Introduction

This document serves as a description of the underlying scientific and technical background behind the iska™ forecast product. As such, it explains what the product is, how it is created and how it should be used.

Short Product Description

The iska weather forecast text messages are delivered daily in the morning and depict the predicted weather for the next 24 hours (from today 06 am until tomorrow 06 am) and the following 24 hours (tomorrow 06 am to 06 am the next day). The messages contain a written forecast for each 24-hour period. For simplicity of use, these are categorical and describe three parameters: likelihood of rain, timing of rainfall and its intensity. This approach means that messages are constructed in such a way that also illiterate farmers are able to extract the useful information even after very little training. All forecasts are specific to the subscriber's location by an automated application that fetches the most common coordinate for each subscriber. Particulars of the product are outlined in greater detail later in Appendix 1.

Scientific background to Ignitia's Numerical Weather Forecast Model

Ignitia has spent 15 man-years on research and development of its high-resolution tropical forecast model regionalized to the West African region. This effort has been necessary, as the global forecast models have inherent difficulties in tropical weather prediction in general (e.g., Laing & Evans, 2011), and in the Sahel and West African region in particular, due to a number of reasons as explained below.

The Problem of the Tropics

Large forecast producing actors (such as NCEP, ECMWF and UKMO) have for a long time provided high quality forecasts to their general public in the US and in Europe, which both are located at the mid-latitudes. Naturally, since none of the actors has as its core mission to make tropical weather forecasts, this has been of secondary rather than primary interest to develop. The underlying dynamics causing most of the day-to-day weather at these mid-latitudes is represented by quasi-geostrophic turbulence theory (Charney, 1971). This theory (and the relevant equations, i.e. the primitive equations set, that are numerically solved) accurately predicts the large scale evolution of the weather, e.g. in predicting low- and high pressure systems, frontal systems, airmass advection and associated parameters such as temperature, humidity and frontal rain distributions. Even as far out as at 5-7 days, forecasts are of generally good quality thanks to the predictability of the synoptic and planetary scales. And in fact, as much of the weather experienced has its origin in these large scales, forecasts are of satisfactory quality 5 days out most of the time.

The problem arises when there is a public perception that the global models are equally skillful in predicting weather in the Tropics. Many weather forecast providers' even present forecasts as far out as 14 days or more through web and mobile applications with apparent detail. However, such provision is not backed up and cannot be defended from a scientific aspect. The reason is that most of the day-to-day weather in the Tropics has its origin in the small-scales where convection is a driving force. The spatial and temporal distribution of convection is inherently difficult to predict due to its stochastic nature, while the envelope of convection (larger scale convective patches) can be predicted fairly well to some degree, much like the synoptic scales at the mid-latitudes. In the Tropics, the quasi-geostrophic approximation collapses at scales smaller than planetary due to the proximity to the equator such that ageostrophic effects become of greater importance, while driving mechanisms at the mid-latitudes such as temperature and pressure differences are not present or of orders of magnitude smaller. Therefore, there is no a priori reason why the global models should perform well in near equator areas. Rather, without any compensating algorithms to predict tropical weather with detail, persistence and statistics-based forecasts would likely perform as good or even better.

Moreover, much of the initialization of a forecast model run, i.e. collection and processing of measurement data, are heavily relying on ground based weather observations from synoptic weather stations. Such infrastructure is severely lacking throughout the Tropics, and given the small-scale nature of many weather events, important data is never retrieved such that large errors are present already at the start of a numerical forecast model run which will quickly degrade the forecast quality. In recent years, satellite retrievals are becoming a much more viable choice for initialization, also thanks to the relatively higher resolution of geostationary satellite data sets in near-equator areas.

Important aspects in modeling tropical weather of West Africa

In order to accurately forecast weather in West Africa, where rainfall is the parameter of particular interest, the most important features in the region need to be represented well in order to hope for any better forecast quality to be achieved. These include, but are not limited to, accurate spatial and temporal representations of (see for example Lafore et al, 2011 and Janicot et al, 2010):

- The seabreeze front (location, intensity and timing)
- Topographic/orographic effects on convection
- Position of low-level convergence areas
- Position and 3D structure of the ITCZ, when it is governing the weather in the domain
- The African Easterly Jet (location, intensity and timing)
- African Easterly Waves (location, intensity and timing. These make up for approximately half the annual precipitation in the region).
- The Tropical Easterly Jet (position and strength)
- The Subtropical Jet (position and strength)
- Presence of dry mid-level air (horizontal and vertical distribution, affects thunderstorm characteristics)
- Correct correlations with equatorial waves and oscillations such as the Madden-Julian Oscillation
- Position of strong low-level humidity gradients (between dry Saharan air and moist maritime air)
- Distribution and intensity of the Harmattan (dry northeasterly trade winds)
- Correct representation of local climatological anomalies such as the Dahomey gap in southeastern Ghana
- Daily cycle effects on position, intensity and timing of convection
- Diurnal and nocturnal land-sea interactions

Above conditions implies that a regional model is important, as it seeks to optimize the predictability of regional phenomena in shape, intensity and timing of such. A complicating factor is that simulating the large scales alone is not enough to produce accurate weather forecasts, as for example the African Easterly Waves have a direct and indirect coupling with convection (Berry & Thorncroft, 2012).

Given that the major portion of the rainfall in the region falls in association with these (Lavaysee et al 2006) and the rest mostly from other types of convective systems, convection needs to be captured from a model perspective. The inherent problem is that much of the research efforts have been devoted to mostly the large scales (meso-, synoptic and planetary scales and mostly at the mid-latitudes) and the microscales (up to a km or so to capture for example microclimatological terrain effects) while the particulars of the convective scales of 1-9 km are much less known. This spatial regime is commonly referred to as the gray scales (e.g. Grell & Freitas, 2014), or can equally well be referred to as physics of no mans land.

How Ignitia has improved weather predictions in West Africa

In order to improve the forecast capabilities, Ignitia has invested over 15 man years on a hybrid approach to forecast the weather in West Africa. At the core, a special high-resolution (3-9 km) edition of the research version of the WRF model was developed. Development within the model has focused on improving the boundary layer representation, the cumulus parameterization and microphysics schemes.

In addition, Ignitia has developed post-processing algorithms for automated probabilistic forecast generation. Due to the relatively stochastic nature of convective origination (modulated by climatology from geographic and diurnal effects that gives a statistical signal), Ignitia has also developed a forecast editing tool and an in-depth manual that are used by a team of specially trained tropical meteorologists with years of experience of West African weather. This approach allows the meteorologists to take into account any departures from the initialization of the numerical forecast model run so to further the accuracy of the forecasts that are issued.

Short technical description of Ignitia's forecasting system

Ignitia's high-resolution tropical forecast model regionalized to the West African region runs on its own supercomputer cluster underground in Stockholm, Sweden. It is a top-end multicore cluster and the code has been parallelized for optimal performance, taking full advantage of the whole system's computing capabilities, leveraged by the use of an Infiniband switch.

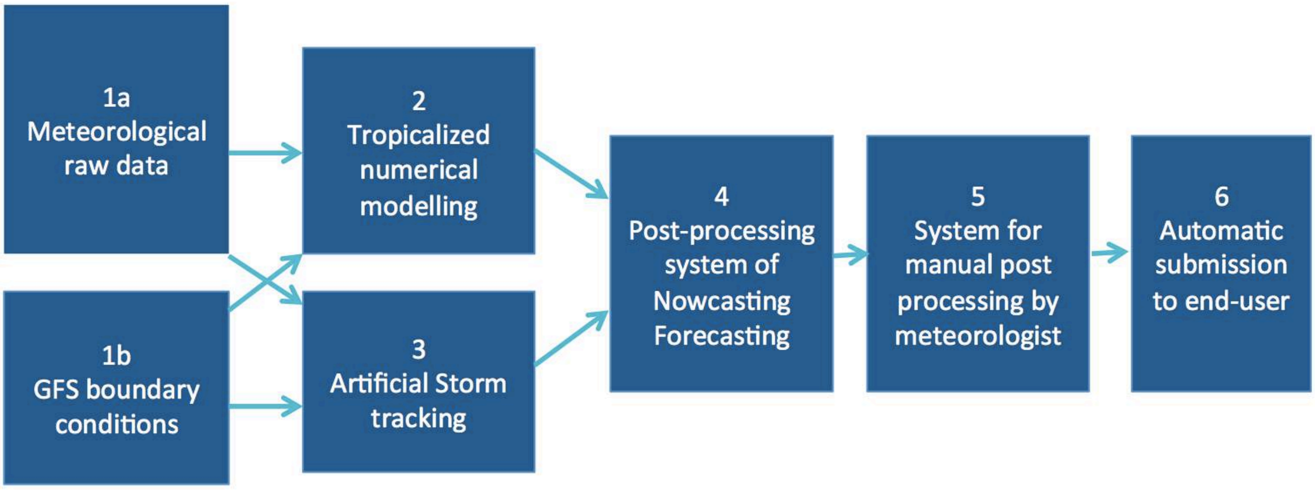

Figure 1. Flow chart of forecasting process from data collection to end-user forecast submission.

(1) Model Initialization

Our forecast model is initialized up to 4 times per day, mostly based on boundary conditions provided by the most recent version of the NCEP GFS model, which captures the synoptic conditions in all essential fairly well. The low resolution of the GFS model also imposes some challenges, but without it having a proper handling of the gray scales and regional/equatorial dynamics anyway, it does not interfere with our own model's downscaling and higher resolution and differential parameterization schemes for boundary layer, cumulus and microphysics among others.

(2) Tropical regional model

The numerical atmospheric model is based on the WRF research version but has been further developed for tropical conditions. It is setup to run with three domains with increasing resolution down to 3 km for areas of particular importance, and covers the whole of tropical Africa, from the Arabic peninsula and the Indian Ocean in the East to the Atlantic and Cape Verde islands in the West. The parameterization schemes have been developed and combined after careful investigations from both a long-term and a case-study perspective on their ability to reproduce the climatology and accurate life cycle of West African rain systems. Front-end research and findings from acknowledged institutions worldwide have been investigated and implemented to make the best possible forecast a reality subject to the general constraints imposed by finite computing power, raw data access and accuracy, and the limitations in the physical understanding of certain processes within the scientific community.

(3) Squall-tracking algorithm

Currently, Ignitia is building an artificial intelligence system that predicts the trajectories of individual thunderstorms in a probabilistic sense that merges the numerical forecast system with observational data from lightning detection, satellite imagery and to be further improved by the utilization of a C-band and X-band radar that covers both maritime and continental areas. This system is self-learning and will be integrated to seamlessly adjust the model results by ensemble selection and adjustments in certain time windows, but is currently running independently aiding the operational forecaster by making use of 15 years of historical thunderstorm tracks in the region.

(4) Nowcasting & forecasting system

In the post-processing of the deterministic forecast model run, we make use of ensemble methods, as well as bringing in the most recent satellite data, balloon soundings, lightning ground strike detection and other monitoring data sets to better capture the atmospheric state at the time of forecast issuance. We use a hybrid geographical and thermo-dynamic probability filter on predicted rainfall to determine areas at risk of rainfall, so to sharpen the probabilistic determination of our rainfall forecasts with regard to spatial, temporal and intensity characteristics.

(5) Manual post-processing and forecast editing

The operational forecaster is presented with a numerically forecasted field on rainfall probability for the West African region. Through the use of our developed GIS application, that handles forecast editing, the forecaster can freely adjust this field given departures from the predicted/suggested field prior to issuing of the weather forecasts. The forecasters have been specifically trained to forecast weather in West Africa since 2010 and have gained considerable experience that further improves the forecast quality in the region.. The forecaster is well aware of common biases and where geographical features tend to take over, why we have gained experience and created tools to be applied manually.

(6) Issuance and submission of forecasts by SMS

Upon completion of forecast editing, the forecast is approved and forecast text messages are created for each subscriber depending on his/her GPS-location. The coordinates are fetched by utilizing our application identifying the most common location of the subscriber during daytime (for a farmer this is where his field is) and are matched to the forecast category at the closest model grid points. Once the text messages have been created, our application binds to the different telecom operators through which we have agreements to submit the SMS forecasts to each subscriber's mobile phone. This is accomplished by using different messaging protocols such as SMPP v. 3.4 and http, for example.

Validation of forecast accuracy

To quantify the accuracy and the model's ability to predict rainfall in the region, a number of performance evaluations have been performed over the years 2013-2014. One set of tests has focused on the full West African region whereas the second set has been focusing on Ghana, where the forecast subscribers have been located during this specific period. Every forecast ever issued is part of Ignitia's database for individual subscriber forecast performance, which is used to identify geographical biases and overall performance.

Due to the lack of ground station observations, complicated also by the mostly local nature of convective rainfall that quickly degrades the validity of a ground observation away from the station, we are using other means to validate the forecasts. Three separate datasets are used to measure the spatial distribution of rainfall and the associated amounts. These are (1) EUMETSAT MPE 24h accumulations, (2) NOAA RFE 2.0 and (3) Vaisala GLD360 ground lightning strike detection data. Some details of these datasets are listed below:

(1) Evaluation of EUMETSAT Multi-sensor Precipitation Estimate (MPE)

This dataset has been operational since many years back and provides instantaneous rain rates based on a combination of brightness temperature from geostationary METEOSAT IR-channels and rain rates derived from the SSM/I-instrument on polar orbiting METOP satellites (Heinemann, 2003). The spatial and temporal resolution is 3 km and 15 minutes, respectively. The product catches rainfall events relatively well, but is subject to large variability in quality over 24h periods due to the presence or absence of recent SSM/I-data from the polar orbiting satellites. In particular, it has a tendency of exaggerating the spatial distribution of rainfall due to anvils from Cumulonimbus being erroneously seen as rain bearing, particularly in cases with strong winds at upper levels, when the anvils can be extended 100's of kilometers away from the actual rain in the cloud system. Also, this dataset has large difficulties in capturing warm cloud rainfall during the monsoonal months of June through to September in the coastal areas.

(2) Evaluation of NOAA Climate Prediction Center African Rainfall Estimation Algorithm (RFE) Version 2.0

The generation of this dataset is in many ways similar to how (1) is generated, i.e. by using microwave sensor (SSM/I and AMSU) as well as geostationary IR-channel data (NOAA, 2012). However, in addition, as it is not a real-time product, the data is processed after the 24-hour collection period and calibrated against any existing ground station measurements of rainfall in the region. This means that nearby to a ground station, the data is very representative of the real rainfall, while away from it, the dataset has large similarities with (1) and shares many of the same deficiencies. However, while its horizontal resolution is coarse (10 km) it has become a de-facto standard dataset for rainfall studies in the region, used by many different actors.

(3) Evaluation of Vaisala Global Lightning Data set 360 (GLD360)

This dataset is a real-time product that detects cloud-to-ground lightning strikes with a uniform global coverage. The location accuracy is around 5 km and more than 70% of all ground strikes are captured. In the West African continental region, any significant rainfall is with basically only one exception associated with lightning strikes, with the highest density of ground lightning strikes being at the leading edge of a propagating storm system, i.e. usually at the western edge. The exception is in coastal areas where maritime warm cloud rain showers (i.e., no icing, only coalescence processes) can give significant amounts during the West African monsoon without any thunder associated with it. Also, any stratiform rainfall (mostly weak intensity with no ground lightning strikes) is almost always occurring in the trailing cloud system from the passage of an active convective cell with lightning strikes.

To create a baseline dataset against which the forecasts are validated, the spatial distribution of rainfall is determined by a combination of (1) and (3) and/or (2) and (3), where unlikely rainfall due to anvil spreading and wind-shear is filtered out if far enough away from any lightning strikes. However, all three datasets are also used separately to estimate their differences for later optimization.

To build the statistics of our forecast performance, each subscriber's forecast history is stored. The statistics are then processed in two ways: one that takes all subscriber locations into account and one that picks one subscriber location in a box that is 3 x 3 km in size. The second is performed to filter out the differential subscriber density effect on forecast performance as such (high density in major towns/cities compared to rural areas, meaning that the statistics would be biased towards cities).
Now, as the forecasts are categorical, i.e.,

1. Dry (<5% chance of rain. Should be close to 0%),
2. Likely dry (5-49% chance of rain. Should be close to 20%),
3. Rain likely (50-75% chance of rain. Should be close to 60%),
4. High chance of rain (75-100% chance of rain. Should be close to 100%),

the forecast accuracy can be evaluated by several different methods, of which the most important ones are listed below, while bearing in mind that these methods apply to categorical forecasts and are not continuous in their nature. For more information regarding these methods for categorical forecasts, see for example Ebert (2008) or WWRP/WGNE (2009).

For each category, we can create a contingency table as below:

Table 3		***Observed rainfall***	
		Yes	**No**
Forecasted rainfall	**Yes**	Hits (a)	False alarms (b)
	No	Misses (c)	Correct Rejections (d)

Note that table 3 is most appropriate for binary categories such as (1) and (4) above, where the aimed ratios are 0% and 100%, respectively. However, it can still be used also for other categories, bearing in mind that the expected percentages is different (20% and 60% rainfall, respectively, with permitted penalty ranges).
From this contingency table, we can for example construct the following common measures:

Accuracy (A) = (a+d)/(a+b+c+d)
A measure of the ratio of correct forecasts compared to all forecasts that have been sent out. Used as-is for category 1 and 4, but differently interpreted for category 2 and 3 as these are not binary in terms of correct/incorrect. Should be 1 if all forecasts were perfect.

Probability of detection (POD) or Hit Rate (HR) = a/(a+c)
A measure of the ratio of rainfall forecasts that were correct. Should be 1 if all forecasts were perfect.

False Alarm Rate (FAR) = b/(a+b)
The ratio of forecasts that said rain but ended up being incorrect, i.e. false alarms. Should be 0 if all forecasts were perfect.

Critical Success Index (CSI) = a/(a+b+c)
A similar measure to the hit rate but that also takes into account the false alarms. Should be 1 if all forecasts were perfect.

Bias = (a+b)/(a+c)
Describes the general ability of our forecasts to form a climatology that should be as close as possible to the observed number of rainfall events. If below 1, it means we are underestimating rainfall and if above, overestimating the same. Should be 1 if all forecasts were in perfect match with climatology.

Probabilistic Hit Rate (PHR) = 1-dryscore-lowscore-likscore-higscore, where

dryscore = abs(1-a/(a+c))*(a+c)/N

where N is the total number of forecasts issued and dryscore represents the category "dry".

```
if c/(a+c)>0.2
    low_dry=c/(a+c)-0.2
else

   low_dry = 0
lowscore=low_dry*(a+c)/N
```

Here low_dry represents the ratio of "likely dry" forecasts that were followed by rain to the total number of such forecasts. This score is 0 if less than 20% of all likely dry forecasts ended up with rain in reality.

```
if a/(a+c)<0.5
   lik_dry=0.5-a/(a+c)
else
   lik_dry = 0
likscore=lik_dry*(a+c)/N
```

Similar to previous category above, if less than than 50% of the rain likely forecasts ended up with rain, the penalty sets in.

```
if a/(a+c)<0.75
   hig_dry=0.75-a/(a+c)
```

else

hig_dry=0

higscore=hig_dry*(a+c)/N

If less than 75% of all "high chance of rain" forecasts ended up with rain, there will be a penalty in the final score.

PHR = 1-dryscore-lowscore-likscore-higscore

The PHR is an integral estimate that calculates, for each category, the hit rate as a combined measure. If any of the categories are outside the targeted ranges, they will be weighted by their number density compared to the total number of forecasts. PHR should be equal to one to represent the best forecast skill.

Heidke Skill Score (HSS) = 2*(a*d-b*c)/((a+c)*(c+d)+(a+b)*(b+d))

A common estimate that also takes into account the influence of chance: pure guessing can also give a correct forecast at times, and should be filtered to represent true skill. Perfect forecasts would yield a value of +1 with this score, whereas 0 means that just chance would give the same result.

During 2014, N=674 028 forecasts were issued and submitted in Ghana and the scores were as follows:

Table 4: Score table 2014 Type of score	Forecast day 1 2014	Forecast day 2 2014	All forecasts 2014
Accuracy	0.91	0.91	-
Hit rate	0.97	0.96	-
False Alarm Rate	0.17	0.26	-
Critical Success Index	0.81	0.71	-
Bias	1.18	1.30	-
Probabilistic Hit Rate	0.99	0.98	-
Heidke Skill Score	0.81	0.78	-
Percent rain: dry	2%	1%	2%
Percent rain: likely dry	21%	24%	23%
Percent rain: rain likely	54%	48%	51%
Percent rain: high chance	83%	74%	80%

The percentages gives an overall view of the performance, but needs to be complemented with statistics on the frequency distribution on how many subscribers have received forecasts that are within the percentual ranges we aim for, for each category. For those outside these ranges, it means there is a systematic bias that needs to be analysed and corrected for in our tuning of the numerical forecast model. For this purpose, we produce decadal bins, and the higher the peak at the aimed percentage (and the shorter the tails) the better the forecasts are as seen from an end-user perspective. As the different seasons are governed by different atmospheric dynamic setups (pre-monsoon, monsoon and Harmattan conditions), the biases can be both spatial and temporal and these scores are therefore continuously monitored to detect any anomalies that may arise under certain atmospheric conditions.

In the below figures 2 and 3 (day 1 and day 2, respectively), it can be seen that for category 1 (dry), most subscribers are in the 0-10% bin, with a few other being in other bins. It should be born in mind though that some subscribers signed up late so that their statistics is not long enough to represent any statistical significance. One erroneous forecast would then degrade the apparent quality for that one subscriber.

For category 2 (previously low, now likely dry), the peak is found in the 30% bin, closely followed by the 20% bin. This is a little bit higher than we would like, but can partly be ascribed to the tendency for satellite rainfall estimates to exaggerate spatial distributions, particularly in cases with strong upper level wind shear.

Category 3, rain likely, is centered around the 60% bin with a quite even distribution between 50 and 70%, while a few percent of the subscribers are in the 30 and 40% bin. The latter is partly due to small-scale climatological features around Tamale and along the sea breeze front some 50-100km inland from the coast.

Category 4 is in all essential the reverse of category 1, with most subscribers receiving rainfall when they get the "high chance of rain" forecast.

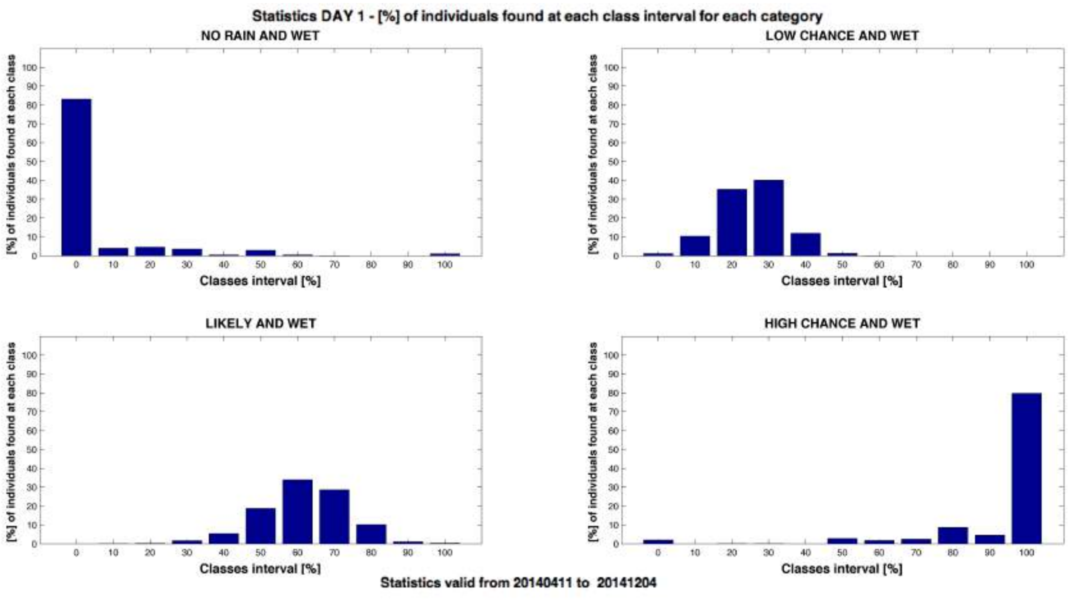

Figure 2. Frequency distribution of forecast categories for day 1. Top left: forecast category 1 "dry", top right: category 2 (likely dry), bottom left: rain likely, bottom right: high chance of rain. The numbers are in percent and represent the number of subscribers receiving a certain forecast category followed by observation of rainfall.

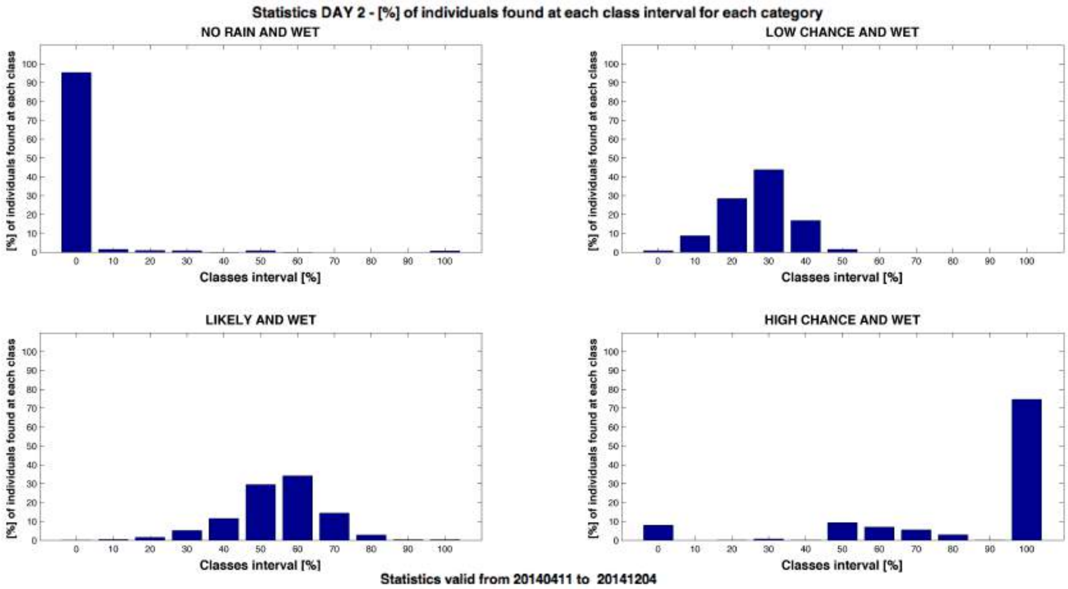

Figure 3. Frequency distribution of forecast categories for day 1. Top left: forecast category 1 "dry", top right: category 2 (likely dry), bottom left: rain likely, bottom right: high chance of rain. The numbers are in percent and represent the number of subscribers receiving a certain forecast category followed by observation of rainfall.

Evaluations from iska users

A number of organisations have used iska over the past three years in Ghana. We here provide some insight into the outcomes of their internal evaluations of the usefulness of the service Ignitia has provided through iska.

One of our clients which is a farmer association in mostly maize conducted a study on 100 randomly located farmers after the 2013 season, and the key findings were as follows (Masara N'Arziki, 2014):

- 100% of the users understood the SMS forecasts they were receiving.
- 92% of the farmers found the service useful in their farming practices.
- 93% of the farmers used iska to take decisions on application of inputs.
- 83% of the farmers were willing to pay for the service themselves would it not be subsidised by the farmer association. The indicated median amount was 5.3 Ghana cedis. (Currently, iska charges about 11 Ghana cedis per year, but can be lowered with increasing subscriber volumes).

The conclusion was that the farmers understand the positive effects of weather forecasts when linking it to both activity planning and optimisation of their input effectiveness.

Another survey (Majeed, 2014) was conducted on 600 farmers by the German agency Gesellschaft fur Internationale Zusammenarbeit (GIZ) for the year 2014, and their key findings were much in line with those from the 2013 study previously presented. In addition, the following results were indicated for the Northern Region farmers:

- 90% of the farmers in the Northern Region said that the forecasts aided them in taking the right decision regarding sowing and planting.
- 90% also indicated that iska aided in taking the right decision regarding application of pesticides.
- 88% found that iska aided in taking the right decision regarding fertilizer input.
- 73% noted that the forecasts helped in taking the right decisions on harvesting.
- 95% of the farmers were willing to pay if GIZ would not pay for them using the service. 63% of these were willing to pay more than 10 GHc per year.
- The most common challenge experienced in using iska was due to network or mobile battery problems.

In addition to above, a lot of anecdotal evidence of the usefulness of iska was reported, showing the importance of having access to accurate weather information and reassuring Ignitia that the service we provide makes a difference to such a degree that a majority of the smallholder farmers are willing to pay for the service. That alone is the best grade we can get from our end-users. If it was not truly useful, no one would pay for it, given the small margins of a smallholder farmer.

Evaluation from third parties is an on-going process, and we get more evidence of the usefulness as more organisations subscribe. During 2015, Yale University is conducting a bi-weekly study, measuring small scale farmers' actions, spread, yield and use iska in 108 villages, with a control group of the same size.

References

Berry, G. J., Thorncroft, C. D., 2012 African Easterly Wave Dynamics in Mesoscale Numerical Model: The Upscale Role of Convection, J. Atmos. Sci., **69-4**, 1267pp

Charney, J. G., 1971 Geostrophic turbulence, J. Atmos. Sci. **28**, 1087pp.

Ebert, E. E., 2008, Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. Meteorological Applications, Royal Meteorological Society, **15**, 51pp.

Grell, G., Freitas, S.R., 2014 A scale and aerosol aware stochastic convective parameterization for weather and air quality modelling, Atmos. Chem. Phys., **14**, 5233pp

Heinemann, T., 2003, MPE validation status report, Technical Report, EUMETSAT.

Janicot, S., Lafore J. P., Thorncroft C., 2010 The West African Monsoon, World Scientific Publishing Company, 608pp.

Lafore J. P., Chapelet P., Mumba Z., Chapelon N., Dufresne M., Agbabu R., Abdoul-Aziz A., Hamidou H., Asencio N., Couvreux F., Nuret M., Garba A., 2011, Forecaster's guide for West African Synthetic Analysis/Forecast, WASA/F, Meteo-France, CNRM, ACMAD, ASECNA, EAMAC

Laing A., Evans J-L, 2011, Introduction to Tropical Meteorology, University Corporation for Atmospheric Research.

Lavaysee, C., Diedhiou A., Laurent H., Lebel T., 2006 African Easterly Waves and convective activity in wet and dry sequences of the West African Monsoon, Climate Dynamics, **27**, 319pp

Majeed, A-R, 2014, A field report on the monitoring and evaluation of AAESC/Ignitia weather forecast by mobile phone, GIZ.

Masara N'Arziki, 2014, Farmer survey on weather forecasting pilot.

WWRP/WGNE Joint Working Group on Verification, 2009, Sponsored by World Meteorological Organization (WMO), http://www.cawcr.gov.au/projects/verification/verif_web_page.html#BS

Appendix 1

Description of the SMS forecast product iska

Each text message comes with a written weather forecast, which includes all the relevant information on rainfall prediction. There is one sentence valid for today and one valid for tomorrow. It is made up by three components:

(a) Likelihood of any rainfall
(b) Timing of rainfall
(c) Intensity indicator

(a) Probability categories: Likelihood of any rainfall

Table 1: Likelihood categories			
1	2	3	4
Dry	Likely dry	Rain likely	High chance of rain

The possible categories are listed above in Table 1. The forecast categories depend on the likelihood of rainfall at each farmer location. The use of these under different weather conditions is briefly explained in the following. In the Tropics, the weather can differ substantially within a short distance and in time. The likely location and timing of isolated thunderstorms and rain showers depend on several factors. The meteorologist will therefore make a qualified assessment of the rain probability based on both forecast ensembles, statistics and experience. To picture how probable it is for a specific location to receive rainfall, we are using the following categories:

(1) Dry

This category represents fair weather, meaning that there is no rain or only a very low chance of rain predicted. The weather can be sunny or cloudy, but the chance of any rainfall is insignificant or low, so that in most cases when this forecast message is sent, it will be dry. We allow for rainfall to occur under this category no more than one out of 20 times.

(2) Likely dry

At times, weather conditions are mainly in favour of dry conditions but with a significant probability of rainfall to occur (isolated storms not predicted to last long or develop into sustained systems). These conditions are common under dry mid-level air situations, shallow monsoonal boundary layer in the Sahel, along sea breeze fronts and isolated storms resulting from topographically enhanced convection. Previously, this category contained the sentence *low chance of rain*, but the word rain draws unnecessary and unproportional attention to rain given the low probability, as in most cases, conditions will be dry. Therefore the category likely dry is used in cases when the probability is less than 50% but above 5%, centered at around 20% in a frequency diagram.

(3) Rain likely

When there is a significant chance of rain, meaning that it is more likely than not that rain will fall (50-75%, centered at 60% in a frequency diagram) while there is still some uncertainty involved, this category is used. This category is common when occasional storms are likely to develop over a wider area, such as during moderately deep monsoonal boundary layer in the Sahel, near an upper level jet entrance area in the Sahel, or when a squall line is approaching coastal areas and the exact timing is uncertain such that the sea breeze circulation is likely to interfere and cause storm decay. This phenomenon is common while rainfall distribution in coastal areas is highly dependent on the exact timing of the system's arrival. The likely category is often used when moderately intense events are predicted to arrive at the coast during unfavourable hours of the day for rainfall to occur.

(4) High chance of rain

When there is a significant chance of rain, meaning that it is much more likely than not that rain will fall, with a relatively large certainty (at least 75%), this category is used. Commonly used when there is a convergence line at low levels along with deep moisture fluxes, an easterly wave is propagating westward towards Ghana or when there is widespread unstable conditions in the troposphere as seen from atmospheric soundings and model-generated instability parameters.

(b) Time specifier for start of rainfall

Table 2: Timing specifiers

1	2	3	4
No specification	Morning time 06 – 12 am	Daytime 12 am – 06 pm	Nighttime 06 pm – 06 am

Our forecasts represent the predicted weather over a 24-hour period, and aim at providing the chance of any rain during this period. Therefore, a forecast indicating rain might be considered as incorrect to a farmer who is just concluding his/her work at the field when the daylight is fading and the day so far has been perfectly dry. However, at 06 pm (evening), only half of the forecast period has passed, meaning that any rain predicted after dark will still govern the forecast text message, even if we know that the whole day will be sunny and that rain is expected only late at night. Consequently, we are using an indicator to tell the farmer when we expect the chance of rain to be the highest by adding a time specification according to table 2 above.

If there is no particular time period that has an enhanced probability of rain, no specification will be used. This is also the case if there are multiple increased rain risk periods during a 24h period. The message will then simply read as indicated in (a) – (g) for "Today, ..." or "Tomorrow, ...", depending on which day the forecast is for. Specifying morning, day and nighttime will help the farmer to plan better his activities, knowing that the rain is expected to occur after work and/or tomorrow morning for example, while the weather could be dry or with less chance of rain otherwise.

Now, with the time specification included, the message could read (the high chance category (d) is given as an example):

i) Today, high chance of rain, morning time.
ii) Today, high chance of rain, daytime.
iii) Today, high chance of rain, nighttime.

Please note that even though this time specification is included, the chance to get rain during other parts of the day is not per definition zero. It simply tells the farmer that the chance of rain is high, and that any rain is expected to occur during the specified time period, while rainfall is less likely but not necessarily of zero probability at other times. As mentioned earlier, the text message is also giving the forecast for tomorrow with the same categories as above.

Intensity specifier

When there is a significant probability of rain, as described above, the descriptor "heavy rain" can be used, if, at the same time, at least one of the following three requirements is met:

i) Large rainfall amounts in a short period of time (more than 40 mm/hour) or persistent rainfall of moderate intensity over an extensive period of time leading to potentially large accumulations.

ii) Potentially destructive winds associated with the storm (more than 15 meters per second). These are predicted based on the presence of moderately thick dry mid-level air and/or an African Easterly Jet core in the vicinity. In the former case, strong localized downdrafts are common, and in the latter, downward flux of horizontal momentum from the jetstream is likely in association with any storms, particularly in connection to African Easterly Waves.

iii) Intense lightning/thunder, as modulated by the predicted instability indices, updraft magnitudes and moderately thick layers of dry air at mid-levels. Lightning detection is also used for existing propagating systems.