, and adapted it to study cyclogenesis in the ECMWF Forecast model.
- We have tested this analysis on 6 different hurricanes and obtained insight on the main diabatic feedbacks helping cyclogenesis to occur.
- We have developed numerical methods to approximate as best we could the physical quantities in the MSE variance budget, based on the interpolated outputs of the ECMWF model on a Gaussian grid.

We already have obtained insight into the physics of cyclogenesis in this model:

- The increase of variance in MSE is mainly due to the increase of the latent heat variance caused by the self-aggregation of moisture.
- The self-aggregation phenomenon is similar to what can be found in idealized RCE simulations, but much faster and happens on a much shorter timescale.
- As MSE variance increases, its diabatic feedback increases and we evolve from an equilibrium state between MSE variance and transport, to a state where the diabatic feedback eventually dominate the dynamical term, helping the cluster to amplify.
- The most important feedback terms are the latent heat flux and the net longwave radiation.
- The Net longwave radiation term appears to be almost always positive, and is entirely due to the cloudy component of the sky while the cluster is forming.
- The latent heat flux term is directly linked to the phase of self-aggregation, and seems to be the trigger of MSE variance increase. It is usually negative or close to zero, and cyclogenesis occurs soon after the term becomes positive. The feedback then amplifies during the cyclone formation, and as we have seen for hurricane Fausto, it inhibit the cyclogenesis if it goes back to being negative.
- Similarly to the numerical RCE simulations, there seems to be a dry and a moist region developing from an extending dry patch. Both regions have different feedback terms and allows us to physically understand the feedback terms.

Future research work linked to that project could be:

- Using reanalysis data to see if the approximation of the advection terms can be improved, and to check the consistency of the current results.
- The study of a rotating RCE simulation, with a special focus on the latent heat flux term.
- Applying the same analysis technique to the Madden-Julian oscillation, for example in the ECMWF model.

Appendix

A The ECMWF data

A.1 ECMWF Forecasts

The ECMWF, established in 1975, is an independent intergovernmental association, known worldwide to provide the most accurate medium-range global weather forecasts to 15 days, monthly forecasts to 30 days and seasonal forecasts to 12 months. Its global operational forecast model, the Integrated Forecast System (IFS), runs twice a day to produce two 10-day forecasts per day, based on the available analysis data at the initial time (produced 4 times per day). The IFS is configured for the following spatial resolution:

- 60 levels in the vertical, with the top level at 0.1hPa.
- T255 spherical-harmonic representation for the basic dynamic field.
- A reduced N128 Gaussian grid with approximately uniform 79km spacing ($\approx 0.7^{\text{o}}$) for surface and other grid point fields.

The atmospheric model is coupled to an ocean-wave model resolving 30 wave frequencies and 24 wave directions at the nodes of its reduced ($1^{\text{o}}\text{lon}, 1^{\text{o}}\text{lat}$) grid. Taking into account the fundamental unpredictability of dynamical systems, the deterministic model is complemented by an Ensemble Prediction System (EPS), which is a set of 50 forecasts at lower resolution, each started from a slightly perturbed state of the initial time analysis data. This system helps quantifying the uncertainty in the prediction of a particular event, which is necessary in the warning of extreme weather events.

As part of the Year Of Tropical Convection project (YOTC) project, aiming at improving modeling and prediction of tropical phenomena, ECMWF has made available to the public ERA-Interim data, containing a very complete set of forecast and reanalysis data from May 2008 to April 2010. In this project, I have only considered the forecast data, as they are physically consistent (ie prognostic variables verify their budgets at every time step).

A.2 Description of the set of data

Concretely, the downloadable forecast products is interpolated on a regular Gaussian grid resulting in:

- 25 interpolated pressure levels in the vertical:

$$p \in \{1; 2; 3; 5; 7; 10; 20; 30; 50; 70; 100; 150; 200; 250; 300; 400; 500; 600; 700; 800; 850; 900; 925; 950; 1000\} hPa$$

- A regular Gaussian grid, with interpolated data, which resolution is chosen to be ($0.5^{\text{o}}\text{lon}, 0.5^{\text{o}}\text{lat}$) to obtain a good compromise between precision and size of the files.

The obtention of the final data set consists of 4 steps:
1) Choosing a 10-days forecast including the studied storm.
2) In that set of data, Downloading the main 4D dynamic variables (at every grid point, every pressure level, and every time step), including:

- The wind velocity:

$$\underline{u} = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$$

- The specific humidity:

$$q = \frac{\rho_v}{\rho_d + \rho_v}$$

- The absolute temperature T.
- The geopotential:

$$\phi = gz$$

- Sometimes, the absolute vorticity (as a way of checking the computation of the gradients):

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$$

3) Downloading 3D surface variables (at every grid point, and every time step) to compute the feedback terms in the MSE budget:

- The Top Of the Atmosphere (TOA) accumulated net LW radiation field:

$$NetLW(p = p_t)$$

- The TOA accumulated net LW radiation field, clear sky (ie the model is run with no clouds):

$$NetLW_{clear\ sky}(p = p_t)$$

- The TOA accumulated net SW radiation field:

$$NetSW(p = p_t)$$

- The TOA accumulated net SW radiation field, clear sky:

$$NetSW_{clear\ sky}(p = p_t)$$

- The Surface accumulated net LW radiation field:

$$NetLW(p = p_b)$$

- The Surface accumulated net LW radiation field, clear sky:

$$NetLW_{clear\ sky}(p = p_b)$$

- The Surface accumulated net SW radiation field:

$$NetSW(p = p_b)$$

- The Surface accumulated net SW radiation field, clear sky:

$$NetSW_{clear\ sky}(p = p_b)$$

- The Surface accumulated Sensible Heat Flux:

$$SHF$$

- The Surface accumulated Latent Heat Flux:

$$LHF$$

4) Finally downloading the Potential Vorticity at 850hPa to track the cyclone:

$$\Pi(p = 850hPa) = \frac{\underline{\omega_a} \cdot \underline{\nabla\Theta}}{\rho}(p = 850hPa)$$

with the absolute vorticity at a point $\underline{x}$ on the Earth with angular velocity $\underline{\Omega}$ defined by:

$$\underline{\omega_a} = \underline{\nabla} \wedge (\underline{u} + \underline{\Omega} \wedge \underline{x})$$

and the virtual potential temperature (conserved during non saturated adiabatic motion) defined by:

$$\Theta = [1 + q(\frac{r_v}{r_d} - 1)]T(\frac{p_0}{p})^{\frac{r_d}{c_p}}$$

B Computation of derivatives by finite difference

The ECMWF model outputs are interpolated data on a Gaussian regular grid at each time step t, parametrized by longitude, latitude, and pressure level (λ, θ, p). We work in the associated local Cartesian-pressure coordinates (x, y, p), and compute all the derivatives using finite difference schemes. The computation will of course not be accurate, since the derivatives of basic dynamic fields are actually computed in spectral space, but we can hope a finite difference scheme with sufficient precision will give a good representation of the derivatives.

B.1 Computation of spatial derivatives on a uniform Gaussian grid

The advection terms involve the gradient of MSE:

$$\underline{u} \cdot \underline{\nabla h} = \begin{pmatrix} u\frac{\partial h}{\partial x} \\ v\frac{\partial h}{\partial y} \\ \omega\frac{\partial h}{\partial p} \end{pmatrix} \approx \begin{pmatrix} \frac{u}{Rcos(\theta)}\frac{\partial h}{\partial \lambda} \\ \frac{v}{R}\frac{\partial h}{\partial \theta} \\ \omega\frac{\partial h}{\partial p} \end{pmatrix}$$

where we are using the mean radius of the Earth:

$$R \approx 6371km$$

B.1.1 Horizontal derivatives on a uniform Gaussian grid using central finite differences

Since these derivatives are the harder to compute, we approximate them using an order 8 centered finite difference scheme. To compute the horizontal finite differences at the point $(\lambda_i, \theta_j, p_k)$, we define the typical coordinates finite differences:

$$\begin{cases} \Delta x = Rcos(\theta_j)\frac{\lambda_{i+1} - \lambda_{i-1}}{2} \\ \Delta y = R\frac{\theta_{j+1} - \theta_{j-1}}{2} \end{cases}$$

and the physical variable typical finite differences:

$$\begin{cases} \Delta_{x,n}h = h(x_{i+n}, y_j) - h(x_{i-n}, y_j) \\ \Delta_{y,n}h = h(x_i, y_{j+n}) - h(x_i, y_{j-n}) \end{cases}$$

so that a Taylor expansion gives:

$$\begin{cases} \frac{\partial h}{\partial x}(x_i, y_j) = \frac{1}{840\Delta x}[672\Delta_{x,1}h - 168\Delta_{x,2}h + 32\Delta_{x,3}h - 3\Delta_{x,4}h] + O(\Delta x^5) = \frac{\delta h}{\delta x} + O(\Delta x^5) \\ \frac{\partial h}{\partial y}(x_i, y_j) = \frac{1}{840\Delta y}[672\Delta_{y,1}h - 168\Delta_{y,2}h + 32\Delta_{y,3}h - 3\Delta_{y,4}h] + O(\Delta y^5) = \frac{\delta h}{\delta y} + O(\Delta y^5) \end{cases}$$

which gives us a numerical way to compute an order 8 approximation of the horizontal advection of the MSE Advection at point (x_i, y_j):

$$\underline{u_2} \cdot \underline{h_2} = \begin{pmatrix} u\frac{\partial h}{\partial x} \\ v\frac{\partial h}{\partial y} \end{pmatrix} = \begin{pmatrix} u(x_i, y_j)\frac{\delta h}{\delta x}(x_i, y_j) \\ v(x_i, y_j)\frac{\delta h}{\delta y}(x_i, y_j) \end{pmatrix}$$

B.1.2 Vertical derivative on non-uniform pressure levels

The term $\omega\frac{\partial h}{\partial p}$ is clearly the harder to compute with precision for multiple reasons:

- The vertical velocity field precision is not as high as other basic fields (horizontal velocities, specific humidity, temperature...).
- The model output is the vertical velocity in Cartesian coordinates w $[m.s^{-1}]$ at the different pressure levels, and computing the vertical velocity in pressure coordinates ω $[Pa.s^{-1}]$ requires several approximations:

1. We assume that the field is hydrostatic:
$$\frac{dp}{dz} = -\rho g$$
2. We assume that the gas is an ideal homogeneous mix of air and water vapor, following the equation of state:
$$p = \rho(r_d + q(r_v - r_d))T$$
with the specific gas constant of dry air and water vapor:
$$\begin{cases} r_d = \frac{R}{M_{air}} \approx 287.058 J.K^{-1}.kg^{-1} \\ r_v = \frac{R}{M_{H_2O}} \approx 461.523 J.K^{-1}.kg^{-1} \end{cases}$$
3. Following these approximations, we apply the following formula for the vertical velocity in pressure coordinates:
$$\omega \approx -\rho w g \approx -\frac{pwg}{(r_d + q(r_v - r_d))T}$$
which allows us to compute it at each pressure level.

- The MSE is only given at 25 pressure levels.
- The differences $p_{k+1} - p_k$ between the pressure levels are all different, which prevents us from using classical Taylor expansions.

We use a rough estimate of the vertical derivative:

$$\frac{\partial h}{\partial p}(p_k) \approx \frac{\delta h}{\delta p} = \frac{h(p_{k+1}) - h(p_{k-1})}{\Delta p_k}$$

with:

$$\Delta p_k = \frac{p_{k+1} - p_{k-1}}{2}$$

A Taylor expansion gives:

$$\frac{\partial h}{\partial p}(p_k) = \frac{\delta h}{\delta p}(p_k) + O((p_{k+1} - p_k)^2) + O((p_k - p_{k-1})^2)$$

making our approximation a first order one.

B.2 Computation of temporal derivatives with non-uniform spacing

The forecast data uses two time intervals:

1. For the 33 first time steps ie roughly for the first 4 days, the time step is $\Delta t = 3h$.
2. Afterward and until the end of the forecast, the time step is $2\Delta t = 6h$.

Consequently, we use a time step-dependent order 5 finite difference scheme to compute the time derivative $\frac{\partial}{\partial t}$ of a physical variable φ:

- On the sides (beginning and end of the time serie), we use a one sided Taylor expansion so that:

  $$\frac{\partial \varphi}{\partial t}(t = t_0) = \frac{\delta \varphi}{\delta t}(t = t_0) + O(\Delta t^5)$$

  with:

  $$\Delta t \cdot \frac{\delta \varphi}{\delta t}(t_0) = -\frac{25}{12}\varphi(t_0) + 4\varphi(t_0 + \Delta t) - 3\varphi(t_0 + 2\Delta t) + \frac{4}{3}\varphi(t_0 + 3\Delta t) - \frac{1}{4}\varphi(t_0 + 4\Delta t)$$

  and:

  $$\frac{\partial \varphi}{\partial t}(t = t_0 + \Delta t) = \frac{\delta \varphi}{\delta t}(t = t_0 + \Delta t) + O(\Delta t^5)$$

  with:

  $$\Delta t \cdot \frac{\delta \varphi}{\delta t}(t_0 + \Delta t) = -\frac{1}{4}\varphi(t_0) - \frac{5}{6}\varphi(t_0 + \Delta t) + \frac{3}{2}\varphi(t_0 + 2\Delta t) - \frac{1}{2}\varphi(t_0 + 3\Delta t) + \frac{1}{12}\varphi(t_0 + 4\Delta t)$$

  The finite differences schemes of $\frac{\partial \varphi}{\partial t}(t = t_{final})$ and $\frac{\partial \varphi}{\partial t}(t = t_{final} - 2\Delta t)$ can immediately be obtained by symmetry (signs are inverted, and since the time step is doubled, Δt is substituted by $2\Delta t$).

- Near the change of time step, which happens at $t_1 = t_0 + 32\Delta t$, we use an adapted centered finite difference scheme for each time step. Defining:

  $$\Delta_{t,n}\varphi = \varphi(t + \alpha(t) \cdot n \cdot \Delta t) - \varphi(t - \beta(t) \cdot n \cdot \Delta t)$$

  with:

  $$\begin{cases} \alpha(t) = 1 & if\ t + n\Delta t \leq t_1 \\ \alpha(t) = 2 & if\ t + n\Delta t > t_1 \end{cases}$$

$$\begin{cases} \beta(t) = 1 & if\ t - n\Delta t < t_1 \\ \beta(t) = 2 & if\ t - n\Delta t \geq t_1 \end{cases}$$

we use the three order 5 following schemes:

$$\Delta t \cdot \frac{\delta \varphi}{\delta t}(t = t_1) = \frac{4}{3}\Delta_{t,1}\varphi - \Delta_{t,2}\varphi + \frac{4}{9}\Delta_{t,3}\varphi - \frac{1}{12}\Delta_{t,4}\varphi$$

$$\Delta t \cdot \frac{\delta \varphi}{\delta t}(t = t_1 - \Delta t) = \frac{3}{10}\Delta_{t,1}\varphi + \frac{31}{150}\Delta_{t,2}\varphi - \frac{13}{120}\Delta_{t,3}\varphi + \frac{7}{330}\Delta_{t,4}\varphi$$

$$\Delta t \cdot \frac{\delta \varphi}{\delta t}(t = t_1 + 2\Delta t) = -\frac{13}{60}\Delta_{t,1}\varphi + \frac{74}{105}\Delta_{t,2}\varphi - \frac{32}{75}\Delta_{t,3}\varphi + \frac{6}{65}\Delta_{t,4}\varphi$$

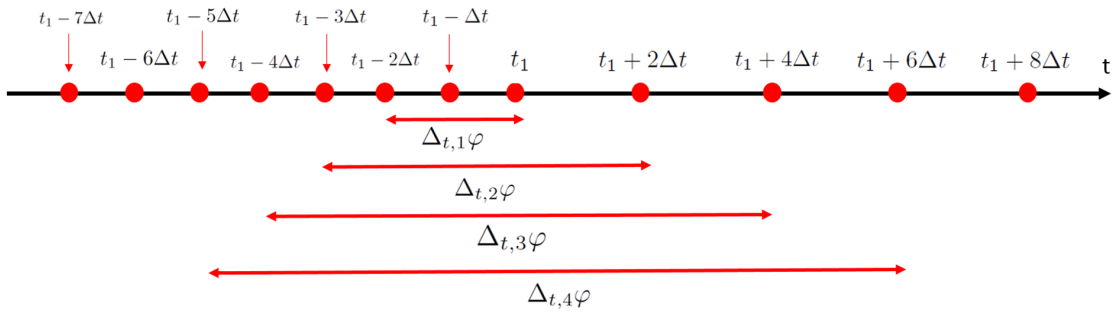

Figure 26: Quantities used for the $t_1 - \Delta t$ finite difference scheme

- Before the change of time step ie for $t \in [t_0 + 2\Delta t\ ;\ t_1 - 2\Delta t]$, we use a classical order 5 centered finite difference scheme. Defining:

$$\Delta_{t,n}\varphi = \varphi(t + n \cdot \Delta t) - \varphi(t - n \cdot \Delta t)$$

we use:

$$\Delta t \cdot \frac{\delta \varphi}{\delta t} = \frac{1}{840}[672\Delta_{t,1}\varphi - 168\Delta_{t,2}\varphi + 32\Delta_{t,3}\varphi - 3\Delta_{t,4}\varphi]$$

- After the change of time step ie for $t \in [t_1 + 4\Delta t\ ;\ t_{final} - 4\Delta t]$, we use the same scheme, substituting Δt by $2\Delta t$. Defining:

$$\Delta_{t,n}\varphi = \varphi(t + n \cdot 2\Delta t) - \varphi(t - n \cdot 2\Delta t)$$

we use:

$$2\Delta t \cdot \frac{\delta \varphi}{\delta t} = \frac{1}{840}[672\Delta_{t,1}\varphi - 168\Delta_{t,2}\varphi + 32\Delta_{t,3}\varphi - 3\Delta_{t,4}\varphi]$$

C Bilinear interpolation of data

Since the "center of the cyclone" is computed as a barycenter using the Potential Vorticity at 850hPa, it is seldom located at a precise grid point. We are thus led to use a bilinearly interpolated grid so that the "center of the cyclone" becomes a point of the new grid, at the center of the domain. The old and new grids are regular with the same homogeneous spacings in both directions: $(\Delta x, \Delta y)$. To compute the value of a variable b at

a point of coordinates (x, y) on the new grid, we use the values of b at the four closest points on the old grid: $\{(x_1, y_1), (x_1, y_2), (x_2, y_1), (x_2, y_2)\}$, with:

$$\begin{cases} x_2 - x_1 = \Delta x \\ y_2 - y_1 = \Delta y \end{cases}$$

We then compute $b(x, y)$ as a barycenter of the four values $\{b(x_1, y_1), b(x_1, y_2), b(x_2, y_1), b(x_2, y_2)\}$ with weight proportional to the area separating point (x, y) to the opposite point, as shown on the figure below:

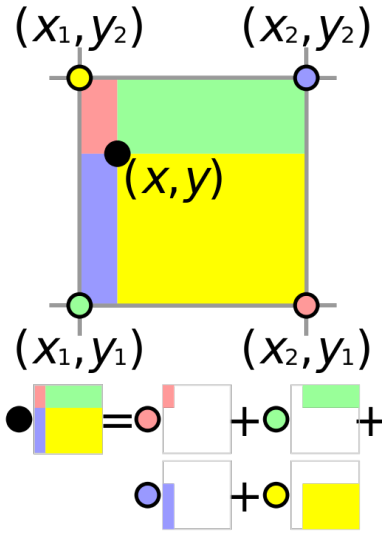

Figure 27: Bilinear interpolation principle

With the appropriate normalization, we obtain:

$$(\Delta x.\Delta y)b(x,y) = (x_2-x)(y_2-y)b(x_1,y_1)+(x_2-x)(y-y_1)b(x_1,y_2)+(x-x_1)(y_2-y)b(x_2,y_1)+(x-x_1)(y-y_1)b(x_2,y_2)$$

We notice than we can write this relation in a more compact matrix form:

$$b(x,y) = \frac{1}{\Delta x.\Delta y} \begin{pmatrix} y_2 - y \\ y - y_1 \end{pmatrix} \cdot \begin{pmatrix} b(x_1,y_1) & b(x_2,y_1) \\ b(x_1,y_2) & b(x_2,y_2) \end{pmatrix} \cdot \begin{pmatrix} x_2 - x \\ x - x_1 \end{pmatrix}$$

D Storm cases

We analyze 6 different storms which all happen in 2008, which characteristics are summarized in the following table:

Storm	Intensity	Region	Day of forecast	Start (h)	End (h)	Initial position (^{o}lon,^{o}lat)
Fausto	Cat. 1 Hurricane	East Pacific	14/07	12	222	(124.6E,23.8N)
Hernan	Cat. 3 Hurricane	East Pacific	01/08	0	240	(29E,14N)
Marie	Cat. 1 Hurricane	East Pacific	29/09	6	228	(122E,18N)
Vongfong	Severe tropical storm	West Pacific	07/08	57	252	(168.5W,44N)
Nuri	Cat. 3 Typhoon	West Pacific	14/08	6	180	(115W,21.2N)
Genevieve	Cat. 1 Hurricane	East Pacific	17/07	18	222	(112W,16N)

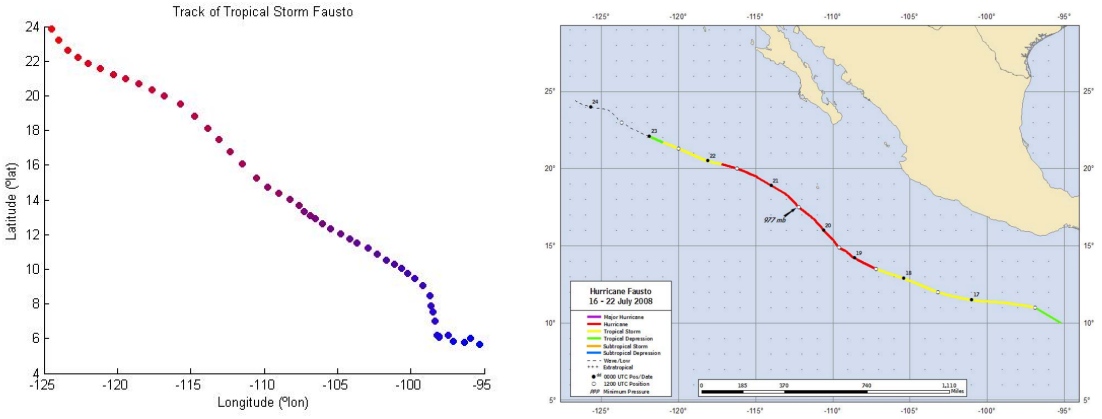

Figure 28: Trajectory of the PV tracker (left) vs Best track of Tropical Hurricane Fausto from the National Hurricane center (right)

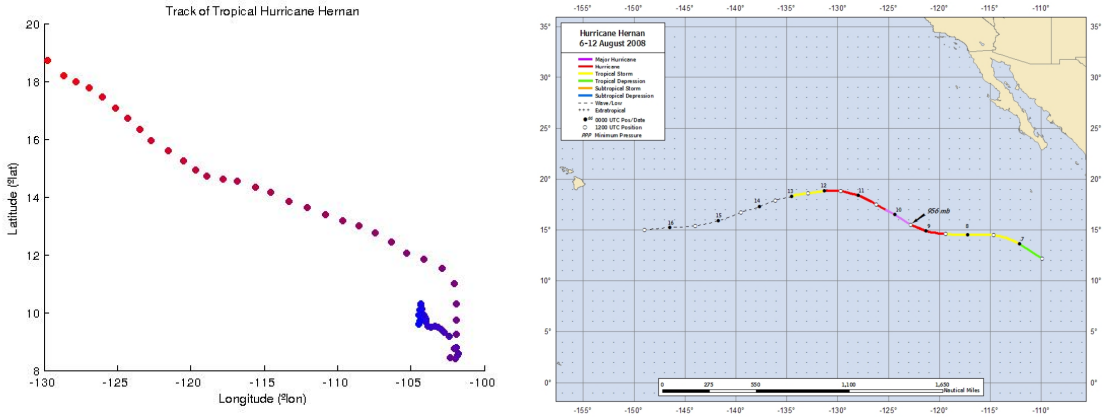

Figure 29: Trajectory of the PV tracker (left) vs Best track of Tropical Hurricane Hernan from the National Hurricane center (right)

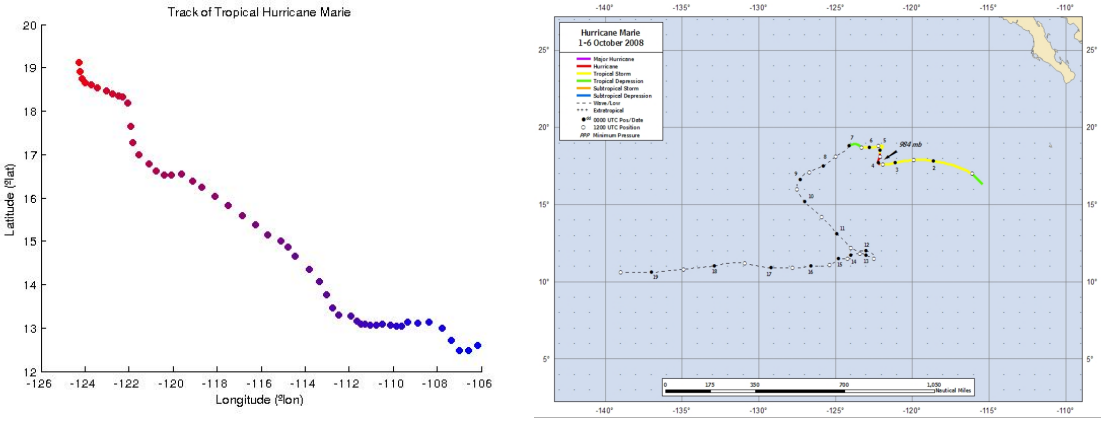

Figure 30: Trajectory of the PV tracker (left) vs Best track of Tropical Hurricane Marie from the National Hurricane center (right)

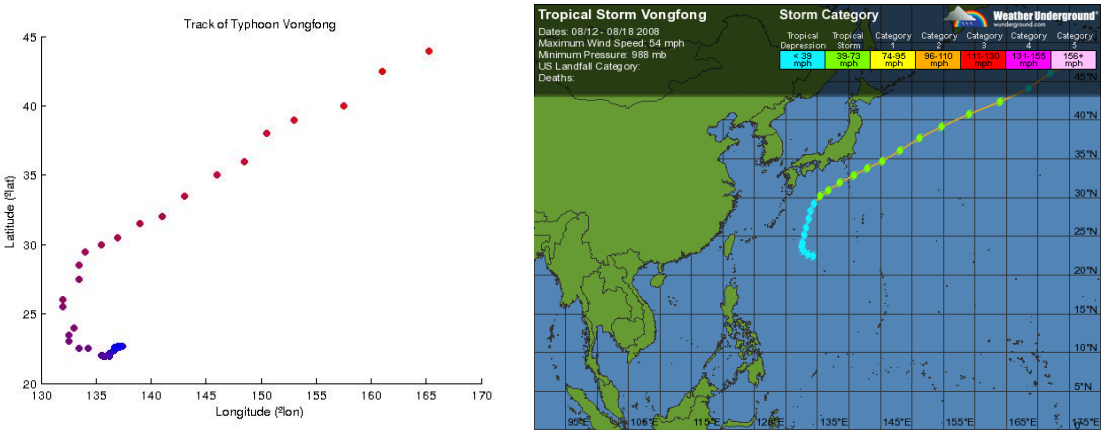

Figure 31: Trajectory of the PV tracker (left) vs Best track of Typhoon Vongfong from the National Hurricane center (right)

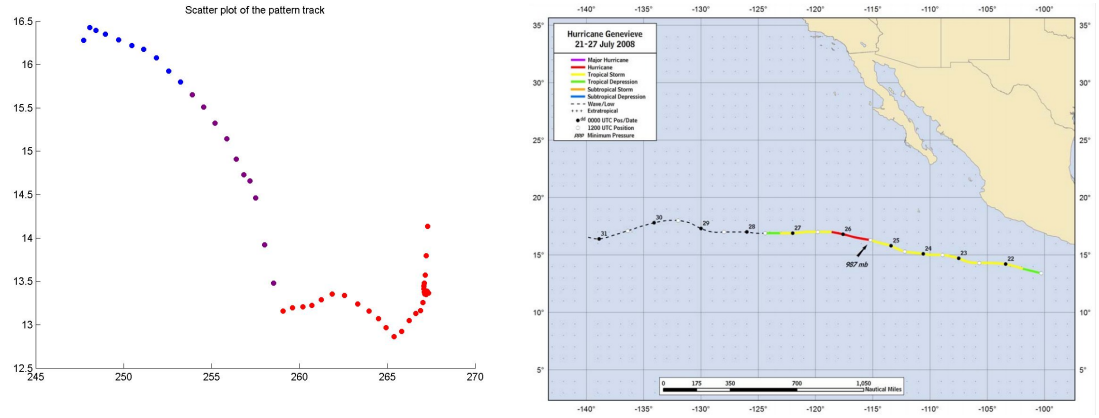

Figure 33: Trajectory of the PV tracker (left) vs Best track of Typhoon Nuri from the National Hurricane center (right)

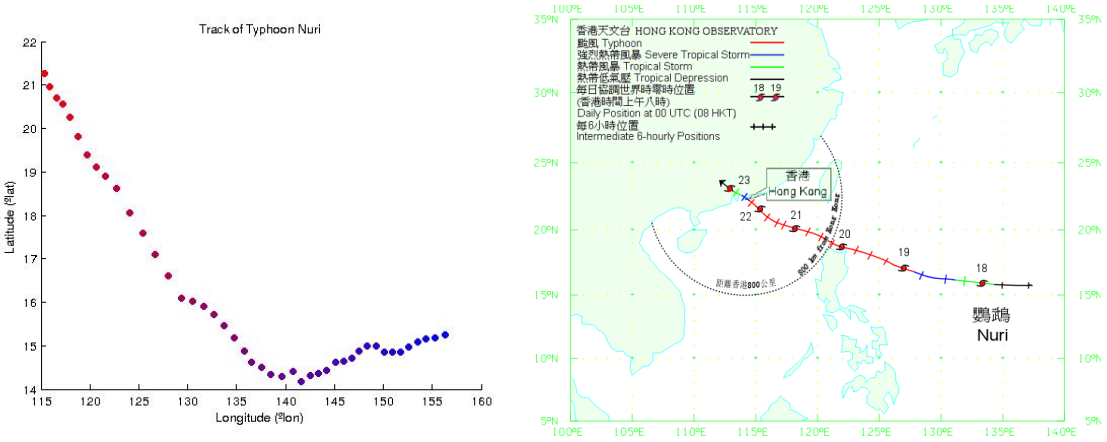

Figure 32: Trajectory of the PV tracker (left) vs Best track of Typhoon Nuri from the National Hurricane center (right)

References

[1] T. J. Dunkerton, M. T. Montgomery and Z. Wang, 2008. Tropical cyclogenesis in a tropical wave critical layer: Easterly waves. Atmos. Chem. Phys. Discuss., 8, 11149-11292.

[2] K. A. Emanuel, 2003. Tropical Cyclones. Annu. Rev. Earth Planet. Sci. 2003:75-104.

[3] K. A. Emanuel, 1994. Atmospheric Convection. Oxford University Press.

[4] K. A. Emanuel, A. A. Wing and E. M. Vincent, 2013. Radiative-convective instability. J. Adv. Model. Earth Syst., 5, doi: 10.1002/2013MS000270.

[5] W. M. Gray, 1968. Global view of the origin of tropical disturbances and storms. Mon. Wea. Rev., 96, 669-700.

[6] Arlene Laing and Jenni-Louise Evans, 2011. Introduction to Tropical Meteorology 2nd Edition, Chapter 8: Tropical Cyclones. Produced by the COMET program and the METED program. Available online at http://www.goes-r.gov/users/comet/tropical/textbook_2nd_edition.htm.

[7] D. S. Nolan, E. D. Rappin and K. A. Emanuel, 2007. Tropical cyclogenesis sensitivity to environmental parameters in radiative-convective equilibrium. Q. J. R. Meteorol. Soc., 133, 2085-2107.

[8] A. A. Wing and K. A. Emanuel, 2013. Physical mechanisms controlling self-aggregation in idealized numerical modeling simulations. J. Adv. Model. Earth Syst., 5, doi: 10.1002/2013MS000269.

[9] A. J. Simmons and D. M. Burridge, April 1981. An Energy and Angular-Momentum Conserving Vertical Finite-Difference Scheme and Hybrid Vertical Coordinates. Monthly Weather Review 109:758-766.

[10] Anders Persson and Erik Andersson, 2013. User guide to ECMWF Forecast products, produced by ECMWF. Available online at http://www.ecmwf.int/publications.

[11] P. Berrisford, D. Dee, P. Poli, R. Brugge, K. Fielding, M. Fuentes, P. Kallberg, S. Kobayashi, S. Uppala and A. Simmons, November 2011. The ERA-Interim archive Version 2.0. ERA Report Series produced by ECMWF. Available online at http://old.ecmwf.int/publications/library/do/references/list/782009.

[12] 2013, IFS Documentation CY40r1, produced by ECMWF. Available online at http://old.ecmwf.int/research/ifsdocs/CY40r1/index.html.

[13] Duane Waliser and Mitch Moncrieff, 2008. The Year Of Tropical Convection Science Plan. WMO/TD-Nº1452.

[14] A. A. Wing, 2014. Physical mechanisms controlling self-aggregation of convection in idealized numerical modeling simulations, PhD thesis, Mass. Inst. of Technol., Cambridge, MA.

[15] John L. Beven, November 2008. Tropical Cyclone Report on Hurricane Fausto (EP072008), National Hurricane Center. Available online at http://www.nhc.noaa.gov/pdf/TCR-EP072008_Fausto.pdf.

[16] Daniel P. Brown, October 2008. Tropical Cyclone Report on Hurricane Hernan (EP092008), National Hurricane Center. Available online at www.nhc.noaa.gov/pdf/TCR-EP092008_Hernan.pdf.

[17] Stacy R. Stewart, November 2008. Tropical Cyclone Report on Hurricane Marie (EP142008), National Hurricane Center. Available online at http://www.nhc.noaa.gov/pdf/TCR-EP142008_Marie.pdf.

[18] G.K. Vallis, 2006. Atmospheric and Oceanic Fluid Dynamics. Cambridge University Press, 745 pp.