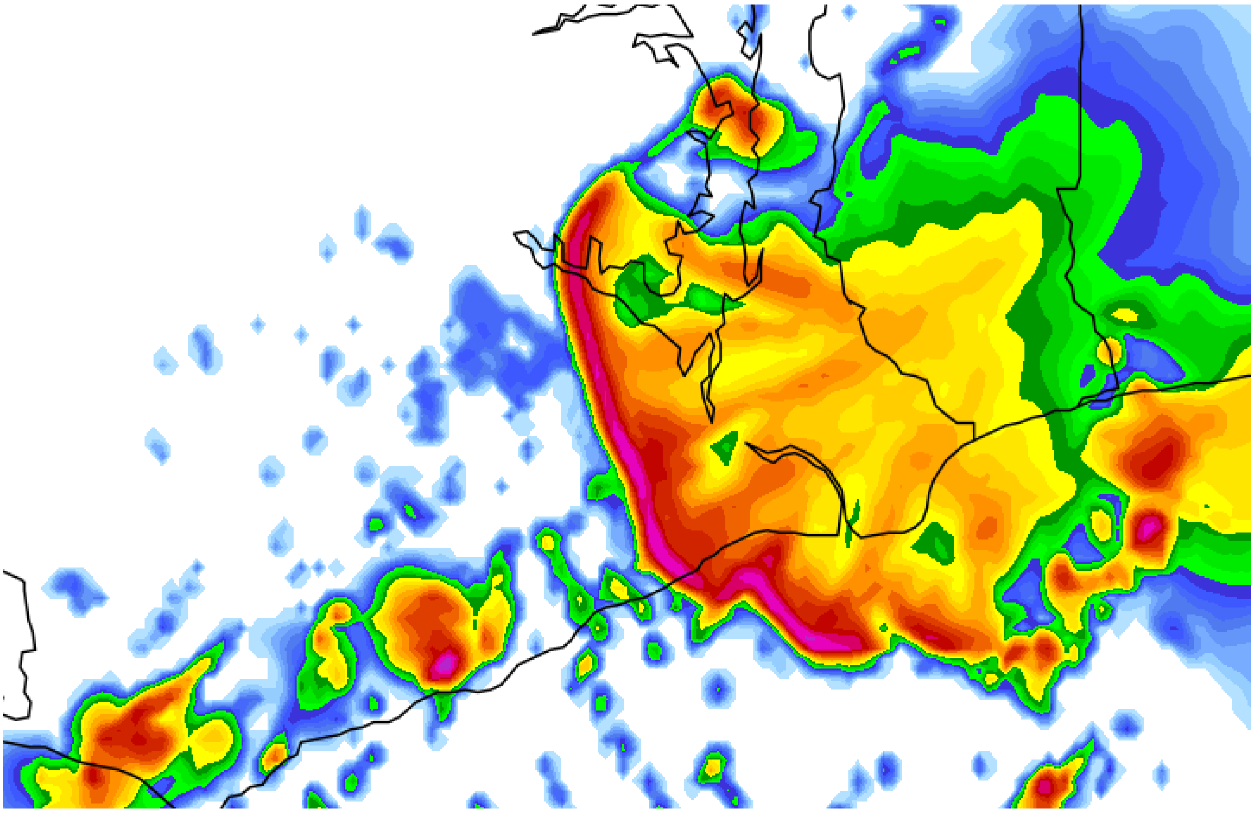
BACHELOR THESIS

High resolution precipitation forecasting for Ghana using the Weather Research & Forecast model



Author:
Wouter MOL
Study: Physics & Astronomy

Supervisors:
Michiel SEVERIN
Infoplaza
and
Aarnout VAN DELDEN
Utrecht University

15 June 2016

Abstract

Accra, the capital of Ghana, is located in a tropical climate where flash floods are a common occurrence, especially during the rain season. A project by multiple companies has been set up which aims to prevent casualties of these flash flood by building a warning system. Warnings will be based on a hydraulic model, which needs precipitation forecast data as input. The Weather Research & Forecast (WRF) model will be used as an operational numerical weather model to try and accurately forecast precipitation at a high resolution. The focus of this research is to find a good model setup for Ghana. This includes designing a domain, deciding what meteorological and geographic input data to use and by which schemes physics will be described. Schemes responsible for precipitation are assumed to be most important. Therefore, the performance of various combinations of cumulus parameterizations and microphysics schemes have been tested in detail by doing two case studies. Verification of the modeled precipitation is limited to satellite observations due to a lack of radar data or surface observations. Results vary a lot for the tested combinations, indicating the model is sensitive to the choice of physics schemes. Overall, the forecast skill is low. However, Thompson microphysics combined with a multi-scale Kain-Fritsch cumulus parameterization scheme consistently performed the best in both cases studied. Most nocturnal convective systems are missed by the model and the too early triggering of convection raises concerns as well. Identifying forecast errors by verifying atmospheric properties such as temperature, wind and moisture could lead the way to improving the model setup.

Contents

Chapter 1

Introduction

Ghana is a country located in western Africa just north of the equator. It has a tropical climate, meaning the atmosphere is warm and moist, especially during the rainfall season. Accra, Ghana's capital, has two distinct rainfall seasons, caused by the Inter-Tropical Convergence Zone (ITCZ), which passes twice a year as it oscillates between the northern and southern tropics (McSweeney, New, and Lizcano, 2012). Warm moist air is separated from hot and dry air along this zone, which is why it is also called the Inter-Tropical Front (ITF). Winds south of the ITCZ are generally south westerly, advecting moisture, while north of the ITCZ north easterly winds transport dry air. Accra's most pronounced rainfall season occurs in the spring, peaking in June. Monthly accumulations are in the order of 150-200mm. Due to the annual timescale of the oscillation, the often not so clear location and the hard predictability, analyses on its position and strength are often done by looking at 10-day periods (AIF).

The amount of moisture available in the atmosphere and the lack of any significant winds can result in slow moving, high precipitating convective storms. Though not a daily occurrence, during the rainfall season, these kind of storms are often present or nearby. Flash floods are thus a common risk. An area directly affected by this is Accra. With no warning system in place and the poor availability of accurate weather forecasts, citizens are often taken by surprise. In order to help prevent fatalities and damage to properties, a project by multiple companies has been set up which aims to build a warning system and a way to reach these citizens. A key part of this project is to set up a numerical weather model to provide short term accurate high resolution precipitation forecasts. The output of the weather model will serve as input to a hydraulic model, which models water levels up to street detail. Based on whether the hydraulic model forecasts flash floods, targeted warnings will be given to residents of Accra so they can prepare. If this concept proves to be successful, it can be expanded to more countries. Infoplaza (IP), a professional weather bureau, is tasked with setting up the numerical weather model, the process of which is documented in this report.

Setting up such a numerical weather model can be done in many different ways, not all of which will be addressed in detail in this research due to the complexity. However, designing a model domain, determining what input and other boundary data will be used as well as a choice of what physics parameterizations to use are important decisions. Additionally, a method of verifying model output needs to be designed in order to learn about the model's performance. Since precipitation forecasting is the main purpose of this model, physics schemes regarding convection and precipitation are of great interest and will be looked at in detail. A perfect combination of parameterization schemes is unlikely to exist, since not only they are all approximations of reality, their performance is subjective and dependent on what type of situation they are applied to. Forecasting flash floods in a tropical setting means the model is supposed to accurately forecast the locations, timing and intensity of convective storms in an atmosphere under weak synoptic forcing, not a simple task. A single combination of schemes is thus unlikely to perform best on all aspects. Moreover, sensitivity to initial conditions plays a big role in forecast uncertainty, especially for high resolution domains (Zheng, Alapaty, Herwehe, Genio, and Niyogi, 2016). The model, once operational, will forecast up to 48 hours, four times a day. When designing the model, it should be kept in mind that, since it is for operational use, there is a limited amount of time available for computing. Based on all information gathered in this research, a baseline model is set up upon which later further improvements can be made.

Chapter 2

Weather Research & Forecast model

2.1 Introduction

The numerical weather model used is the Weather Research & Forecast (Wei Wang et al., 2016) model (v3.8, April 2016), developed by UCAR[1]. Used specifically is the Advanced Research (ARW) core. It is a non-hydrostatic (with a hydrostatic option), highly modifiable open source model, used in many different situations. Examples are operational high resolution forecasts, hurricane case studies and even climate analyses. Two main types of simulations can be run using this model, namely ideal or real cases. The former is typically used to simulate a specific phenomenon in an ideal (2D or 3D) setting, where the atmosphere is horizontally homogeneous and the lower boundary is flat. This way the model can be tested to see if, for example, it manages to solve tropical deep convection, which it does well (Costantino and Heinrich, 2014). Simulating real world weather is done using so called real cases, requiring 3D information about the atmospheric setting and the Earth's surface, further discussed in section 2.4.

2.2 Model physics and dynamics

The atmosphere is a very complex system to model, as it consists of many components which either directly or indirectly affect one another. Each of these components is described in the model by a set of physical laws and parameterizations, but there is more than one way to do this. Some schemes are simple and fast, usually at the cost of accuracy, others are complex and computationally intensive, but more accurate. The choice of schemes depends on what is considered relevant and important for the case at hand. Since this study is centered around precipitation forecasting, components directly dealing with convection and precipitation will be solved using advanced schemes. While other components are still very important, their influence is probably not as direct, hence less advanced or even simpler schemes can be used to save calculation time at little loss of accuracy. Recommendations of the WRF documentation (Dudhia, 2008) are followed. Each component and its chosen scheme will be discussed now. Nested domains will have identical physics as their parent domains, unless mentioned otherwise.

2.2.1 Radiation schemes

Shortwave and longwave radiation both have their own scheme. A few examples are surface heating by direct sunlight, cloud albedo and the absorption by greenhouse gases. Used in this model, for both radiation schemes, is Rapid Radiative Transfer Model for General Circulation Models, or RRTMG for short. Both the longwave and shortwave versions of this scheme are used in well known models like ECMWF IFS[2] and NCEP GFS[3]. Radiation schemes are too computationally intensive to call every integration step. Here, it is called every 12 simulation minutes, a recommended setting[4].

[1] University Corporation for Atmospheric Research

[2] European Center for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS)

[3] National Center for Environmental Prediction (NCEP) Global Forecast System (GFS)

[4] Recommended by the WRF documentation is the same amount of minutes as the model resolution Δx in km's of the parent domain. I.e. $\Delta x = 12$ means a radiation call every 12 minutes.

2.2.2 Surface schemes

The atmospheric surface layer and the land surface model are both described by surface schemes and provide information to the planetary boundary layer (PBL). The atmospheric surface layer scheme is responsible for heat, moisture, momentum fluxes and surface friction forces. It provides exchange coefficients to both the land surface model and PBL and also includes water surface fluxes of heat and moisture. The land surface model is used to model land up to a few layers deep, unlike water which is regarded as just a surface[5]. Snow cover, soil moisture fluxes and soil temperature are examples of things managed by the land surface scheme. Monin-Obukhov's similarity theory scheme is used for the atmospheric surface layer and a Noah Land Surface scheme for the land surface model. Four land surface layers are used, which is a default setting. Optional urban physics are not used, due to the amount of computational power it requires compared to its benefits for the current application. It would be more applicable to very high resolution simulations and large urban areas. Urban effects are still accounted for by the model though, as surface schemes utilize geographical data such as land use. This is more thoroughly discussed in section 2.4. Sea surface temperatures are left constant as they hardly vary in short range forecasts.

2.2.3 Planetary Boundary Layer schemes

Moisture, heat and momentum fluxes in the PBL are managed by a PBL scheme. It is responsible for local and non-local mixing of air, entrainment with the stable, free atmosphere above and vertical diffusion. The choice of a PBL scheme is coupled to that of the atmospheric surface layer, i.e. not all parameterizations of these schemes can work together, thus the choice is restricted. Used here is the Yonsei University PBL scheme.

2.2.4 Cumulus parameterization

Explicitly solving cumulus[6] clouds requires a high enough spatial resolution. For $\Delta x \geq$ 10km, this process needs to be parameterized, since at least in the early stages the size of cumulus clouds are smaller than the model's grid boxes. Parameterization is usually unnecessary for $\Delta x \leq$ 3km, since convection can be solved well using this resolution, but might still help to initiate convection in tough cases. No cumulus scheme was designed specifically for the 3 to 10km range, so which one to use within this range, if at all, is not obvious.

Cumulus schemes parameterize convective cloud processes like condensation, precipitation and the redistribution of mass. These processes are important in the early development of convective storms. It produces sub-grid scale clouds and surface precipitation, but only when certain conditions are met. A convection trigger, which keeps track of these conditions and decides whether they are met, is therefore an essential part of these schemes. Some schemes have multiple convection triggers to choose from, differing in the approach, criteria and coefficients. A decision of which scheme and trigger to use will be made based on test results.

2.2.5 Microphysics

A microphysics scheme is essential as it is responsible for atmospheric heat and moisture tendencies and the production of precipitation. Also the microphysics of hydrometeors[7] like the formation, what type and their fall rates (precipitation) are managed by the microphysics scheme. Depending on how advanced the scheme is, classes, also called mass variables, will include water vapour, cloud water, cloud ice, snow, rain and graupel. Whenever a model is to solve explicit up drafts, i.e. solve convection, a graupel class should be included. Another distinction between microphysics schemes is the amount of moments they work with. Most are single moment, which stands for the mixing ratio of a species. Multi moment schemes also include the number density and in some cases also the reflectivity of particles. Unsurprisingly, this requires more computational power. Even though cumulus parameterization and microphysics schemes overlap in what they do, they can work side by side. The total precipitation of

[5] It still has physical properties, e.g. it can evaporate water and slow down surface winds. A more sophisticated ocean model is available for Hurricane WRF

[6] Convective clouds

[7] Frozen or liquid bits of water in the atmosphere

a grid cell is then the sum of what both schemes have produced. Just like the cumulus schemes, a decision of what microphysics scheme to use will be decided later on.

2.2.6 Model dynamics

Model dynamics are options mostly concerning the numerical details of integration. Everything is left on default, except for a few damping options. As recommended, very strong vertical motion is slightly damped, as this could destabilize the model. Moreover, the model's initial setup had some issues with waves at the model top, caused by large scale deep convection not diffusing properly. This resulted in strange artifacts along the border of the model's parent domain. Hence, a diffusion option has been enabled which adds a Rayleigh relaxation layer to the model top, damping vertical motion.

2.3 Model domain

To provide numerical forecast data for the hydraulic model, it has to at least include the whole river domain which extends as far north as Burkina Faso. Also important to include is the Gulf of Guinea, part of the ocean south of Ghana, as this is the main source of low level moisture. Long lived convective systems affecting Ghana often originate from the east, initiating as far as Nigeria, and propagate westward over the warm water, along the coastline and in some cases further landward.

Keeping the main area of interest centered in the domain makes it large. At a high resolution, it quickly becomes computationally intensive, especially with a domain of the size just described. To solve this problem, only the most important area will be run at a high resolution, nested within a coarser parent domain which spans the whole region, shown in figure 2.1. Feedback between these two domain is turned on, i.e. they don't run independently. The nested domain has a 4x4km resolution, i.e. $\Delta\mathrm{x} = 4\mathrm{km}$, and is 148x130 grid points large. It is nested within a 140x140 grid points parent domain with $\Delta\mathrm{x} = 12\mathrm{km}$. Both these domains have 41 vertical levels, distributed from the surface to 50hPa by the model itself. The 4km domain includes Accra, a relatively small part of the Gulf of Guinea, a larger part of the river domain and other land north of the city (figure 2.1).

2.4 Input and geographical data

The cases studied in this research use the archived NCEP GFS 0.50 degree, 27 vertical levels analyses data as atmospheric initial condition for both domains. Real-time GFS datasets include 47 vertical levels, but can not be used for the case studies. The boundaries of the parent domain are fed with corresponding forecast data of the same resolution on a 3 hour interval. The nested domain receives its boundary data from the parent domain. Initialization is done at 12Z[8] on the day prior to the case study. This is to let the model spin up and also deal with overnight large scale convective systems, which will be further discussed in section 4.1. The GFS itself has a spin up time too, of course, so a forecast dataset could also be used to initialize the model with, but the influence of this will not be looked at in this research.

Geographical input data has been provided by UCAR. It is a large package of geographical data such as land use, vegetation albedo and topography, mostly available in a 30 arcsecond[9] resolution. Figure 2.2 illustrates the orography of both model domains. The gradient of the surface height and mountain peaks are sharp in the 4km domain, whereas they are more smoothed out in the parent domain. Such a high detail benefits the model's ability to solve local effects such as orographic lifting, a push that might trigger thunderstorms.

The type of model surface provides essential information to the surface schemes and therefore requires a closer look. While the distinction between land and water is obvious, dense forests and urban areas are vastly different in the way they deal with e.g. radiation and precipitation as well. An overview of the land use, expressed in terms of vegetation fraction, is shown in figure 2.3. The 12km domain draws a general picture of land being less vegetated further away from the ocean. In higher detail, local features such as urban areas (e.g. the dark gray area nearby 1.5W, 6.5N) and the artificial lake become more pronounced. Light gray spots scattered across the land are bodies of water.

[8] 12Z (Zulu) is 12:00 UTC, which in case of Ghana is equal to local time

[9] 0.00833 degrees, or about 900 meters

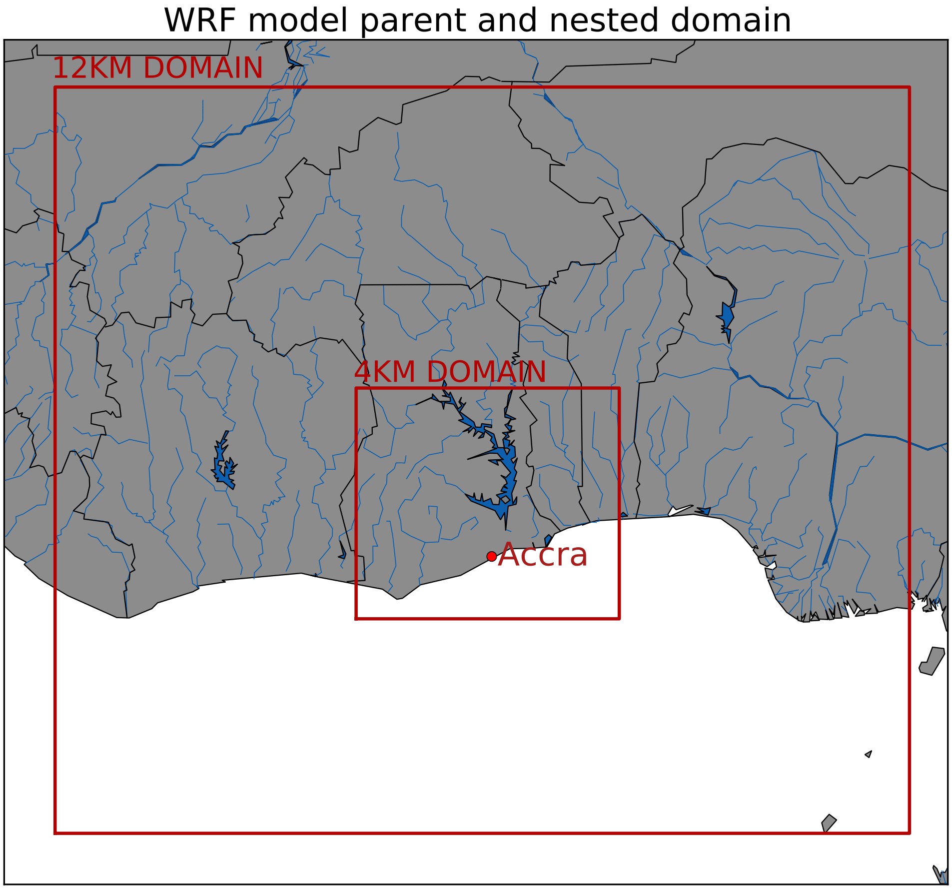


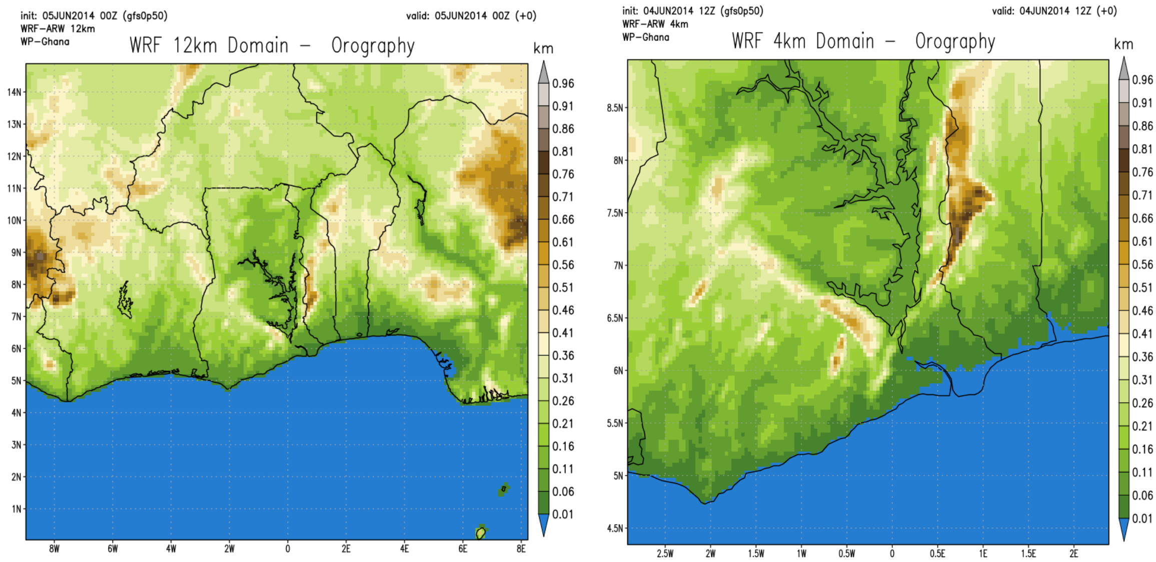
Figure 2.1 – The 4km domain nested within a 12km parent domain, covering all important areas

More detailed information is available to the model. In fact, there are 24 different categories of land use, each with their own physical properties. Every grid point has a dominant category, based on the individual fractions of land use type in the grid point. Dominant categories are used by the model as land use, illustrated by figures 2.4 and 2.5. Urban areas are visible in both domains, but again, details are much better resolved in the 4km domain. Despite lakes being labeled as barren tundra, they behave like water. I.e., they have no vegetation (figure 2.3), the soil moisture content is 100 percent and 2m dew points are high due to evaporation, illustrated in figure 2.6. This error in the geographical data is therefore considered to be of no influence.

2.5 Model integration time step

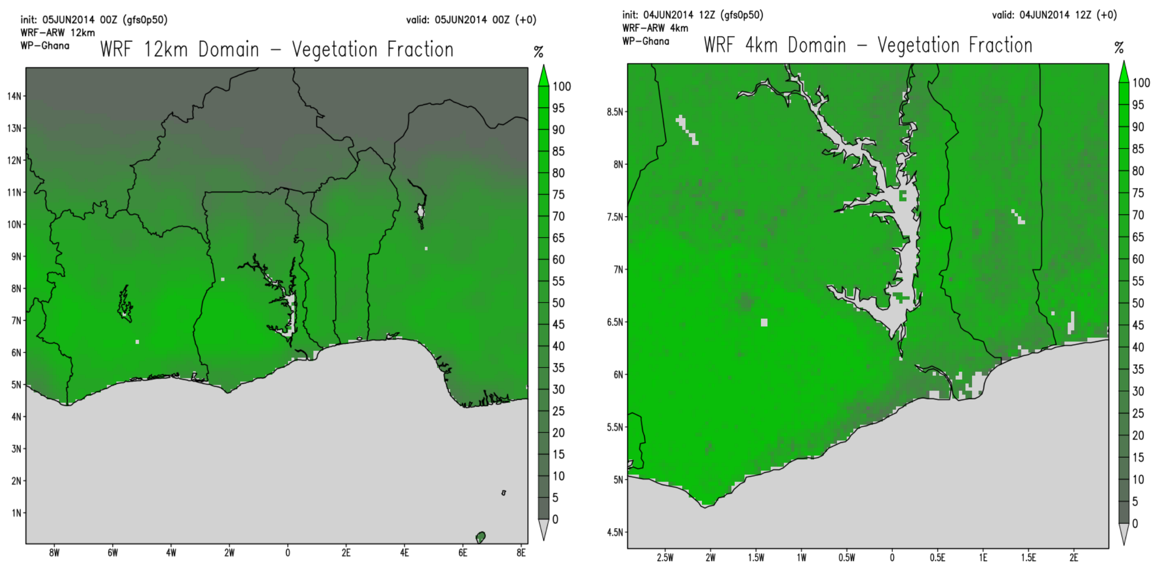
In order to have a numerically stable running model, a sufficiently small integration time step Δt is one of the main requirements. E.g. instabilities arise when air moves beyond a neighboring grid point due to its velocity being too high for the given Δt. This can, if it persists, crash the model. In order to prevent such events from happening, the Δt setting would need to be lowered. A smaller Δt, however, means it will take longer to complete a full run. Another downside in this case is Δt will be set to a value at which the model runs stable in case of extreme velocities. Often, this means Δt is unnecessarily small for a large part of the simulation and thus is not an efficient method. A way to deal with this problem is to use adaptive time stepping. This means letting the model decide during the simulation whether the time step needs to be lowered or if it can be raised, based on a CFL[10] condition. This lets the model adjust Δt, within a specified range, according to the situation. Adaptive time stepping is used in this model. Δt in seconds is related to the domain resolution in km. This means the integration time step range for the 12km domain is $24\mathrm{s} \leq \Delta t \leq 96\mathrm{s}$ and $8\mathrm{s} \leq \Delta t \leq 32\mathrm{s}$ for the 4km domain.

[10] Courant-Friedrichs-Lewy condition, which describes numerical stability



(a) (b)

Figure 2.2 – Orography of the 12km **(a)** and 4km **(b)** model domains



(a) (b)

Figure 2.3 – Surface vegetation of the 12km **(a)** and 4km **(b)** model domains

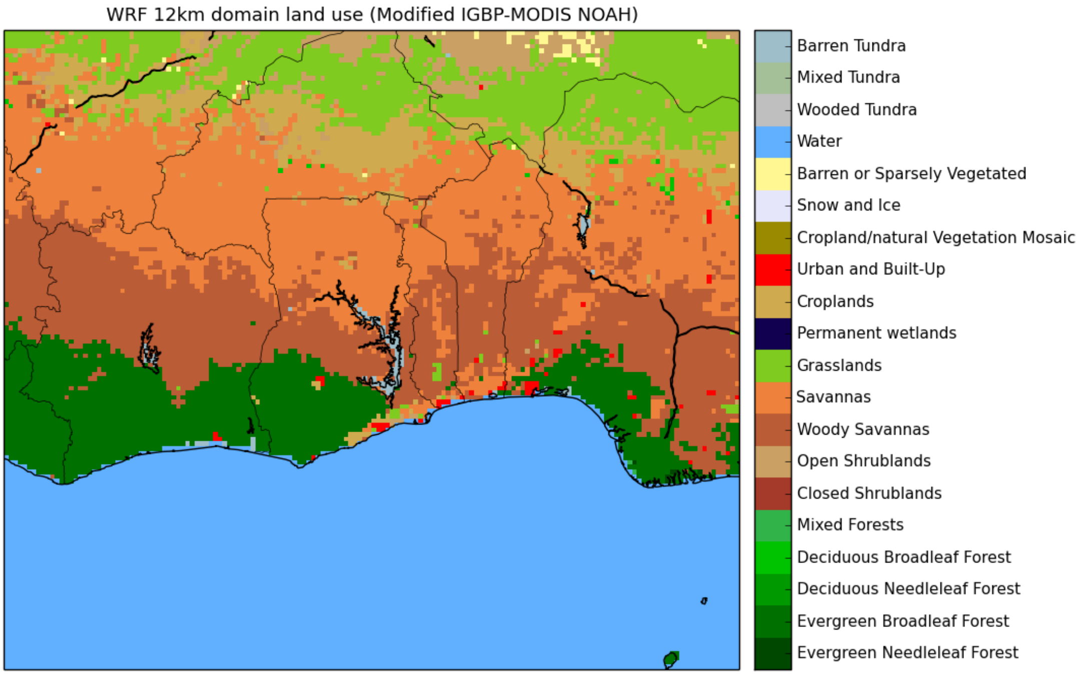


Figure 2.4 – Modified IGBP-MODIS NOAH categories, most dominant land use of the 12km domain

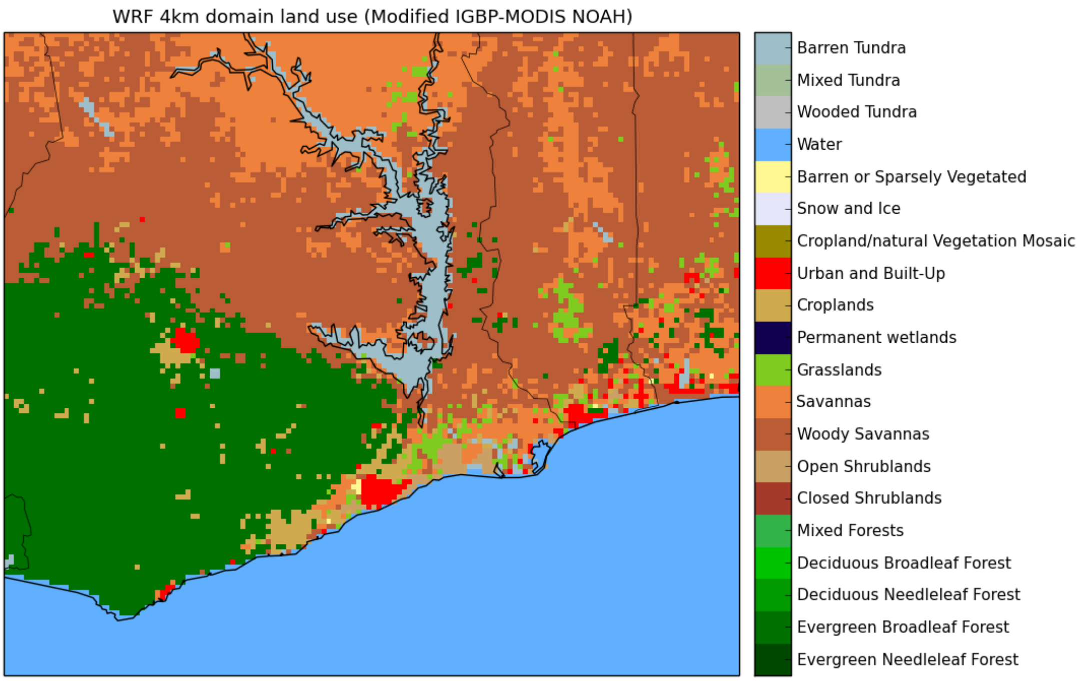
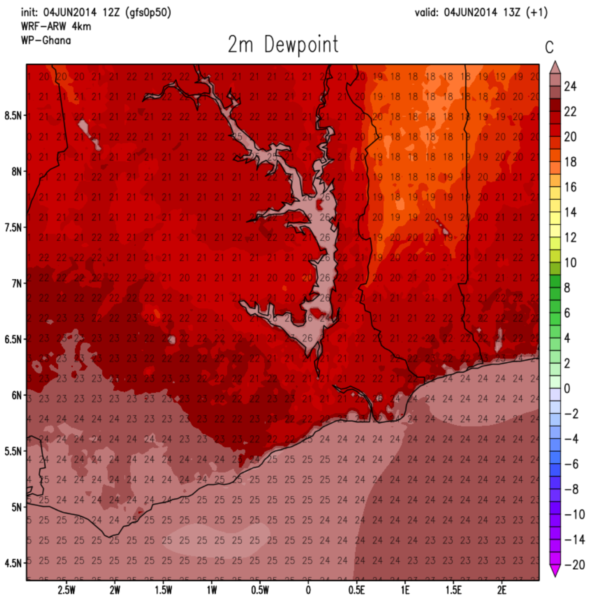
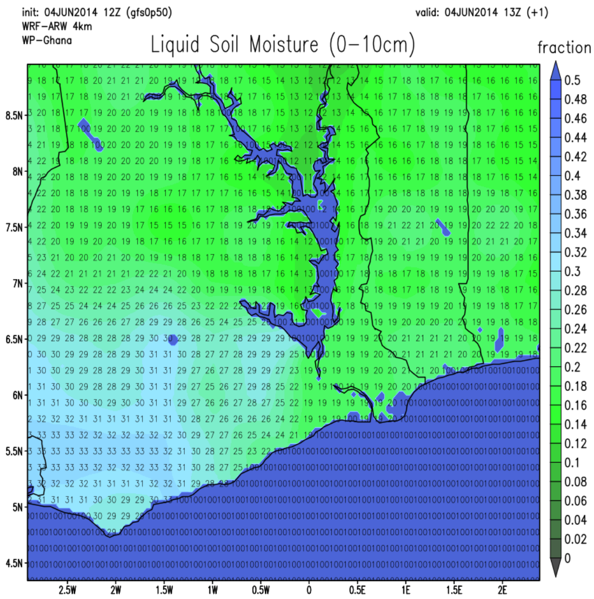


Figure 2.5 – Modified IGBP-MODIS NOAH categories, most dominant land use of the 4km domain



(a) 2m dew point temperatures of the 4km domain as forecast by WRF



(b) 0-10cm soil moisture fraction of the 4km domain as forecast by WRF

Figure 2.6 – **(a)** and **(b)** illustrate lake areas behave as one would expect. Dew points over water are significantly higher than surrounding areas and the soil moisture content is at a 100 percent.

Chapter 3

Observations

3.1 Introduction

Surface observations in Ghana and surrounding areas are scarce. With precipitation measurements almost non existent and no operational precipitation radar, model verification is restricted to remote sensing via satellites. At first, this seems like it will not make verification easier, but it does have its advantages. Satellites can cover large regions at, usually, a much higher resolution than what is available from surface observations while also covering remote areas such as oceans.

Satellite derived precipitation products are, however, only indirect estimates using algorithms and bias corrections, not direct measurements. Since convective cloud tops are often larger than the area producing rainfall, these products generally overestimate the area of precipitation. Infrared based products in particular, since it is not possible to see through cloud tops with infrared wavelengths. Visible products have their downsides too, as will be discussed momentarily. These products are therefore to be treated with caution. To not rely on just one technique, two different types of satellite products are used. If the results turn out not to be dependent on which product is used, confidence in the conclusions will increase.

3.2 Precipitation observations derived from satellites

3.2.1 MSG-CPP

One of the precipitation observation products used in this research is MSG-CPP (Meirink, 2012). The algorithm, named Cloud Physical Properties (CPP), is being developed mainly at the Royal Dutch Meteorological Institute. This algorithm uses satellite sensors such as SEVIRI[1] which are on board of the Meteosat Second Generation (MSG) satellites to estimate properties such as water droplet size, cloud optical thickness, incoming surface shortwave radiation and cloud top temperature (Brasjen, 2014).

This product is only available during daytime, as the algorithm depends on the backscattering of incoming solar radiation. The product is masked in the areas where the angle between the sun and earth's surface is too low. However, the product is still too unreliable in areas close to mask border, since tops of convective storms cast shadows onto other parts of the storm, influencing the derived precipitation signal. Therefore, the used temporal domain is scaled down slightly further and ranges from 8 to 16 UTC, available once every 15 minutes. The spatial resolution of this product in the area of interest is just under 3km, therefore it shows details of large scale convective systems quite well. The calibration of this product is done only once, so it may underestimate or exaggerate the intensity and area of rainfall, usually the latter.

3.2.2 GPM-IMERG

Global Precipitation Measurement (GPM) is a joint mission of NASA and the Japanese Aerospace Exploration Agency (JAXA) to provide worldwide precipitation observations (Huffman, Bolvin, and Nelkin,

[1] Spinning Enhanced Visible and Infrared Imager

2015). Using many available satellites and one specifically launched for this mission, a precipitation product with a 0.1x0.1° spatial resolution has been developed. The algorithm used is the Integrated Multi-satellitE Retrievals for GPM (IMERG), which combines satellite imagery and calibrates the product using, if available, surface observations such as rain gauges and precipitation radars but also radar sensors in satellites. The algorithm continuously re-calibrates with the information available to improve reliability.

There are multiple versions available, which differ in the amount of observational data used. The dataset used here for verifying WRF output is the half hourly 'final' product version, which includes all possible observational data into the algorithm. Despite having a coarser spatial resolution compared to MSG-CPP, it's availability at night and continuous recalibration is an advantage.

Chapter 4

Method

To set up a usable model for Ghana, many things can be researched to find optimal settings for precipitation forecasts. Very important are, as mentioned in the introduction, the cumulus parameterization and microphysics schemes. These two components of the model will be looked at in detail in order to find a suitable combination. In order to do this, a method has been developed described in the following section.

4.1 Case studies

The search for suitable cases started by looking for days on which fatal flash foods have occurred. These would be equivalent to operational cases in which the model has to perform well. The selection has been narrowed down to two cases, chosen to be different in atmospheric setup so the model can be tested on more than one aspect. Found in both cases, though, in an unstable airmass covering an area larger than Ghana. Surface based and mixed layer CAPE[1] values in both cases are in the order of 2000-3000 J/kg. Additionally, PWAT[2] ranges from 45 to 65mm, so there is potential for severe convection and rainfall and thus flash floods.

Firstly, 5 June 2014 (Davies, 2014), from now on named 'case A', features a mesoscale convective system (MCS) originating as far east as Nigeria. At approximately 9 UTC, what is left of the MCS enters Ghana from the east and strengthens due to the diurnal cycle. It passes right over Accra causing flash floods and then propagates further westward. New storms get set off on the outflow boundary which leaves the country around 20 UTC. During all this time, just north and south of the coastline, scattered, less severe storms are present. Overnight convective systems, leftovers from the day before, dissipated or left Ghana before sunrise.

3 June 2015 (Silver, 2015), hereafter named 'case B', is a different story. Around 12 UTC, a zone of converging winds caused by a sea breeze forms and starts triggering convection along its length. About 2 hours later it becomes more widespread as the boundary layer heats up and outflow boundaries start triggering new cells further north. Activity continues throughout the evening.

4.2 Combinations of cumulus parameterizations and microphysics schemes

A selection of three cumulus parameterizations and two microphysics schemes will be tested in all possible combinations (table 4.1). Additionally, based on results of these initial 12 runs, a few simulations are run without the use of a cumulus parameterization in the 4km domain. This is to gain more insight into the behavior and necessity of such schemes. 'a' and 'b' suffixes will be used to label those runs. E.g. 4 would become 4a, 4b is then the additional run.

WRF Single Moment 6-class (WSM6) and Thompson microphysics, both single moment 6-class schemes including graupel, are the selected microphysics. These two will initially be tested in combination with three cumulus schemes. All three are variations of Kain-Fritsch (KF) schemes, with different

[1] Convective Available Potential Energy in J/kg, a way to quantify atmospheric instability

[2] Precipitable WATer, the amount of moisture in a column of air expressed in terms of how much rain would fall if it all precipitated

convection triggers. The first of three is the default WRF cumulus scheme, but set to use a modified version of the KF trigger, named KFMT, designed to better function for convection under weak synoptic forcing (Ma and Tan, 2009). This is applicable to tropical convection and thus worth testing. The second option is a multi scale trigger for the KF scheme (MSKF), which adjusts its convection criteria coefficients based on the numerical grid resolution. It has shown improvement for high resolution grids compared to the default KF trigger (Zheng et al., 2016). Lastly, WRF v3.8 features a newly introduced variation of the default KF scheme. It replaces the ad hoc trigger by one linked to the boundary layer properties via probability density functions, a cumulus potential (CuP) method (Berg, Gustafson, Kassianov, and Deng, 2013). It has adjustable coefficients to do with criteria like bin frequency and size, but the default ones will be used.

Table 4.1 – Cumulus parameterization and microphysics combinations for case A and B by run number

Case A	KFMT	MSKF	CuP
WSM6	1	2	3
Thompson	4	5	6
Case B			
WSM6	7	8	9
Thompson	10	11	12

4.3 Verification method

Verifying the model's precipitation output will be done using various statistical scores to objectively measure its performance (Doswell, Davies-Jones, and Keller, 1990, Hogan, Ferro, Jolliffe, and Stephenson, 2010). Interpreting them will be done in combination with looking at model outputs and also placing them in the context of the case at hand.

4.3.1 Performance scores

The first set of statistical scores will be calculated using a 2x2 contingency table. An event is said to occur when, in a grid point, precipitation above a certain threshold is reached, resulting in a 'yes' or 'no'. A model grid point can then be compared to an observation grid point, resulting in one of four possibilities as shown in table 4.2.

Table 4.2 – 2x2 contingency table

		Observed		
		Yes	No	
Forecast	Yes	a	b	a+b
	No	c	d	c+d
		a+c	b+d	n = a+b+c+d

The bias scored then defined as follows,

$$\mathrm{bias} = \frac{\mathrm{a} + \mathrm{b}}{\mathrm{a} + \mathrm{c}} \tag{4.3.1}$$

where a value of 1 means unbiased, $\mathrm{bias} > 1$ means the amount of events is overforecast and $0 \leq \mathrm{bias} < 1$ means underforecast. d does not appear in this equation, because precipitation is a low frequency event and 'no' observation or forecast is therefore overrepresented. Eq. (4.3.1) is not scaled by n, so it might be misleading when there is very little precipitation forecast or observed. A score which indicates how well the forecast overlaps with observations in terms of area is the Gilbert Skill Score, also known as the Equitable Threat Score. It shows the fraction of correctly forecast events, correcting for hits by chance,

$$\text{GSS} = \frac{a - a_{ref}}{a - a_{ref} + b + c} \quad (4.3.2) \qquad a_{ref} = \frac{(a+b)(a+c)}{n} \quad (4.3.3)$$

where a_{ref} are hits expected by chance. Its values lie between $-1/3$ and 1, 1 being perfect forecast and anything ≤ 0 shows no skill.

Another way of comparing observations and forecasts is by calculating statistics based on absolute precipitation values instead of 'yes' or 'no' events. This is done by calculating the following quantities,

$$\text{MAE} = \frac{1}{n}\sum_k |m_k - o_k| \quad (4.3.4) \qquad \text{MSE} = \frac{1}{n}\sum_k (m_k - o_k)^2 \quad (4.3.5) \qquad \text{ME} = \frac{1}{n}\sum_k (m_k - o_k) \quad (4.3.6)$$

Summation is done over k, where k is the k-th grid point. Model and observation values are named m and o respectively. The Mean Absolute Error (4.3.4) and Mean Square Error (4.3.5) are both positive numbers, where a higher outcome means a worse performance. Eq. (4.3.5) emphasizes large differences. A quantitative bias, as opposed to Eq. (4.3.1) is given by Eq. (4.3.6), where $\text{ME} > 0$ means the model overestimates the total rainfall. Values resulting from these equations can be compared to grid averaged precipitation values to get an idea of the relative size of the errors.

4.3.2 Reprojection of datasets

Both GPM, MSG and the WRF domain have different spatial resolutions and grid projections, so before comparison they need to be reprojected onto a common grid. After reprojection, each grid point of each dataset should represent the same location and area. To avoid interpolation, data will be downsampled to the coarsest grid resolution, which is GPM. For a given input domain, each grid point with a corresponding latitude and longitude will be linked to the closest grid point of the destination domain (figure 4.1). Since the input domains have a higher resolution, more than one grid point will be linked to a destination location. Each of these input grid points will have a precipitation value, which in turn are averaged (unweighted) to arrive at a single value for the new grid point. It should be noted that averaging smooths out extremes.

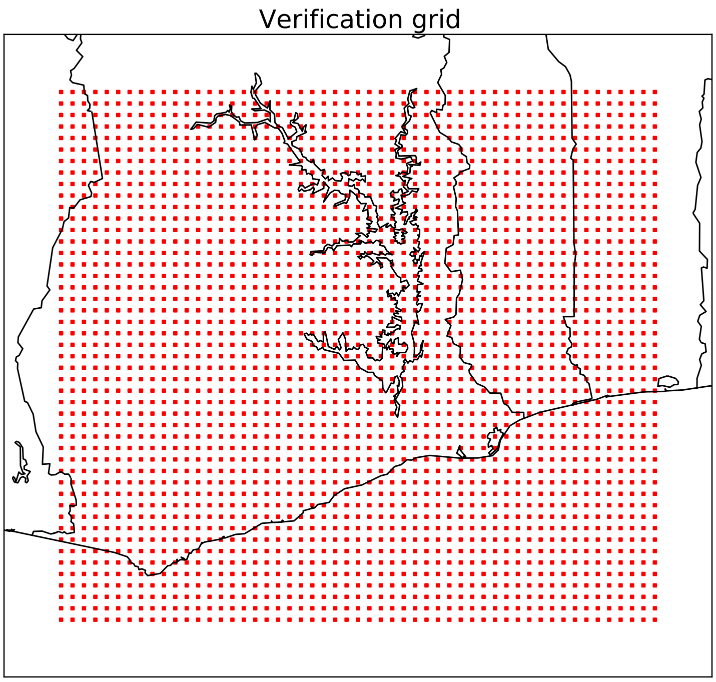


Figure 4.1 – The verification grid, where WRF and MSG are projected on. Red dots represent grid points, each point will have a precipitation value.

Chapter 5

Results and Discussion

Both cases will be looked at separately, due to their significant difference. Data points in figures are plotted at the end of their respective precipitation accumulation range, i.e. 10 UTC means the score based on 8-10 UTC precipitation totals. For identifying run numbers, refer to table 4.1. Runs 4, 5, 10 and 11 have also been run without cumulus parameterization in the 4km for reasons explained in the following sections.

5.1 Case A

The observed MCS, or at least something close to it, is visible in all runs in both domains. Problematic for some runs, especially 1 and 2, is, however, the velocity at which it propagates. As shown in figure 5.1, runs 1 and 2 peak significantly earlier, causing a lack in overlap with observations which has a negative impact on the performance scores. Other runs seem to estimate the peak precipitation rather well, with an offset of about an hour. A reason for this might be the difference in how well structured the convective system is. An example is illustrated in figure 5.5 by simulating radar reflectivity, which shows run 1 has a smooth bowed structure, whereas run 5a is unstructured and consists of multiple cells.

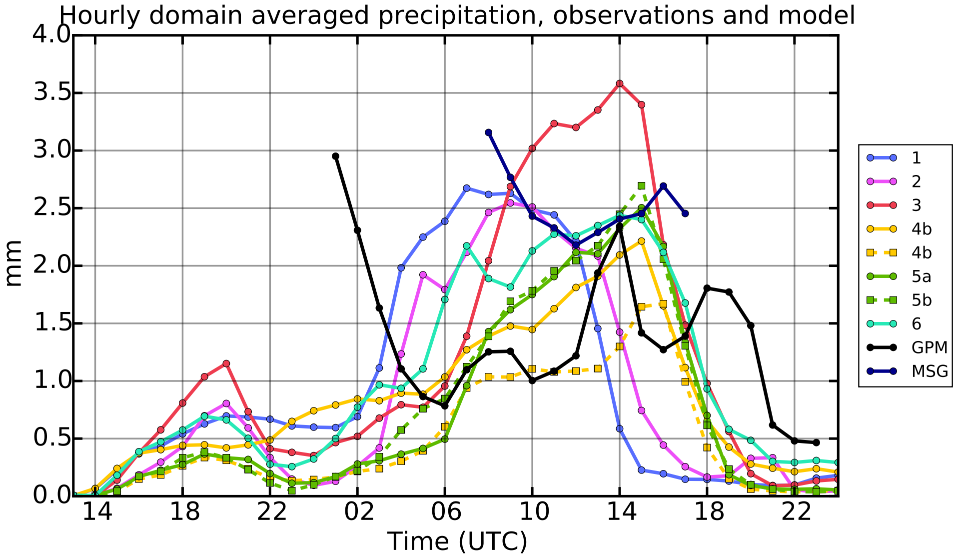


Figure 5.1 – Hourly domain averaged precipitation in mm for both observations and model runs of case A

The overproduction of rainfall is most significant in runs 3 and 6, though the distribution throughout the day is different. Clear becomes the fact that overnight precipitation, as shown by GPM, was missed by all runs. Moreover, while all runs cease to produce significant precipitation early in the evening, GPM shows still considerable amounts of activity at 18-21 UTC. An overview of rainfall accumulations including all datasets relevant to this case can be seen in figure 5.2. It should be noted that MSG and

GPM look very different in figure 5.1. While it is mostly a difference in intensity, judging by their rainfall accumulations shown in figure 5.2, the location of precipitation is also not identical.

Statistical scores have been calculated from 8 to 16 UTC, a time frame in which both the GPM and MSG datasets can be used. The interval is 2 hours, in order to reduce the influence small differences in timing have and also smooth out GPM's calibration. The relative performance of runs based on these scores appear not to be dependent on whether GPM or MSG was used to verify the model. Due to their differences in total rainfall though, the bias score can not easily be interpreted. Runs closest to a bias of 1 also differ per threshold, highlighting the difference in performance between forecasting light and heavy precipitation. Evident from figure 5.3 is the fact that run 4a produces a lot more light precipitation than 4b, but does much less so for heavy precipitation. 4b, which has no cumulus parameterization but is otherwise the same, has a bias of less than half of 4a for a 1mm threshold, hinting the cumulus parameterization scheme too easily produces light rainfall. Another consistent feature is run 1 and 2 having one of the highest bias scores, declining down to almost 0 towards to end of the afternoon, indicating they let the convective system move too fast, as was noted earlier.

The Gilbert Skill Score is the highest for run 5a and 5b for both light and heavy precipitation (figure 5.4), independent of what observation source has been used. It ranges from about 0.0 to 0.2 for a 1mm threshold and 0.0 to 0.06 for a 10mm threshold. Especially the 10mm threshold shows low scores overall and there appears to be a declining trend throughout the day. All other runs perform worse, with many scoring below 0, indicating they have no forecast skill.

By looking at errors in the total rainfall between 8 to 16 UTC (table 5.1), it becomes evident that errors are large relative to the mean precipitation that falls. Also, the bias (4.3.6) is consistently lower when compared to MSG instead of GPM. Furthermore, absolute and squared differences are much larger for MSG, but as before, the relative performance between runs remains mostly affected by this. Runs 3 and 6 show the most significant deviation from observations due to the overproduction of precipitation. Runs 1, 4a and 4b appear to be the most careful with producing precipitation. 4b and 5b produce slightly less rainfall than their counterparts (4a, 5a), but their error relative to the total rainfall does not show a consistent improvement.

Table 5.1 – Case A MSE and mean squared precipitation (mm^2), MAE, ME and mean precipitation (mm) of 8 to 16 UTC accumulated rainfall

8 - 16 UTC	MSE		Mean squared pcp	MAE		ME		Mean pcp
run	GPM	MSG	WRF4km	GPM	MSG	GPM	MSG	WRF4km
1	570	1113	497	16.1	21.6	0.7	-7.3	12.2
2	702	1207	724	16.6	21.9	2.6	-5.4	21.9
3	1388	1744	1797	23.5	27.6	12.8	4.8	24.4
4a	592	1058	563	16.1	20.4	2.7	-5.3	14.2
4b	540	1005	446	15.0	19.7	-1.6	-9.6	10.0
5a	863	1171	977	18.5	21.2	4.9	-3.2	16.4
5b	745	1050	849	18.2	20.5	5.1	-2.9	16.7
6	1074	1512	1247	19.6	24.0	6.2	-1.8	17.7

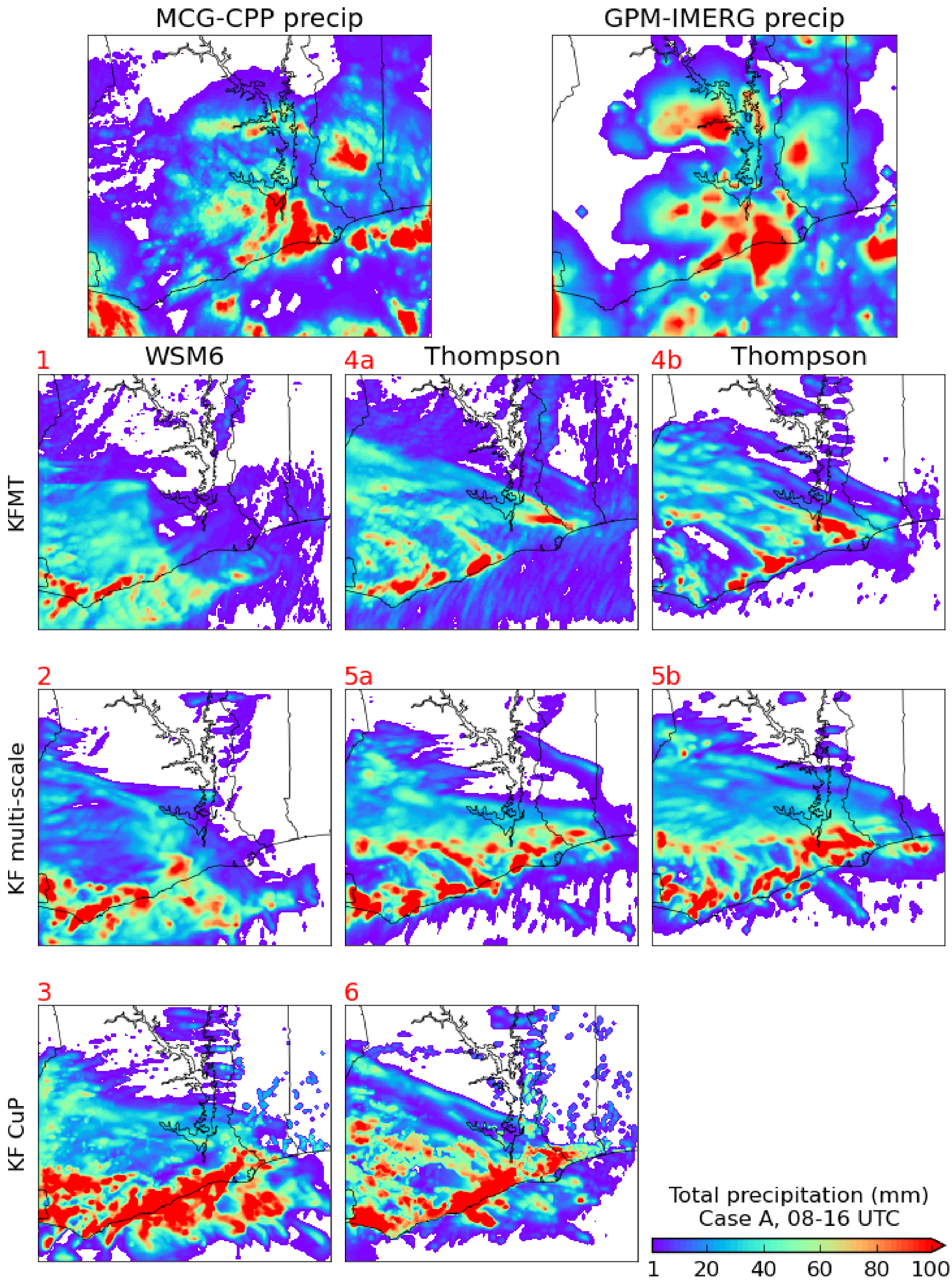
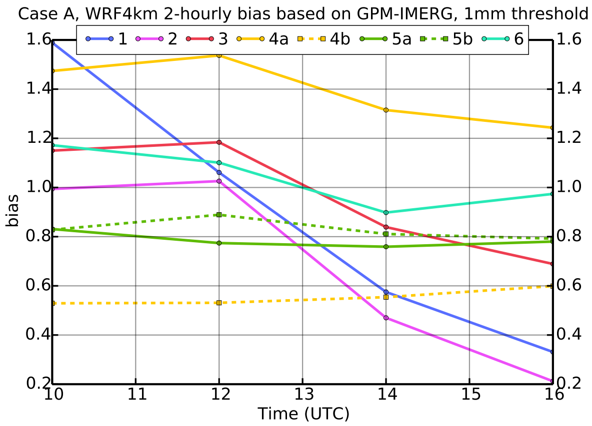
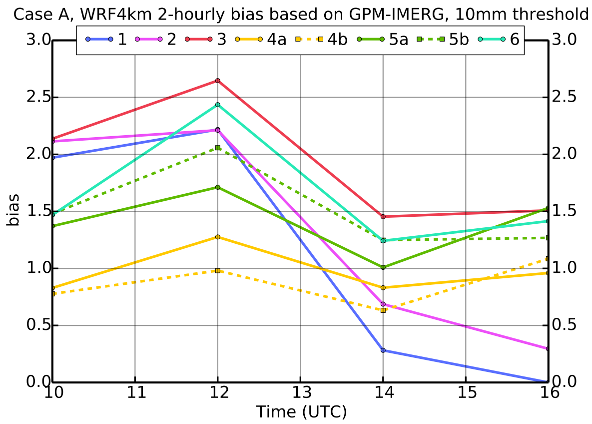


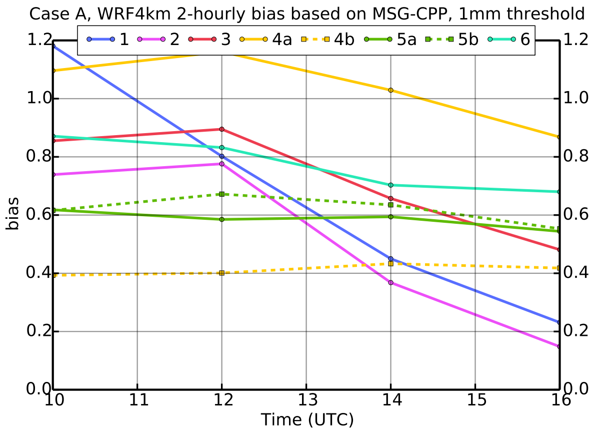
Figure 5.2 – Overview of rainfall totals from 8 to 16 UTC of both observational datasets and the eight WRF runs with different physics settings. Rows have the same cumulus parameterization, columns are identical in microphysics. Two exceptions are 4b and 5b, these do not have a cumulus parameterization scheme themselves, but their parent domain does.



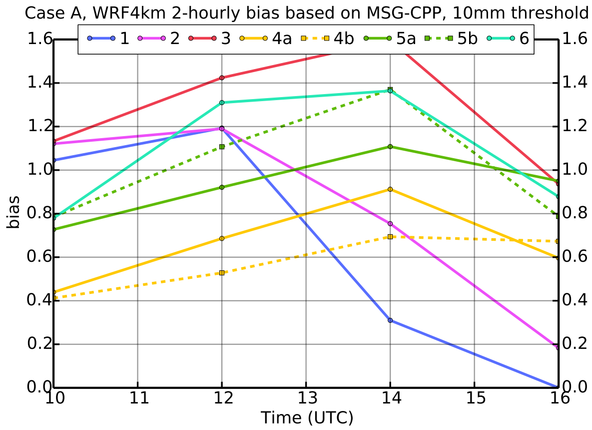
(a)



(b)



(c)



(d)

Figure 5.3 – Bias score of case A, based on 2-hourly accumulated precipitation, verified with **(a, b)** and MSG **(c, d)** datasets

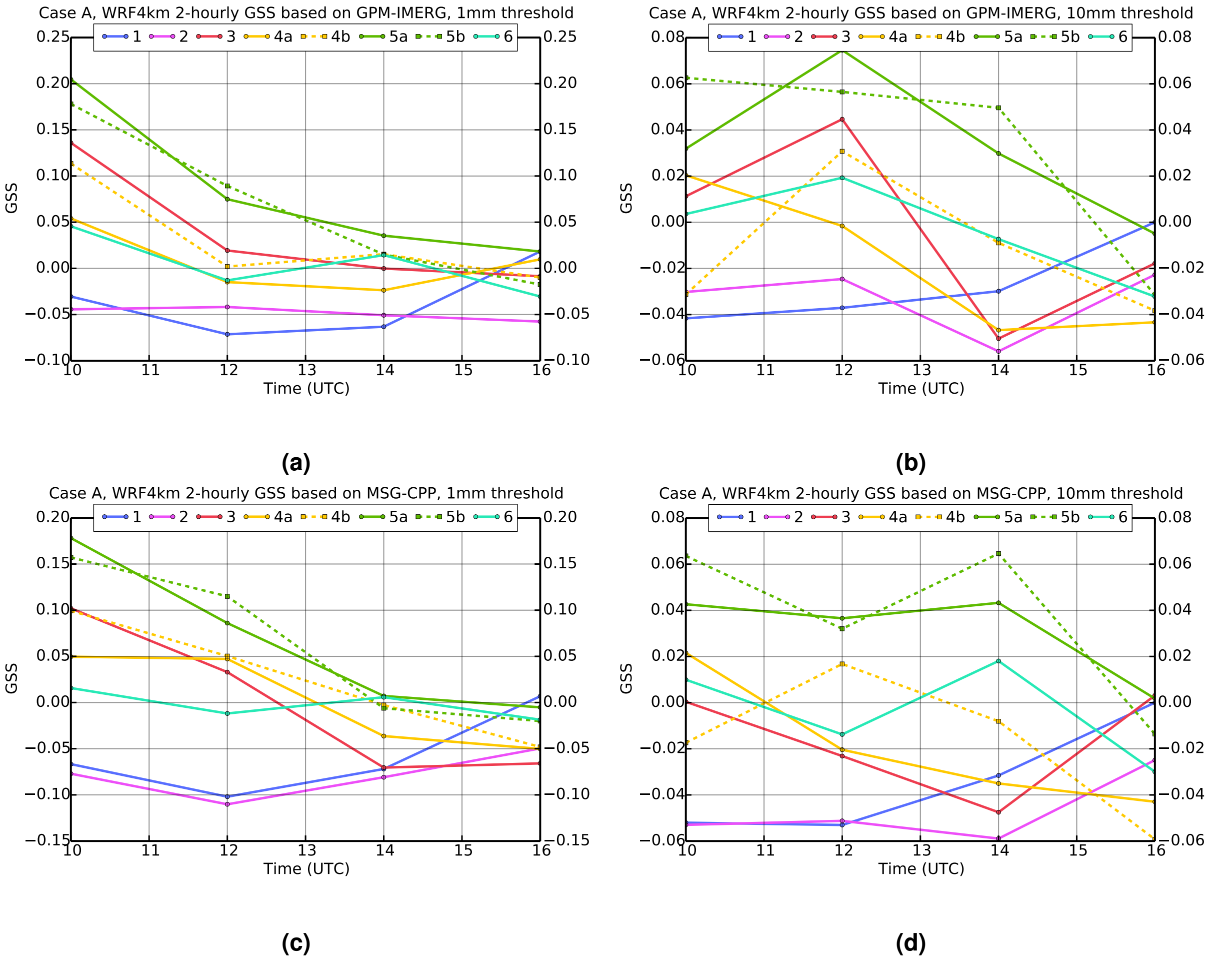


Figure 5.4 – Gilbert Skill Score of case A, based on 2-hourly accumulated precipitation, verified with GPM **(a, b)** and MSG **(c, d)** datasets

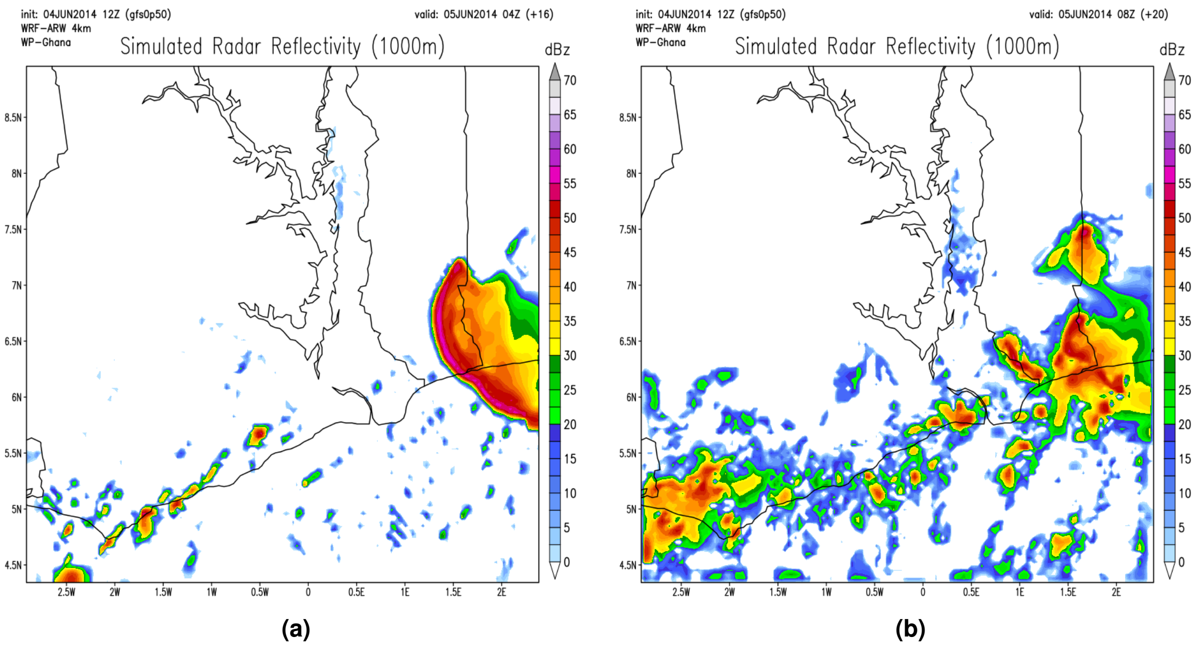


Figure 5.5 – Simulated radar reflectivity at 1000m above ground level of run 1 **(a)** and run 5a **(b)**

5.2 Case B

Convection of this case initiates along a sea breeze front during the second half of the afternoon (14 UTC) and continues throughout the evening. Thus MSG will not be used for the second half of this case, as the product is not usable after 16 UTC. Bias and skill scores are calculated from 8 to 22 UTC, using only GPM this time, for the same reason. MSG and GPM lie close together in terms of location and intensity, unlike in case A, illustrated in figure 5.6. The moment of convective initiation of all model runs seem to be too early, but the strong increase of activity around 14 UTC coincides with observations across all runs.

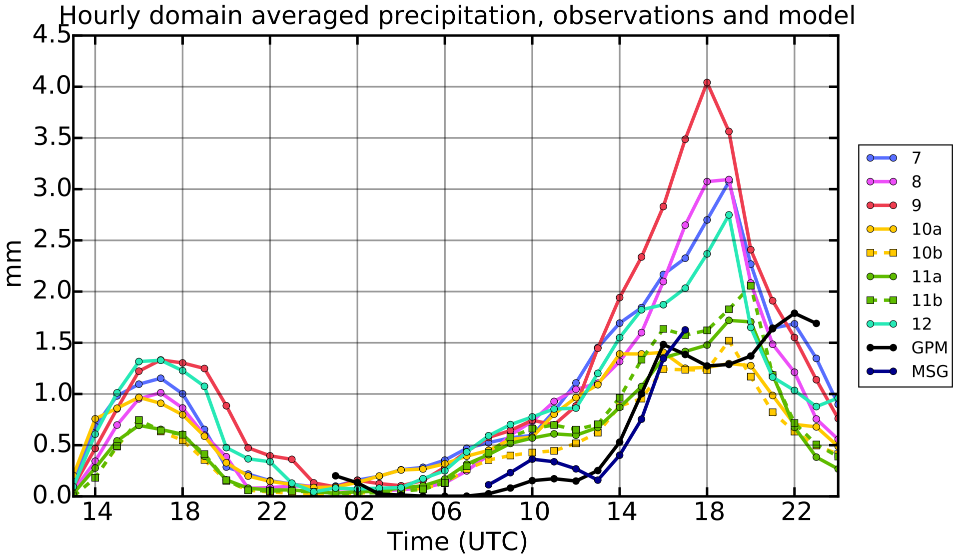


Figure 5.6 – Hourly domain averaged precipitation in mm for both observations and model runs of case B

According to observations, it is mainly dry until 13 UTC, but all runs show rainfall, evident from the rather large positive bias (figures 5.6, 5.8) which declines as the day progresses. Once again, like in case A (5.1), the cumulus parameterization in run 10a seems to cause excessive light precipitation, whereas run 10b shows a much improved bias of 1 to 2 instead of 5 to 7. The difference is again small in case of a 10mm threshold. Runs 10b, 11a and 11b are the most unbiased of all compared to others. Based on just the GSS (figure 5.9), it seems the model has a lot of trouble forecasting this case. Though likely caused by the local nature of the scattered convective storms, no single run performs well throughout the day. Runs 11a and 11b do relatively well in the earlier stages, but break down below 0 in the evening.

By looking at errors in accumulated precipitation between 8 and 16 UTC, table 5.2, it is once again visible how all runs are positively biased. 10a,b and 11a,b show the least errors, which was also true in case A. Due to the fact most convection occurs in the second half of the afternoon and early parts of the evening according to observations, another time range has been analyzed. 12 to 22 UTC errors based on GPM, shown in table 5.3, are the smallest for 10a,b and 11a,b. Most noticeable is their bias being close to 0, which combined with their respective poor skill scores indicate the model has trouble locating convective storms, but does rather well in forecasting precipitation accumulations for this case. Further evidence for this claim is illustrated in figure 5.7, where runs 10a,b and 11a,b show a line of excessive precipitation (100mm or more) along the coastline instead of slightly further landward. A possible cause is the line of converging winds not making it properly past the coastline, which is illustrated in figure 5.10 for run 11a specifically. At 8 UTC there is a clear zone of convergence just south of the coast line, but once the sea wind makes it past the coastline the convergence has mostly vanished. As a result, convection triggers too early and in the wrong place. Additionally, scattered thunderstorms still develop in the afternoon due to a very unstable atmosphere. As shown in figure 5.11 **(b)**, CAPE values are in the order of 2000-3000 J/kg with CIN values mostly -10 to 0 J/kg, indicating free convection can occur. In absence of the zone of converging winds, free convection is indeed what happens in this model run.

Table 5.2 – Case B MSE and mean squared precipitation (mm^2), MAE, ME and mean precipitation (mm) of 8 to 16 UTC accumulated rainfall

8 - 16 UTC	MSE		Mean squared pcp	MAE		ME		Mean pcp
run	GPM	MSG	WRF4km	GPM	MSG	GPM	MSG	WRF4km
7	298	347	319	10.0	10.3	6.4	6.3	10.2
8	280	326	331	9.0	9.4	56	5.6	9.4
9	445	493	498	11.4	11.8	7.7	7.7	11.6
10a	201	240	206	8.1	8.3	4.3	4.3	8.2
10b	235	266	206	7.0	7.2	1.7	1.6	5.5
11a	224	305	243	7.1	7.6	2.4	2.4	6.3
11b	287	331	294	8.1	8.3	3.3	3.3	7.2
12	442	498	483	10.1	10.5	5.9	5.9	9.7

Table 5.3 – Case B MSE and mean squared precipitation (mm^2), MAE, ME and mean precipitation (mm) of 12 - 22 UTC accumulated rainfall

12 - 22 UTC	MSE	Mean squared pcp	MAE	ME	Mean pcp
run	GPM	WRF4km	GPM	GPM	WRF4km
7	694	782	19.1	8.7	20.7
8	709	820	18.9	7.6	19.6
9	1016	1329	22.0	13.5	25.5
10a	367	268	13.2	0.0	12.0
10b	478	318	14.1	-1.8	10.2
11a	534	380	16.0	0.1	12.1
11b	605	469	16.7	1.5	13.5
12	584	689	16.0	5.3	17.3

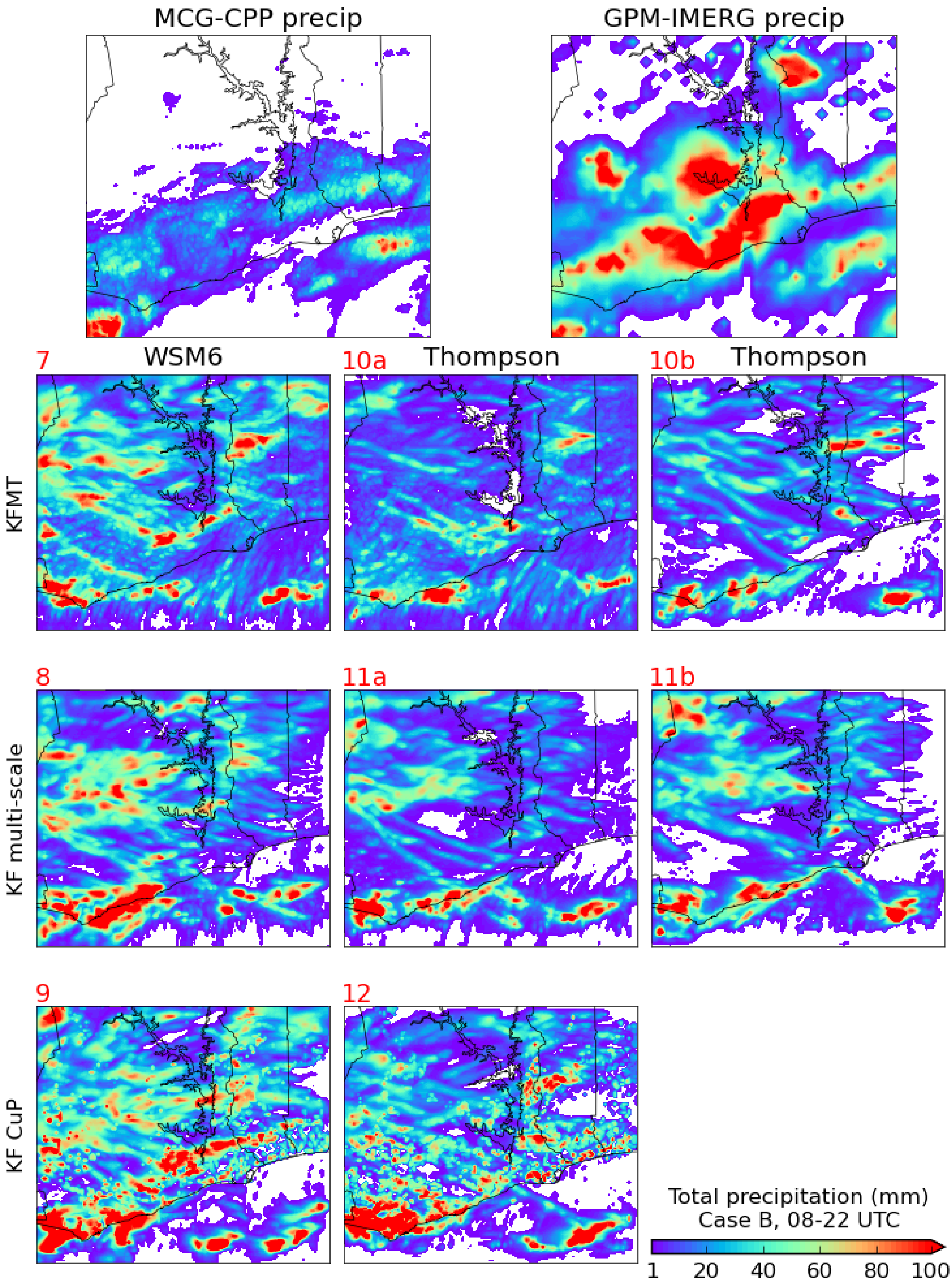
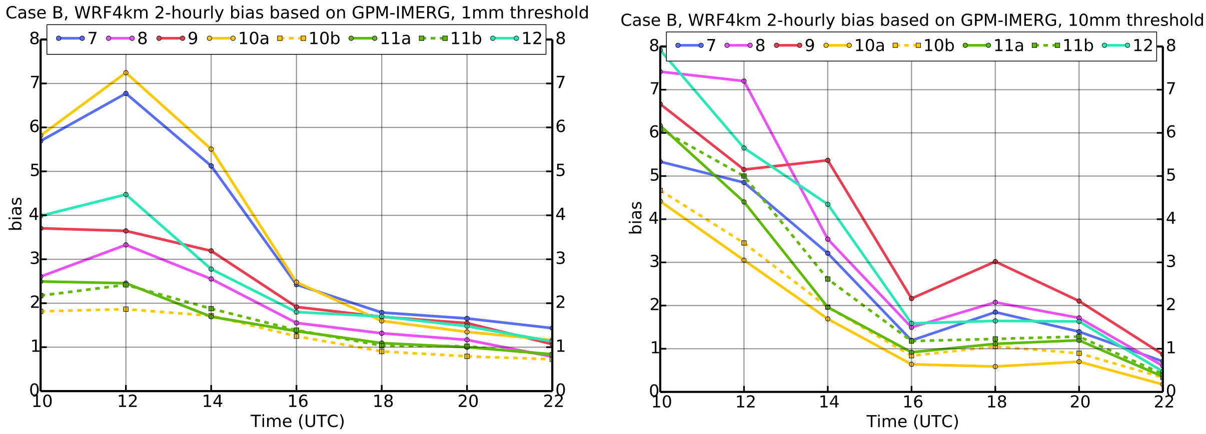
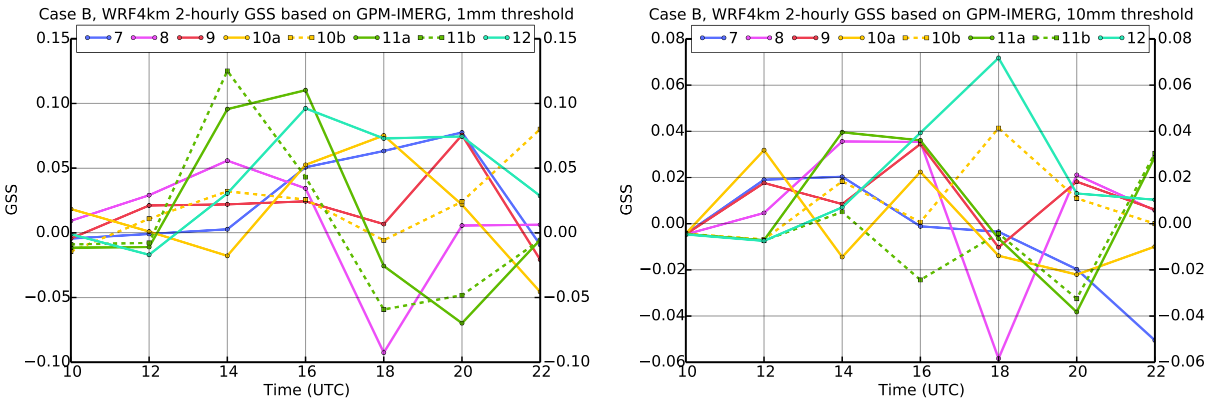


Figure 5.7 – Overview of 8 to 22 UTC rainfall totals of both observational datasets and eight WRF runs with different physics settings. Rows have the same cumulus parameterization, columns are identical in microphysics. Two exceptions are 10b and 11b, these do not have a cumulus parameterization scheme themselves, but their parent domain does. The MSG-CPP plot falls behind the others due a lack of availability in this time range.



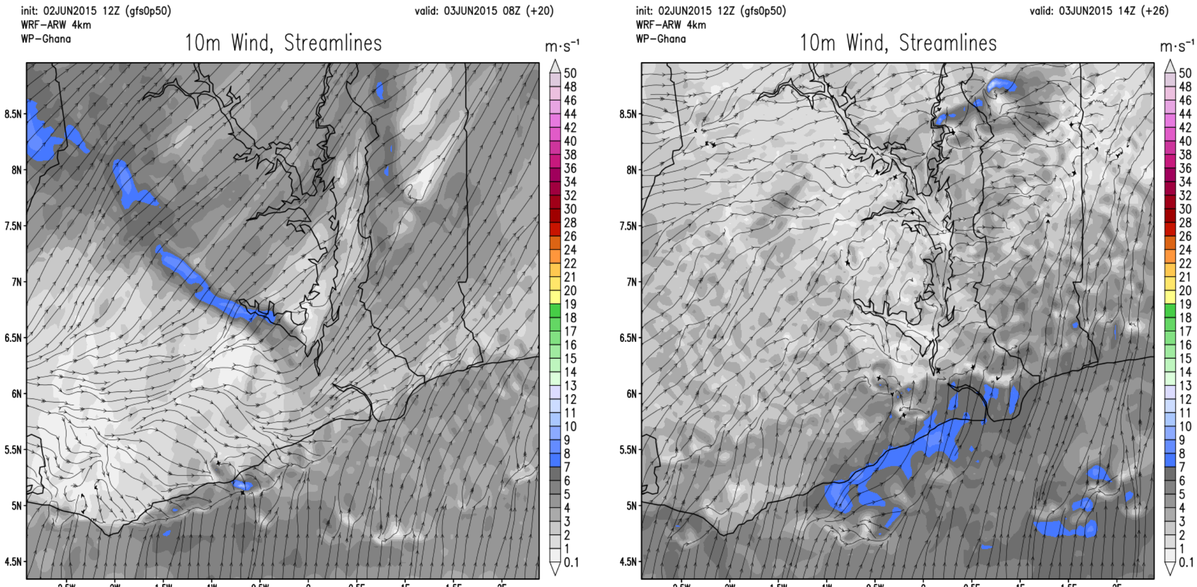
(a) (b)

Figure 5.8 – Gilbert Skill Score of case B for a 1mm **(a)** and 10mm **(b)** threshold, based on 2-hourly accumulated precipitation, verified with GPM)



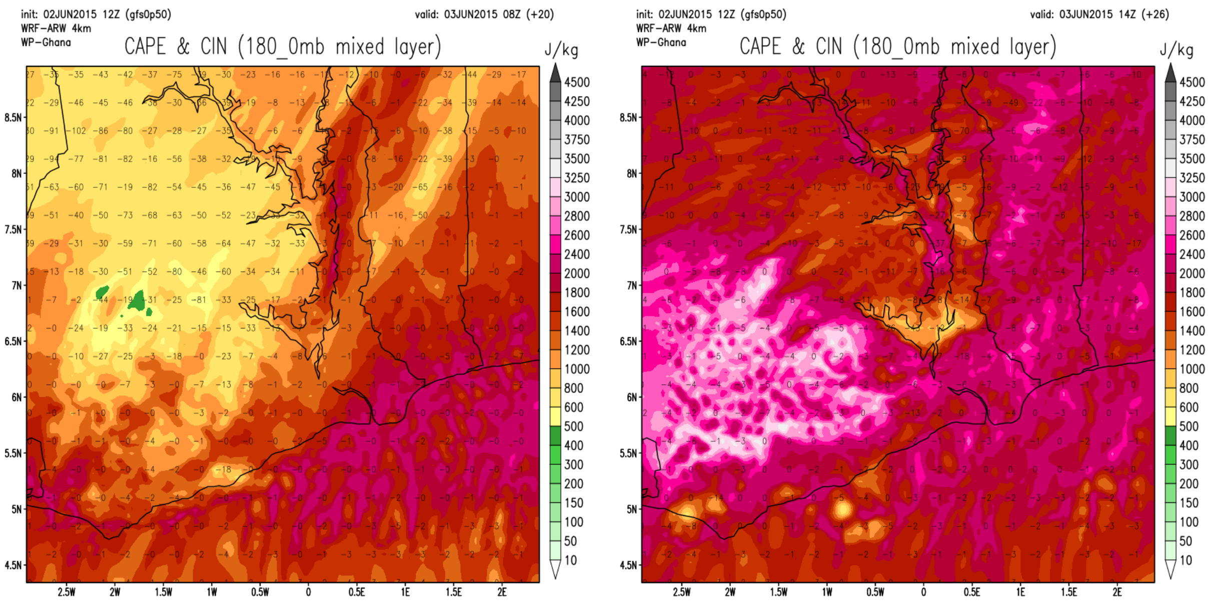
(a) (b)

Figure 5.9 – Gilbert Skill Score of case B for a 1mm **(a)** and 10mm **(b)** threshold, based on 2-hourly accumulated precipitation, verified with GPM



(a) (b)

Figure 5.10 – 10m wind direction and magnitude of run 11a. A zone of converging winds is visible just south of the coastline at 8 UTC **(a)**, whereas is has mostly vanished six hours later as the sea wind has moved further landward **(b)**



(a) (b)

Figure 5.11 – Mixed layer (lowest 180hPa) CAPE (shaded) and CIN (text) at 8 UTC **(a)** and 14 UTC **(b)** of run 11a.

Chapter 6

Conclusion

A 12km resolution domain with a 4km resolution nest, centered around Ghana, has been designed. While the choice of most physics schemes have been based on recommendations, a closer look was taken at the schemes regarding precipitation. Even though only two cases have been studied, a lot has become clear about the model's performance in forecasting precipitation under different combinations of cumulus parameterization and microphysics schemes. By looking at bias and the Gilbert Skill Score as well as errors in rainfall accumulations over longer periods, sensitivity to which combination has been used appears to be large. The relative performance of combinations are mostly independent of whether GPM or MSG was used as a verification dataset. GPM and MSG were generally the same in case B, but case A showed a clear difference between the two. In absolute terms, no run performs well in forecasting both the location and intensity of convective storms. Also the propagation velocity of large scale convective systems is off for some combinations. Moreover, all combinations missed the convection of the night leading into case A. Case B shows a poor performance after sunset too, as all runs underestimate rainfall activity.

There are, however, certain combinations that performed relatively and consistently well in both case A and B. Initially, all runs included a cumulus parameterization scheme in both domains. Case A showed in runs 1 and 2 the MCS propagated too fast, thus the WSM6 microphysics in combination with either the multi-scale or tropical trigger is not good. These two combinations also overestimate precipitation in case B by a lot. Overestimation of precipitation happens most in runs 3, 6, 9 and 12, which have the KF CuP scheme in common, therefore this scheme will not be used. Runs 3 and 9 are more positively biased than 6 and 12 respectively, i.e. WSM6 microphysics tends to overproduce precipitation more than the Thompson scheme. Additionally, runs 4a, 5a, 10a and 11a perform consistently better than their contourparts using WSM6 in terms of forecast skill, bias and MSE, making the Thompson microphysics consistently better than WSM6 for this application.

At this point, there were two cumulus schemes left to choose from. In order to gain more insight into their particular behavior, four more simulations have been run, named 4b, 5b, 10b and 11b, which do not include cumulus parameterization in the 4km domain. 4a and 10a show a significantly lower bias at a threshold of 1mm compared to 4a and 10a, indicating the KFMT trigger causes the scheme to produce light precipitation too easily. The MSKF scheme does not have this problem. A reason for this might be the fact the MSKF scheme adjusts its convection trigger parameters based on the grid resolution and is therefore more suitable for a high resolution grid than other schemes. At a 10mm threshold, no clear difference was found between KFMT and MSKF based on the extra runs. From this, it can be concluded Thompson microphysics with the multi scale Kain-Fritsch cumulus parameterization is the best performing combination of all those tested, in both case studies.

It is unclear whether the overall low scores will turn out to be a problem for the hydraulic model. Low scores due to small differences in timing and location might still result in accurate enough warnings, but answering this question requires case studies in combination with the hydraulic model and a definition of what is considered to be 'accurate enough'. Needless to say, the current weather model and thus the warning system is not at an operational level yet.

6.1 Further Research

The method of this research has been very focused on just two cases and a limited set of varying physics schemes. A straightforward approach would be to look at more cases, try out different initializing times, and test additional cumulus parameterization and microphysics schemes. Adjusting the KF CuP's coefficients to make the trigger more suitable for a tropical environment might also be an option. To gain more insight into why forecasts are off, a closer look can be taken at atmospheric properties like pressure, moisture and wind direction verified with surface observations and soundings. Possible mistakes by the model might be attributed to, for example, surface and boundary layer schemes. The distribution of vertical levels in the model domain is done automatically by the model, which is not necessarily the best. This could be set manually, for example to include more levels in the lower part of the model domain, possibly leading to improvement boundary layer processes. Raising the model top beyond 50hPa could also help diffuse deep convection better. An under highlighted aspect in this research are the input data. Initialization at noon might worse than at midnight, but also whether analyses or spinned up GFS forecast data are used might be of great influence and worth taking a look at.

Bibliography

Africa itcz monitoring. URL http://www.cpc.ncep.noaa.gov/products/international/itf/itcz.shtml. Accessed: 2016-05.

Infoplaza: Expertise nu ook in ghana. URL http://www.infoplaza.nl/archives/2213/.

L. K. Berg, W. I. Gustafson, E. I. Kassianov, and L. Deng. Evaluation of a modified scheme for shallow convection: Implementation of CuP and case studies. *Mon. Wea. Rev.*, 141(1):134–147, jan 2013. doi: 10.1175/mwr-d-12-00136.1. URL http://dx.doi.org/10.1175/MWR-D-12-00136.1.

A. M. Brasjen. Precipitation estimation from infrafred satellite imagery. Master's thesis, Delft University of Technology, aug 2014.

L. Costantino and P. Heinrich. Tropical deep convection and density current signature in surface pressure: comparison between WRF model simulations and infrasound measurements. *Atmospheric Chemistry and Physics*, 14(6):3113–3132, mar 2014. doi: 10.5194/acp-14-3113-2014. URL http://dx.doi.org/10.5194/acp-14-3113-2014.

R. Davies. Flood and fire disasters in accra, ghana. jun 2014. URL http://floodlist.com/africa/flood-accra-ghana-fire-explosion.

C. A. Doswell, R. Davies-Jones, and D. L. Keller. On summary measures of skill in rare event forecasting based on contingency tables. *Wea. Forecasting*, 5(4):576–585, dec 1990. doi: 10.1175/1520-0434(1990)005⟨0576:osmosi⟩2.0.co;2. URL http://dx.doi.org/10.1175/1520-0434(1990)005<0576:OSMOSI>2.0.CO;2.

J. Dudhia. WRF physics options. Technical report, National Center for Atmospheric Research, 2008. URL http://www2.mmm.ucar.edu/wrf/users/tutorial/200807/WRF_Physics_Dudhia.pdf. Accessed: 2016-03.

R. J. Hogan, C. A. T. Ferro, I. T. Jolliffe, and D. B. Stephenson. Equitability revisited: Why the "equitable threat score" is not equitable. *Wea. Forecasting*, 25(2):710–726, apr 2010. doi: 10.1175/2009waf2222350.1. URL http://dx.doi.org/10.1175/2009WAF2222350.1.

G. J. Huffman, D. T. Bolvin, and E. J. Nelkin. Integrated multi-satellite retrievals for gpm (imerg) technical documentation. Technical report, Mesoscale Atmospheric Processes Laboratory, NASA Goddard Space Flight Center,Science Systems and Applications, Inc, jun 2015.

L.-M. Ma and Z.-M. Tan. Improving the behavior of the cumulus parameterization for tropical cyclone prediction: Convection trigger. *Atmospheric Research*, 92(2):190–211, apr 2009. doi: 10.1016/j.atmosres.2008.09.022. URL http://dx.doi.org/10.1016/j.atmosres.2008.09.022.

C. McSweeney, M. New, and G. Lizcano. UNDP Climate Change Country Profiles, Ghana. feb 2012. URL http://ncsp.undp.org/sites/default/files/Ghana.oxford.report.pdf.

J. F. Meirink. MSG Cloud Physical Properties (CPP) product decription, 2012. URL http://msgcpp.knmi.nl/mediawiki/index.php/MSGCPP_product_description. Accessed: 2016-05, Last modified: 2012-11-26.

C. Silver. Accra floods 2015 pictures: Ghana death toll climbs as capital is inundated with water. jun 2015. URL http://www.ibtimes.com/accra-floods-2015-pictures-ghana-death-toll-climbs-capital-inundated-water-1953224.

Wei Wang et al. Arw version 3 modeling system user's guide. Technical report, National Center for Atmospheric Research, jan 2016.

Y. Zheng, K. Alapaty, J. A. Herwehe, A. D. D. Genio, and D. Niyogi. Improving high-resolution weather forecasts using the weather research and forecasting (WRF) model with an updated kain–fritsch scheme. *Mon. Wea. Rev.*, 144(3):833–860, mar 2016. doi: 10.1175/mwr-d-15-0005.1. URL http://dx.doi.org/10.1175/MWR-D-15-0005.1.

Acknowledgements

A special thanks goes to Wilfred Janssen and Noud Brasjen at Infoplaza for helping with understanding how WRF works and figuring out ways to verify the model output. Also a thank you to Jonas Schuitemaker and Yoni Stap who could spare some of their time to help spell check and fix the grammar of this report.