# Mapping of potential fishing zones in support of fisheries management in West Africa

**Kwame Adu Agyekum[1,2]\*, George Wiafe[1,2] and Francis Kofi Ewusie Nunoo[2]**

[1]Coastal & Marine Resources Management Centre, University of Ghana, Legon.
[2]Department of Marine & Fisheries Sciences, University of Ghana, Legon.
*Corresponding author: kaagyekum@gmail.com

## ABSTRACT

The distribution pattern of three tuna species *viz.* Skipjack (*Katsuwonus pelamis*), Yellowfin (*Thunnus albacares*) and Bigeye (*Thunnus obesus*) and key environmental factors were analyzed to develop potential fishing zone (PFZ) maps to support fisheries resource management in West Africa. The PFZ maps were developed from a Generalized Additive Model built from a binomial distribution with a probit link function. Inputs into the model were presence- absence data generated from tuna catch records from Ghana as well as sea surface temperature (SST), sea surface heights (SSH), geostrophic currents (UV) and salinity (SSS) datasets covering latitudes 10°S to 40°N and longitudes 35°W to 15°E in the period 2014-2015. Results from the analyses showed a non-linear response pattern with varying relationships among the environmental factors. The major fishing grounds identified from the outputs of the predictive model corresponded with regions where SST was between 21 and 28.5 °C, SSH above 0.05 m, geostrophic current speed above 0.25 m/s and SSS levels above 33.5 ppt. High tuna distribution was consistent with the major upwelling regions, suggesting that the dynamics of the upper ocean may account for the observed spatial pattern. Outputs from the model can provide a quick overview of the possible areas where tuna is likely to aggregate and the potential fishing grounds where monitoring and surveillance need to be intensified.

**Keywords:** Tuna; Generalized Additive Model; Potential fishing zone maps; fisheries management, Guinea Current

## Introduction

Tropical tunas occupy the warm upper oceanic waters, with temperatures between 20-30°C (Holland *et al.,* 1990). In the eastern equatorial Atlantic, the optimal oceanographic conditions support high abundance of tuna. There are many fishing companies operating in the high seas targeting mainly Skipjack (*Katsuwonus pelamis*), Yellowfin (*Thunnus albacares*) and Bigeye (*Thunnus obesus*) tuna in the region.

The tuna industry of West Africa makes a considerable contribution to food security and provides economic benefits. The region serves as an important spawning ground and migratory path (Fromentin and Powers, 2005), and catch information from surface gears indicate that the eastern equatorial Atlantic is a major nursery ground for tuna (Mensah and Quaatey, 2002). Trends in catch over many decades in the major tuna fishing countries in West Africa show a steady increase, from about 1,000 tonnes in the 1960s to over 80,000 tonnes in 2000 (Miyabe and Nakano, 2004). The increased catch is not only a result of the high occurrence of tuna species but can also be attributed to increased fishing pressure from the tuna fisheries industry. This can result in the collapse of the tuna fishery, hence the need to adopt management measures for sustainable exploitation of the resource in line with current trends in global climate change.

Current increases in sea surface temperature associated with global climate change and the increasing fishing pressure are expected to affect fish stocks (IPCC, 2007).

Generally, warming climate will increase stratification of the upper ocean, which will reduce nutrient regeneration for increased plankton growth. This may affect migration, feeding behaviour and spawning, and in the long term, the existence of marine living organisms. Even more threatening is the irreversible impact of overfishing of many ecosystems which affects not only the balance of life in the oceans, but also the social and economic well-being of coastal communities. Adopting conservation methods for effective management of the tuna fisheries requires an in-depth understanding of the impact of changes in the habitat of fish (Holland *et al.,* 1990). Such changes have been observed to affect foraging behaviour and migratory patterns of pelagic fishes (Gutenkunst *et al.* 2007).

The eastern tropical Atlantic is the spawning ground for these tunas (Mensah and Quaatey, 2002). They are known to undertake large-scale migration over long distances, often staying outside the Exclusive Economic Zones of countries in West Africa (Blackburn and Williams, 1975), which makes them vulnerable to illegal fishing practices. There is high incidence of illegal fishing activities in the Gulf of Guinea (Nunoo *et al.,* 2015). For instance, the unselective fishing of juveniles and small tunas by purse seiners through the use of fish aggregating device (FAD) and other sophisticated equipment like sonars, radio buoys, echo sounder, fish finder, etc. appear to make this fisheries unsustainable (Adinortey *et al.,* 2014). This, coupled with weak enforcement of fishing and maritime laws, makes marine living resources vulnerable and likely to be overfished. This raises concerns about management approaches that need to be adopted for sustainable harvesting of such commercially important fish species. Currently, many coastal states including developing countries have adopted electronic vessel monitoring systems (VMS) in their fisheries management strategies. These include the use of Automatic Identification System data to provide near-real time traffic, navigation and vessel information for fisheries application (Natale *et al.,* 2015).

This systems can be implemented at national to regional scale to effectively identify the major fishing grounds and monitor the activities of fishing vessels in those areas.

This, when implemented operationally, can provide near real-time and forecasts of fishing grounds to fisheries managers and other national agencies such as the coast guards and the navy.

## Methodology

Tuna catch data from 2014 to 2015 were obtained from logbooks of purse-seine tuna fleets, courtesy of the Fisheries Scientific Surveys Division of the Fisheries Commission of Ghana. The data comprised tuna catches (in tonnes) for three dominant species in the eastern equatorial Atlantic, the geographic coordinates where they were fished, and the dates of each catch. In our analysis, monthly catch-per-unit effort (CPUE) were calculated from the total weight of tuna caught and the number of fishing days in the month.

Four environmental parameters, i.e. sea surface temperatures (SST), sea surface heights (SSH), geostrophic currents (UV components) and salinity (SSS) datasets with a spatial resolution of 9 km over a spatial extent of 10°S to 40°N and 35°W to 15°E in the Atlantic Ocean were extracted from the global Mercator Ocean products for the same period. Monthly variations in the environmental parameters were analyzed to ascertain any potential impact on the distribution of tuna in the eastern equatorial Atlantic in a predictive model.

### *Construction of Generalized Additive Model (GAM)*

A spatial predicting tool for empirically estimating the distribution of fish was constructed from a generalized additive model. Predictors for the model were SST, SSH, geostrophic current and SSS. A presence/absence variable was constructed from the tuna data by assigning a value of 1 to every grid from which any of the three tuna species was caught and a value of 0 to every grid with no catch. The presence-absence tuna data followed a binomial distribution. Consequently, the GAM used a binomial distribution as the family associated with a probit link function. The GAM was implemented in R data analysis statistical package (version 0.97.551) as:

$$y \sim \alpha + s(\text{SST}) + s(\text{SSH}) + s(\text{Currents}) + s(\text{SSS}) + \varepsilon$$

where y is the response variable, α is a constant and ε is the error term; s is the spline smooth function of the predictors.

## Results

### Tuna abundance

The trend in tuna catch data for the study period shows monthly variation in abundance of tuna in the eastern tropical Atlantic (Figure 1). There is considerable reduction in abundance during January to March. A closed season is enforced each year between February and

March in the equatorial Atlantic on tuna vessels flagged to Ghana. The Fisheries Enforcement Unit (FEU) of the Monitoring, Control and Surveillance (MCS) Division of the Fisheries Commission has the national mandate to ensure sustainable fisheries management. The FEU is equipped with electronic monitoring systems that can track all fishing vessels participating in the vessel monitoring programme. The catch composition from the logbook records indicates that Skipjack was the dominant species, accounting for about 59% of the weight of tuna landings in Ghana (Figure 2). The two other common species, Bigeye and Yellowfin accounted for 15% and 27% of the landing, respectively.
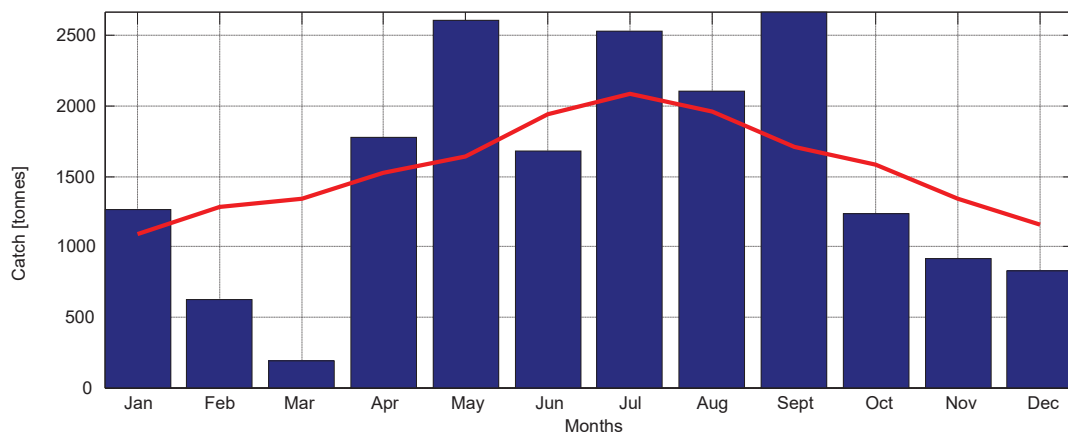


Fig. 1: Catch per unit effort (weight in tonnes/number of fishing days – blue bars) and average catch (tonnes – red line) for Skipjack, Yellowfin and Bigeye tuna in the eastern equatorial Atlantic. (The values are averages for each month for the two years, 2014 and 2015).
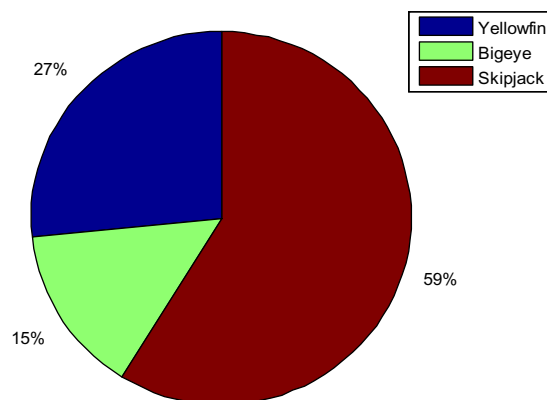


Fig. 2: Percentage composition of Skipjack, Yellowfin and Bigeye tuna in the eastern equatorial Atlantic from 2014 to 2015.

Fishing effort also varied at a trend similar to the observed seasonal trend in abundance during the study period. High catch per unit effort (CPUE) was recorded from May to September. During periods of reduced fishing intensity the percentage distribution of Skipjack increased slightly, which may suggest it was a response to reduced effort coupled with improved biological productivity within the region. Tuna abundance for all the three species peaked during September/October which is the end of the major upwelling season.

### Seasonal variability of oceanographic conditions

Monthly variability of catch and oceanographic conditions (SST, SSS, current velocity and SSH) between 2014 and 2015 closely followed a seasonal pattern (Figure 3a-d). High tuna catch was found in June to September where SST was coldest (Figure 3a), SSS was at its peak (Figure 3b), velocity of surface currents in the equatorial Atlantic was highest (Figure 3c) and SSH showed negative anomalies (Figure 3d). The peaks of catch occurred during the periods when the seasonal upwelling (upward flow of cold nutrient-rich bottom water) had intensified. SST during that period was at its minimum, and ranged between 25.5 and 26.5°C. Low catches were recorded during months when SST exceeded 27.5‰. In the Guinea Current region, salinity levels were lower in January but steadily increased in April, and remained dense from June till August, after which they began to dip (Figure 3b). Salinity ranged between 34 and 35°C in the Guinea Current region during the study period. The trend in surface current velocities showed that the Guinea Current experienced minimum current velocities from January to April and from October to November (Figure 3c). Maximum velocities were observed from May to August, a period which has a strong connection with increased wind speeds during the upwelling season.
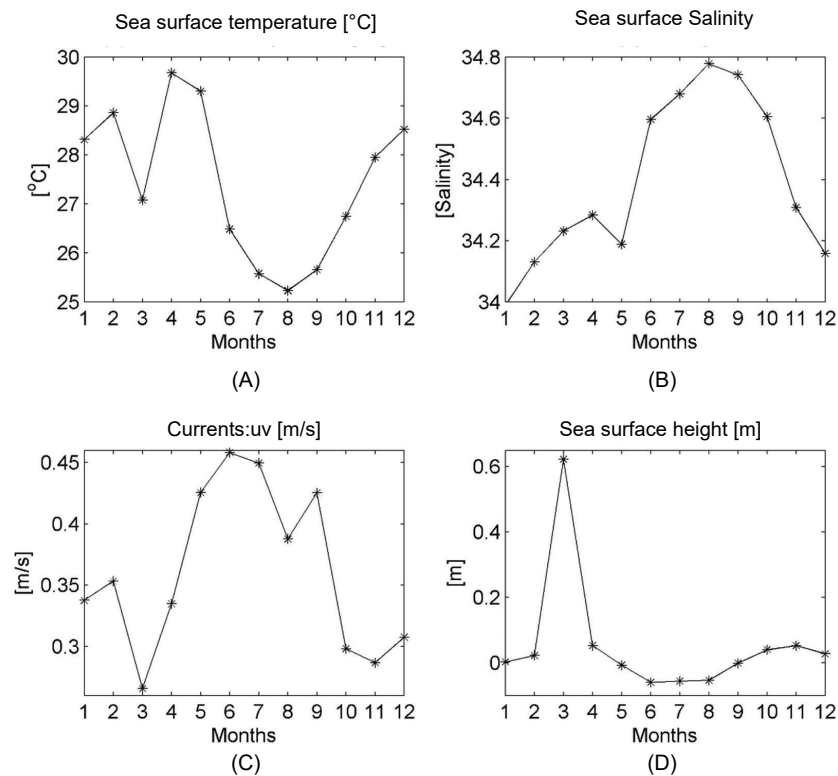


Fig. 3: Monthly variation in (A) sea surface temperature (B) sea surface salinity (C) currents and (D) sea surface height from tuna catch locations. (The values are averages for each month for the two years, 2014 and 2015).

## Preferred oceanographic conditions from tuna catch data

Generalized Additive Model (GAM) analyses using SST, SSS, surface current velocity and SSH as predictors yielded a significant reduction in the residual deviance using a χ2 test (Table 1). When the four predictors were used alone, SST had the lowest AIC. Subsequently, SSH, currents and SSS were added to the model, which further reduced the AIC and residual deviance. This indicates that the strong combined effect of these environmental parameters influenced tuna distribution in the eastern equatorial Atlantic. Among all the relationships shown in Table 1, predictor variables of Model 7 were significantly non-linear (χ2-tests: p < 2.2*10-16) with the least residual deviance value.

**Table 1: Analysis of deviance of three-variable GAM of presence-absence data of tuna catch.** (Changes in AIC and residual difference, degrees of freedom and associated significance from a Chi-square test)

| Model | AIC | Residual deviance | χ² p-value |
|---|---|---|---|
| 1. Occurrence ~ SST | 3404.6 | 3782.1 | <<0.00001 |
| 2. Occurrence ~ Current velocity | 3416.9 | 3774.8 | <<0.00001 |
| 3. Occurrence ~ SSH | 4062.9 | 4623.3 | <<0.0001 |
| 4. Occurrence ~ SSS | 4289.3 | 4933.9 | <<0.0001 |
| 5. Occurrence ~ SST + SSH | 3263.2 | 3615.9 | <<0.00001 |
| 6. Occurrence ~ SST + SSH + Currents | 2105.9 | 2014.9 | <<0.00001 |
| 7. Occurrence ~ SST + SSH + Currents + SSS | 2100.0 | 1981.3 | <<0.00001 |

Temperature had a predominant influence on tuna distribution. It had the lowest residual deviance when compared with SSH, currents and SSS. The loess plot revealed a non-linear response pattern with a strong positive relationship between temperature and tuna distribution from about 21°C to 28.5°C (Figure 4a-d). From the plot, temperatures between 22.2°C and 27.6°C provided the best thermal condition that influenced tuna aggregation.

Sea surface height data revealed that the tunas were mostly concentrated in regions with surface height anomalies above -0.03 m but not exceeding 0.05 m (Figure 4b). These SSH were observed in the shelf regions off Ghana, the equatorial region and northern parts of the Guinea Current. Ocean surface heights for these ranges were predominantly observed from July to September and associated with increased primary productivity.

The loess plot showed that surface current velocities associated with tuna catch distribution occurred in regions of the ocean with current speeds above 0.28 m/s (Figure 4c). The plot further indicates that there is almost a direct linear relationship with current speed and the abundance of tuna. Spatially, most regions of the Guinea Current often do not have magnitude of surface currents exceeding 0.4 m/s except during the upwelling season from July to September.

In the eastern equatorial Atlantic, salinity levels recorded from the tuna catch position did not exceed 35 ppt (Figure 4d), which was relatively low compared to the coastal waters. However, the regions off the northern coast of Gambia and the upwelling regions in the Gulf of Guinea had relatively high salinity levels almost all year round.

(A)                                    (B)



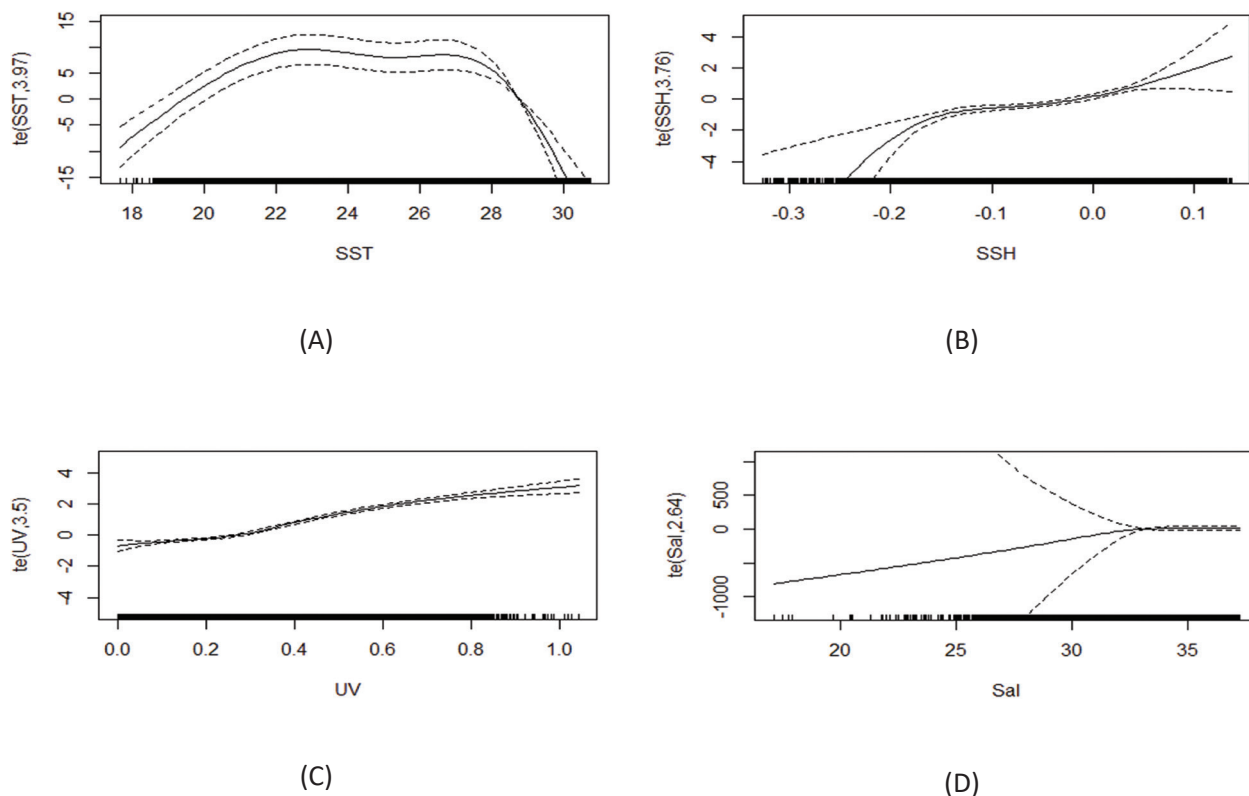(C)                                    (D)

Fig. 4: Loess plots of (A) sea surface temperature (B) sea surface height (C) currents (D) sea surface salinity used in the Generalized additive model (GAM). (Dashed lines indicate 95% confidence intervals. The relative density of data points is shown by the inner ticks on the x-axis)

An output from the GAM using SST, SSH, surface currents and SSS as inputs ensures fisheries managers can now forecast potential fishing zones (PFZ) with some degree of certainty. Figure 5 shows a forecast product that indicates that the optimal conditions conducive for fishes to aggregate were in the equatorial regions and north of latitude 6°S which are the major upwelling regions. Low probability of tuna occurrences was observed in the coastal waters with riverine discharge from the Niger and Congo rivers, as well as regions traversed by the westward flowing North Equatorial Counter Current (NECC). Changes in the spatial distribution of tuna are expected to vary with processes that drive phytoplankton growth and changes in riverine discharge and current flow systems in the region.
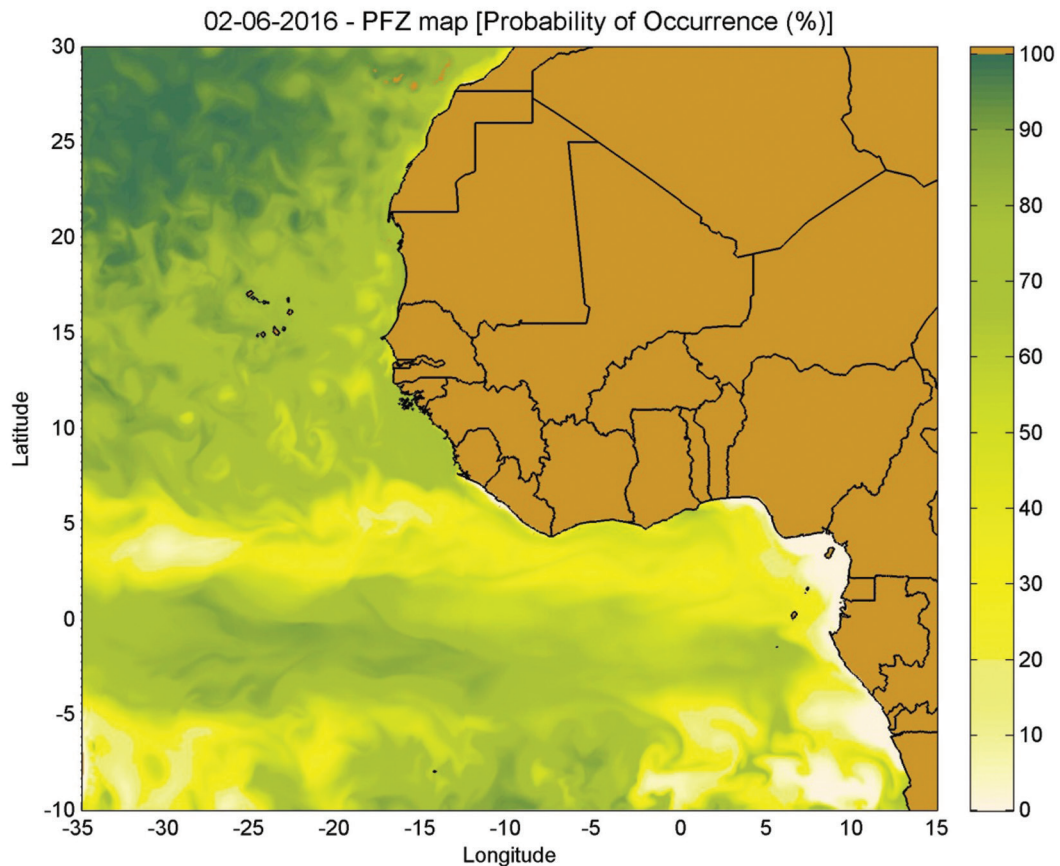
Fig. 5: Potential fishing zone map derived from Generalized additive model (GAM) with oceanographic parameters (sea surface temperature, sea surface height, sea surface currents and sea surface salinity) as inputs.  The scale (0-100) gives an indication of favourable conditions that promote fish aggregation.

## Discussion

Tuna catches in the Gulf of Guinea largely occur in two major upwelling areas, i.e. off the shelf waters of Ghana and Cote d'Ivoire, and within the equatorial waters. In the Gulf of Guinea, the seasonal upwelling which peaks during July to September ensures an abundance of plankton (Wiafe *et al.,* 2008) that is preyed on by small pelagics, especially *Sardinella*, which are in turn preyed on by large pelagics (Koranteng, 2002). These biological interactions in the Gulf of Guinea increase food availability for large pelagics, which explains the rise in tuna abundance. Sardines have been observed

to be an important component of the diet of juvenile Pacific Bluefin tuna, *Thunnus thynnus orientalis*, in the productive waters off Japan (Baillif, 1991; Polovinia, 1996). It was observed in this study that an increase in tuna catch occurred after the upwelling season. During this period in the equatorial Atlantic, there is a slight rise in temperature. Increased temperature after upwelling may have improved temperature preference for tropical tuna. The sub(surface) vertical movements bring nutrients and subsequent local enrichment that enhance phytoplankton growth with a corresponding

increase in zooplankton (McGlade *et al.*, 2002) and small pelagics which are the preferred tuna prey (Sevchenko, 1996). Results of PICOLO I and II experiments in the tropical Atlantic indicate increased zooplankton biomass (copepod) is associated with increased Chl-a concentrations and decreased SST during periods of increased upwelling (Champalbert and Pagano, 2002).

The low abundance of Bluefin can be attributed to the conservation efforts being implemented by the regional fisheries management organizations (e.g. International Commission for the Conservation of Atlantic Tunas - ICCAT) mandated to ensure the sustainable use of the resource. These management strategies specifically target Bluefin, which prohibits fishing in the East Atlantic by large-scale baitboats and trolls. These industrial fishing fleets in recent years are permitted to fish from 1 July to 31 October. Again, fishing vessels flagged to Ghana are prohibited from fishing from January to February, further reducing fishing efforts and catch during those periods.

Fishing effort of tuna followed a seasonal pattern similar to tuna abundance in the eastern Atlantic. There is intense fishing during the period of the year when upwelling intensifies resulting in high phytoplankton abundance and its associated cold surface temperatures. Changes in temperatures have also been observed to affect fish activity and distribution. Studies of temperature effect on metabolic rate and swimming speed of Pacific Bluefin tuna suggest there is reduced metabolism and activity during intermediate ambient temperatures and considerably high metabolism at very low or high temperatures (Blank *et al.,* 2007). There is also the tendency for tuna to vertically migrate in response to changing temperature in the upper ocean surface (Kitigawa *et al.,* 2000). Copepods, the dominant zooplankton, peak in numbers during the upwelling season (Wiafe *et al.,* 2008), and are the preferred prey of *Sardinella spp.* (Koranteng and McGlade, 2001). Increased primary and secondary production, inadvertently, will result in food availability for tuna in the equatorial and shelf regions in the Gulf of Guinea. During this period there is intense fishing by both the artisanal and industrial fisheries for small pelagics and high value commercial species including

tuna. The warming of the Guinea Current large marine ecosystem peaks in January to May. This creates a highly stratified conditions with a deep and strong thermocline which locks the nutrient-rich water at deeper depths.

The GAM analyses revealed strong environmental influence on the distribution of tropical Atlantic tuna. The optimal range for the four parameters associated with fish catch locations which were observed are also linked to the intensification of the upwelling in the Gulf of Guinea. This strong connection between the distribution of the tuna and upwelling gives credence to the huge importance that regional and international fisheries management bodies attach to fishing activities in that area. The output from this model will be useful for monitoring the temporal and spatial changes in the upper ocean and their influence on tuna distribution. A reliable and timely generation of this habitat mapping information will further enhance the sustainable management of tuna resources. Inference made from the forecast PFZs can ensure that activities of fishing fleets that target Atlantic tuna can be predicted to identify areas which are likely to be targeted for fishing. Hence, monitoring and surveillance patrols can be planned based on hotspots identified from the PFZ maps.

High AIC and residual deviance observed during the initial setup of the GAM with only SST, SSH, SSS and surface currents clearly indicates that the distribution of tuna, like many other fishes, responds better to a suite of environmental parameters. However, the lowest AIC or residual deviance for surface temperature, when compared with surface height, salinity and currents, further indicates its major role in fish assemblages in the major upwelling centres. Walsh and Kleiber (2001) noted the importance of temperature on tuna based on the physiological impact and ecological influence on the species. Temperature exerts a strong effect on energetic demands of endothermic species, and it influences biological production in marine ecosystems. Regions in the Guinea current LME with high tuna distribution are the upwelling centres off the shelves of Ghana-Cote d'Ivoire and the equatorial Atlantic which have moderate to low surface temperatures, negative SSH anomalies,

relatively high surface currents and low salinity. Wexler *et al. (*2011) provides various thermal preferences for survival, development and growth of Pacific Yellowfin tuna larvae. For instance, yolk sac and first-feeding Yellowfin larvae exhibited lethal limits for their survival at temperatures less than 21°C and greater than 33°C. These figures are within observed thermal ranges in the eastern equatorial Atlantic, supporting studies that indicate the area is a spawning and nursery area.

## Conclusions

The spatial distribution of tuna followed the cold tongue of water that emanated from the southwestern coast of Africa, which moved westward along the equator.  This has a persistent connection with the upwelling regions of the west coast of Africa. Tuna catch locations were in less warmer waters with relatively less thermal variation. Those areas in the Gulf of Guinea are also highly productive, especially during the upwelling seasons. Outputs from the model can provide a quick outlook of the possible areas where tuna is likely to aggregate and the potential fishing area where monitoring and surveillance needs to be intensified.

## Acknowledgement

## References

Adinortey, A. E., Ofori-Danson, P. K., Nunoo, F. K. E., Abdulhakim, A., & Tayib, Y. (2016). Trends and effects of Gears on the catches of Tuna landed in Ghana. Elixir Bio Diver, 93, 39593-39598.

Bayliff, W. H. (1991). Status of northern bluefin tuna in the Pacific Ocean. In IATTC, world meeting on stock Assessment of Bluefin tunas: strengths and weaknesses. *Inter-American tropical Tuna Commission Special Report* (Vol. 7, pp. 29-88).

Blackburn, M. (1975). Distribution and ecology of skipjack tuna, *Katsuwonus pelamis,* in an offshore area of the eastern tropical Pacific Ocean. *Fish. Bull*, 73, 382-411.

Blank, J. M., Morrissette, J. M., Farwell, C. J., Price, M., Schallert, R. J., & Block, B. A. (2007). Temperature effects on metabolic rate of juvenile Pacific bluefin tuna *Thunnus orientalis. Journal of Experimental Biology*, 210(23), 4254-4261.

Champalbert, G., and Pagano, M. (2002). Copepod feeding in a tuna fishery area of the tropical Atlantic Ocean. *Comptes rendus biologies*, 325(2), 171-177.

Fromentin, J. M., & Powers, J. E. (2005). Atlantic bluefin tuna: population dynamics, ecology, fisheries and management. *Fish and Fisheries*, 6(4), 281-306.

Gutenkunst, R., Newlands, N., Lutcavage, M., and Edelstein-Keshet, L. (2007). Inferring resource distributions from Atlantic bluefin tuna movements: An analysis based on net displacement and length of track. *Journal of Theoretical Biology*, 245(2), 243-257.

Holland, K. N. (1990). Horizontal and vertical movements of yellowfin and bigeye tuna associated with fish aggregating devices. *Fish. Bull.,* 88, 493-507.

IPCC (2007). Climate Change Mitigation, Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007.

Kimura, S., Kato, Y., Kitagawa, T., and Yamaoka, N. (2010). Impacts of environmental variability and global warming scenario on Pacific bluefin tuna (*Thunnus orientalis*) spawning grounds and recruitment habitat. *Progress in Oceanography*, 86(1), 39-44.

Kitagawa, T., Nakata, H., Kimura, S., Itoh, T., Tsuji, S., & Nitta, A. (2000). Effect of ambient temperature on the vertical distribution and movement of Pacific bluefin tuna *Thunnus orientalis. Marine Ecology Progress Series*, 206, 251-260.

Koranteng, K. A., & McGlade, J. M. (2001). Climatic trends in continental shelf waters off Ghana and in the Gulf of Guinea, 1963–1992. *Oceanologica Acta*, 24(2), 187-198.

McGlade, J. M., Cury, P., Koranteng, K. A., & Hardman-Mountford, N. J. (2002). Fish species assemblages on the continental shelf and upper slope off Ghana. *The Gulf of Guinea Large Marine Ecosystem: Environmental Forcing & Sustainable Development of Marine Resources*, 11, 173.

Mensah, M. A., & Quaatey, S. N. K. (2002). An overview of the fishery resources and fishery research in the Gulf of Guinea. *The Gulf of Guinea large marine ecosystems*, 227-240.

Miyabe, N., & Nakano, H. (2004). *Historical trends of tuna catches in the world* (No. 467). Food & Agriculture Organization.

Natale, F., Gibin, M., Alessandrini, A., Vespe, M., & Paulrud, A. (2015). Mapping fishing effort through AIS data. *PloS one*, 10(6), e0130746.

Nunoo, F. K. E., Asiedu, B., Olauson, J., & Intsiful, G. (2015). Achieving sustainable fisheries management: A critical look at traditional fisheries management in the marine artisanal fisheries of Ghana, West Africa. *Journal of Energy and Natural Resource Management*, 2, 15-23.

Polovina, J. J. (1996). Decadal variation in the trans-Pacific migration of northern bluefin tuna (*Thunnus thynnus*) coherent with climate-induced change in prey abundance. *Fisheries Oceanography*, 5(2), 114-119.

Shevchenko, N. F. (1996). Feeding of the Oceanic lightfish *Vinciguerria nimbaria* (Gonostomatidae) in the dynamically active zones of the Eastern Equatorial Atlantic. *Journal of Ichthyology*, 36, 476-478.

Walsh, W. A., and Kleiber, P. (2001). Generalized additive model and regression tree analyses of blue shark (*Prionace glauca*) catch rates by the Hawaii-based commercial longline fishery. *Fisheries Research*, 53(2), 115-131.

Wexler, J. B., Margulies, D., and Scholey, V. P. (2011). Temperature and dissolved oxygen requirements for survival of yellowfin tuna, *Thunnus albacares*, larvae. *Journal of Experimental Marine Biology and Ecology*, 404(1), 63-72.

Wiafe, G., Yaqub, H. B., Mensah, M. A., and Frid, C. L. (2008). Impact of climate change on long-term zooplankton biomass in the upwelling region of the Gulf of Guinea. *ICES Journal of Marine Science*, 65(3), 318-324.

# The microbiological quality of imported frozen chicken drumsticks from retail meat shops in Accra, Ghana

**Angela Parry-Hanson Kunadu[1], Derrick Brian Amoako[1] and Kwaku Tano-Debrah[1*]**

[1]Department of Nutrition and Food Science, College of Basic and Applied Sciences, University of Ghana
*Corresponding author: ktano-debrah@ug.edu.gh

## ABSTRACT

There has been an increase in chicken meat consumption in Ghana due to the affordability and convenience of imported frozen chicken parts compared with fresh chicken meat. However, there is little information on the quality of frozen chicken meat. The aim of this study was to assess the microbiological quality and safety of imported frozen chicken drumsticks from supermarkets, cold stores and open markets. A total of 24 chicken meat samples were analyzed to determine the Total Viable Count (TVC), Total Coliform Count (TCC), concentration of *Escherichia coli* and prevalence of *Salmonella*. The mean TVCs were 6.34, 5.92 and 5.42 Log CFU/g respectively for open markets, supermarkets and cold stores. *E. coli* and *Salmonella* prevalence was 66.7% and 91.7% respectively. There was no significant difference in the microbiological quality and prevalence of *E. coli* and *Salmonella* on frozen chicken meat purchased from all three retail outlets. Results of this study suggest that raw frozen chicken drumsticks from Ghanaian retail outlets have poor microbiological quality and are highly contaminated with potential pathogens including species of *E. coli* and *Salmonella*.

**Keywords:** Chicken, Microbiological quality, food safety, *Salmonella, Escherichia coli*

## Introduction

Chicken meat is an important component of the Ghanaian diet. Its consumption has risen steadily since the year 2000 (GAIN, 2011) and now comes second to fish in the supply of animal proteins in the Ghanaian diet. In 2011, poultry consumption was 115,000MT, up from 108,000MT the previous year. Yet, domestic supply of broilers meets less than 10% of the local demand. According to the GAIN (2011) report, 90% of chicken meat consumed locally is imported mainly in the frozen state. In the urban and peri-urban areas, imported frozen chicken parts are highly patronized by operators of hotels, restaurants, fast food joints and for home consumption. The high patronage is due to the lower cost of such imported chicken parts compared to locally produced broilers. Besides, they are sold 'pre-cut', making them convenient to use. The drumsticks are particularly popular.

Freezing is usually the method of choice for long term preservation of meat, and is known to cause 1 to 2 Log CFU/g reductions in microbiota, especially during slow freezing (Adam and Moss, 2008; Adu-Gyamfi *et al.,* 2014). However, freezing can give a false sense of quality and safety if the cold chain is not properly maintained. The microbiology of chicken meat at the point of sale depends on several factors including method of slaughter, sanitation during processing, packaging and dispatch, hygienic handling at retail, and maintenance of the cold chain from slaughter to retail (Selvan *et al.,* 2007). Contamination of raw chicken meat along the value chain can be aggravated during temperature abuse.

In Ghana, frozen chicken parts are marketed in supermarkets, cold stores and open markets that have the infrastructure to support cold storage. Typically, frozen meat is purchased in bulk and transported at ambient

temperatures to retail shops. As frozen meat begins to thaw, condensation on the meat surface supports the multiplication of both spoilage and pathogenic microorganisms, which affect the microbiological quality of the chicken meat. Other microbiological risk factors at retail points include frequent opening and closing of freezers, which increases the freezer temperature and thus compromises the cold chain. Direct contact of frozen meat with ungloved hands, and often after handling retail items such as money, could lead to non-source contamination. Furthermore, with a poorly resourced power sector, even small and medium scale retailers who own generators do not power their freezers throughout the period of power outages. All these have implications for the quality and safety of frozen chicken meat at the point of sale.

Presently, limited information exists on the quality of frozen chicken in the Ghanaian retail market. Arhin Sackey *et al.* (2002) assessed the presence of *Campylobacter*, *Salmonella*, *Shigella* and *Escherichia coli* in live and dressed poultry from Accra, reporting <15% prevalence of the targeted pathogens. Adu-Gyamfi *et al.,* (2014) also reported the presence of *E. coli* on raw chicken meat. Therefore, frozen chicken meat on the Ghanaian market needs to be frequently monitored to safeguard the safety of the populace and design policies and interventions to control such meat and prevent enteric foodborne diseases.

The aim of this study was to assess the microbiological quality and the prevalence of *E. coli* and *Salmonella* on imported frozen chicken drumsticks from Ghanaian retail shops.

## Materials And Methods

### Sample collection

Twenty-four packs of frozen chicken drumsticks were purchased from twelve retail shops in Accra and Tema metropolis. The retail shops comprised four supermarkets, four cold stores and four open markets with functional freezers. For each of the shops, a pack of chicken drumsticks was purchased on two different days. Each sample was placed in a sterile plastic bag and transported immediately to the laboratory in a disinfected thermos flask. Frozen chicken was allowed to thaw at refrigeration temperatures prior to microbiological testing.

### Microbiological analyses

A 25g portion of chicken was excised aseptically and placed in a sterile stomacher bag containing 225ml of sterile buffered peptone water (Oxoid, England). The sample was homogenized for 2 minutes in a stomacher and the homogenate used for further serial dilutions with sterile buffered peptone water as diluent. Microbiological tests were conducted using standard microbiological methods (ISO 4833-1 for total viable count, ISO 4832:2006 for total coliform count, ISO 16654:2001 for *Escherichia coli*, and ISO:6579:2002 for *Salmonella* spp.). Total aerobic mesophiles were determined on Plate Count Agar (Biolab, USA) incubated at 37˚C for 24 h. To enumerate Total coliforms, Violet Red Bile Glucose Agar (Oxoid, England) was used and cultured at 37˚C for 24 h. *Escherichia coli* was detected using Eosin Methylene Blue Agar (Oxoid, England) incubated for 24h at 45˚C. Colonies with metallic sheen were enumerated as presumptive *E. coli*. Presumptive *Salmonella* was detected using Brilliant Green Agar (Oxoid, England) at 37˚C for 24 h. Red colonies with imparted red/pink colour in the surrounding medium were considered typical colonies of *Salmonella*. For all microbiological testing, the pour plate method was used to determine microbial contamination on frozen chicken samples. Characteristic colonies enumerated were reported as CFU/g. Presumptive colonies of *E. coli* and *Salmonella* were confirmed with the following biochemical tests: Triple Sugar Iron (TSI) test, Indole test, catalase test, citrate test, and Gram test for *E. coli*; and Triple Sugar Iron (TSI) test, Catalase test, Citrate test, urease test, lysine decarboxylase test and Gram test for *Salmonella* spp.

## Statistical analysis

One-way Analysis of Variance (ANOVA) was used to determine whether there were statistical differences between the microbiological quality of frozen chicken drumsticks from the different categories of retail shops. P value was set at ≤0.05. Samples were analyzed in duplicate. ANOVA was carried out to detect significant differences between counts from the three types of retail outlets using Minitab 14.

## Results and Discussion

The mean bacterial counts and prevalence of *Salmonella* in raw frozen chicken drumsticks are presented in Table 1. The total viable bacterial counts ranged from 5.40-7.48 Log CFU/g, 3.78-6.67 Log CFU/g and 3.00-6.92 Log CFU/g for open markets, supermarkets and cold stores respectively. There were no significant differences (P> 0.05) between mesophilic counts between all three types of retail outlets. Summaries of confirmatory tests for *Salmonella* and *E. coli* are presented in Tables 2 and 3 respectively. With the exception of one open market sample, most of the counts were within the national specification of <7 Log CFU/g. The level of compliance of the frozen chicken samples to national standards is presented in Table 4. The upper limits for mesophilic bacteria on frozen chicken drumsticks were generally higher than those reported by Alvarez-Astorga *et al.* (2002) who recorded counts in the range of 4.97- 5.79 Log CFU/g for raw chicken legs. Higher mesophilic bacterial counts have been reported in other studies. For instance, vacuum packaged refrigerated ostrich fillets were found to contain 6.20-7.78 Log CFU/g bacteria, while a mean aerobic count of 7.70 Log CFU/g was reported for frozen chicken meat (Alonso-Calleja *et al.,* 2004: Javanmard *et al.,* 2006). High bacterial loads were attributed to possible temperature abuse during processing, storage, transportation, distribution or retail.

Similarly, high mesophilic counts recorded for frozen chicken drumsticks in this study could be due to a lack of integrity of the cold chain from dispatch, transportation, and especially storage, and cross-contamination from handling at retail. Transportation at ambient temperatures is a common practice in Ghana, which, in such a tropical environment, could reach well beyond 30˚C. Several freeze-thaw cycles could have resulted in microbial proliferation. The poor microbial quality of meat observed could also be due to cross-contamination from other unpackaged meats usually stored together in the freezers. Cross-contamination with hands of personnel handling different types of meat and other non-meat retail items, including money (currency notes), are all possible sources of microbial contamination. Moreover, repeated opening and closing of freezers can result in fluctuations of freezing temperatures and may allow phychrophiles such as *Listeria monocytogenes* which has a growth temperature range of -1.5 to 45°C to proliferate, albeit slowly (FSANZ, 2017; Ryser and Donnelly, 2015).

The presence of mesophilic bacteria is an index of the hygienic quality of meat and also represents a significant portion of spoilage bacterial populations on the meat. Biochemical changes such as sliminess and putrefaction which are evidence of microbial spoilage of meat are often masked when meat products are frozen. Such changes become evident when microbial populations reach 7 log CFU/g and above. More than 66% of the frozen chicken samples in this study had total viable bacterial counts >6.0 Log CFU/g (Table 1). Temperature fluctuations due to frequent opening and closing of freezers and multiple freeze thaw cycles because of power fluctuations could lead to quality deterioration of chicken meat as a result of microbial growth of psychrophiles during temperature abuse.

**Table 1: Bacterial counts and prevalence of *Salmonella* on frozen chicken drumsticks**

| Market type | Market samples | $Log_{10}$ CFU/g (±Standard deviation) | | | |
|---|---|---|---|---|---|
| | | TVC | TCC | E. coli | *Salmonella* detection |
| Open Air Market | OP1 | 6.18 (±0.94) | 5.10(±2.55) | 2.88(±1.56) | ++ |
| | OP2 | 6.44 (±1.47) | 4.00(±0.99) | 4.00(±0.99) | ++ |
| | OP3 | 6.18(±0.52) | 5.89(±0.41) | 3.12(±2.23) | ++ |
| | OP4 | 6.55(±0.08) | 2.90(±0.57) | 2.54(±0.51) | + |
| Supermarket | SP1 | 5.22(±2.04) | 3.69(±3.38) | 0.59(±0.83) | ++ |
| | SP2 | 6.04(±0.37) | 2.32(±0.40) | 0.85(±1.20) | ++ |
| | SP3 | 6.37(±0.27) | 4.47(±2.03) | 4.54(±2.31) | + |
| | SP4 | 6.05(±0.44) | 4.85(±0.21) | 0.77(±1.09) | ++ |
| Cold store | CS1 | 6.65(±0.37) | 3.51(±2.78) | 0.50(±0.71) | ++ |
| | CS2 | 4.24(±1.75) | 2.36(±0.31) | 2.70(±3.82) | ++ |
| | CS3 | 5.70(±0.06) | 3.57(±0.24) | 1.74(±2.46) | ++ |
| | CS4 | 5.09(±0.12) | 3.20(±0.28) | 1.06(±1.49) | ++ |
| | P value | 0.181 | 0.271 | 0.499 | |

TVC: Total Viable count; TCC: Total Coliform Count; +detection on one sample, ++ detection on 2 samples from the same market

**Table 2: Summary of confirmatory test results for *Salmonella* isolated from frozen chicken samples**

| Sample ID | Triple Sugar Iron | | | | Catalase | Citrate | Urease | Lysine De-carboxylase | Gram's test | Morphology |
|---|---|---|---|---|---|---|---|---|---|---|
| | Butt | Slant | $H_2S$ | Gas | | | | | | |
| OP1 | + | + | + | + | + | + | - | + | - | Rods |
| OP2 | + | + | + | + | + | + | - | + | - | Rods |
| OP3 | + | + | + | + | + | + | - | + | - | Rods |
| OP4 | + | + | + | + | + | + | - | + | - | Rods |
| CS1 | + | + | + | + | + | + | - | + | - | Rods |
| CS2 | + | + | + | - | + | + | - | + | - | Rods |
| CS3 | + | + | + | + | + | + | - | + | - | Rods |
| CS4 | + | + | + | + | + | + | - | + | - | Rods |
| SM1 | + | + | + | + | + | + | - | + | - | Rods |
| SM2 | + | + | + | - | + | + | - | + | - | Rods |
| SM3 | + | + | + | + | + | + | - | + | - | Rods |
| SM4 | + | - | - | + | + | + | - | + | - | Rods |

OP1-4: Open markets at four locations in Accra; CS1-4, Samples from Cold Stores at four locations in Accra; SM1-4: Samples from supermarkets at four locations in Accra. From each sample location, six presumptive colonies were tested. Presented here are summary results for at least one positive colony for Salmonella from frozen chicken sampled from each of the sample locations.

**Table 3: Summary of confirmatory tests for *Escherichia coli* isolated from frozen chicken obtained from markets in Accra**

| Sample ID | Triple Sugar Iron | | | | Catalase | Indole | Citrate | Gram's test | Morphology |
|-----------|------|-------|--------|-----|----------|--------|---------|-------------|------------|
|           | Butt | Slant | $H_2S$ | Gas |          |        |         |             |            |
| OP1 | - | + | - | + | + | + | - | - | Rods |
| OP2 | - | + | - | + | + | + | - | - | Rods |
| OP3 | - | + | - | + | + | + | - | - | Rods |
| OP4 | - | + | - | + | + | + | - | - | Rods |
| CS1 | - | + | - | + | + | + | - | - | Rods |
| CS2 | - | + | - | - | + | + | - | - | Rods |
| CS3 | - | + | - | + | + | + | - | - | Rods |
| CS4 | - | + | - | + | + | + | - | - | Rods |
| SM1 | - | + | - | + | + | + | - | - | Rods |
| SM2 | - | + | - | + | + | + | - | - | Rods |
| SM3 | - | + | - | + | + | + | - | - | Rods |
| SM4 | - | + | - | + | + | + | - | - | Rods |

OP1-4: Open markets at four locations in Accra; CS1-4, Samples from Cold Stores at four locations in Accra; SM1-4: Samples from supermarkets at four locations in Accra. From each sample location, six presumptive colonies were tested. Presented here are summary results for at least one positive colony for *E. coli* from frozen chicken sampled from each of the sample locations.

Javanmard *et al.* (2006) reported a mean coliform count of 7.00 Log CFU/g for frozen chicken meat, which was higher than the values obtained in this study. On the other hand, coliform counts from this study are supported by similar studies by Adu-Gyamfi *et al.* (2012) who reported mean coliform counts of 3.80 Log CFU/g for supermarkets, 3.46 log CFU/g for local markets and 3.14 log CFU/g for chicken carcass obtained from farms. Coliforms are good indicators for the general sanitary condition of the food and may suggest hygiene gaps in food processing environments and in processed foods.

**Table 4: Compliance of frozen chicken drumsticks to national microbiological standards**

| Market type | N | Mean counts ($Log_{10}$ CFU/g) (% Compliance to Standard) | | | |
|-------------|---|------|------|---------|------------|
|             |   | TVC  | TCC  | E. coli | Salmonella |
| Open market | 8 | 6.34 (87.5%) | 4.47 (50.0%) | 2.56 (12.5%) | 7* (12.5%) |
| Supermarket | 8 | 5.92 (100%) | 3.90 (50.0%) | 1.69 (37.5%) | 7* (12.5%) |
| Cold Store | 8 | 5.42 (100%) | 3.16 (87.5%) | 1.50 (50.0%) | 8* (0%) |
| National Standard | | < 7.0[a] | <4.0[b] | ND[c] | ND[d] |

TVC: Total Viable count; TCC: Total Coliform Count, *Number of samples positive for *Salmonella*, Ghana Standard reference numbers a: GSS955/2013; b: GS 236:1997; c: GS 955/2009; d: GS 955/2013.

The presence of *E. coli* (fecal coliforms) on the other hand suggests recent fecal contamination. In this study *E. coli* had a prevalence rate of 67% while *Salmonella* prevalence reached 92% (Table 2). There were no significant differences in prevalence of *E. coli* and *Salmonella* between open markets, cold stores and supermarkets (P> 0.05). Similar prevalence has been reported for *E. coli* on chicken (Arhin Sackey *et al.*, 2001; Adu-Gyamfi *et al.*, 2012; Odwar *et al.*, 2014; Sharma and Chattopadhyay, 2015). The presence of *E. coli* on frozen chicken may suggest poor personnel hygiene and unsanitary contact with carcasses along the value chain.

*Salmonella* prevalence generally exceeded figures reported in the literature. Adu-Gyamfi *et al.* (2012) reported a *Salmonella* prevalence of 7% on raw chicken meat in Accra, Ghana, whereas Arhin Sackey *et al.* (2001) reported a 13% prevalence in freshly dressed chicken carcasses in the same city. These discrepancies could be due to the fact that fresh or chilled poultry carcasses, which the above authors investigated, have shorter shelf-lives, and thus experience less handling and storage during processing and retail compared to frozen meat. Although the association of *Salmonella* with chicken has been well established (Vindigni *et al.*, 2007; Adam and Moss, 2008), the high prevalence obtained in this study raises concerns about food safety.

## Conclusions

The results of this study highlighted the issue of poor microbiological quality of frozen chicken drumsticks at the point of sale. The high prevalence of potential foodborne pathogens is of great concern. It is recommended that regulatory authorities work closely with stakeholders in the frozen chicken trade to set performance objectives at different stages in the value chain and employ control measures that will effectively preserve chicken meat until use.

## References

Adam, M. R. & Moss, M. O. (2008). Low temperature storage - Chilling and freezing. In *Food Microbiology* (3rd Ed., pp. 96-98). Cambridge, UK: The Royal Society of Chemistry.

Adu-Gyamfi, A., Torgby-Tetteh, W. & Appiah, V. (2012). Microbiological quality of chicken sold in Accra and determination of $D_{10}$-value of *E. coli*. *Food and Nutrition Sciences* 3, 693-698.

Alonso-Cajella, C., Martínez- Fernández, B., Prieto, M. & Capita, R. (2004). Microbiological quality of vacuum-packaged retail ostrich meat in Spain. *Food Microbiology* 21, 241-246.

Álvarez-Astorga, M., Capita, R., Alonso-Calleja, C., Moreno, B. & Garćia-Fernández, M del C. (2002). Microbiological quality of retail chicken by-products in Spain. *Meat Science* 62, 45-50.

Arhin Sackey, B., Mensah, P., Collison, E. & Sakyi Dawson, E. (2001). *Campylobacter, Salmonella, Shigella* and *Escherichia coli* in live and dressed poultry from metropolitan Accra. *International Journal of Food Microbiology* 71, 21-28.

Cetinkaya, F., Cibik, R., Soyutemiz, G. E., Ozakin, C., Kayali, R. & Levent, B. (2008). *Shigella* and *Salmonella* contamination in various foodstuffs in Turkey. *Food Control* 19, 1059-1063.

FSANZ. (2017). *Listeria monocytogenes*. In: Agents of Foodborne Illness. Food Standards Australia New Zealand online technical series. Available at http://www.foodstandards.gov.au/publications/Documents/Listeria%20monocytogenes.pdf. Accessed June 21st, 2018.

Global Agriculture Information Network, USDA. (2011). Ghana: Poultry and products brief annual. Foreign Agricultural Service. Retrieved from http://gain.fas.usda.gov/Recent%20GAIN%20Publications/Poultry%20and%20Products%20Brief%20Annual_Accra_Ghana_10-7-2011.pdf. Accessed April 5th 2015.

Ghana Standards Authority (2009). GSS 955/ 2009. Ghana Standards Authority, Accra, Ghana.

Javanmard, M., Rokni, N., Bokai, S. & Shohhosseini, G. (2006). Effects of gamma irradiation and frozen storage on microbial, chemical and sensory quality of chicken meat in Iran. *Food Control* 17, 469-473.

Odwar, J. A., Kikuvi, G., Kariuki, J. N. & Kariuki, S. (2014). A cross-sectional study on the microbiological quality and safety of raw chicken meats sold in Nairobi, Kenya. *BMC Research Notes* 7, 627.

Ryser, E. T. & Donnely, C. W. (2015). Listeria. . In Y. Salfinger & M. L. Tortorello (Eds.) *Compendium of Methods for the Microbiological Examination of Foods*. (5th Ed., pp 425-443). American Public Health Association, APHA Press, Washington, DC.

Sharma, K. P. & Chattopadhyay, U. K. (2015). Assessment of microbial load of raw meat samples sold in the open markets of city of Kolkata. *IOSR Journal of Agriculture and Veterinary Science* 8, 24-27.

Selvan P. R., Narendra B., Sureshkumar S. & Venkatamanujam V. (2007). Microbial quality of retail meat products available in Chennai City. *American Journal of Food Technology* 2, 55-59.

Vindigni, S. M., Srijan, A., Wongstitwilairoong, B., Marcus, R., Meek, J., Riley, P. L. & Mason, C. (2007). Prevalence of foodborne microorganisms in retail foods in Thailand. *Foodborne Pathogens and Disease* 4, 208-215.

# The contribution of enterprise establishment factors to employment growth of the pineapple industry in the Nsawam-Adoagyiri District in the Eastern Region of Ghana

## Seth Dankyi Boateng[1*] and Kafui Kwesi Agbe[1]

[1]Department of Agrucultural Extension. School of Agriculture. University of Ghana, Legon
*Corresponding author: sdboateng@ug.edu.gh

### ABSTRACT

The establishment of enterprises is influenced by various factors identified in the theories of enterprise establishment. However, most entrepreneurs do not pay particular attention to such factors when establishing and managing their enterprises. The aim of this study is to examine how enterprise establishment factors contribute to employment growth of the pineapple industry in Nsawam-Adoagyiri District. The enterprise establishment factors investigated in this study are the basic tasks identified by the entrepreneur, availability of funds for enterprise establishment, business information and entrepreneur's expertise from previous education and previous work. A survey was conducted using mixed methods research. Data were collected from 172 entrepreneurs in the pineapple industry from eight farmer based organizations randomly selected from Nsawam-Adoagyiri District of Ghana. The data were analyzed with SPSS software using Chi square and multiple regression analysis. The results show that availability of funds for enterprise establishment contributed significantly to employment growth. The study found that business information did not contribute to employment growth of the pineapple industry in the study area. While entrepreneurs' expertise acquired from educational background did not contribute significantly to employment growth of the industry, entrepreneurs' expertise from previous work background contributed significantly to employment growth. The regression analysis showed that the enterprise establishment factors investigated in this research explain 52% of the variations in employment growth in the pineapple industry in the study area. It can therefore be concluded that enterprise establishment factors do not only influence the enterprise establishment process but also contribute to employment growth of the pineapple industry in the study area.

**Keywords:** Employment growth, Enterprise establishment factors, Expertise and employment growth, Basic needs, Business information

## Introduction

The importance of Small and Medium Scale Enterprises (SMEs) as a key driver of economic growth and employment is now widely acknowledged in both developed and developing countries (UNCTAD, 2005; EU, 2011). In Ghana, almost 80% of the economy is made up of SMEs (NDPC, 2010). According to the Korean Development Institute (2008), SMEs in Ghana constitute more than 90% of business units and provide employment for more than 60% of the employed labour force. SMEs account for about 22% of GDP, principally in the agricultural and transport sectors.

SME establishment involves the organization of resources in an economic environment: intellectual, human, social and financial. Entrepreneurs organize these resources to pursue market opportunities. The activities of these entrepreneurs lead to the creation of firms or enterprises.

The establishment and growth of SMEs is important for monetary or social gains, not only to the individual entrepreneur, but also to the nation. De Kok *et al.* (2011) assert that the establishment and growth of enterprises play a very important role in the creation of

jobs. However, they have indicated that only 50 percent of newly established enterprises survive beyond 5 years. Factors affecting the growth and survival of enterprises include inadequate or unavailable funds, inadequate business information, poor business management and marketing skills, and poor corporate governance (Mbugua *et al.*, 2013; Wong and Merrilees, 2005; Fjose *et al.*, 2010). For instance, Mbugua *et al.* (2013) determined that inadequate finance affects employment growth of enterprises.

In the agricultural sector, production and value addition to agricultural produce are vital for agricultural and rural development. Growth in agribusiness and agro-industries in Ghana is fundamental to agricultural transformation and the nation's quest to become an agro-based, industrialized economy. The agricultural industry offers a broader spectrum of activities that comprises production, retailing, processing, exporting and marketing. Growth in agricultural enterprises will be characterized by increase in the number of people employed in the sector, increase in the supply of products resulting in an increase in sales volume of the enterprises, and the advancement in technology usage. Aryeetey *et al.* (2008) also noted that the agricultural industry is key to wage employment and economic development.

Enterprise establishment is perceived as a way of tackling the problem of unemployment and poverty (Robson *et al.*, 2009). The entrepreneurial process of recognising opportunities and acquiring the necessary resources to create enterprises results in the formation of micro, small and medium scale enterprises. These enterprises employ a very large percentage of the workforce in Ghana (The New Legon Observer, 2008). In Ghana, SMEs are principally in the agricultural and transport sectors (AFDB/OECD, 2005).

The agricultural sector in Ghana employs more than half (55.8%) of the workforce (Otoo *et al.*, 2009), most of whom are engaged in both production and processing of agricultural products. This makes the sector important in employment generation. The establishment of enterprises in the agricultural sector will not only create more jobs for people in the economy, but also help control post harvest losses in the country.

Post harvest losses lead to wastage of agricultural output in the country. The waste generated in agricultural output is due to the glut of produce during peak seasons, and inadequate storage and processing infrastructure along the value chain. This glut of produce observed along the value chain can be solved by the establishment of effective and efficient agro enterprises. However, the agro enterprises in the country are few, and while some of them are struggling to grow, others have collapsed. De Kok *et al.* (2011) assert that only fifty (50) percent of newly established enterprises survive after 5 years. In Chile, less than forty-two (42) percent of small businesses survive five years and less than fifty (50) percent survive ten (10) years (Cabrera *et al.*, 2002). Also, Barringer *et al.* (2005) established that of the estimated seven hundred thousand (700,000) new ventures started each year in the United States, only 3.5 percent grow sufficiently to evolve into large firms.

The growth of enterprise is seen as its development from small through medium to large scale (Mao, 2005). Numerous problems have been identified by various researchers in the field of enterprise growth as contributing to the slow growth or collapse of enterprises (Krake, 2005; Wong and Merrilees, 2005; MoFEP, 2012; Tawaih *et al.*, 2013; Fjose *et al.*, 2010). Access to credit, market and land, and poor corporate governance have been identified as factors affecting enterprise growth (MoFEP, 2012). Tawiah *et al.* (2013) has argued that growth is not only affected by credit accessibility of SMEs but also by the unfavourable business environment. Fjose *et al.* (2010) have posited that lack of knowledge and relevant competencies constrain the growth of micro enterprises. However, these researchers have typically measured the growth of firms at the business level (Kruger, 2004), with little or no focus on factors that lead to the establishment of the enterprises.

A review of theories of enterprise establishment (Bull *et al.*, 1995; Schoonhoven and Romanelli, 2001) shows that enterprises are created or formed through the occurrence of certain phenomena. It is expected that when the entrepreneur identifies the basic task that the enterprise is established to perform, has the requisite expertise, operates in a favourable business environment and secures the relevant resources to establish the enterprise,

then enterprise growth in terms of employment growth may be achieved. Also, entrepreneurs around the world attribute much of their success to past work experience and knowledge of the industry, in addition to business contacts obtained through personal social networks and educational background (Nitcher *et al.*, 2005).

Apart from entrepreneurs acquiring experience and knowledge from their previous work and educational background, many public agencies are established to assist entrepreneurs in the acquisition of skills. Some of these agencies in Ghana include the National Board for Small Scale Industries, the Association of Ghana Industries and GRATIS Foundation. In addition, other agencies which provide funding have been set up to provide the needed financial assistance to entrepreneurs for the establishment and development of their enterprises. Examples of such agencies include the Enterprise Development Fund, the Export Development and Agricultural Improvement Fund and the Skill Development Fund. However, despite the large number of agencies and funds instituted in the country for enterprise establishment, the enterprises are not experiencing the expected growth and development (UNCTAD, 2005, Ahiawodzi and Adade, 2012). It is possible that enterprise establishment factors are generally overlooked by both development agents and researchers who focus almost all their attention on post establishment factors while providing support for the establishment and growth of micro enterprises.

Enterprise growth is mostly viewed in the context of internal and limited external factors (Pitelis, 2009; Mao, 2005), with little consideration for the factors that contribute to the establishment of the enterprises. However, Delanoë (2013), Gilbert *et al.* (2006), Masakure *et al.* (2006), and Akoten *et al.* (2006), have determined that factors associated with enterprise establishment (basic task identified, availability of funds, business information and expertise) have a relationship with the establishment and employment growth of enterprises. Yet most entrepreneurs do not pay particular attention to these factors when establishing enterprises (Tushabomwe-Kazooba, 2006). This study is therefore designed to investigate the contribution of enterprise establishment factors to employment growth of the pineapple industry in the Nsawam-Adoagyiri District in the Eastern Region of Ghana.

The study seeks to answer the question; how do enterprise establishment factors contribute to employment growth of the pineapple industry? The specific objective of the study is to determine the contribution of enterprise establishment factors to employment growth of pineapple enterprises in the study area.

## Methods

The study used mixed methods for the data collection. A survey of 172 entrepreneurs in both the production and processing of pineapple were selected by simple random sampling from eight (8) farmer based organizations in the Nsawam-Adoagyiri District. A questionnaire containing both closed and open ended questions was administered for the collection of quantitative and qualitative data. Variables investigated under enterprise establishment factors include enterprise characteristics, availability of funds for enterprise establishment, business information and entrepreneur's expertise.

Employment growth was measured as change in the number of employees within the past three years (2012, 2013 and 2014). The data collected were analyzed using chi square test to determine the relationship between employment growth and enterprise establishment factors. Multiple regression analysis was used to determine the extent to which enterprise establishment factors explained the variations in employment growth in the pineapple industry in the study area. Narratives by respondents were used to support the discussion of the results from the data analysis.

## Results and discussions

### *The contribution of enterprise characteristics to employment growth*

The contribution of enterprise characteristics to employment growth was investigated under two sections: the type of pineapple enterprise (Table 1) and the customer needs that products meet (Table 2).

**Table 1. The contribution of enterprise establishment factors to employment growth**

| Variables | Employment growth | | Chi Square results |
|---|---|---|---|
| | No change | Positive change | |
| **Type of Agro enterprise** | | | |
| **Agro producing enterprise** | 154(96.3%) | 6(3.7%) | $\chi^2$= 47.60; p value = |
| **Agro processing enterprise** | 5(41.7%) | 7(58.3%) | 0.00; df = 1. Sig. |
| <u>Customers' needs</u> | | | |
| **Consumption need** | 42(82.4%) | 9 (17.6%) | $\chi^2$= 10.56; p value = |
| **Raw material need** | 117(96.7%) | 4(3.3%) | 0.01; df = 1. Sig. |
| <u>Funds availability for enterprise establishment</u> | | | |
| **Available** | 4(66.7%) | 2(33.3%) | $\chi^2$= 5.91; p value = |
| **Not available** | 155(93.4%) | 11(6.6%) | 0.02; df = 1. Sig. |
| <u>Access to business information</u> | | | |
| **Access** | 157 (92.4%) | 13(7.6%) | $\chi^2$= 0.17; p value = |
| **No access** | 2(100.0%) | 0(0%) | 0.69; df = 1. NS |
| <u>Business information sources</u> | | | |
| **Business advisory centres** | 1(33.3%) | 2(66.7%) | |
| **Family/Friends** | 70(86.4%) | 11(13.6%) | $\chi^2$= 0.17 |
| **AEA/NGOs** | 86(100.0%) | 0(0%) | p value = 0.69 |
| **No source** | 2(100.0%) | 0(0%) | df = 3. NS |
| <u>Entrepreneur's expertise from educational background</u> | | | $\chi^2$= 2.60 |
| **Expertise** | 99(90.0%) | 11(10.0%) | p value = 0.11 |
| **No Expertise** | 60(96.8%) | 2(3.2%) | df = 1. NS |
| <u>Entrepreneur's expertise from previous work background</u> | | | $\chi^2$= 7.95 |
| **Expertise** | 70(86.4%) | 11(13.6%) | p value = 0.01 |
| **No Expertise** | 89(97.8%) | 2(2.2%) | df = 1. Sig |

Source: Field Survey, 2014

The relationship between the types of pineapple enterprise and employment growth was significant ($\chi^2$ = 47.60; df= 1; p = 0.00). The pineapple processing enterprises are more likely to demonstrate positive employment growth than the producing ones. The finding agrees with Schaffnit-Chatterjee's (2014) findings that increased activity in the value chain, especially in the processing of raw produce, would lead to increase in employment.

The processing of more produce to market requires the employment of more workers to carry out the various activities. This was shown in a narrative typical of pineapple processing entrepreneurs as captured below:

*"You need to do a lot of activities when you are processing. You need people to off load the pineapples, operate the machines, bottle the drinks etc. So you see I need more people to help me"* (Pineapple processor from Adoagyiri).

The narrative shows that the pineapple processing entrepreneur will employ additional workers due to the enormous activities involved in the enterprise. Other

pineapple processing entrepreneurs made mention of the development of capacity in terms of finance and machinery to increase processing activities as shown in a typical narrative below:

*"I have just bought a new machine for juice extraction and I need to add two people to my work force"* (Pineapple processor from Nsawam).

The excerpt above shows that some of the pineapple processing entrepreneurs employ more workers due to the acquisition of machinery to expand processing capacities. Meanwhile the pineapple producing enterprises recorded no employment growth because many of the entrepreneurs and not have the capacity in terms of finance and machinery to expand production to employ more workers. The following excerpts are examples of how some entrepreneurs of agro producing enterprise expressed their concerns:

*"Formerly I had permanent workers on my field but now I have none because I do not have the money and machinery to expand my farm"* (Pineapple producer from Fotobi).

*"I am the only one working on my farm due to money problems"* (Pineapple producer from Pokrom).

It can be deduced from the above statements by entrepreneurs of agro producing enterprises that lack of money is the main reason for not employing people to work on their field. This affects the number of people they employ, and hence employment growth.

On how satisfying customers' needs contributes to employment growth of pineapple enterprises, the chi square test result from the table 1 shows that the relationship was significant, ($\chi^2$ = 10.56; df = 1; p = 0.01). The result indicates that pineapple enterprises that supply products for consumption needs are likely to experience more positive employment growth than pineapple enterprises that supply products for raw material needs (Table 1). Thus the satisfaction of the consumption need of customers will promote employment growth of the pineapple enterprise, wheras the satisfaction of the raw

material need of processors, which is largely undertaken by the pineapple producing enterprises, will not.

### The contribution of availability of funds to enterprise growth

Availability of funds for enterprise establishment was investigated on the basis of entrepreneur's knowledge that such funds were available from government institutions and other organizations for the establishment of the enterprise. The relationship between availability of funds and employment growth was significant, ($\chi^2$ = 5.91; df = 1; p = 0.02) (Table 1). Entrepreneurs who know about available funds were more likely to obtain a higher proportion of positive growth than those who do not. Entrepreneurs with such knowledge are likely to access the funds to employ workers for their pineapple enterprises. This result was in contrast with the findings of Nganda *et al.* (2014), who indicate that financial knowledge of the owner has a marginal weak association with growth of SMEs. The variation in results obtained is explained by constraints imposed by the business environment, the firm, and market characteristics in measuring business environment variables on the size of African firms (Olukunle, 2013).

### Access to business information and employment growth

Access to business information was categorized into "access" and "no access". Entrepreneurs who obtained business information were classified as "access" and those who did not obtain business information were classified as "no access". The relationship between access to business information and employment growth was not significant, ($\chi^2$ = 0.17; df = 1; p = 0.69). This means that there is no significant difference between the growth of employment by entrepreneurs that had access to business information and those who did not. This contrasts with results from Kenya by Bunyasi *et al* (2014) which found that access to business information had a significant effect on the growth of SMEs. It was determined that the information received by the entrepreneurs was more of a technical

than a managerial nature, which did not contribute much to the management of employees in the enterprises.

Business information sources were classified as information received from business advisory centres, Agricultural extension agents/Nongovernmental organizations, and Family/friends. Analysis of the results shows that the relationship between business information sources and employment growth of pineapple enterprises was not significant, ($\chi^2 = 0.17$; df= 3; p = 0.69) (Table 1). There was no significant difference between employment growth of entrepreneurs who accessed business information through business advisory centres, family/friends, AEAs/NGOs and those who did not. There is the possibility that the business information received by the entrepreneurs was used in areas of the business other than the employment of workers.

### Entrepreneurs' expertise and employment growth

The examination of the contribution of entrepreneurs' expertise acquired from educational background to employment growth of agro enterprises shows that the relationship  was not significant, ($\chi^2 = 2.60$; df= 1; p = 0.11) (Table 1). This means that expertise of entrepreneurs acquired from educational background does not contribute to the growth of enterprise in terms of employment. Most of the skills acquired from entrepreneurs' educational background are technical in nature and not managerial, especially in the area of employment.

The contribution of expertise acquired by entrepreneurs from their previous work background to the growth in employment of the enterprise was also examined. The result shows that the relationship  was significant, ($\chi^2 = 7.95$; df = 1; p = 0.01). This means that there are differences in employment growth of entrepreneurs who had previous work experience before establishing such enterprises and those who did not. It was observed that entrepreneurs who acquired expertise from their previous work background had a higher proportion of positive employment growth than entrepreneurs without such expertise. Entrepreneurs acquire management,

communication and negotiation skills from their previous work that are used to employ and maintain their staff. The result confirms the findings of Nitcher *et al.* (2010) that successful entrepreneurs benefited from marketing, administration and negotiation skills developed in previous jobs. These skills acquired by the entrepreneur are a source of competitive advantage within the enterprise (Hernández-Maestro *et al.*, 2009). The following excerpts are typical of some statements obtained from the interactions with entrepreneurs during the field survey:

> "My previous work as a manager helps me in taking
>
> managerial decisions" (Pineapple processor from Nsawam).
>
> "In my previous work, I was employing people for the enterprise, so I  know who to employ"(Pineapple processor from Adoagyiri).
>
> "I know a lot of people I can employ to work for me"
>
> (Pineapple producer from Fotobi).
>
> "I was a supervisor so I learnt how to deal with employees"
>
> (Pineapple producer from Pokrom)

From the above statements it can be observed that some of the entrepreneurs were managers who were employing people and supervising others in their previous work place. Their work, according to them, has given them the experience and skills to manage, communicate, and negotiate with people for employment into their agro enterprises.

### Enterprise establishment factors and employment growth

A multiple regression analysis was used to determine the percentage of the variation in employment growth explained by the combined effects of the enterprise establishment factors (Table 2). For this analysis, employment growth (dependent variable) was measured as the number of workers employed within a period of 3 years. The predictor variables were the factors considered

by the entrepreneurs during the process of enterprise establishment. They were: type of agro-enterprises, customer needs, funds availability, access to business information, expertise from educational background, and expertise from previous work experience. From the analysis the predictor variables explain 52% of the variations in employment growth of the pineapple industries in the study area (Table 2).

**Table 2. Regression analysis of enterprise establishment factors and employment growth**

| Variables | Mean | Std | Multiple regression weight | |
|---|---|---|---|---|
| | | | b | B |
| Type of agro-enterprise | 1 | 0.00 | 0.99** | 0.08** |
| Customer needs | 1 | 0.00 | 0.96** | 0.06** |
| Funds availability | 0.04 | 0.18 | 1.67** | 0.21** |
| Access to business information | 0.99 | 0.11 | 0.25 | 0.12 |
| Expertise from educational background | 0.64 | 0.48 | 0.28 | 0.06 |
| Expertise from previous work experience | 1.54 | 0.50 | 0.01* | 0.41** |

*$p<0.05$.  **$p<0.01$  R=0.66    $R^2$= 0.52   Adjusted $R^2$= 0.41

## Conclusion

This study has addressed only the question of how enterprise establishment factors contribute to employment growth of the pineapple industry. The study found that the characteristics of the enterprise, that is, the basic needs identified by the entrepreneur, contribute to employment growth of the enterprise. The findings support the fact that increased activity in the value chain, especially in the processing of raw produce, would lead to increase in employment.

Availability of funding for enterprise establishment was found to have a relationship with growth in employment. This suggests that if entrepreneurs are educated on where to get funds for agro enterprise establishment, they could access such money to acquire expertise that could lead

to the employment of more workers. Therefore, funding is vital for employment growth. Enterprise establishment factors explain 52% of the variations in employment growth in the pineapple industries in the study area.

There was no significant relationship between entrepreneurs' expertise from educational background and employment growth of the pineapple industry. This was attributed to the fact that entrepreneurs had forms of training other than formal education that enhanced their skills in establishment and management of agro enterprises. Entrepreneur's expertise from previous work background was found to contribute to employment growth. This was due to the human resource management that entrepreneurs acknowledged receiving from their previous work environment. It can be concluded that the enterprise establishment factors do not only influence the entrepreneurial or enterprise establishment process but also contribute to employment growth of the pineapple industry.

## Recommendations

On the basis of the outcomes of the study, it is recommended that would-be entrepreneurs should be given knowledge about the availability of funds for enterprise establishment since such knowledge could lead to the employment of workers when funds are accessed. With the revelation that enterprise establishment factors investigated in this research explain 52% of the variations in employment growth in the pineapple industry, it is recommended that educational institutions include in their curriculum content that helps prospective entrepreneurs to acquire expertise related to enterprise establishment factors such as expertise in sourcing funds, identification of basic needs that can be met by new enterprises, identification of customer needs, and an understanding of the types of enterprises they wish to engage in.

## References

Ahiawodzi, A. K., & Adade, T. C. (2012). Access to credit and growth of small and medium scale enterprises in the Ho Municipality of Ghana. *British Journal of Economics, Finance and Management Sciences,* 34(6).

Akoten, J. E., Sawada, Y., & Otsuka, K. (2006). The determinants of credit access and impacts on micro and small enterprises: the case of garment producers in Kenya. *Economic Development and Cultural Change, 54*(4), 927- 44.

Bull, I., Thomas H., & Willard, G. (1995). *Entrepreneurship Perspectives on Theory Building.* Elsevier Science Ltd: The Boulevard, Langford Lane, Kidlington, Oxford OX5, UK.

Bunyasi, G. N. W., Bwisa, H., & Namusonge, G. (2014). Effect of access to business information on the growth of small and medium enterprises in Kenya. *International Journal of Business and Social Science* 5(10), 121-128.

MoFEP. (2012). *Consultative Group Paper*, 2012. Retrieved January 24, 2014, from www.mofep. gov.gh/CG2012/userfiles/.../private_sector_development.pdf.

Cabrera, A. & Cabrera, E. F. (2002). Knowledge sharing dilemmas. *Organisation Studies*, 23(5), 687-710.

De Kok, J., Vroonhof, P., Verhoeven, W., Timmermans, N., Kwaak, T., Snijders,J., & Westhof, F. (2011). *Do SMEs create more and better jobs?* EIM Business and Policy Research  Zoetermeer, The Netherlands.

Delanoë, S. (2013). From intention to start-up: the effect of professional support. *Journal of Small Business and Enterprise Development, 20*(2), 383-398.

European Commission. (2011). *Policies in support of high-growth innovative SMEs.* An INNO-Grips Policy Brief by empirica Communication and Technology Research. Principal author: Stefan Lilischkis. Bonn.

Fjose, S., Grünfeld, A. L., & Green, C. (2010). *SMEs and growth in Sub-Saharan Africa.* Menon Business Economics, Oslo.

Gilbert, A. B., McDougall, P. P., & Audretsch, B. D. (2006). New venture growth: a review and extension. *Journal of Management, 32*(6), 926 - 950.

Hernández-Maestro, R. M., Muñoz-Gallego, P. A., & Santos-Requejo, L. (2009). Small-Business Owners' Knowledge and Rural Tourism Establishment Performance in Spain. *Journal of Travel Research, 48* (1), 58-77.

Korea Development Institute (KDI). (2008). *Building the Foundation for the Development of SMEs in Ghana.* Ministry of Strategy and Finance, Republic of Korea Government publications: Korea

Krake, F. B. G. J. M. (2005). Successful brand management in SMEs: a new theory and practical hints. *Journal of Product and Brand Management, 14*(4), 228-238.

Kruger, M. E. (2004). *Creativity in Entrepreneurship Domain.* Unpublished Phd. Thesis. University of Pretoria, South Africa. Retrieved from http://upetd.up.ac.za/thesis/available/etd-08242004-145802/.../03chapter3.pdf.

Masakure, O., Henson, S., & Cranfield, J. (2009). Performance of microenterprises in Ghana. A resource-based view. *Journal of Small Business and Enterprise Development, 16* (3), 466 – 484.

Mbugua, J. K., Mbugua, S. N., Wangoi, M., Ogada, J. O., & Kariuki, J. N. (2013). Factors affecting the growth of Micro and Small Enterprises: A case of Tailoring and Dressmaking enterprises in Eldoret. *International Journal of Business and Social Science, 4(5),* 285-293.

Moa, H. (2009). Review on enterprise growth theories. *International Journal of Business and Management*, 4(8), 20-23.

Nganda, J. W., Wanyonyi, K. W., & Kitili, E. M. (2014). Determinants of growth of small and medium enterprises in Kakamega central sub-county, Kenya**.** *Journal of Business Administration and Management Sciences Research, 3*(3), 022-031.

Nitcher, S. & Goldmark, L. (2005). Understanding small and micro enterprise growth. *USAID Development Alternative Inc.* Microreport #6.

Otoo, K. N., Osei-Boateng, C., & Asafu-Adjaye, P. (2009). *Information and Knowledge for the workin class the labour market in Ghana: A descriptive analysis of the labour market component of the Ghana Living Standard Survey* (V). Research paper Number 2009/01.

Pitellis, C. (2009). *Edith Penrose's theory of the growth of the firm fifty years later.* Oxford Press, 4th edition, 2009. Retrieved from http//mpra.ub.uni-muenchen.de/23180/.

Robson, P., Wijbenga, F., & Parker, S. (2009). Entrepreneurship and policy, challenges and directions for future research. *International Small Business Journal, 27*(5), 531-535. Romanelli, E., & Schoonhoven, C. B. (2001). The local origins of new firms. In Schoonhoven, C. B. & Romanelli, E. (eds.), *The entrepreneurship dynamic: Origins of entrepreneurship and the evolution of industries* (pp). Stanford, California: Stanford University Press.

Tawiah, S., Ennin, S., Fosu, K., Ghansah, L., & Oppong, C. K. (2013). *The Impact of Microfinance on Small and Medium Scale Enterprises in Ghana.* (Unpublished Dissertation). Christian Service University College, Ghana.

The New Legon Observer, (2008*). The Private Sector in Ghana: Contributing to Employment, Growth and Social Responsibility.* A Ghana Society for Development Dialogue Publication, 12 June, 2008, vol 2, No. 10.

Tushabomwe-Kazooba, C. (2006). Causes of small business failure in Uganda: A case study from Bushenyi and Mbarara towns. *African Studies Quarterly, 8*(4), 27-36.

United Nations Conference on Trade and Development. (2005). *Improving the competitiveness of SMEs through enhancing productive capacity.* United Nations, New York and Geneva. Retrieved from http//www.unctad.org/meeting.

Wong, H. Y. and Merrilees, B. (2005). A brand orientation typology for SMEs: A case research approach. *Journal of Product and Brand Management, 14* (3), 155- 162.

# Isotopic compositional characterization of siderite nodules: Implications for the reconstruction of the depositional paleo-environment of the Takoradi Shale Formation, around Essipon, Ghana

**Jennifer Edzordzinam Agbetsoamedo[1]\*, Daniel Kwadwo Asiedu[1], David Atta-Peters[1] and Chris Yao Anani[1]**

[1]Department of Earth Science, University of Ghana, Legon
*Corresponding author: jennyagbey@gmail.com

### ABSTRACT

The late Devonian to early Carboniferous Takoradi Shale Formation (TSF) of the Sekondian Group outcrops along the western coast of Ghana is typically composed of hard, compact, black or dark grey fissile shale or sandy shale, rich in organic matter. The upper part of the formation is characterized by inclusions of large discoidal nodules of compact, siderite or clay ironstone. The geological history of this formation, however, has not been robustly established. We carried out petrographic and stable isotopic studies on these siderites in order to constrain and understand the depositional environment of the TSF. The results obtained will add to the reconstruction of the geologic history of the formation. A total of 20 samples of the siderite nodules hosted in the shales of the TSF were analyzed for their carbon and oxygen isotope compositions using the laser spectroscopy technique. The $\delta^{18}O$ (VPDB) values of the analyzed samples range from -7.04 to 3.72, whereas $\delta^{13}C$ (VPDB) values typically range from -6.07 to -9.4 per mil. The observed $\delta^{18}O$ values are typical of carbonates precipitated from $^{18}O$ depleted marine pore waters, suggesting that precipitation of the siderite took place under reducing conditions. The C and O plots place the nodules in the boundary between marine and continental environments, with a larger percentage being marine. The textural fabric of the siderite is marked by separate episodes of calcite veining which reflect the different stages in the diagenetic history of the nodules and the TSF concurrently. The study therefore suggests that the siderite nodules formed during early stages of diagenesis in a marine to near shore environment followed by encroachment of meteoric water. Thus, marine siderite precipitation ceased and calcite became the more stable phase occurring as late forming veins in the siderite nodules, evidenced in the textural fabrics.

**Keywords:** siderite, stable isotope, geochemistry, depositional environment, Takoradi Shale Formation, Sekondian Group.

## Introducton

The Takoradi Shale Formation (TSF) is one of the formations of the Ordovician to early Cretaceous sedimentary rocks of the Sekondian Group. It is well exposed along the coasts of the Western and Central regions of Ghana (Figure 1). These formations have over the years been studied by many researchers (e.g. Atta-Peters and Annan-Yorke, 2003; Asiedu *et al.,* 2005; Asiedu *et al.,* 2010), resulting in diverse ideas and conclusions about its depositional history. Previous studies on the TSF have utilized micropaleontological and bulk rock geochemical data of the shale to infer its depositional environment and paleoclimate (Atta-Peters *et al., 2003,* Asiedu *et al., 2010*). These studies suggest a marginal marine depositional environment for the formation. Recent sedimentological studies on the TSF, however, suggest that deposition took place in a deltaic environment (Agbetsoamedo, 2014).

Geochemical parameters have been used by various authors to understand the paleo-oxygenation and depositional conditions of ancient sediments (e.g. Jones and Manning, 1994). Noteworthy amongst these geochemical parameters are the stable isotope compositions of siderite and pyrite mineralization in sediments and sedimentary rocks. Reducing environments tend to favor siderite precipitation (Berner, 1971), thus it is a useful indicator of oxidation-reduction conditions in sediments. Furthermore, it is reasonable to assume that siderite, as a common early diagenetic mineral, is more likely to be influenced by the chemistry of the depositional setting than later minerals that form from more evolved pore waters. Hence, Mozley (1989) demonstrated a strong relation between the geochemistry (including isotopic composition) of early diagenetic siderite and depositional environment. Because siderite is often one of the earliest minerals to precipitate in sediments, there has been considerable interest over the years in using its geochemistry as an indicator of its depositional environment.

The isotopic data should reflect the diagenetic environment of mineral precipitation and this can be based on the present day understanding of the formation of siderite. In places where petrographic evidence indicates early siderite formation, the geochemical data may be used to determine the environment of sediment deposition (i.e. whether it is marine, brackish or fresh water).

In this study, the depositional environment of the TSF has been inferred from the carbon and oxygen isotopic compositions of the siderite nodules present in the shale part of the TSF. The stable isotopic ($\delta^{13}C$ and $\delta^{18}O$) compositions of the diagenetic minerals, siderite and pyrite, which are common mineral constituents in the TSF, have also been used to interpret the evolution of the pore water within the sediments.

### Geology of the Study Area

On the basis of lithology and stratigraphy, Mensah (1973) proposed seven formations for the Sekondian Group. This Group is about 1200m thick and consists of a predominantly fine-grained basal unit, Ajua Shale, overlain by six predominantly arenaceous lithologic units (Figure 2). The rock types vary in proportions within each formation, but in all sandstones dominate (about 55%), followed by shale/siltstone (40%) and conglomerate (5%). The Sekondian Group rests unconformably on the Paleoproterozoic Birimian Supergroup. The rocks are extensively faulted but unmetamorphosed (Mensah, 1973).

The Takoradi Shale Formation is sometimes classified with the underlying sandstone formation as Takoradi Beds. The Formation is about 200 m thick and it is typically composed of hard, compact, black or dark gray fissile shale or sandy shale, rich in organic matter. The base of the Takoradi Shale Formation is composed of thin bedded micaceous sandstones which intercalate with grey shales. The upper part of the succession shows interbedded grey-green grit bands and nodules of fine, grey lenticular sandstones up to several meters thick. Also included are siderite nodules and gypsum which occurs as thin veins traversing the shale and trends parallel to the bedding planes (Atta-Peters et al., 2003).

### Siderite Occurrence in the Field

Along the coast of Essipon, the TSF is characterized by large discoidal nodules of compact, finely granular siderite or clay ironstone. The nodules range in diameter from 10cm to 40cm and occur along the bedding planes in the fissile carbonaceous shale portions of the formation (Figure 3). Oxidation of the iron content in the nodules results in reddish colorations on the surfaces of the nodules. Pyrite also occurs as microscopic disseminations in the upper parts of the exposures. Paleontological studies have shown that the TSF are characterized by terrestrial miospores and marine acritarchs. Interpretations based on these microfossils have indicated Late Devonian/Early Carboniferous age for the formation which is deposited in a near shore-coastal (deltaic) environment (Atta-Peters et al., 2003).
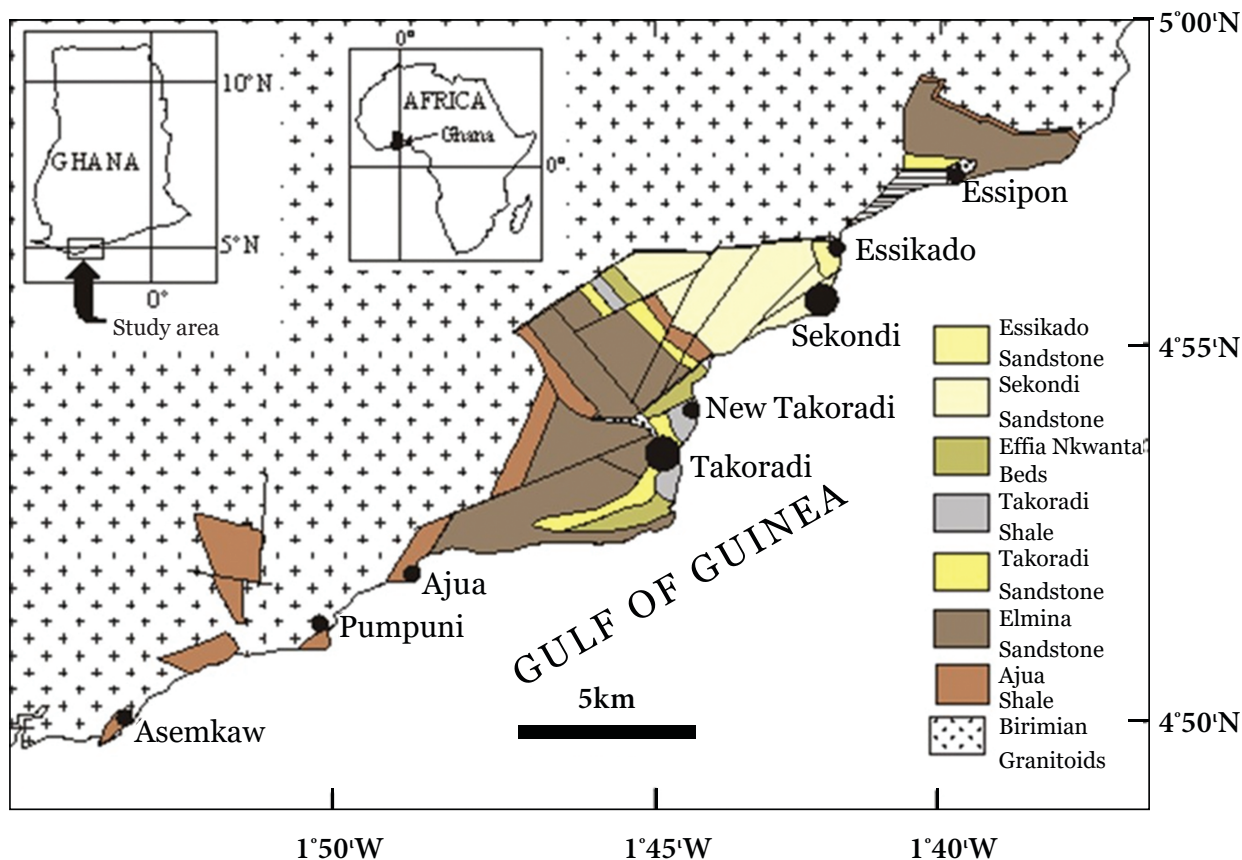
Fig. 1: The simplified geological map of the Sekondian Group in the Sekondi–Takoradi area, Ghana (after Asiedu *et al., 2005*; inset showing the map of Ghana and Africa).

## Methods

A total of 30 samples of the siderite nodules were collected from the TSF outcrops along the Essipon beach, Takoradi. Twenty (20) of these were prepared (cleaned and cut) at the laboratory of the Department of Earth Science, University of Ghana. Samples were collected for isotopic measurement and analysis in a commercial laboratory (ALS minerals, Canada). Laser Spectroscopy was adopted in the analysis for the C and O isotopic composition of the siderite nodules. A prepared sample (0.035g) of the siderite nodule was reacted with concentrated phosphoric acid at 72°C for 1 hour. The $CO_2$ gas produced was then measured for carbon and oxygen isotopes by off-axis integrated cavity output laser

spectroscopy. Results for $\delta^{13}C$ were calculated relative to Vienna Pee Dee Belemnite (VPDB) and results for $\delta^{18}O$ were calculated relative to Vienna Standard Mean Ocean Water (VSMOW). Analytical errors were 0.3‰ for $\delta^{13}C$ and 0.5‰ for $\delta^{18}O$. Fractionation factors were used for calcite ($CaCO_3$) where necessary. A total of 30 samples, 20 samples of the siderite nodules and 10 samples of the siderite lenses/bands) were prepared for petrographic analysis aimed at determining their mineralogical composition. This was carried out in the laboratory of the Department of Earth Science in University of Ghana, Legon. Mineral identification and modal percentages of minerals under thin section were estimated by the use of the petrographic microscope and visual estimation charts. Photomicrographs were then

taken for each thin section examined. All petrographic observations made were employed for determining or inferring the diagenetic history of the siderite nodules.

## Results

### Siderite petrography and Stable Isotope Composition

Thin section analysis revealed that the siderite nodules contained about 90-99% siderite oolitic grains with minor occurrences of late diagenetic calcite veins, rhodochrosite disseminations and silica grains. The veins vary in size from about 0.1mm to 2mm (Figure 4). The mode of formation of siderites is often characterized by peculiar textural variations. The siderite nodules in the TSF occur as an assemblage of well-sorted, silt-sized grains (up to 0.1 mm) in about 90% of the sample slides, to zones of predominantly sideritic nodules, lenses and bands up to 5 cm thick. The concentration of the siderite in the nodules ranges from sediments with about 60 volume percent (in a few instances) and can reach 95 volume percent of the nodule (in many instances). The samples that showed less than 60% siderite composition were typical of the lenses and bands occurring in the TSF.
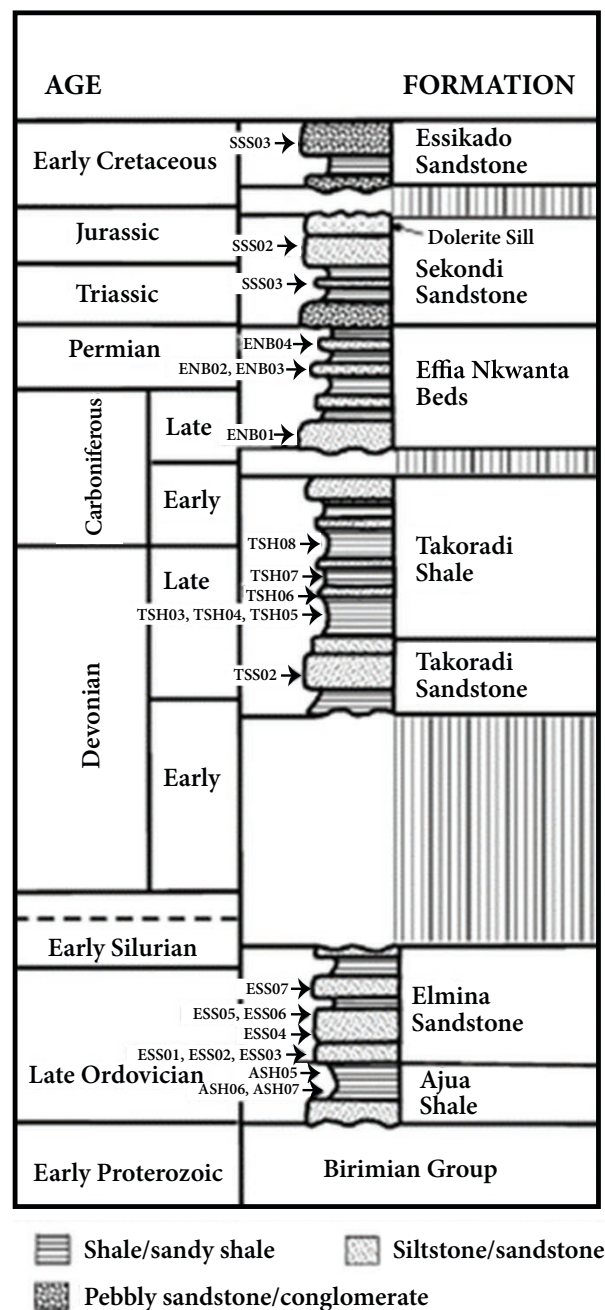


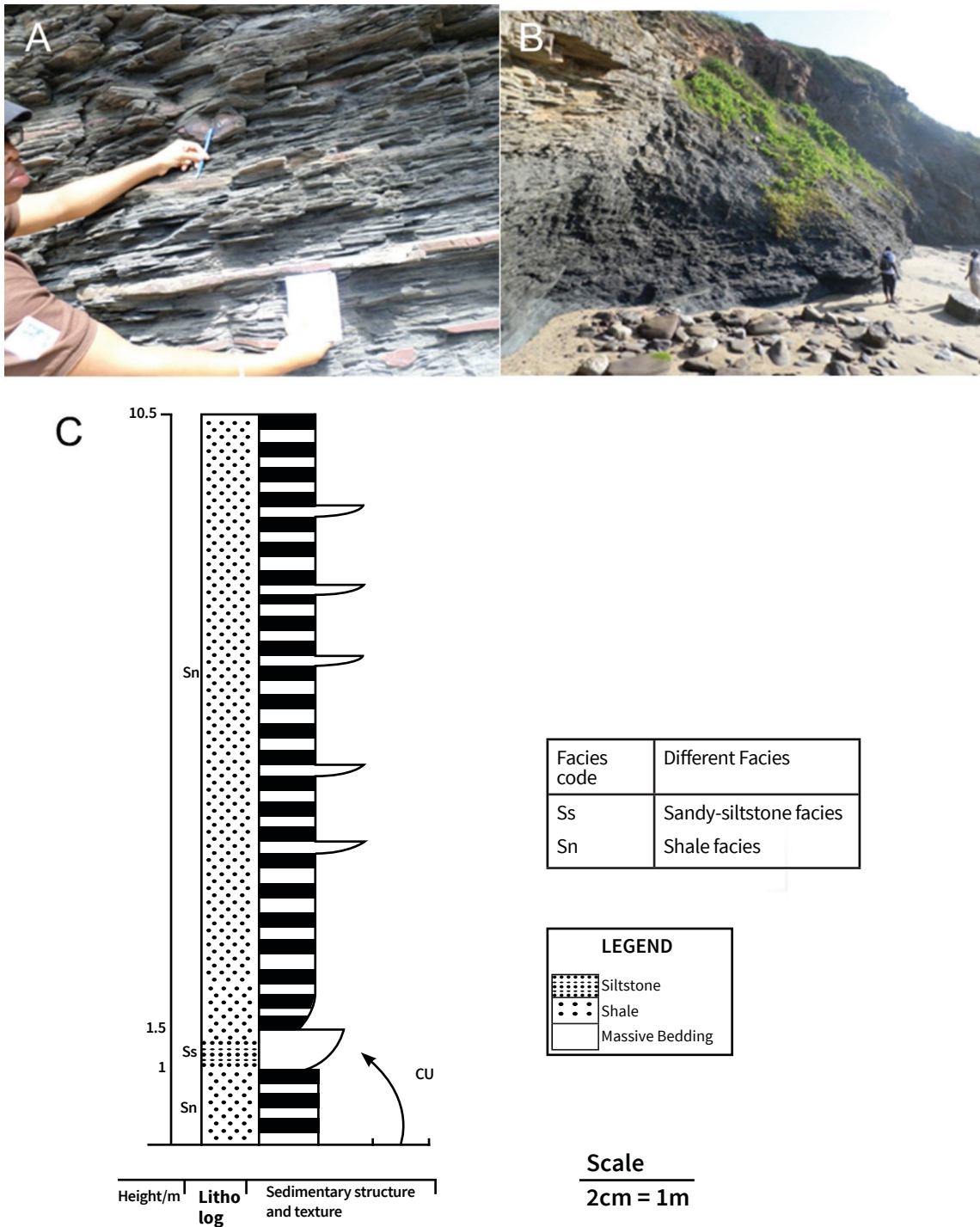Fig. 2: Stratigraphic column of the Sekondian Group (after Asiedu *et al., 2005*)

Fig. 3: (A) Siderite nodules embedded in the dark grey carbonaceous shale of the TSF (B) Vertical section of the TSF showing the dark grey, siderite nodule-bearing carbonaceous shale (C) Lithostratigraphic log of the dark grey carbonaceous section of the TSF.
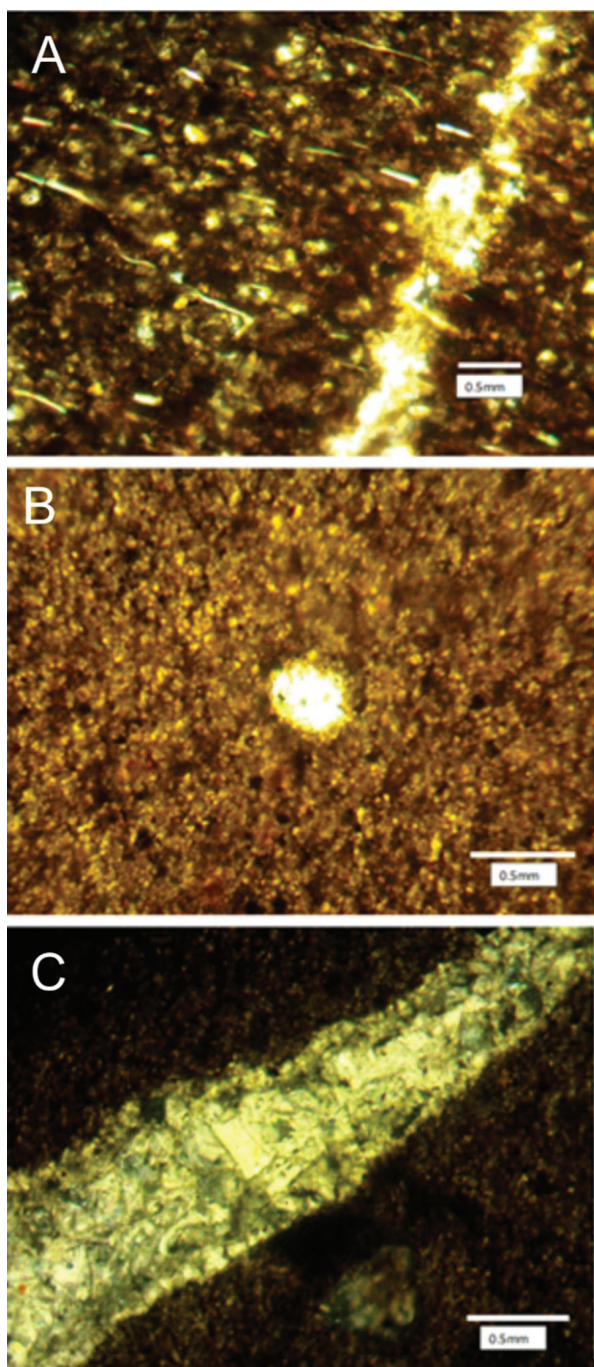
Fig. 4: Photomicrograph of siderite showing A) Silica and Rhodochrosite dissemination in siderite B) Fine grained oolitic siderite, about 98% composition C) About 1mm thick calcite vein.

The $\delta^{18}O$ values range from 23.6 to 34.7 per mil (VSMOW) or -7.04 to 3.72 (VPDB) while those of the $\delta^{13}C$ range from -6.07 to -9.4 per mil (VPDB) with two exceptions of -14.75 and -11.9 per mil (Table 1).

**Table 1. Carbon and oxygen isotopic compositions of siderite nodules from the TSF**

| SAMPLE DESCRIPTION | Delta $^{18}O$ (VSMOW) Per mil | Delta $^{13}C$ (VPDB) Per mil | Delta $^{18}O$ (VPDB) Per mil |
|---|---|---|---|
| SN001 | 27 | -8.87 | -3.75 |
| SN002 | 28.3 | -6.07 | -2.48 |
| SN003 | 30.1 | -6.26 | -0.74 |
| SN004 | 27.2 | -8.52 | -3.55 |
| SN005 | 32.5 | -14.75 | 1.58 |
| SN006 | 30.4 | -7.44 | -0.45 |
| SN007 | 34.7 | -9.4 | 3.72 |
| SN008 | 30 | -8.72 | -0.83 |
| SN009 | 27.5 | -8.18 | -3.26 |
| SN010 | 25.9 | -7.39 | -4.81 |
| SN011 | 24.7 | -5.01 | -5.97 |
| SN012 | 26.6 | -8.79 | -4.13 |
| SN013 | 25.4 | -8.09 | -5.30 |
| SN014 | 24.2 | -6.75 | -6.46 |
| SN015 | | | -29.94 |
| SN016 | 26.5 | -7.26 | -4.23 |
| SN017 | 29.8 | -11.9 | -1.03 |
| SN018 | 26.8 | -8.64 | -3.94 |
| SN019 | 23.6 | -8.17 | -7.04 |
| SN020 | 24.8 | -6.99 | -5.88 |

Pee Dee Belemnite (PDB) = 0.970065smow - 29-94 (Friedman and O'Neil, 1977).

## Discussion

### Petrography

The siderite can be said to be authigenic or diagenetic, which suggests precipitation from no calcite precursor, (Baker, 1996; Chamley, 1989), considering their compositional make up and field relations with respect to the host rock (where they tend to form within the beds/laminae). The TSF is carbonaceous and thus rich in organic matter, hence the bacterial reduction of organic matter in the sediment most likely produced the required carbonate for siderite precipitation. Bacterial reduction of organic matter, which often takes place during shallow burial, generates water, carbon dioxide, and biogenic methane. This increases the pH of the pore fluids, creating a more alkaline fluid which can lead to precipitation of carbonates. Carbonate precipitation of this kind often occurs patchily as concretions which develop intermittently along bedding surfaces. In addition, the iron (high Fe/Ca ratio) necessary to produce the siderite could have been adsorbed on pelagic clays in the sediments (Basov *et al.,* 1979; Matsumoto, 1989). The siderite nodules are also characterized by calcite ($CaCO_3$) veins, rhodochrosite ($MnCO_3$) disseminations and some silica ($SiO_2$). These are late stage diagenetic minerals which formed after the precipitation of the siderites. Pyrite is another mineral observed in the TSF which indicates that the sediments have passed through a stage of sulphate reduction.

### Carbon and Oxygen Compositions

The carbon and oxygen isotopic compositions (Table 1) of the TSF siderite nodules were plotted on a discriminatory diagram to help determine the possible environment of deposition. The $^{13}C$ and $^{16}O$ values for the siderite nodules show depleted values with somewhat minimal variations that plot closely to the marginal marine environment (Figure 5). The small variation in $\delta^{13}C$ values is attributable to the less extensive bacterially facilitated reactions. These reactions involve the decomposition of organic matter which is known to control pore water $\delta^{13}C$ compositions in organic rich sediments. The negative carbon values

for the diagenetic siderite can be said to be produced during ferric iron reduction (suboxic conditions) and during or immediately after sulfate reduction in a marginal marine environment (which has a mixture of marine and non-marine pore waters) as suggested by some earlier researchers (Atta-Peters and Annan-Yorke, 2003). The oxidation of methane near the base of the zone of sulfate reduction could also contribute to the negative carbon values (Raiswell and Berner, 1987). The negative values also rule out any possibility that the carbon content of the TSF siderite nodules could have been formed from the input of $^{13}C$-enriched bicarbonate derived from methanogens, $^{13}C$-enriched volcanic gas entering the water body, bicarbonate from dissolution of biogenic $^{13}C$-enriched calcite, or the high productivity of $^{12}C$-enriched planktonic algae, which would have resulted in $^{13}C$ enrichment of the residual carbon pool (Fritz *et. al.,* 1971).

The $\delta^{18}O$ values for the siderite nodules are generally negative (depleted) and show minimal variation. The depleted oxygen values are characteristic of marginal marine originated pore water, and the possible explanations for this depletion are recrystallization, precipitation at anomalously high temperatures, water-sediment interaction, mixing with meteoric water and variation in seawater composition. The most likely explanation for this instance would be either mixing of meteoric water and marine waters in marine shelf sediments or mineral water interaction during early diagenesis (i.e. precipitation of $^{18}O$ enriched minerals in the nodules or host rock which would cause a net depletion of $\delta^{18}O$ in the pore water). However, the latter can be ruled out of the possible cause, considering the fact that the minerals in the TSF (host rock) are $^{18}O$ depleted, and thus should rather have made the siderite nodules $^{18}O$ enriched.

The depleted oxygen values of the TSF siderite nodules also place them in an anoxic or euxinic environment. Euxinic environments are those in which the circulation of water is restricted, leading to highly depleted oxygen levels or anaerobic conditions in the water.
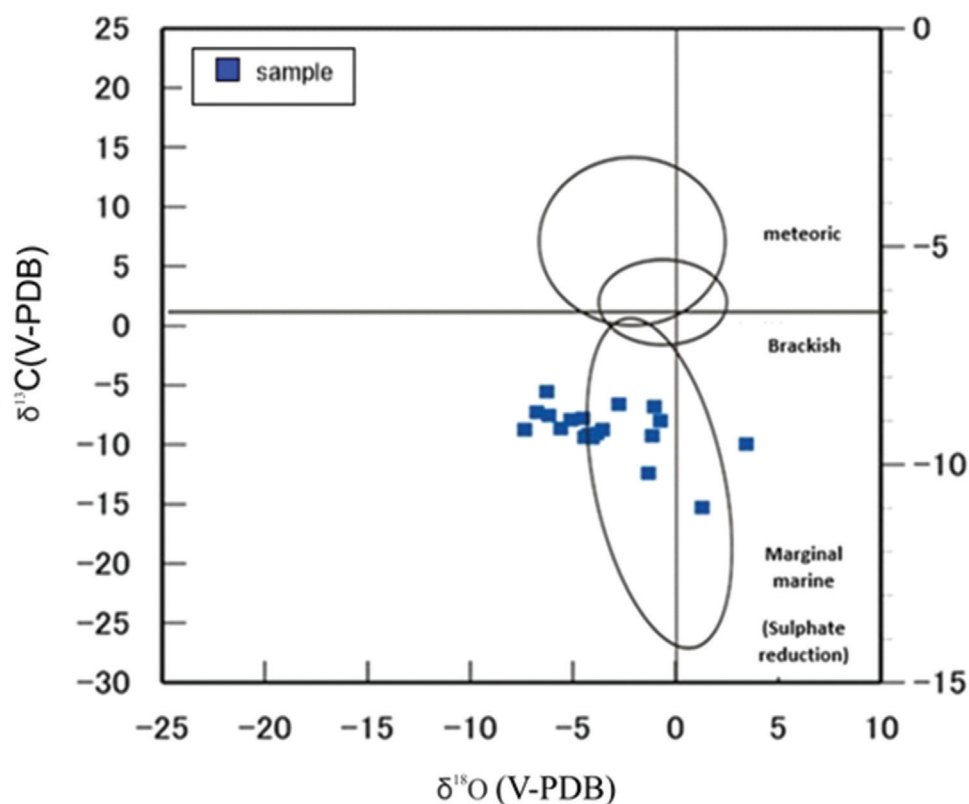
Fig. 5: Plot of δ¹³C against δ¹⁸O values for siderite nodules from the TSF, modified after Mozley and Wersin (1992).

Euxinic also implies anoxic conditions with the presence of excess sulphur that binds with hydrogen to form hydrogen sulfide ($H_2S$), and this is evidenced in the siderite nodules and the TSF. Such conditions may also develop in swamps, barred basins, stratified lakes, and fiords. Euxinic sediments are those deposited in such conditions, and are usually black and organic-rich.

Both the oxygen and carbon isotopic compositions of the siderites in the TSF were used in predicting the environments of deposition of the TSF. Zymela (1996) used assumptions to conclude that at relatively normal temperatures (typical of tropical climates today) the **δ¹⁸O** values of early precipitated carbonate would be near 0 to -2 °/$_{oo}$ (PDB). Mineral precipitation in freshwaters would result in more negative values due to the lighter isotopic composition of meteoric water. In brackish water, which represents mixed waters of the two environments mentioned above, early formed carbonates should

display transitional values between the two end members (Baker, 1996). Late stage marine siderites which formed at greater depths may exhibit more negative δ¹⁸O values. This is a reflection of increasing temperatures with depth of burial. Therefore, zoned concretions which formed over a wide depth range should reveal increasingly lighter δ¹⁸O from the core to the edge (assuming that the system remained more or less closed). Such decreasing oxygen isotope trends in various concretions have been observed by several authors (Timofeyeva *et al., 1976*; Irwin, 1980; Hennessy and Knauth, 1985; Mozley and Carothers, 1992).

A compilation of both C and O isotopic compositions plot the siderite nodules in the transition between marine and continental environments (near shore to brackish environments). However, a larger percentage of the samples exhibit higher marine signatures than continental when compared to the plots of continental

and marine isotopes by Mozley and Wersin (1992). They showed that marine and non-marine siderites are characterized by distinctive compositional trends. Marine siderites generally plot in a field ranging from very low $\delta^{13}C$ values and relatively high $\delta^{18}O$ values to relatively high $\delta^{13}C$ values and low $\delta^{18}O$ values. Non-marine siderites generally have higher $\delta^{13}C$ values than marine siderites. This feature is evident in the samples of the siderite nodules as well. They show low $^{13}C$ values and very low $^{18}O$ values, and this thus places the siderite nodule samples in a marginal marine environment, evidenced by the plots.

## Conclusion

Results from the isotopic analysis of siderite nodules in the Takoradi shale formation (TSF) suggests a mixture of terrestrial and marine pore fluids which resulted in the precipitation of the siderites. The oxygen and carbon isotopic data for siderites from the TSF showed minimal variation in compositional fields. The $\delta^{18}O$ values range from 23.6 to 34.7 per mil (VSMOW) or -7.04 to 3.72 (VPDB) while those of the $\delta^{13}C$ range from -6.07 to -9.4 per mil (VPDB) with two exceptions of -14.75 and -11.9 per mil. The siderites have $\delta^{18}O$ values characteristic of carbonates precipitated from $^{18}O$ depleted marine pore waters. The depleted isotopic values plot the siderite nodules in the marginal marine environment. It can also be concluded that the siderites formed during the early stages of diagenesis in a subaqueous environment, which is evident in the mineralogical and textural features of the siderite nodules. The high siderite composition (60-90%), with very minimal late stage occurrences of calcite veins and silica disseminations, suggests very minimal alteration of iron bearing minerals in an environment with reducing conditions. It is suggested that the TSF was deposited in a marginal marine (to brackish) environment, under early diagenetic processes.

## References

Agbetsoamedo, J. E. (2014). Depositional environment of the Takoradi Shale Formation, south western Ghana. (Unpublished Master's thesis). University of Ghana, Accra, Ghana.

Asiedu, D. K., Hegner, E., Rocholl, A. & Atta-Peters, D. (2005). Provenance of Late Ordovician to Early Cretaceous sedimentary rocks from southern Ghana, as inferred from Nd isotopes and trace elements. *Journal of African Earth Sciences*, 41, 316–328.

Asiedu, D. K., Atta-Peters, D., Hegner, E., Rocholl, A & Shibata, T. (2010). Palaeoclimatic control on the composition of Palaeozoic shales from southern Ghana, West Africa. *Ghana Mining Journal*, 12, 7–16.

Atta-Peters, D. (2000). Early Cretaceous miospores from the Essikado Sandstone Formation of the Sekondian Group of Ghana. *Revista Espanola de Micropaleontologia*, 32, 245–257.

Atta-Peters, D. & Anan-Yorke, R. (2003). Latest Devonian and Early Carboniferous pteridophytic spores from the Sekondi Group of Ghana. *Revista Espanola de Micropaleontologia*, 35, 9-27.

Baker, J. C., Kassan, J. & Hamilton, E. J. (1996). Early diagenetic siderite as an indicator of depositional environment in the Triassic Rewan Group, southern Bowen Basin, eastern Australia. *Sedimentology*, 43, 77-88.

Basov, V. A., Lopatin, B. G., Gramberg, I. S., Danjushevskaya, A. I., Kaban'kov, V. Y., Lazurkin, V. M. & Patrunov, D. K. (1979). Lower Cretaceous lithostratigraphy near Galicia Bank. *Initial Reports of the Deep Sea Drilling Project*, 47, 683-717.

Berner, R. A. (1971). *Principles of sedimentology*. New York: McGraw-Hill.

Chamley, H. (1989). *Clay Sedimentology*. Berlin: Springer.

Claypool, G. E., Holser, W. T., Kaplan, I. R., Sakai, H. & Zak, I. (1980). The age curves of sulphur and oxygen isotopes in marine sulphate and their mutual interpretation. *Chemical Geology*, 28, 199-260.

Friedman, I., & O'Neil, J. R. (1977). *Data of geochemistry: Compilation of stable isotope fractionation factors of geochemical interest* (Vol. 440). Washington, D.C.: US Government Printing Office.

Fritz, P., Binda, P.L., Folinsbee, F.E., & Krouse, KR. (1971). Isotopic composition of diagenetic siderites from Cretaceous sediments in Western Canada. *Journal of Sedimentary Petrology*, 41, 282-288.

Hennessy, J., & Knauth, L. P. (1985). Isotope variations in dolomite concretions from the Monterey Formation, California. *Journal of Sedimentary Petrology*, 55, 120- 130.

Irwin, H. (1980). Early diagenetic carbonate precipitation and pore fluid migration in the Kimmeridge Clay of Dorset, England. *Sedimentology*, 27, 577-591.

Jones, B., & Manning, D. A. (1994). Comparison of geochemical indices used for the interpretation of palaeoredox conditions in ancient mudstones. *Chemical Geology*, 111, 111-129.

Matsumoto, R. (1989). Isotopically heavy oxygen-containing siderite derived from the decomposition of methane hydrate. *Geology*, 17(8), 707-710.

Mensah, M. (1973). On the question of the age of the Sekondi Series, Upper Devonian or Lower Carboniferous rocks of Ghana. *Ghana Journal of Science*, 13, 134–139.

Mozley, P. S. (1989). Relation between depositional environment and the elemental composition of early diagenetic siderite. *Geology*, 17, 704-706.

Mozley, P. S., & Carothers, W. W. (1992). Elemental and isotopic composition of siderite in the Kuparuk Formation, Alaska: Effect of microbial activity and water/sediment interaction on early pore-water chemistry. *Journal of Sedimentary Research*, 62, 681-692.

Mozley, P. S., & Wersin, P. (1992). Isotopic composition of siderite as an indicator of depositional environment. *Geology*, 20, 817-820.

Raiswell, R., & Berner, R. A. (1987). Organic carbon losses during burial and thermal maturation of normal marine shales. *Geology*, 15, 853-856.

Timofeyeva, Z.V., Kuznetsova, L.D. & Dontsova, Y.I. (1976). Oxygen isotopes and siderite formation. *Geochemistry International*, 13, 101-112.

Zymela, S. (1996). Carbon, oxygen and strontium isotopic composition of diagenetic calcite and siderite from the Upper Cretaceous Cardium formation of western Alberta. (Unpublished doctoral dissertation). McMaster University, Ontario, Canada.

# Redefining the Concept of Big Data: A Ghanaian Perspective

Eleanor Afful[1], Kofi Sarpong Adu-Manu[2], Grace Gyamfuah Yamoah[2], Jamal-Deen Abdulai[2], Nana Kwame Gyamfi[2], Edem Adjei[2], Isaac Wiafe[2] and Ferdinand Apietu Katsriku[2*]

[1]KACE AITI
[2]Department of Computer Science, University of Ghana
*Corresponding author: fkatsriku@gmail.com

## ABSTRACT

The world is on the verge of a data tsunami. Voluminous amounts of unstructured data are being generated using different technologies. To manage the huge amounts of data being generated, a new concept of 'Big Data' has evolved. The emergence of 'Big Data' is leading to real transformation in the business world. Governments and commercial enterprises on the African continent are beginning to take an interest in the use of technologies associated with Big Data for the analysis of enormous amount of data they currently generate and they wish to do so in real time. The advances being made in big data technologies have fuelled this uptake. Until recently companies in Ghana did not realize the utility of big data analytics due mainly to lack of knowledge and the limited penetration of these technologies. Increasingly, however, these companies now realize the difference in value that data analytics could make to their decision making process and to develop strategies that will give them competitive advantage. It has become clear to many of these corporate organizations that they are in possession of large volumes of data which, if properly analysed, can provide them with a wealth of knowledge to run their businesses more efficiently and productively. The analytic necessary to the understanding of these wealth of data are provided by big data technologies. This paper seeks to redefine the concept of big data and reviews its development, the potential impact that big that can have on a developing economy, the sectors of the economy of Ghana that stand to gain most from adoption of big data technologies and how these can be achieved. We propose that big data concept be defined more objectively by the use of a function. The paper shows how big data can be leveraged for rapid economic advancement. The paper additionally examines the investment prospects of adopting big data technologies for the economic environment of Ghana and some of the issues that organizations must resolve to successfully implement the technologies in Ghana.

**Keywords:** Big Data, Analytics, Economic development, Big Data Architectural, Ghana

## Introduction

Traditional data modelling and organization methods have proved useful and appropriate for varied functions in the past few decades and this is attested to by the phenomenal success of relational database systems. These traditional methods are coming under huge strain with the exponential growth in data and in most cases the traditional systems are unable to cope. This growth in data has been fuelled in particular by the success of internet companies such as Google and Facebook.

Huge volumes of information are still available and unexploited because the existing data modelling and management tools are not well suited to handling such information. Data aggregation, transforming data scattered across multiple sources into a new summary, is one of the key features used in databases, especially for business intelligence (e.g. extract, transfer and load (ETL), online analytical processing (OLAP) and analytics /data mining).

For databases built on Structured Query Language (SQL), aggregation is used to prepare and envision data for a more profound level of analysis. Such an operation is however difficult and often impossible to perform on

enormous volumes of data in terms of the memory and time requirements (memory-and-time-consumption).

Database maintenance and optimization is a key activity for relational databases. As the number of queries from across multiple sources increases, optimizing query execution becomes difficult to handle. For databases bigger than relational ones, a key requirement is that they be maintained and optimized for continuous optimal performance; such a task thus becomes less than trivial.

Additionally, the data residing in the database must be highly structured and cleansed. Businesses spend significant effort to extract, transform and load the data between data warehouse and relational databases. Enormous costs are involved in doing these and greatly limits the breadth of data available for analysis. The current systems are not easily scalable and do not scale up to the combined increase in velocity, volume and variety as defined for big data.

This paper proposes a more objective definition of the concept of big data, looks briefly at its development and what impact these technology can have on a developing economy. The paper examines those sectors of the Ghanaian economy which could possibly benefit the most from application of big data technologies and how the aforementioned benefits might be achieved, based on new architectures.

## Background

Initially Big data is normally defined using the three V's, Volume, Velocity and Variety. However of late, two other parameters have been included: Veracity and Value. The big data concept may then be depicted as shown in Figure 1. Variety is assured through the numerous and diverse data sources, each generating some quantity of data per unit time to the data volume. The amount of data generated per unit time may not be static but dynamic and subject to change over time. The data being generated from these sources may either be structured or unstructured. Volume is the summation of all the data coming from diverse sources per unit time and arriving

at one processing centre. In the literature volume has been defined as how much data there are, velocity, the rate at which new data are created and how quickly the data are processed and variety is defined in terms of the format of the data, whether structured, semi-structured or unstructured.

The two new dimensions are veracity, used to refer to the trustworthiness of the data and value, which refers to what gain businesses' can derive from the data have been added lately. Other dimensions have also been used, notably volatility and validity; however these have not gained widespread acceptance and use. It is noteworthy that these definitions only provide a qualitative view of what is described as big data. Some researchers have sought to define 'big data' in terms of a fixed volume such as petabytes or zettabytes. There is however no consensus on exactly what quantity of volume would constitute big data. Velocity may be defined as the rate of change in the volume of data generated and transferred to the enterprise office. Value is derived from the analytics performed on the data. Ultimately, the volume and velocity are intimately linked to the processing capacity of the system under consideration and hence the business needs. Assigning a numerical value to what will qualify in terms of volume as big data is thus not very useful. What may qualify as big for one enterprise may not be so for another enterprise. A helpful definition will be "when the data arriving begins to exceed the processing capacity of the conventional database and data warehouse solutions available to an enterprise". For many businesses therefore, big data becomes a moving target as they need to constantly evolve new solutions for the data they process. The volume of data will depend on the rate at which new data are being generated and the rate at which there are arriving. As such volume and velocity are intricately linked. What has not been mentioned as far as velocity is concerned is the rate at which the data are being processed. This constitutes another aspect of velocity not intricately linked to volume. Even though it may be argued that the processing rate affects the volume of data yet to be processed, it does not affect the total volume of data an organization has.

Fig. 1: Big data concept and interaction between the v's

$$velocity = \frac{dv}{dt}$$

$$Volume = \sum_{1}^{N} sources$$
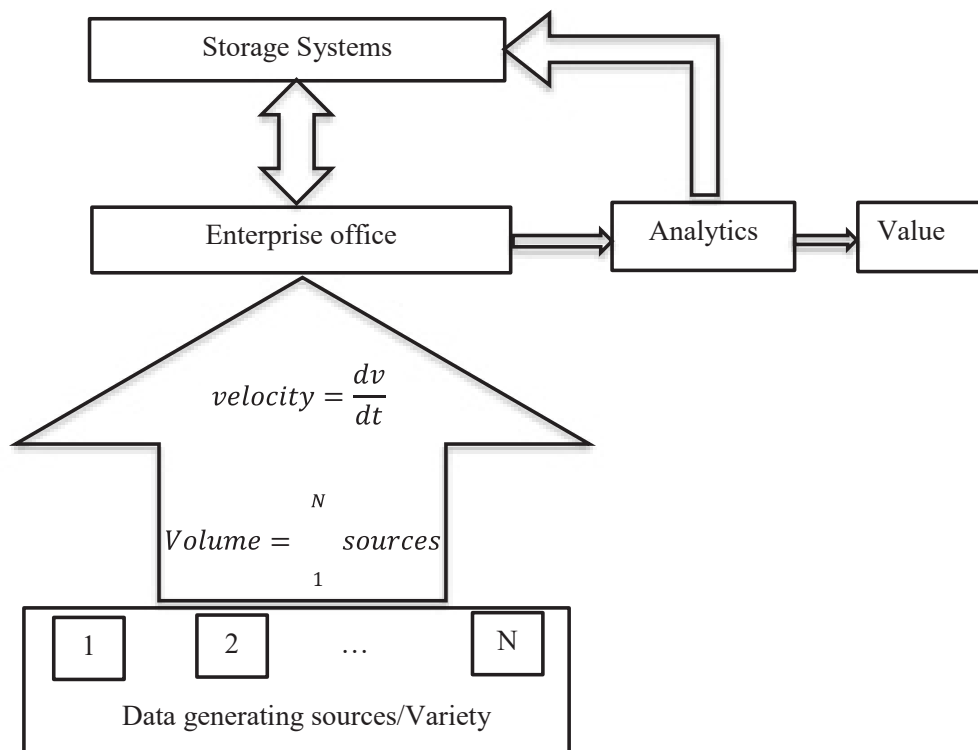
| 1 | 2 | … | N |

Data generating sources/Variety

Fig. 1: Big data concept and interaction between the v's

Veracity of data is also not a new concept. Businesses have always sought to ensure the quality of data they obtain. What is new is the fact that with big data, a substantial proportion of the data is from external sources and hence additional measures now need to be taken to ensure the quality of the data. This, coupled with the wide variety of data, necessitates new requirements. Data cleansing thus becomes an additional and important element of the processing cost.

Creating business value out of data has always been a driving motivation for many businesses, so this is not a new concept. What provides a business with a competitive advantage is what insights it can derive from its data warehouses to enable it make better, real time, and smart decisions. This implies a real need for detailed and in-depth knowledge and hence the demand for new analytical tools.

The key new dimension now is therefore variety. Businesses have always been interested in data for the value that may be derived from insights such data provide on analysis. When the data grow beyond their current processing capability, they simply upgrade their systems to more powerful ones and in most instances the same relational database technology. As such, growing data volumes has always been an issue that businesses have to contend with. The rate at which new data are being generated and arriving and the variety of formats the data take are hence the key issues. Whilst in the past businesses had to deal with data of the same type and form, the explosion in new data sources mean that data are constantly being generated from new sources and taking different forms. Businesses now have to integrate different forms of data; unstructured, graph data, voice, images, video, etc. What is the relation between the value we derive from big data and the other dimensions namely volume, velocity and variety? It is proposed that a relation be established between the value derived from big data and the other dimensions, namely volume, velocity and variety. In our opinion, it would be helpful to

view big data and the value obtained from it as a function dependent on the parameters of volume, velocity and variety. Such a relation may be expressed as a function dependent on the parameters of volume, velocity and variety.

Given some of these challenges, it is necessary to develop new technological means of managing, analysing, visualizing and extracting meaningful information from the large and complex heterogeneous data sets that are being generated from diverse and distributed sources. Progress in this area will enhance scientific innovation and provide new paths to scientific inquiry. The development of new data analytic tools and algorithms will also be an important outcome of any progress made. Other benefits will include the development of scalable data infrastructure and architecture which will ultimately lead to a better understanding of social processes and interactions for greater security, economic growth, and in general, an improved quality of life and the wellbeing of people.

To efficiently model the data requirements for business intelligence and analytics, a new technology has emerged, called NoSQL (Non-Structured Query Language), a distributed non-relational database with variations in implementation. NoSQL was designed to cater for demands of data that were being generated by the web. The origins of NoSQL can be traced to the work done at Google in building a proprietary database, Big Table. The Big Table was designed to overcome some of the inherent limitations of Relational Database Management Systems (RDBMS). Some of these limitations are the need for specialized and robust servers that were less prone to failure; the length of time required to process queries; and the need to have structured data. Since this early work, many companies have also turned their attention to developing such systems that promise low implementation cost in terms of the hardware requirements and the ability to massively scale up horizontally i.e. by adding thousands of nodes so that storage and retrieval would be distributed across them using parallel processing techniques, thereby reducing storage and retrieval time. All these allow the setting up of

server farms very quickly and cheaply. These systems also have a high level of fault tolerance since the same data are stored on different machines and they do not have any limitations on structure, which means one could store almost anything together.

With NoSQL it is possible to efficiently and cost-effectively build massive computing systems capable of handling the exponential growth in the volume of large data sets as is currently experienced. Independent of its format structured, unstructured or semi-structured the techniques underpinning NoSQL, ensure that the limitations imposed by RDBMS on data size, format and speed are eliminated, leading to fast and efficient ways of processing and analysing variety of data in real-time, bringing real benefits to businesses.

With regards to the velocity with which new data are being generated, NoSQL has the capability to process terabytes and exabytes amounts of data in real time. The new techniques implemented in NoSQL process, extract, load and transform data in the database eliminating the need for the data to be transferred in and out of the database of the data warehouse. The advances in processing and storage capabilities in computing technology in the last decade with increased speed has effectively eliminated data size as a constraint.

### Platforms and tools for big data

Even though arguments have been made to the effect that most of the data generated today are either unstructured or semi-structured, emerging big data technologies could be divided into two categories- structured and unstructured data. To handle structured big data, a number of customized technologies have emerged. These technologies are aimed at storing and retrieving the large amounts of data associated with big data. The Google File System (GFS) is an internet scale file system, a robust and scalable system which provides the sort of reliability required for internet applications. Object-store techniques aim to improve on redundancy and data availability. The Amazon Simple Storage System, OpenStack Swift and Nirvanix cloud storage are examples

of this approach. Underlying many of these solutions is the massive parallel processing (MPP) technique. MPP is based on a distributive processing architecture consisting of a series of nodes controlled by a master. When engaged the master distributes a query across the nodes for maximum processing efficiency. Similarly, the system can do autofast data import and export through the same underlying mechanism. Almost all the vendors operating in this domain use either software or hardware combined into a single compliance. This ensures consistency in the hardware and that is crucial to obtaining optimum performance. Data locality plays an important role in obtaining high performance in big data analytics. By processing the data as close as is possible to its generating source, we minimize the highly prohibitive costs of data transfers. MapReduce exploits the concept of data locality to give an improved performance. A variant of MapReduce is Hadoop. Hadoop is an open source implementation of MapReduce. It is based on the Hadoop Distributed File System (storage) with distributed processing architecture consisting of a series of nodes controlled by a master. When engaged the master distributes a query across the nodes for maximum processing efficiency. This programming paradigm allows for massive scalability across hundreds or thousands of servers in any Hadoop cluster of nodes much like Google core infrastructure, which requires different skills sets. Building analytic solutions requires knowledge of a new set of Application Programming Interfaces (API) and this is one major drawback. MapReduce is typically controlled by Java programming language, the term is used to refer to two separate and distinct jobs that Hadoop programs perform. In the first task the program does a mapping of input data and then processes it to produce key/value pairs. The reduce function takes those key/value pairs and then combines or aggregates them to produce the final results. The name, MapReduce gives a clue of the order in which the tasks are carried out, the reduce job is always performed after the mapping has been done. Combining the use of data warehousing, data mining and relational database alongside techniques such as optimization, simulation, visualization and predictive analysis for big data sets provide better strategies to obtain insight from massive data sets enabling better decision.

## *Implementation Areas of Big Data*

**Key/Value Store** is a fundamental data model used for example in Hadoop, Voldemort, DynamoDB and Memcached. Key-value databases are lightweight, schema-less, relationship-less and transaction-less data stores used primarily for storing temporary data in memory. Examples of such formats of key value database used for very large scale storage systems include Riak, Redis and MemcachedDB. The key can be synthetic or auto-generated while the value can be String, JavaScript Object Notation (JSON), BLOB (basic large object) etc. The key value type basically, uses a hash table in which there exists a unique key and a pointer to a particular item of data. There can be matching keys in different containers which are made up of logical group of keys. Performance is enhanced to a great degree because of the cache mechanisms that accompany the mappings. Key/Value pairs however fail to offer ACID (Atomicity, Consistency, Isolation, Durability) capability, as they fail on consistency. This capability must be provided for by the application itself. To read a value one needs to know both the key and the bucket because the real key is a hash (Bucket + Key). The Key Value Store database model is popular because it is easily implemented. A major weakness of this scheme is that it becomes increasingly difficult to maintain unique value keys as the volume of data grows. To address this challenge, complex schemes are introduced to generate unique character keys for very large sets.

## *Document Oriented Database*

The idea here is to aggregate the data, mainly in the form of key value pairs, this is then compressed into a searchable record format. XML (Extensible Markup Language), JSON (Java Script Object Notation) and BSON (which is a binary encoding of JSON objects) are some of the typical encoding schemes available. One significant distinction between a key-value store and a document store is that a document store has associated with it the attribute metadata related to the stored content. This provides a means of querying the data based on the stored content. Unlike traditional relational

databases where data and relationships are stored in tables, in this scheme they are simply a collection of documents independent of each other. Document style databases are schema-less and this makes a simple task of adding fields to JSON documents without having to first define the required changes. The most commonly used document-based databases are CouchDB, Apache and MongoDB. To store data CouchDB employs JSON with JavaScript as the querying language and MapReduce and Hypertext Transfer Protocol (HTTP) to implement the application programming interface.

### Column Family Database

A column-family database provides the capability to organize the rows as groups of columns. This capability implies that each single row of a column-family database now has the capacity to contain several columns. All the columns which are related are grouped together as column families providing the capability to retrieve the columnar data for multiple entities. This is achieved through an iterative process. The flexibility that column family provides applications enables a wide range of complex queries and data analyses to be performed. This is reminiscent of the functionalities supported by a relational database. This design enables them to store massive volumes of data running into billions of rows with each row containing hundreds and possibly thousands of columns. Significantly, a column family database can still provide very fast access to these vast quantities of data due mainly to a most efficient storage mechanism. If a column-family database is well-designed then it will be fundamentally faster and have greater scalability than an equivalent relational database holding the same volume of data. This performance is achieved at a cost, it can only support a specific set of queries unlike the queries in a relational database which are more generalized. Designers of column-family databases must ensure that column families are designed optimizing for the most commonly use queries for the applications under consideration. In contrast, majority of relational Database Management Systems (DBMS) store their data in rows. Storing data in columns as done in column families has the benefit of allowing fast search/access as well as providing for

efficient data aggregation. In relational databases a single row is stored as a continuous disk entry. As a result, different rows of data may be stored as different entries on the disk. On the other hand, columnar databases store all the cells which correspond to a particular column as a contiguous disk entry; this makes the search/access time much faster than can possibly be achieved in a relational database.

### Graph Database

The graph database is the final variant of NoSQL database management systems that is considered in this work. Unlike the other models, the graph based DBMS models represent the data based on tree-like structures and using edges to connect the various nodes such as is used in graphs. Just as in mathematics, certain operations are much simpler to perform using these types of models. These databases are commonly used by applications where it is necessary or required to establish boundaries for connections. For example when you register on a social network of any sort, your friends' connection to you and their friends' relation to you are much easier to work with using graph-based database management systems. An example is Neo4J, the most widely used graph store apart from RDF (Resource Description Framework) triple stores.

### Architectural (Conceptual) framework

In order for Ghana to leverage big data for economic development, a conceptual framework supporting the activities of all stakeholders (individuals, private and public sector) should be developed. This architectural or conceptual framework should take into consideration the role of companies or organizations, policy makers, institutions, and individual users towards the adoption of big data for economic development.

Several frameworks have been discussed in the literature (Manyika *et al.*, 2011; Wamba *et al.*, 2015; Global Pulse, 2012). We propose in this paper an architectural framework that can support the use and implementation of big data to boost the Ghanaian economy as depicted

in **Figure 2**. This framework seeks to point out the benefits of the use of big data in driving the economy of Ghana. From the framework, a collaboration between the public and private sectors in Ghana is a step towards an integrated economy and this can boost productivity significantly with the implementation of big data. Companies/industries in Ghana are expected to provide incentives to enhance the economy and also for users in the form of rewarding innovation. Big data analytics offer a huge economic impact for organizations (Gangadharan, 2014).

Policy makers produce and use data to facilitate enhanced policymaking processes, it is encumbered upon them to also play their part by promoting and fostering data-driven innovation and growth throughout economies (Andrade *et al.*, 2014). For big data to realize its potential in Ghana, innovation which is driven by advances in technology, policymakers need to articulate coherent guidelines, standards and policies on the use of data and the associated technologies. A possible way of achieving this objective is through openness and transparency; ensuring that using open data formats public data is accessible, promoting legislation which is balanced and which takes into consideration the competing needs of all sectors of the economy; and supporting education that focuses on equipping students with data science skills and competencies (Andrade *et al.*, 2014).
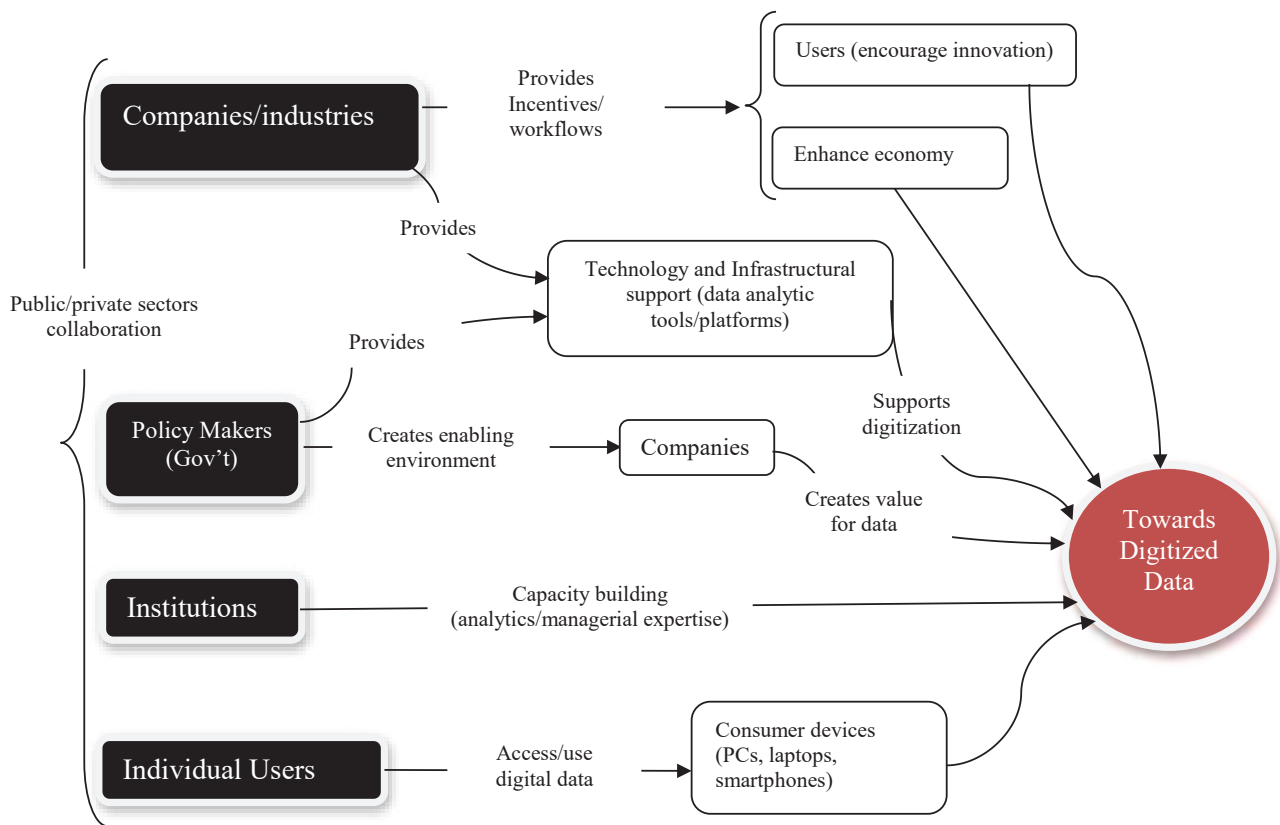


Fig. 2: Architectural framework for big data implementation in Ghana

Institutions also have a role to play in the progressive development and implementation of big data in Ghana. One of the major challenges is lack of expertise. Individuals with technical know-how in big data analytics are in short supply. According to Manyika *et al.*, (2011), a major limit on realizing value from big data is the acute shortage of skills and talent, especially of people with profound capability and proficiency in statistics and machine learning, and managers and analyst who know how to operate companies by using insights from big data. This places a huge responsibility on educational institutions and Information Technology training centres to build the right capacity for the nation to have a cream of talents ready to take up the task of analytics. This requires that the relevant departments be strengthened to enable them to fulfill those expectations. The role that Computer Science can play in transforming economies in the 21st century and beyond is well known and the arguments in support of this have been well made. This must however be driven from the highest echelons of government. A conscious effort must be made on the part of government to move towards an economy driven by advances in computing technology and hence the need to support computer science education in the country.

Global Pulse (2012) discussed the foremost apprehensions and challenges raised by big data for development and suggested probable ways of addressing some of them. They went ahead to discuss the sources of data for development in a growing environment such as Ghana. For a developing country like Ghana to benefit from the full potential of big data, these data sources should be taken into consideration. Most sectors of the economy in Ghana still depend to a great extent on paper-based record-keeping, as such, the data source is not automated and hence, easily digitized. The data sources can be digitally generated, passively produced, automatically collected, geographically accessible and continuously analysed (Global Pulse, 2012). These data sources are relevant for big data for economic development. Ghana generates massive amounts of digital data from different streams of the economy (online data) from different organizations and online platforms.

This is however nowhere near what is possible. In a big data environment, the data are expected to be generated or created digitally such that they could be processed or manipulated by an electronic device. The recent Police and Fire Service e-Recruitment drive is an example of how such migrations could be achieved (Ghana Police Service 2016 e-recruitment, Ghana Fire Service e-recruitment 2016). The data produced should have the ability to interact with other digital services. They should be collected automatically after they have been produced. The location or time span for one operation should be available/accessible and should be analysed in real-time with no difficulty. Until the country meets some or all the sources of big data for development, Ghana will not be able to leverage big data for economic development.

### Big Data and Ghana's Economy

The world has reached a stage where data are all around us. This data can be obtained from digital images, social media streams, financial and banking transaction records, wired and wireless sensors, GPS signals, and a myriad of other sources. Today, approximately 12 terabytes of data are generated from tweets alone on a daily basis. The flow is quickening and shows no signs of abating; with nearly 90% of the data in the world today created in only the last two years. We are truly facing a data tsunami; and there will be 44 times more of the data currently available by the year 2020 (Manyika *et al.*, 2011 as qtd by Gobble 2013). The advent of disruptive technologies such as Internet of Things plays a huge contributory role in the phenomenal growth in data that we now witness.

Recently, the Ghanaian economy has seen great boost in the emergence of companies and organizations that collect increasing amounts of digitized data from clients and employees. Some of these sectors of the economy are the oil and gas industry, healthcare industry, financial services (banks), telecommunication industries, government agencies, retail shops and other data driven businesses. In this paper, only a few of these areas will be discussed along with the potential impact of big data analytics. The increase in telecommunication network providers in the country is an indication that the

majority of Ghanaians have subscribed to one or more of these telecom networks. Ghana can take advantage of the opportunities big data offers that can be leveraged to create a better environment for its citizens and organizations.

In 2011, the Government of Ghana introduced some of its services online. The online services are made available at the Government of Ghana web portal. The web portal promises to serve as one-stop window for services and information offered by all Ministries, Departments and Agencies (MDA), MMDAs and other relevant government of Ghana agencies. The portal consists of four sub-portals, categorized as Citizens, Non-Citizens, Businesses and Governments as shown in **Figure 3**. This is a clear indication that substantial structured and unstructured data will be obtained by the government (Ghana Government e-Services Portal, 2011). Since Ghana performs these services online it is possible that

most of these data that are being generated and collected are from different devices and are of different formats (photos, videos, text, audio, etc.). This makes the data unstructured.

In recent times, Ghana has seen a major shift from paper-based to electronic record keeping in most of the agencies and ministries. For example, recently the National Health Insurance Authority (NHIA) introduced biometric data collection for all clients on their scheme. Other agencies such as the National Identification Authority (NIA), the Electoral Commission (EC), the Ghana Education Service (GES), the Social Security and National Insurance Trust (SSNIT) and the Ghana Health Service (GHS) are all transitioning from the traditional data collection and progressing to electronic data processing and collection. **Figure 4** depicts how these agencies access their individual databases for their day to day transactions.



Fig. 3: e-Services Portal of the Government of Ghana

*Segmented databases*



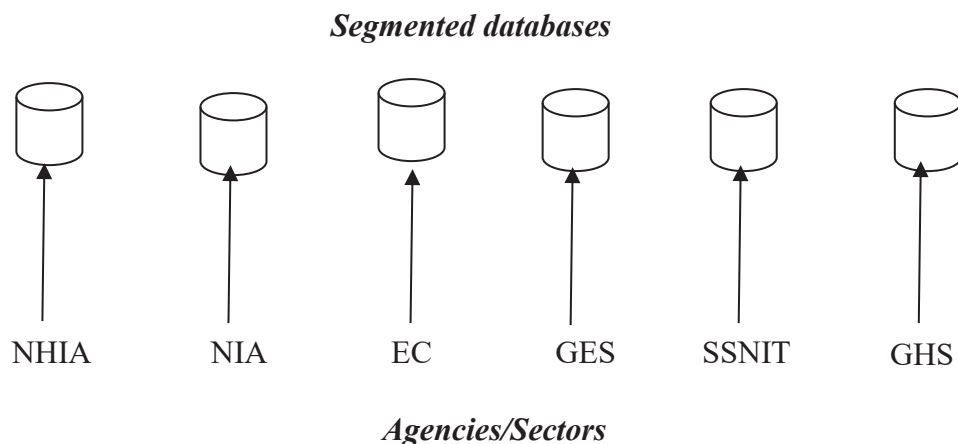NHIA    NIA    EC    GES    SSNIT    GHS

*Agencies/Sectors*

Fig. 4: Shift from paper based records to electronic records

One major challenge here is that these agencies do not share their data. Thus though they produce and collect a lot of data, they are sitting in data warehouses without being put to much use. Users are unable to gain remote access to this data, and when they do get access, there are no analytical tools for any meaningful information to be extracted from it. It is imperative that policies be put in place to regulate sharing and accessing data from a common platform within the public sector; the government could create high value for such data. Rather than look for solutions outside of the country, departments of computer science should be strengthened to provide the research leadership necessary for its realization.

The financial services arena has also seen a sharp increase in its online transactions. The use of mobile communication devices has become routine for personal communications and also for financial and business transactions including money transfer, job search, buying and selling of goods, as well as for the transfer of data such as school grades, examination results, stock levels and prices of various commodities, medical information (Global Pulse, 2012). The introduction of new technologies is helping to drive a wave of innovation across the African financial services sector as banks create new and accessible banking channels and take banking services to previously unbanked parts of society

(Suresh, 2012). In the health industry, the introduction of biometric registration of patients and employees has attested to the data acquisition within that sector. Each of these agencies generate huge amount of data about their clients especially when there are transactions to be processed. A shift in the Ghanaian economy from paper-based to digital data is in the right direction. This shift cannot be accomplished solely on the basis of imported technologies but must be driven by local content achieved through research leadership by our academic institutions.

### Healthcare

For any health information system to be effective it is important that it has access to all health data pertinent to the case under consideration in real time. In many countries of the world this data would come from many different and unconnected systems, Hammond *et al.* (2010). Lewis *et al.* (2012) reported that in low and middle income countries 42% of health institutions use ICT to extend geographic access to health care, whilst 38% use it to improve on data management. According to Raghupathi (2010) as cited by Raghupathi and Raghupathi (2014), the healthcare industry historically has generated large amounts of data, driven by record keeping. Ghana is not very different. The health industry

in Ghana generates millions of data records, but most of these are stored in hard copy form, whereas the current trend is toward rapid digitization of these large amounts of data. A number of health facilities are now moving towards digital records, but currently all these efforts are segmented and disjointed. To derive benefit from the digitization process, these efforts by individual facilities have to be coordinated and centralized (Asangansi and Braa, 2010). Effort must be made to implement an architectural platform onto which individual agencies can simply 'plug in and play'. Health policy makers in Ghana must provide vision and develop the required strategies necessary to achieving a fully integrated health information system for the country.

### Financial Services (Banking)

There is convincing evidence that business has now recognized the ascendancy of data in the business sphere. In a survey recently conducted by Capgemini and *the Economist* of over 600 global business leaders, three-quarters of business leaders agreed that their organizations were data driven, and 90% of them, besides land, labour and capital, recognized information as the fourth factor of production (Gobble, 2013).

Ghana's banking sector has transformed from traditional walk-in and operate transactions to online and electronic banking operated venture where the presence of the customer is not really needed. The sector has expanded substantially over the last decade. The financial sector generates and stores massive amounts of data about customers. According to Suresh (2012), data from the banking industry indicate that banks in the Ghanaian markets spend up to10 percent of their operating income on data management.

In Ghana, despite the challenges in managing and securing customer data, Fidelity Bank, a mid-sized financial firm that has grown over the last ten years to become one of Ghana's leading financial institutions, has invested in a comprehensive, Big Data solution (Suresh, 2012).

Another key consideration associated with the significant growth in data volume of financial institutions is risk in the form of fraud. Constant vigilance and deterrence through technology is the key to protection and employing big data technology is a key measure to prevent attacks (Kothai, 2015). In implementing Big Data in the financial industry, the proposed framework in Figure 5 is proposed.

### Challenges in implementing Big Data in Ghana

In practice, Big Data as a technology faces many challenges, one of which is heterogeneity and incompleteness. Since computer systems work most efficiently if they can store multiple items that are all identical in size and structure, the efficient representation, access and analysis of unstructured or semi-structured data poses analytical and storage difficulties. Another challenge is with the volume of data to be worked on within an organization. Managing large and rapidly increasing volumes of data can be challenging and requires that faster processing components and storage systems be designed and built. Also, with large data sets to be processed, speed could be an issue to deal with.

This is because the larger the data set to be processed, the longer it will take to analyse. Another challenge is privacy. For instance, there are strict laws governing what can and cannot be done with electronic health records. Big data raises concerns and fears regarding the inappropriate use of personal data, particularly through linking of data from multiple sources.

The implementation of Big Data in Ghana comes with its own challenges apart from the ones discussed above. Big Data has much potential for development. Big Data for development has been defined by Global Pulse (2013), to mean the identification of sources of big data relevant to policy and planning of development programmes. This concept is distinct from both "traditional" development data concept and what the private sector and mainstream media refer to as Big Data.
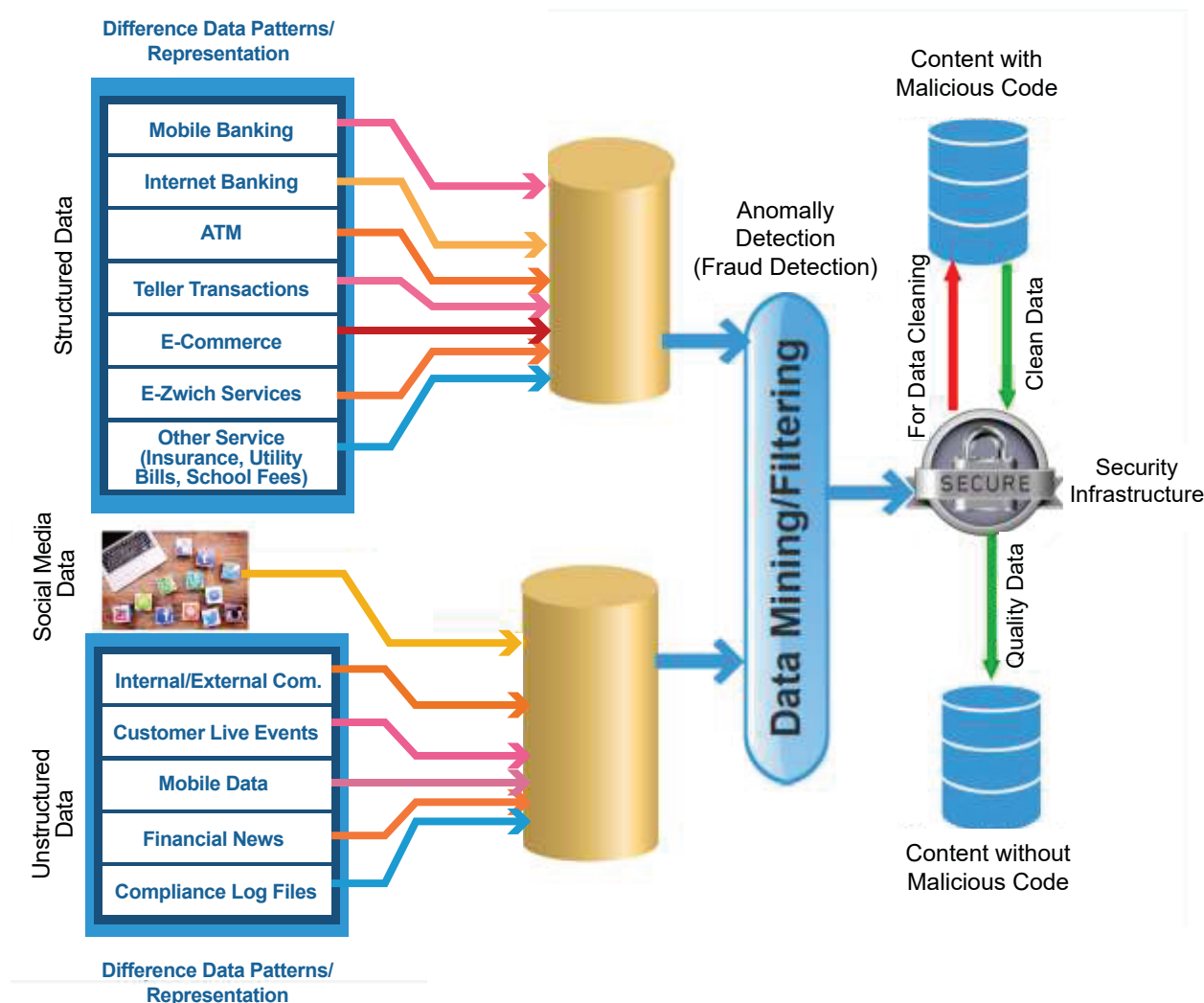
Fig. 5: Framework for big data implementation in the health sector

The lack of infrastructural support and the right technology is also a challenge to the implementation of big data in Ghana. Another challenge is the unavailability of the skilled personnel with the knowledge and skills related to big data analytics. The Commonwealth of Australia in 2013 suggested that since there is a shortage of university degrees that have a curriculum focused on big data analytics, it is important for education providers to design courses geared towards education and training in the area of big data scientists.

Research leadership has a key role to play in realization of the benefits big data has to offer for the Ghanaian economy. Research funding should be provided to educational institutions to run courses and training programmes aimed at producing the cadre of personnel with core skills to drive development in the field. Collaboration between government agencies and research/academic institutions will bring more opportunities for skills development and training and also the Government agencies should create procedures and practices that provide an enabling environment for responsible data analytics (Commonwealth of Australia, 2013).

Privacy and Security is an essential part of every society. Individuals and organizations have data that they protect. For anyone wishing to explore Big Data for development, privacy and security is a primary concern, since it has implications for all areas of work, from data acquisition and storage to retention, use and presentation (Global Pulse, 2013). Data anomalies are normally not detected at the early stages of data analysis and very often they are not discovered in real time. It is imperative to take note of the type of technology used in order to combat the anomalies.

## Conclusion

This paper aimed to identify the potential for development and use of big data in current information administration in Ghana. The adoption of such practices by various institutions or organizations and government agencies in Ghana will enable these organizations to take full advantage of Big Data technologies. This will permit agencies to deliver better-quality and integrated services, improve policy development and identify new services and opportunities to make use of the national information assets, that is, Ghana government data and other data collected by the various agencies in the country. We have reviewed some of the technologies currently being used and proposed a functional definition of big data. We conclude that if the government harnesses the potential of big data to analyse data sets that are generated by the different agencies in the Ghanaian economy, this could improve government operations, policy development and service delivery for rapid economic development. There is the need also to strengthen research institutions to provide the leadership required to drive this effort.

## References

Andrade P.L., Hemerly, J., Recalde G. and Ryan P.S. (2014). From Big Data to Big Social and Economic Opportunities: Which Policies Will Lead to Leveraging Data-Driven Innovation's Potential? The Global Information Technology Report – World Economic Forum. p.81-86

Asangansi I, Braa K. (2010). The emergence of mobile-supported national health information systems in developing countries. Stud Health Technol Inf 2010;160(Pt 1):540–4. [PubMed]

Commonwealth of Australia. (2013). Big Data Strategy — Issues Paper. Available at: https://www.finance.gov.au/files/2013/03/Big-Data-Strategy-Issues-Paper1.pdf

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., and Gnanzou, D. (2015). "How 'Big Data' Can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study," *International Journal of Production Economics*. 165, p.234-246

Gangadharan J. (2014). 7 Ways to Leverage the Data Goldmine with Big Data and Analytics. *Global Practices, Domain Expertise, Customer Experience*

Global Pulse, (2012). Big Data for Development: Challenges and Opportunities. Available at http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf

Global Pulse. (2013). Big Data For Development: A Primer Harnessing Big Data for Real-Time Awareness. Available at: http://www.unglobalpulse.org/sites/default/files/Primer%202013_FINAL%20FOR%20PRINT.pdf

Gobble M.M. (2013). Big Data: The Next Big Thing in Innovation. *Research-Technology Management*, Vol. 56, No. 1. p. 64-66

Government of Ghana e-Services Portal. (2011). Available at: http://www.eservices.gov.gh/SitePages/Portal-Home.aspx

Hammond W.E, Bailey C, Boucher P, Spohr M, Whitaker P. (2010). Connecting Information To Improve Health. Health Aff (Millwood) 2010. Feb 1;29(2):284–8.[PubMed]

Kothai M. (2015). How to use big data to combat fraud. *Big Data Science and Technology. World Economic Forum.* https://agenda.weforum.org/2015/01/how-to-use-big-data-to-combat-fraud/

Lewis T, Synowiec C, Lagomarsino G, Schweitzer J. (2012) E-health in low- and middle-income countries: Findings from the center for health market innovations. Bull World Health Organ; 90(5):332–40. [PMC free article] [PubMed]

J. Manyika J., M. Chui M., B. Brown B., J. Bughin J., R. Dobbds R., C. Roxburgh C. and A.H. Byers A.H. (2011). Big data: The next frontier for innovation, competition and productivity. *McKinsey Global Institute*. Available at: (http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation)

Raghupathi W. (2010). Data Mining in Health Care. In *Health care Informatics: Improving Efficiency and Productivity*, (Edited by Kudyba S.) Taylor & Francis p.211-223

Raghupathi W. and Raghupathi V. (2014). Big data analytics in healthcare: promise and potential, *Health Information Science and Systems* 2 (1) (2014), p. 3

Suresh K L. (2012). Managing the Big Data Challenge in Ghana's Banking Sector. *Building a Smarter* Planet. Available at http://asmarterplanet.com/blog/2012/11/21189.html

# Design and Implementation of an Integrated Web Application for the Motor Traffic and Transport Directorate of the Ghana Police Service

**Emmanuel Opoku Debrah[1], Jamal-Deen Abdulai[1], Isaac Wiafe[1] and Ferdinand Apietu Katsriku[1*]**

[1]Department of Computer Science, University of Ghana
*Corresponding author: fkatsriku@ug.edu.gh

## ABSTRACT

A major challenge facing most organisations is how to share data and services in a timely and cost effective manner to simplify business processes. Integrating new application modules or devices with an existing system smoothly and without any discernible errors or complications is a major issue. A framework for creating rapidly loosely coupled service applications components, the Service Oriented Architecture (SOA), which meets time and cost constraints, has been proposed. In this paper, we report on the adoption of an iterative approach to implement an integrated web based information system for the Motor Transport and Traffic Directorate (MTTD) of the Ghana Police Service. The system is developed as a set of independent web applications sharing a database to provide a single and easy point of information access. In each iteration, a composite of the system is designed and tested, thus meeting the project design objectives.

**Keywords:** Ghana Police Service, Service Oriented Architecture, database, web services

## Background

The Motor Traffic and Transport Directorate (MTTD) is a division under the Ghana Police Service. It was formerly known as the Motor Traffic and Transport Unit. The directorate is responsible for all road safety in the Country (Motor Transport and Traffic Directorate). The Directorate became part of the Ghana police Service in 1952 and is headed by the Commander of MTTD who reports to the Inspector General of Police (IGP).

The main duties of the MTTD are to educate the public about accident precautions, train personnel to arrest drivers who violate the traffic law and regulations, and work with other stakeholders such as the National Road Safety Commission (NRSC) and Driver Vehicle and Licensing Authority (DVLA) in ensuring deriver safety on the roads. They are also responsible for the compilation and publication of all road accident statistics in the country, as well as traffic control and management.

The activities of the MTTD are classified into four categories in terms of its operations, ticket processing, insurance records, accident records, and driver licensing units. Current methods used by the MTTD are outdated, slow and provide very little possibility of integration, causing duplication of data throughout their processes (Alonso, 2004). In this work, we demonstrate how the agile development module principles may be used together with web service as underlying implementation technology to develop and integrate a system for the MTTD.

*Web based information systems:* The information system of the MTTD may be viewed as consisting of individuals, computer devices and equipment with the relevant software programs, a dynamic database, and organisational practices that interact in a recommended systems configuration. The system is designed to record, store, update, and expedite the automation of

data usage on a persistent basis. The data thus stored and its administration are interconnected to members of the organisation, police officers, clients (in this case the drivers), external users (of which the insurance companies are example) and other law enforcement agencies who depend on the MTTD information systems for information. The main purpose of the information management system is to meet operational needs of the unit and to facilitate information reporting and data analytics relevant to promoting road safety. The information system also permits the exchange of information among the various stakeholders (Whisenand, 1971).

### Service Oriented Architecture (SOA)

Whilst there is no universal agreement in the literature on the meaning of Service Oriented Architecture (SOA), it is generally acknowledged that the underpinning notion is service-centric distributed computing. A service may be viewed as a self-contained logical representation of a business activity with a specified outcome. Such an activity must be repeatable and may be made up of other sub services. In this paper, we adopt the definition by Organization for the Advancement of Structured Information Standards (OASIS), which is believed to be more inclusive and complete than other definitions proposed in the literature. SOA is defined by OASIS as *A paradigm for organizing and utilizing distributed capabilities that may be under the control of different ownership domains. It provides a uniform means to offer, discover, interact with and use capabilities to produce desired effects consistent with measurable preconditions and expectations* (W3C, 2007). We may then infer that the term capability is central to this concept. A capability may denote a set of functionalities or techniques implemented for a specific business process or sub-process, implying that a group of capabilities may also be viewed as a service. Therefore, from the perspective of SOA, a service not only refers to a set of capabilities, but also to all related capabilities which can be grouped together to be used by diverse business processes. From the foregoing, three types of services under SOA implementation may be identified (Erl, 2008a):

- An **Entity Service:** information system entities like staff, ticket, license and reports provide the foundation for this kind of activity and place a practical limitation on the context within which these entities might be used.
- **Task Service:** this is an information system service whose functional boundary is directly related to an explicitly defined business task or process.
- **Utility Service:** this is an application service providing reusable service functions of a non-business-centric nature. Such service functionalities may include event logging, reports and exception handling to other entities and task services.

The design goals of SOA are to achieve the following, among other things: loose coupling, reusability, standardized service contract, abstraction, composability, autonomy, statelessness and discoverability of services (Endrie, 2004; Erl, 2008b). The benefits associated with SOA include its ability to leverage the use of existing assets, the ease with which it can integrate and manage complexity, the more reactive and faster time-to-market that it provides and its cost reduction and increased reusability.

The SOA framework is not without its challenges and should not be seen as a silver bullet for systems design (Fuerlicht, 2006; Erl, 2008a). The adoption of this framework increases design complexity. It also requires an agreed design standard, as this increases uniformity and compatibility throughout the diverse segments of the system. There also would be the need for governance structures to be put into place to ensure that laid down processes are adhered to.

The reported work was aimed at integrating the database systems for the Eastern Regional Motor Transport and Traffic Directorate of the Ghana Police Service. A consolidated service bus was used to simplify the interoperability of applications across platforms (Web application, Mobile Application), enhance communication within the MTTD infrastructure and create an enabling ecosystem for a new layer of abstractions to be added without the need to modify the

complete system. The three services (Figure 1b) were developed as freely coupled services and dependability attributes, namely security, embedded within them to include authentication and encryption as well as reliability. An evaluation process was conducted to show that by implementing systems as web services, MTTD is able to obtain the declared advantages of an SOA model, particularly those linked with the flexibility of web services to the underlying business environment.

The challenge confronting the MTTD is to have an information system to control their operations and provide road duty officers access to the system using a mobile application to verify drivers' information. To tackle the problem, the following two solutions were proposed:

i.   First, design and implement an integrated information system for the organization, which handles all the four departments of the organisation (Fig. 1a).

ii.  Implement web services with mobile application to connect to the MTTD information systems. Web-services implementation was suggested as an option which facilitates the amalgamation of mobile applications with other heterogeneous web applications. Figure 1b shows the proposed integrated system.
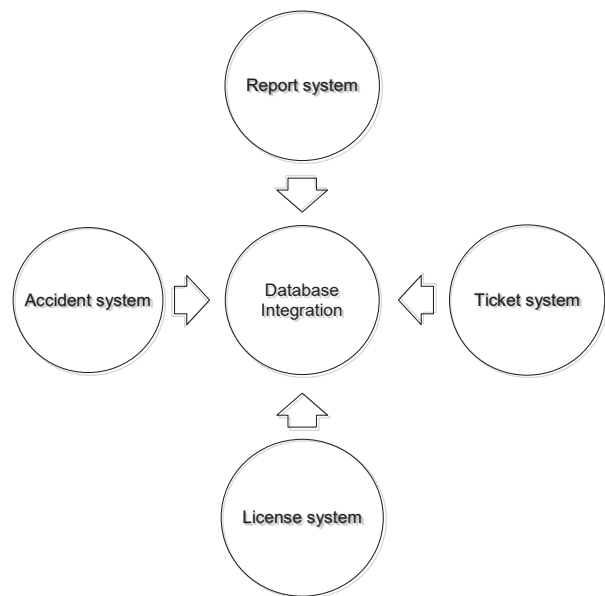


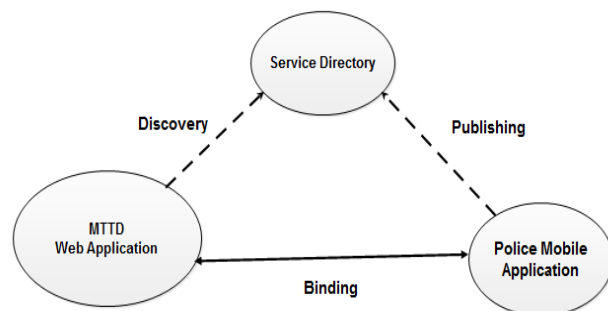Fig. 1a: Proposed information system architecture for MTTD



Fig. 1b: Proposed Integrated Systems

**Table 1: Non-functional Requirements**

| SN | Non-functional Requirements | Relevance | Description |
|---|---|---|---|
| 1 | System Performance | Key | A key requirement is system performance; this should be adequate and fast to provide a good user experience. |
| 2 | User Friendly Environment | Key | A simple and user friendly system interface is an essential requirement. |
| 3 | Application Maintenance | Key | Clear documentation on the application and its maintenance needs to be provided. |
| 4 | Paging | Required | When dealing with large chunks of data, system should be capable of splitting these data using data paging and present the information in a user-friendly format. |
| 5 | Application Scalability | Required | System should be easily scalable; such flexibility will enable future growth. |
| 6 | Platform Independence | Required | System should be capable of being deployed on any platform. |

## Non-Functional Requirement

The Non-functional requirement listed in Table 1 applies to all the four systems that were designed.
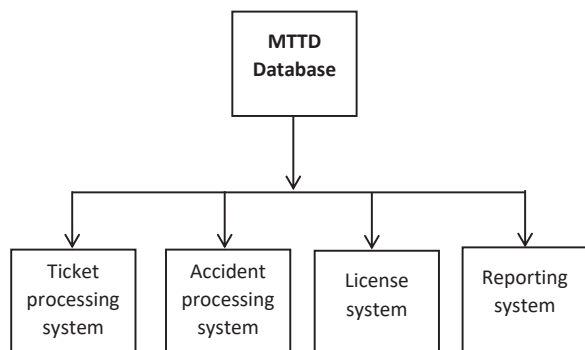


Fig. 2: Context diagram for the MTTD system

Figure 2 shows the context diagram for the entire MTTD system. These systems were designed to be hosted on the same webserver to share a common database.

Even though each system operates separately from the other, they are integrated in a way to share the same database. This means that any of the systems can run from a different webserver; once the database integration connection points to the database, it should be able to work perfectly without any problem. Any of the above systems can run independently with or without the other systems as long as the database is up and running, even though some of them depend on each other for data to

work with. As an example, the license system must issue drivers' licenses and capture drivers' details before the ticket processing system can record a ticket for a driver. This doesn't mean that the ticketing processing system cannot run without the license system; it will, but it will not get the data needed to function effectively.

Again, the reporting system depends on the ticketing processing system, the accident processing system and the license system to generate the needed report for third parties including insurance companies. The reporting system can run even if all the three systems are down, but will not be consistent with Schmidt's (2005) claim. In this case, even though the system is available, it cannot be reliable (Schmidt, 2005). The entire system is managed by one administrator who assigns roles and permissions.

## System Implementation and Testing

The implementation phase involved modelling the solution, coding the solution using PHP script language and MySQL database for the various web applications within the MTTD information system, and adopting a REST architectural style for building web services. The MTTD web based integrated system is implemented as several services interacting with each other in a loosely coupled fashion to achieve the intended objectives. The service is implemented on a web server which provides access to the service using HTTP protocol. Figure 3 depicts the overall system architecture.
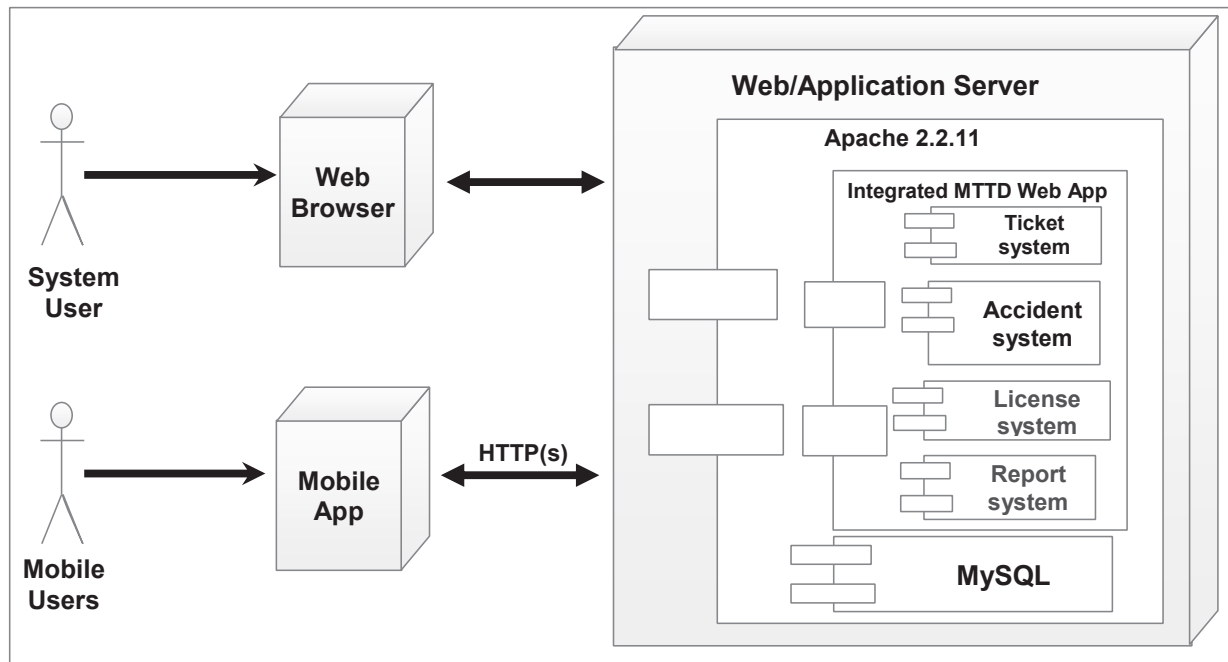
Fig. 3: Overall system architecture

The web applications were developed using PHP 5.3.0 which is a common server scripting language that offers developers the capability to develop web applications based on the Model, View and Controller (MVC) paradigm. The MVC paradigm offers a development environment that tends to disconnect the application logic "controller" from User-interface "view" and also from the data manipulation layer "model". This separation presents developers with the flexibility to alter any of the layers "controller, view or model" without requiring them to make changes to other layers. For complex systems this flexibility has great advantages. It also gives the developer a rich and varied set of scripts that helps to generate a number of common functions automatically. On the other hand, mobile application WS was developed using PHP Mysql with RestFul Web Service.

### Database implementation

Storing and recalling data is one of the key elements within the integrated MTTD web applications. The MTTD web service URI acts as a resource which must contain information from the database. This information will be stored within a MySQL database, located on the same server that hosts the MTTD Information system. MySQL is an open source Relational Database Management System (RDBMS) and this informed its choice. In addition to the fact that it is freely available, it is also relatively light compared to other RDBMS such as ORACLE. The MTTD application database will consist of a number of tables to manage the ticket, accident, license and report processing.

### Web Applications implementation details

As shown in Figure 2, there are four (4) main components of the MTTD web applications developed: the ticket processing system, license processing system, accident processing system and report system.

### First Iteration: Implementation of Ticket processing system

At the first iteration, we implement the ticket processing system. The ticket processing system has three (3) user permission levels: the police clerk, the court clerk and the administrator. The administrator is responsible for managing user accounts and permission.

### Second Iteration: Implementation of Accident processing system

At the second iteration the accident processing system is implemented. The accident processing system is the first independent web application that was integrated into the ticket processing system to share the same database. Figure 4 shows a screenshot of the database connection string that was used for the integration.

```
// Database connection info

var $Host = '127.0.0.1';
var $Port =         ;
var $Username = 'accident';
var $Password  = '@accident';
var $DbName = 'spbase';
```

Fig. 4: Database connection for accident processing system

### Third Iteration: Implementation of License processing system

At the third iteration we implement the license processing system. The license processing system is the second independent web application that was integrated into the existing integrated system, which comprises the ticket processing system and accident processing system, to share the same database. The database integration was done using a PHP database connection string similar to that used for the accident processing system.

### Implementation of Report system

The fourth and final iteration was the implementation of the reporting system. The report processing system is the third independent web application that was integrated into the existing integrated system to share the same database.

### Implementation details of WS

To implement web services, a WS stack and APIs are required. The WS used in this work was created using RestFul web service with PHP. REST is a simple stateless architecture that generally runs over HTTP. The REST web services system produces a status code response in JSON or XML format. The main purpose for implementing the web service is to authenticate and search for drivers' license information outside the domain of the MTTD information system.

### MTTD Mobile Application

In order to implement and test the web service, a small mobile application was developed to make a system call to the MTTD web services application developed in this project. This mobile application was developed on Android platform and allows an officer to check the validity of driver license information.

### Testing

Driver information and officer details were verified during the tests conducted. Google chrome Advance REST client was used to test the WS created. Login request and response web services were tested and Figure 5 shows the output test result. It may be seen that the details were successfully verified within 365ms.

200 Success Loading time 365 ms

**CSP:** active
**Origin**: chrome-extension://hgmloofddffdnphfgcellkdfbfbjeloo
**User-Agent:** Mozilla/5.0(Windows NT 6.3; WOW64)AppleWebKit/537.36 (KHTML, like Gecko) Chrome/43.0.2357.134 Safari/537.36
**Content-Type:** multipart/form-data; boundary=ARCFormBoundarycqdapu96zb0529
**Accept:** */*
**Accept-Encoding:** gzip, deflate
**Accept-Language**: en-US, en;q=0.8
**Cookie:** PHPSESSID=oi32m455915bomjbfgoh8mbip5

**Date:** Mon, 20 July 2015 12:05:35 GMT
**Server:** Apache/2.2.11 (Win32) PHP/5.3.0
**X-Powered-By:** PHP/5.3.0
**Content-Length:** 50
**Keep-Alive:** timeout=5, max=100
**Connection:** Keep-Alive
**Content-Type**: application/json

Fig. 5: Test result of login to the web service and response time.

A driver license search using web service was also conducted and Figure 6 shows the output results. As can be seen, the response times are reasonably good. By using an online free open source testing tool called Agilload (Optimise your load and performance testing costs, 2015), all services and their functionalities were tested and shown to be working. It can be seen that a driver license search was successfully conducted in 324ms.



200 Success    Loading time 324 ms

**CSP:** active
**Origin**: chrome-extension://hgmloofddffdnphfgcellkdfbfbjeloo
**User-Agent:** Mozilla/5.0 (Windows NT 6.3; WOW64)AppleWebKit/537.36 (KHTML, like Gecko) Chrome/43.0.2357.134 Safari/537.36
**Content-Type:** application/x-www-form-urlencoded
**Accept:** */*
**Accept-Encoding:** gzip, deflate
**Accept-Language**: en-US, en;q=0.8
**Cookie:** PHPSESSID=oi32m455915bomjbfgoh8mbip5

**Date:** Mon, 20 July 2015 12:10:07 GMT
**Server:** Apache/2.2.11 (Win32) PHP/5.3.0
**X-Powered-By:** PHP/5.3.0
**Content-Length:** 50
**Keep-Alive:** timeout=5, max=100
**Connection:** Keep-Alive
**Content-Type**: application/json

Fig. 6: Test result of search for license details web service and response time

Figure 7 depicts the output of the web service from the database in JSON format to be consumed by the mobile application and any other future application which will need driver information from the MTTD information system for implementation. A free online tool, Chrome Advance REST client, was used for this testing and evaluation.

```
            dateIssue: ''2015-06-16''
            drvName: "SAMUEL APPIAH MENSAH"
            drvDOB: "1993-04-20"
            drvAddress: "OBOC WEWEW MMER"
            drvRestriction: "Active"
    }
    -3: {
            drvLicenseNo: "LIGH10202"
            dateIssue: ''2015-05-12''
            drvName: "EMMANUEL APPIAH"
            drvDOB: "1989-02-07"
            drvAddress: "BOX 98111"
            drvRestriction: "Active"
    }
    - 4: {
            drvLicenseNo: "LO23124"
            dateIssue: ''2015-06-26''
            drvName: "MARTIN OFFEI"
            drvDOB: "1978-06-14"
            drvAddress: "KOFORIDUA"
            drvRestriction: "Active"
    }
    - 5: {
            drvLicenseNo: "LO23126"
            dateIssue: ''2015-06-27''
            drvName: "TIE LINCOLN"
            drvDOB: "1973-12-12"
            drvAddress: "P. O. BOX 88232
            KOFORIDUA"
            drvRestriction: "Active"
    }
    - 6: {
            drvLicenseNo: "6"
            dateIssue: '0000-00-00''
            drvName: ''2015-06-27''
            drvDOB: "0000-00-00"
            drvAddress: "1973-12-12"
            drvRestriction: "P. O. BOX 88232
            KOFORIDUA"
    }
```

Fig. 7: JSON Output of the implemented web service

## Evaluation

A number of key parameters were identified at the start for the purposes of evaluation.

*Secure Login*: A secure login was implemented successfully for all the various web applications, allowing multiple user accounts to access the system. Hash md5 algorithm was used to ensure secure password storage within the database and a PHP function was put together to validate user sessions upon accessing any page within the system. The requirement for a secure login with a separate user account was implemented satisfactorily. An additional security session checking script was an added advantage to the system.

A password change in the script was implemented successfully to enable users change their passwords when necessary.

*File Management*: One of the key functional requirements was file management. This included the ability to view, edit and add client tickets, accident files and driver licence information with ease. All of these requirements were implemented into the system along with the addition of policy categorisation, meaning that this area of functionality was met fully. It is also worth noting that the addition of touches such as a JavaScript function to change which policy type to insert, means that the final system actually delivers more than the original requirements laid out.

*Reports*: The reports were implemented into the system in such a way that both an HTML format (for printing or storing) and a CSV format (for use in a spread sheet program) report could be generated. This was useful for MTTD in terms of the flexibility offered with the data the system outputted, as their requirements for the layout and format could change in the future.

### Data Integrity and Security

An area of concern was the future migration of data from the new system. However, choosing to utilise a MySQL database meant that exporting the data would not be a particularly challenging task – obviously conversion of this data to fit a new systems format would require some work, but the main requirement was that the data could be accessed and moved if needed – which it can. To ensure the integrity of the systems data, validation was implemented onto user input forms. Through the use of JavaScript to immediately alert the user to any incorrectly formatted data, the system was able to remain

user friendly whilst enforcing data format rules. System testing ensured that this validation worked correctly throughout and as a result, data integrity within the database was assured.

The security of the system and data was enforced through two measures.

1.  A secure user login and session authentication script ensured system access was only granted with the correct credentials.

2.  SQL injection and hacking attempts were prevented through the implementation of a character stripping and validation function.

## Conclusion

This paper has reported on the implementation of an integrated usable web application for the various operations of the MTTD of the police service of Ghana using the SOA approach. We reviewed the concept of Service Oriented Architecture and also key web services principles, standards and components, and have shown how they could be used in application development. Four independent web applications were implemented and successfully tested. The agile development methodology was adopted for this work. To date, an industry standard software development methodology tailored specifically for SOA is non-existent. We have thus shown that it is possible to combine an iterative process with an agile methodology to achieve development goals. SOA/ Web services remain an important and one of the most promising areas in computer science. Notwithstanding the huge effort that has been exerted in the development of the fundamental standards and technologies, there remains a gap in what web services promise and what can realistically be implemented. With the current trend towards cloud based computing and services, the SOA approach may be an important methodology to consider.

## References

Alonso, G., Casati, F., Kuno, H. & Machiraju, V. (2004). *Web Services. Concepts, Architectures and Application*. Springer, Berlin.

Endrie, M., Ang J., Arsanjani, A., Chua, S., Comte, P., Krogkahl, P., Luo, M. & Newling, T. (2004). *Service Oriented Architecture and Web Services.* IBM Red Book, 2004. [E-book].

Erl, T. (2008a). *SOA Principles of Service Design*. Boston, MA: Prentice Hall.

Erl, T. (2008b). *Service-Oriented Architecture Concepts, Technology and Design.* Boston: Pearson Education Inc.

Fuerlicht, G., (2006). "Enterprise SOA: What are the benefits and challenges," presented at System Integration Conference, (pp. 1-8), Czech Republic.

Motor Transport and Traffic Directorate. Available online: http://en.wikipedia.org/wiki/Motor_Transport_and_Traffic_Directorate [Accessed: June 14, 2015].

OASIS, "*Web Services Distributed Management: Management Using Web Services (MUWS 1.0) Part 1*". 2005. Available online: http://www.oasis-open.org [Accessed: June. 05, 2015].

Optimise your load and performance testing costs. Available online: http://www.agileload.com/ [Accessed: 20th, July 2015].

Schmidt, M.T., Hutchison, B., Lambros, P. & Phippen, R.(2005). The Enterprise Service Bus: Making service-oriented architecture real. *IBM Systems Journal*, 44 (4), 781-797.

W3C, (2007). "*SOAP Version 1.2 Part 2*". 2007. Available online: http://www.w3.org [Accessed: June 15, 2015].

Whisenand, P. M. (1971). Automated Police Information Systems: An Argument for Vertical and Horizontal Integration. *Journal of Criminal Law and Criminology* 62 (3).