ENVISIONING CHANGE IN THE STATISTICS-EDUCATION CLIMATE

DAVID STERN
IDEMS International
d.a.stern@idems.international

ROGER STERN
Statistics for Sustainable Development
r.d.stern@stats4sd.org

DANNY PARSONS
IDEMS International
danny@idems.international

JAMES MUSYOKA
Maseno University
jkmusyoka@maseno.ac.ke

FRANCIS TORGBOR
AIMS Ghana
f.torgbor@aims.edu.gh

ZACH MBASU
African Maths Initiative
zmbasu@gmail.com

ABSTRACT

The African Data Initiative started as a crowd-sourced campaign to improve the teaching of statistics in African universities. The analysis of climate data provides one suitable context to illustrate ideas that lead to a radical new form of teaching. The problem within the context comes first, the technicalities are largely reduced – mathematics is supported by meta knowledge and backed up by modelling; calculations are reduced by user-friendly software that is also used by experts. The problems are treated similarly to research questions and the results are often easier to interpret, making sense as potential answers in their context. The criteria of this approach are compared to the framework proposed by W. G. Cobb to reform statistics education in the light of the latest developments in statistics, driven by the huge increase of data. Implementation details are presented around three components: case studies, data, and the required skills. Together, these three components describe an alternative education pathway centred around statistical problem solving. The focus on interpretations of results within a real context enables software, mathematical thinking and modelling to play a supportive role, which flattens the prerequisites of complex methods and encourages their use across all levels of education.

Keywords: *Statistics education research; Climate Data; Technological innovation; African context*

Statistics Education Research Journal, 19(1), 206–225, http://iase-web.org/Publications.php?p=SERJ

1. INTRODUCTION

Within the statistics education community there is general agreement on the importance of student experiences with real and realistic data and prioritisation of teaching statistical understanding over traditional mathematical formulations. Across the world, great progress is being made in this direction through guidelines (e.g., Franklin et al., 2007), curricula (e.g., Ministry of Education, 2012; Forbes, 2014), examinations (e.g., Pearson Education, 2018), competitions (ISLP, n.d.), and consulting experience (e.g., Jersky, 2002; Bilgin & Petocz, 2013; Fletcher, 2014; Borovcnik, 2018). However, as data become more central, statistics remains misused at least as often as it is used well, and even scientific publications are prone to misinterpretations (Amrhein, Greenland, & McShane, 2019).

Across Africa many of these internationally established norms of modern statistics education are yet to be introduced; students are often still trained in highly theoretical ways with little exposure to data other than to illustrate a statistical method. Despite a huge variability of circumstance in different countries across the African continent, the educational systems are united by the challenge they face to train the next generation of statisticians to be able to serve their countries' data needs.

This article emerges from a broader collaboration, the African Data Initiative (ADI; Stern, 2017) that was founded to implement improvements in the teaching of statistics across Africa. ADI started in 2015 with a crowd-funding campaign to develop free, easy-to-use statistical software. The first visible product of this initiative was the development of R-Instat, a menu-driven front-end to R, which includes tailored components that facilitate National Meteorological Services (NMS) to analyse their historical climate data. The next steps are about reaching out and transforming the way statistics is taught and learned.

Climate data has been recognised as a valuable resource in statistics education (Witt, 2014; or Pearson Education, 2018). In this paper, we move the discussion beyond affirming the value of climate data to teach today's statistics by investigating the potential role climate data can play in enabling statistics education to transform itself to meet current and upcoming needs.

The idea that the training of future statisticians should change is linked to the rise of data science (Davenport & Patil, 2012) and the growth of data in our society (IEAG, 2014). Statisticians tend to take a broad view of statistics, which encompasses all aspects of working with data, as is highlighted in a definition of statistics from the Royal Statistical Society (2019):

> "Statistics changes numbers into information. Statistics is the art and science of deciding what the appropriate data to collect is, deciding how to collect them efficiently and then using them to answer questions, draw conclusions and identify solutions."

This raises the question of what we should be teaching future statisticians related to ideas, which are not currently prominent in the curriculum, and what the implications are, of considering statistics as both an art and a science.

We present a vision related to the training of future statisticians and investigate how climate data and related studies can be used to engage students and develop a statistical skill base that prepares them for a range of areas, which they may encounter in the future. We start with the framework of five principles and two caveats, put forward by Cobb (2015), to test the suitability of the proposed approach of using climate data. We then present case studies that follow such an innovative approach and discuss the difficulties with data and the required skills to illustrate the specific ideas of how ADI intends to use the field of climate for innovations in the teaching of statistics even in more difficult environments as we encounter in wide areas of Africa.

Section 2 presents Cobb's framework (Cobb, 2015) for a radical reform in statistics teaching. The ADI approach is benchmarked against this.

Section 3 presents two case studies. The first uses meteorological data for African farmers through an initiative called Participatory Integrated Climate Services for Agriculture (PICSA). The second case study concerns renewable energy.

Section 4 discusses several sources of African meteorological data and how to read it into R. We also point out that "Each country has historical climatic data."

Section 5 is concerned with skills. The package R-Instat is discussed, which does not require programming skills. This is important – not only in the African context – since many students, particularly those taking service courses in statistics, neither have exposure to data nor to programming skills.

Section 6 presents our conclusions related to the suggested approach, which implies that a radical reform of statistics teaching using climatic-data examples and problems should enable students to be better prepared to become future statisticians.

2. COBB'S FRAMEWORK FOR THE FUTURE

Cobb makes a strong case for radical reform in statistics teaching. Although presented from a U.S.-centric perspective, the framework is valid more generally. It puts forward five challenging principles and two caveats. We use this framework to test the ADI approach and the versatility of climate data, particularly in the African context. Of course, there should be data from multiple areas of application, with climate being one of many, but the field is cross-cutting enough to have widespread relevance.

The five principles of Cobb are (Cobb, 2015, p. 273–278, formulates them as imperatives):

- Flatten prerequisites,
- Seek depth,
- Embrace computation,
- Exploit context,
- Teach through research.

The two caveats of the suggested framework are, according to Cobb (2015, pp. 278):

- "All curriculum is local",
- "For change to endure, you have to institutionalize it".

This framework provides a powerful structure to investigate the potential of climate data to facilitate this transformative reform. We focus on the potential within the context of ADI, which aims to use open tools to support widespread African reform, and also to stimulate discussion with a wider international audience.

2.1 CHALLENGE 1: FLATTEN PREREQUISITES

The idea of flattening prerequisites links to the problem-solving approach, which lies at the heart of the ADI approach. Traditional statistics courses have largely been hierarchical with topics introduced in their order of computational prerequisites. A problem-solving approach, which embraces technology, can reverse this sequence and introduce methods required for the analysis of a task on demand. This approach ties in well with the description of statistics as an art as well as a science. In many arts subjects, the same task can be set for students at almost any level of study; tasks are differentiated by the complexity and depth of the answer, rather than the difficulty in undertaking the task.

Within the ADI approach, there have already been examples of this way to teach, with postgraduates engaging high-school students in their areas of interest at mathematics camps. Most notable are instances whereby postgraduates, involved in defining and creating the climate menus in R-Instat, have delivered sessions where students work with full climate data sets and software to do the same analyses, which experts could be doing. Our experience shows high-school students can engage with large climate datasets and real research questions, in an extra-curricular environment. This supports that flattening the prerequisites can be implemented and students can learn what they need, through working on the problem, rather than relying on substantial prerequisite knowledge.

2.2 CHALLENGE 2: SEEK DEPTH

The principle of seeking depth is about creating deep understanding of statistical ideas without focusing on the formalism or formulas used in calculation. This is related to flattening-the-prerequisites. The challenge is to break the rule that it is not possible to use and understand the concepts if one has not built up the formalism through prerequisites. However, it remains essential to present this content and build an understanding of the ideas without depending on the formalism.

A problem-driven approach, which starts with real data in context and has learners using the same (software) tools as professionals, can enable ideas to emerge naturally when needed. To design tools that play a role across a range of users, from beginners to experts, is central to ADI; i.e., enabling complex problems to create opportunities for deeper thinking rather than being stuck by the technicalities of the required tools.

Climate data offers many opportunities for this approach, while staying within the learners' intuitive understanding. Even data as simple as daily precipitation can provide a rich source of concepts such as seasonality corresponding to periodicity, climate change relating to trends, and storms relating to the asymmetry of distributions, extreme events, and outliers. Climate data also includes the concept of data at different levels, with sub-daily, daily, monthly, yearly, and station levels, all natural to consider.

Analysing data at different levels is usually seen as an advanced skill but, in this context, insights can be gained by thinking of data at levels and combining data across levels. People have extensive experience of weather and climate and are often able to remember specific historical events and relate to components they find in the data. This familiarity with the context facilitates an understanding of the complexities of the statistical methods needed to study the phenomena.

2.3 CHALLENGE 3: EMBRACE COMPUTATION

This principle links to technological progress and the emergence of data science. Embracing computation requires recognition that the skills needed to work with tomorrows' data differ from the traditional topics taught in the past and it is therefore not clear where reform would end. Radical reform could enable African countries not only to catch up with developed countries in terms of education, but also to achieve advantages related to innovative skills.

This opportunity arises because technology allows the curriculum to be turned upside down; it is no longer necessary to teach alongside mathematical complexity as many concepts can be understood with meta knowledge rather than with mathematical details. This facilitates teaching advanced data skills with conceptual depth, building from the practical skills of analysis rather than from theoretical considerations. Use of appropriate software with substantial data sets also permits interactive data exploration, with learners quickly trying many presentations and where wrong directions do not matter because they have not taken much time on computation. This promotes a more creative skill set with thoughtful experimentation being highly valued.

Software-based exploration of climate data is very rich. Some of the tailored analyses developed for ADI can facilitate the exploration of complex summaries, such as the start of the rainfalls across seasons and locations. Traditionally, such tailored summaries would not be used in teaching; but making the computation easy, enables students to imagine other possibilities. Climate data can also facilitate the introduction of advanced statistical concepts, which require computationally complex methods such as fitting harmonic regression, generalised linear models, or considering the order of Markov-chain models that may describe the data well. These methods arise naturally through a standard breakdown of daily precipitation into the probability of rain and the amount of rain on rain days. Modelling the change between states in precipitation from day to day raises the question "How does today's weather relate to the recent past?" – which is both easy to understand and hard to answer.

Students intuitively have assumptions and questions about the climate that enable them to appreciate the value of exploring the data. The mathematical complexity is no barrier to exploration; on the contrary, the potential of sophisticated methods can provide motivation to engage in the mathematical details, which can differ for students at different levels. We have repeatedly engaged school students at mathematics camps with complex rainfall models with good results, showing that these models can be introduced in introductory statistics courses. We are not proposing that this will lead to students automatically gaining a depth of understanding; yet, through early exposure, students may find the motivation to gain deeper understanding in subsequent courses.

2.4 CHALLENGE 4: EXPLOIT CONTEXT

Engaging students in explicit contexts gives meaning to the methods used for analysis. To teach with real data provides an opportunity for students to experience statistics within real-world contexts. However, there is a "tension between abstraction and meaning-in-context", which "energizes our subject" (Cobb, 2015, p. 276). The tension comes from the balance between understanding a single context and the skill of abstracting common ideas from problems that only seem different because of their different contexts. The skill of abstraction is critical for statisticians to transfer methods developed for one context to others. However, with any given problem, a statistician also needs to find solutions for this specific context, which requires understanding the peculiarities and needs that arise.

Traditionally, a curriculum is based around the abstract methods being taught, and examples are constructed around these methods; such examples can be contrived and disjoint from reality. By beginning with a rich and engaging context, it is possible to introduce methods and approaches that arise naturally when students find ways to tackle the problem. Problem-led learning of this sort has many advantages but is less predictable in terms of the content covered.

We claim that climate data provides a context, which students at all levels can grasp immediately. From this familiar start, there are many opportunities for abstracting general concepts. A classic example is daily precipitation data, which has a natural structure of dry days and rainy days. Without considering this structure, even simple summaries such as the mean can prove misleading, while including the structure leads to more useful summaries like the percentage of rain days and the mean precipitation on rainy days. This simple but important concept often perplexes students who are used to turning the handle of calculation without thinking further. Considering the natural structure of data before performing an analysis is a general principle, which can be abstracted from this easily understandable example.

Starting with context before methods, provides more interesting, realistic exercises for students that better equip them with the skills they need to work with data in the real world. They start with a problem without knowing, which method they 'are expected' to use. The same context and exercise can also be used with students from school level through to postgraduate, and with both statistics students and non-specialists. While the content remains the same, the statistical methods and approaches can differ depending on the capabilities of the students. This relates back to the first principle of flattening prerequisites.

2.5 CHALLENGE 5: TEACH THROUGH RESEARCH

According to Cobb (2015), teaching through research is the most important of the five principles. Our role as statistics educators should not be "to prepare students to use data to answer questions, it should be to get them to answer questions that matter with data" (p. 277). Such an experiential approach goes substantially further than has been suggested in the previous principles, while relying on those principles to make it possible. Teaching through research breaks down the barriers of discipline and encourages learners to think in a problem-based way trying to ask genuine questions and learn how to use data to answer such questions.

In many areas the concept of learn-by-doing is known to provide motivation to engage students, and this can be applied across academic levels. However, the idea of teaching through research goes further, by implying we should teach statistics in a transdisciplinary way. Taken to its extreme, could statistics specialists be taught through joint courses with service students? Imagine a statistics curriculum whereby statistics students do joint research with agriculture students in one course and health students in another. This would give the real-world skills a practicing statistician needs, and provide a different approach to service teaching, but it would be a big leap from how the curriculum is currently devised.

A more modest change could be to use climate data as a step towards this transdisciplinary approach. Students at all levels have knowledge of climate, as well as questions and hypotheses about it. There are many research questions around climate and its relation to other disciplines that are within the reach of students to explore. A single research scenario can be used to teach a substantial set of skills, possibly across courses so that students gain depth of understanding from repeatedly being exposed. One such detailed and rich research scenario is given in Section 3. A research topic can, if made central to a sequence of courses, build student awareness of their role in supporting research. Applied statisticians enable researchers of all disciplines to achieve better results through improved use of their data. Training statisticians through genuine roles in research can prepare students with the statistical skills to contribute to research, and provide the transferable skills needed to work in interdisciplinary teams in the future.

2.6 ALL CURRICULUM IS LOCAL (BUT WE NEED TO BE GLOBAL)

Cobb (2015) discusses two caveats aligned with the set of principles presented above. The first is that university curricula are fundamentally local. Unlike secondary schools where the curriculum is usually determined by a centralised process (often linked to a centrally administered examination), universities have the freedom and responsibility to identify their focus and design and develop their own curriculum accordingly. Hence, curriculum decisions are made at every institution and innovations only apply locally.

Across Africa there has been an explosion of new universities within the last twenty years. This has created a need for a more global effort to support development and change of curricula as well as the delivery, particularly in those universities who have not yet the necessary expertise. A coordinated effort to support many institutions in changing their teaching, to enable more impactful learning is behind ADI. We are supporting a gradual improvement of teaching and curricula within any given institution (Stern, Ongati, Agure, & Ogange, 2010). With an incremental approach, change can grow from individual lecturers teaching their courses differently. This takes the local condition of a curriculum down to the contents of a single course, which is the essence of any teaching. Being local puts the power of change into the hands of each motivated individual.

2.7 FOR CHANGE TO ENDURE, YOU HAVE TO INSTITUTIONALISE IT

The second caveat is that change is only sustainable if it can be institutionalised; otherwise the changes risk being lost when individuals teach different courses or move between institutions. This indicates the danger of simply focussing on the individuals who are motivated to innovate. Two ways to institutionalise change are either through formal structures or through the accepted working culture. Awareness of both approaches is valuable to achieve sustainable change, particularly in African contexts where the local working culture can play a key role.

Institutions change their formal structures following mainly internal processes. Even when external stimuli and funding supports a change, the results are driven primarily by the internal staff. External actors can play an essential role in steering an internal process to improved outcomes if, and only if, there are staff to do the work needed to drive the process from within. A few highly motivated internal staff can transform formal structures if they have the right support. Changing local working culture is a different matter. Within any institution, the working culture cannot be defined by an individual, but relates to how groups of people act. Of course, individuals can affect groups, either positively or negatively, but often group behaviour is driven by incentives. Different people respond to different incentives, some mainly to financial, others through ambition, and some have intrinsic motivation where the incentives are harder to articulate. Most are driven by a mixture of incentives so that workplace culture is often defined primarily by local incentives.

In the African context, many staff still travel abroad for their doctoral studies. The biggest challenge they face when coming home is the workplace culture. Often the 'culture shock' of coming home is much harder than that of 'a new environment'. Internationally, these staff would take on post-doctoral positions, through which they would have a supportive environment to grow into full academic positions. However, in many African institutions, they return immediately to positions of responsibility where there is often no support and their personal working culture is at odds with that of the institution. This can be an isolating experience, which often demotivates capable individuals and holds them back from achieving their potential. Those who, instead, do all their studies locally, have other difficulties. Hence, within African institutions, there are many individuals who have the potential to achieve much more than is possible within their current working cultures.

Institutional changes in working culture can come from networks across universities and across countries. Innovations could spread virally through such networks, if tools and external incentives can be provided that enable staff innovate and grow through their involvement (Stern, 2014). ADI attempts to provide the tools and support this process, but the working culture needs to grow locally. Such networks can develop into associations or societies, which can create incentives supporting their working culture.

3. CASE STUDIES

Cobb's framework aligns well with the goals of ADI and we found that climate data and problems can be useful in achieving the radical change proposed by these principles. His caveats raised issues around implementation, which clarify the cross-institutional approach that ADI is taking to support African institutions to move in this direction. Cobb's framework has proved useful to deepen reflections on where we are heading, and it remains to communicate the practicalities of implementation. Giving learners the opportunity to use data to explore relevant questions is also aligned with the ideas to work in a context, as emphasised by the Gaise guidelines (Franklin et al., 2007). This section presents case studies of climate-focused research, which are relevant in African contexts. The case studies illustrate how to start with meaningful contexts, which provide rich learning paths. The first case study is about the use of climate information for farmers' decision making and the second is about the use of renewable energy.

Case study 1: Investigation of the variation of rainfall for better decisions in agriculture

Across the African continent, rain-fed agriculture in small-holder agricultural systems remains the norm. The Participatory Integrated Climate Services for Agriculture Initiative (PICSA; University of Reading, 2019) provides access to climate information and decision-making tools and has been shown to support 'farmer innovation' (Clarkson, Dorward, Osbahr, Torgbor, & Kankam-Boadu, 2019). PICSA – as illustrated diagrammatically in Figure 1 – makes extensive use of historical climate data. As an educational programme, it also provides numerous opportunities to bring in different statistical approaches leading to a broad range of learning outcomes. The climate information is shared with farmers for three different time periods. The first panel in Figure 1 is headed *Long before the season* and refers to the historical climate information (mainly precipitation).

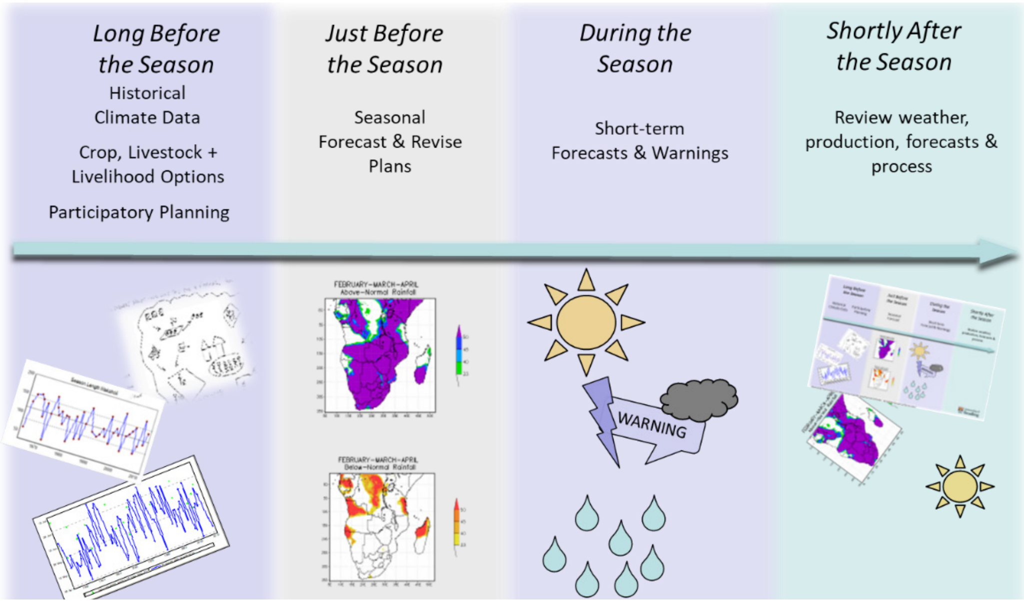

Figure 1. Overview of the PICSA process

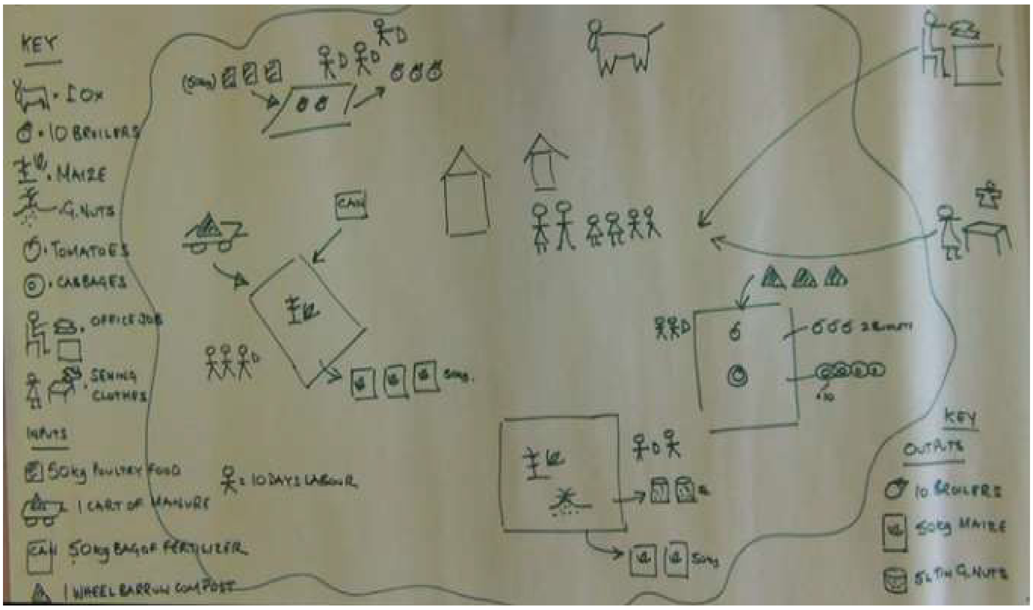

Figure 2. An example of a resource-allocation map produced in PICSA

The main capacity-building instrument in PICSA is a one-week workshop, designed for people (often agricultural extension officers) who work with farmers. PICSA is also a popular case study in the one-week mathematics camps, organised for school children as part of the African Data Initiative. There the statistical tasks include the production of time-series graphs. The children also gain ideas of the other, largely participatory, activities, which include constructing resource-allocation maps (see Figure 2) for families to assess, which of their activities could benefit most from the knowledge of climate risks.

One challenge in PICSA is, from the daily records of rainfall and temperature, to produce graphs that are shared with farmers. The production of the graphs is a task for the staff of the National Meteorological Service (NMS) in each country, and this is partly through in-service training for their staff. The diagrams need to be sufficiently clear so that they can be shared with farmers, some of whom have no formal education (Figure 3, Left). These diagrams are not just shown to the farmers, but are used by them to judge risks of alternative farming practices. These graphs have been important in helping farmers to consider changes in their current practices. There is (correctly) widespread belief in climate change, consistent with the results from analyses of temperature data. However, in the tropics, the essential element is rainfall. People also believe that rainfall has changed substantially due to climate change. Such a view may result in famers' feeling that there is little they can do (as it is due to climate change), whereas the information from the data can differ from people's perceptions (Rao, Ndegwa, Kizito, & Oyoo, 2011; Stern & Cooper, 2011).

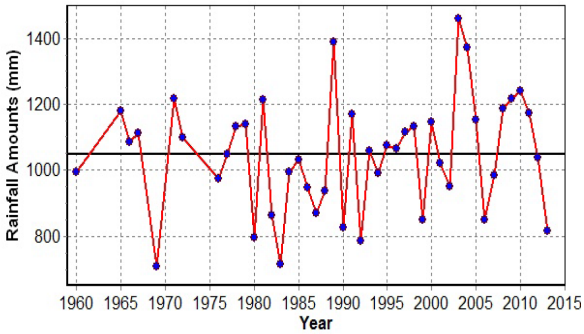

Figure 3. Farmers interpreting a rainfall graph (Left) like the time-series diagram (Right)

Such graphs (see Figure 3, Right), have not been widely shared with farmers or extension-workers in the past. They indicate that the major current issue with rainfall is variability and that this variability has not really changed from the issues that affected their parents or grandparents. Clearly, if temperatures have changed, the climate elements are interlinked, and hence changes in rainfall will become evident in the future. But, for now, coping with the variability is the main concern; the issues concerned with rainfall are local, and hence PICSA tries to help farmers to cope better with the current variability in rainfall. This realisation opens the door for many to seek better options for their farming practices. And statistics by itself is all about studying and coping with variability (Cooper, Dimes, Rao, & Shapiro, 2008).

Graphs are prepared for various aspects of precipitation, including the start and length of the rainy season, the total rainfall within a season, and the length of the longest dry spell. These graphs are also prepared for various stations that are close to the farming communities under scrutiny. Preparing such diagrams starts with daily data, which is often tens of thousands of rows, and requires choices. There are open research topics, which relate to the production of the graphs. Although easy to produce, once the data are available and in shape, getting the required data, with a long record without too many missing values and close to each farming community, is challenging. In many African countries, the official records are relatively sparse even when they have not been interrupted for social or environmental reasons. There are networks of volunteer stations and opportunities for satellite data to complement the official data. Section 4 provides more detail on the challenges to get or clean data.

The PICSA process has been evaluated using a mixture of qualitative and quantitative data (Clarkson, Dorward, Osbahr, Torgbor, & Kankam-Boadu, 2019). The approach has proven to be exceptionally effective in encouraging farmers to change their current practices to improve their crops. Some of the details that can be found in the data relate to complex social questions such as who is making what sort of innovation and why? Analysing the data from these surveys is also possible as it has been anonymised and made available as open data.

Many statistical concepts can be linked to this case study. Climate data is an example of routinely collected data; it forms a time series and has elements of periodicity (seasonality) and trend (climate change), which are intuitively understandable. There is also the possibility of including spatial components in an analysis. The idea of investigating climate change includes the collection of survey data to compare the reality with people's perceptions. Ideas of communicating data to non-expert audiences are involved, with the graphs produced to enable farmers to evaluate risks. The subject of monitoring and evaluation is introduced through the qualitative and quantitative work to evaluate the effectiveness of the approach. It is also possible to introduce ideas on experimentation through the comparison of different options available to farmers.

This wide range of topics related to a single context with open research questions provides an example to rethink how we could structure courses. A case study like this gives opportunities to create interesting activities or project work from school to postgraduate students.

Case study 2: Use of renewable energy This case study illustrates analyses of data NMS have to investigate the potential of renewable energy. Key variables include sunshine hours (and other measures of radiation) as well as wind speed and direction. This is a broad area and most studies would focus on electricity generation. There are other local energy problems that are relevant in the African context.

The importance of solar cooking is an example of a local renewable energy problem, which may not otherwise gain global attention. The context of this case study is well provided through information from Solar Cookers International (n.d.) whose website states that "3 in 7 people today lack modern fuel to cook food." Figure 4 shows a simple solar cooker that reflects sunlight onto a cooking pot. It serves as a slow cooker that can sterilise water in one hour and cook a meal in about three hours. Currently, many rural families in Africa still use wood for their cooking. This cocker could be a substitute – but only on days when the length of sunshine is sufficient. Hence, with sunshine data, evaluating the proportion of days that may be used for cooking is a useful task. When this cooker was presented at a workshop in Australia, many people thought it could also become an innovative and educational addition to a barbecue!

Figure 4. A solar cooker constructed from cardboard with a pot centred inside a plastic bag

In Africa, questions on climate are often dominated by an analysis of rainfall data. This case study is an illustration of the importance of analysing other climatic elements. Questions on risks are also often ignored. The solar cooker can only be used when there is sufficient sunshine. Hence, the main topic considered here is an evaluation of the proportion of days when the cooker can be used.

One potential problem is the lack of data. Sunshine records are usually available from the NMS but the data is not always computerised and then often just recorded as the daily number of sunshine hours. Where such records are available, they are often patchy and for only relatively few sites in each country. However, satellite data has recently become available as gridded data of 5-10 km every 15-30 minutes. Could this be used instead? If so, we could answer more detailed questions such as how often lunch or dinner or both could be prepared? The idea of bringing in other data sources is relevant here. Satellite estimates are known to be reliable sources of information for radiation and hence sunshine. They are freely available for more than 30 years. This makes it possible to introduce a whole new set of skills related to merging data sources and combining data at different levels. These questions should not divert attention too much from the central questions of using the solar cooker; however, given the gridded spatial nature of the satellite data, it is now possible to introduce other ideas such as producing maps of the results into the curriculum.

These two case studies illustrate some of the rich research activities that can use climate data and have local relevance across many African environments. They are a long way from some of the existing examples of how climate data has been used with students (Pearson Education, 2018; Witt, 2014) partly because of the skills they engage students with, but also because of the size and richness of the data we propose that should be used for the tasks. The depth of these case studies illustrates how the same case study can be revisited by multiple activities across different courses.

4. DATA

The value of the climatic case studies comes partly from the richness of the questions and contexts surrounding them but primarily from the substantive climatic data that is needed to tackle them. All countries have historical data on climate variables and a growing amount of such data is now freely available. There are other sources of climate-related data becoming freely available that complement and add value to the traditional station data. All these data types challenge our traditional teaching of statistics as they are interesting, relevant, and accessible that can help address questions that interest learners but they do not correspond well to a traditional basic statistics curriculum. This section presents and relates data, which is available for the contexts provided in the case studies. It raises questions about how we should be using and integrating data from various sources.

Historical station data is the key data source for studying climate. Australia provides an example of how easy it can be to access historical rainfall data. For example, data for Sydney over 130 years, from 1885 to 2019, is freely available online (Australian Government, Bureau of Meteorology, n.d.). Many African NMS are also prepared to provide daily data for at least one station in their country. As an example, Figure 5 shows daily data from Dodoma, Tanzania, supplied by the Tanzania Meteorological Authority. This is for four variables, rainfall, temperature, and sunshine duration, and dates back to 1935 for the rainfall data. This data is in the R-Instat data library, as are data sets from various other countries. Most African countries have many volunteer stations in addition to the official ones managed by the National Meteorological Service. Through ADI, records were collected on daily rainfall and computerised for over 50 stations in Western Kenya. This resulted in more than 600,000 records of daily rainfall plus data on other variables from seven of the stations. Students computerised the data using a climate data management system. Computerised copies were returned to the stations and the owners agreed that this data can be used freely. They were stored on the Harvard Dataverse system (Musyoka, 2015) and are also available in the R-Instat data library.

Data View

	Date (D)	Year	month	doy_	Rain	Tmax	Tmin	Sunh
14260	1974-01-15	1974	Jan	15	0.2	25.4	17.3	2.7
14261	1974-01-16	1974	Jan	16	0.0	29.5	17.4	10.1
14262	1974-01-17	1974	Jan	17	0.0	30.8	18.2	11.3
14263	1974-01-18	1974	Jan	18	0.0	32.0	18.1	11.6
14264	1974-01-19	1974	Jan	19	0.0	31.0	18.4	8.7
14265	1974-01-20	1974	Jan	20	0.0	30.2	19.5	6.4
14266	1974-01-21	1974	Jan	21	0.9	31.5	18.0	10.8
14267	1974-01-22	1974	Jan	22	0.0	31.8	18.4	11.6
14268	1974-01-23	1974	Jan	23	0.0	30.6	20.4	7.5
14269	1974-01-24	1974	Jan	24	0.0	32.0	18.7	10.7
14270	1974-01-25	1974	Jan	25	0.0	31.0	19.0	7.0
14271	1974-01-26	1974	Jan	26	0.0	31.0	18.0	4.2

Figure 5. Daily climate data for a site in Tanzania

Data View					
	lon	lat	time	time_date (D)	SDU
20	35.8	-6.15	19	1983-01-20	8.083
21	35.8	-6.15	20	1983-01-21	5.830
22	35.8	-6.15	21	1983-01-22	9.523
23	35.8	-6.15	22	1983-01-23	9.910
24	35.8	-6.15	23	1983-01-24	7.635
25	35.8	-6.15	24	1983-01-25	8.607
26	35.8	-6.15	25	1983-01-26	10.876
27	35.8	-6.15	26	1983-01-27	NA
28	35.8	-6.15	27	1983-01-28	9.815
29	35.8	-6.15	28	1983-01-29	8.952
30	35.8	-6.15	29	1983-01-30	11.009

Figure 6. Satellite data on sunshine duration (SDU)

Data from satellite observations is a second key source. It is also available for various variables since 1983. Such data is freely available for all European and African countries (Schulz, Albert, Behr, & Caprion, 2009) and provide an addition and possibly an alternative to the station data. Examples have been added to the R-Instat library. They are in a format called NetCDF, a format, which is accessible via an R package, which has been extended in R-Instat to transform the data into the same tidy shape as shown for the station data in Figure 5. Figure 6 shows the data for the grid point closest to the station at Dodoma, Tanzania.

These initial two data sources are both big and complex. Having access to both opens up interesting and important questions related to our cases studies. For example, for renewable energy such as in the previous section, having access to satellite estimates of sunshine duration rather than the daily station records, changes the data picture substantially. Questions arising could include: Should just the station data be used, or could the satellite-based data be used instead? Perhaps there should be a preliminary study to compare the quality of the two sources? Such types of questions are currently rarely asked in statistics courses but are growing in importance as the sources of data increase.

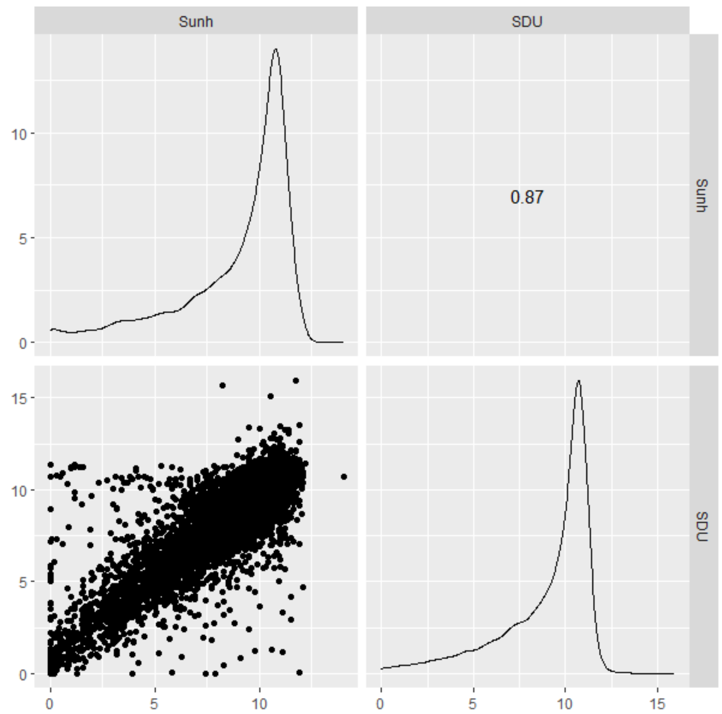

Figure 7. Relation between station (Sunh) and satellite (SDU) estimates of daily sunshine duration (hours) at a location in Tanzania – Scatter plot and density plots

Based on the results of the comparison, then maybe the satellite-based data could be used to quality-control and fill the gaps in the station data. Then the augmented station data could be used. Or perhaps the comparison would justify the use of the satellite data alone. That would have the big advantage that suddenly each degree square has the equivalent of about 400 stations! Many African countries complain that they lack data, while the availability of satellite data provides an example of the converse. The danger is more of being overwhelmed by the mass of data. Coping with large data sets is an important skill for future students and this includes understanding, which questions the big data can answer and which it cannot.
For our sunshine example, an initial comparison of the two data sets is shown in Figure 7. This is promising. The shape of the two distributions of sunshine hours is similar and the correlation is 0.87. The scatterplot indicates a few days to be investigated more closely, where one data set has indicated a lot of sunshine, while the other has not. These cases are relatively few, however.
This is just a start of the comparison, because any seasonality is hidden in the displays in Figure 7. Breaking down the data to monthly level, Figure 8 shows that the distribution of sunshine duration remains similar for the two sources with correlations (not shown) ranging from 0.50 in the dry season to 0.90 in the rainy season.
These discussions on the appropriate data to use are part of the broader teaching we would encourage. The discussion above also shows how this problem is one that students could return to, in several of their courses.

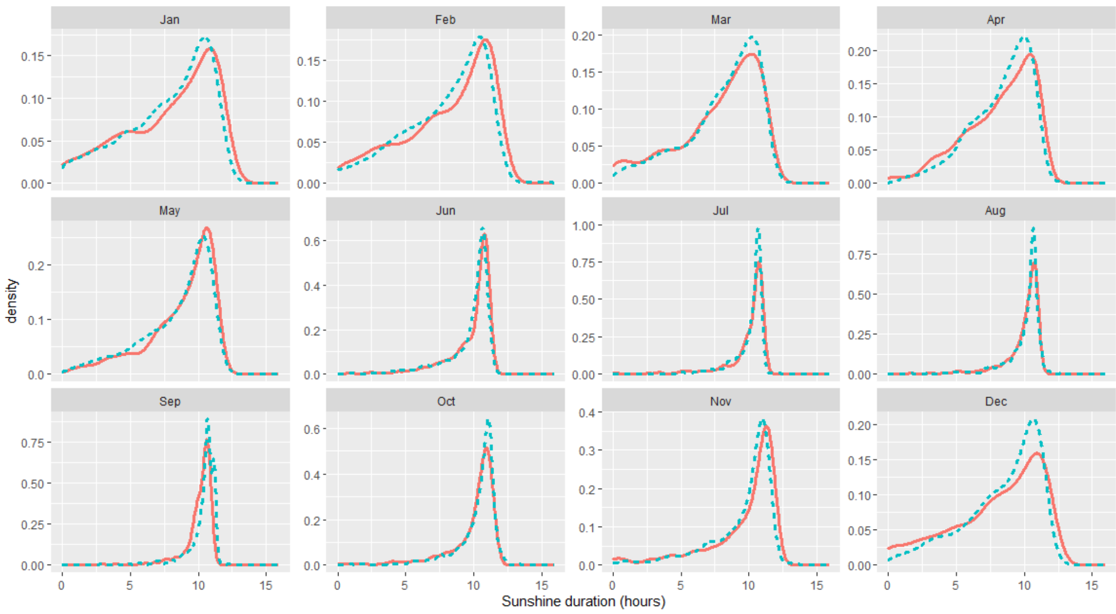

Figure 8. Monthly distribution of sunshine hours at a location in Tanzania – station (red solid curve) and satellite estimates (turquois dashed curve)

The ERA5 data (Copernicus, 2018) is a further potential game changer for many climate problems. These are re-analysed data, resulting from the use of global climate models used to provide short-term forecasts. They are at about a 30 square kilometres grid, hence there are only about 9 or 16 "stations" per square degree. But, on the positive side, these data sets are available for a very wide range of variables. They are on an hourly basis and they are global. They date back to 1979. The data sets are freely available.

The ERA5 data includes hourly sunshine data, so there are now three sources to compare for the example provided above. The ERA5 data also includes precipitation, which is so important for many applications in Africa, including the PICSA project. Rainfall estimates are not available from the EUMETSAT data, though the data base includes cloud temperature data, which is used independently by other project teams to estimate rainfall (e.g., TAMSAT, see Maidment et al., 2017; CHIRPS, see Funk et al., 2015).

Many other climate datasets are also available, including through R packages; for example, the open-air package (Carslaw & Ropkins, 2012) provides 65,000 rows of hourly pollution variables, plus wind speed, and wind direction for 7 years for a site in the United Kingdom. Such datasets are usually in a similar format to the one shown in Figure 6; this is ideally primary data, or at least as daily records. Avoiding semi-processed data, e.g., as monthly totals, or annual extremes, for teaching and research, is important as semi-processed data sets usually do not provide enough context. Often, it is useful for students to include the initial data summary step. Otherwise, both the primary and the summary data can be provided. A bonus is that outliers in the summary data can then be investigated via an examination of the primary data.

The presentation of data in this section illustrates that large rich datasets are freely available and locally adapted to the climate case studies even in the African context. These can provide research-focused context-specific education as advocated in Cobb (2015). Access to such rich and varied data also opens the door to educational opportunities that expose students to a wide variety of statistical concepts and data skills.

5. SKILLS

Statisticians of tomorrow will require a broader repertoire of skills than they currently have. This is one of the reasons Cobb (2015) claims that radical reform, not just small changes, is needed. Through the case studies in context, and varied related sets of data, we illustrate some of the broader skills that can be acquired through working with climate data.

The provision of real problems in context provides an opportunity to expose students to a range of non-technical and soft statistics-related skills that are harder to include in a traditional setting.

Communication of results is a highly important skill, particularly given the need to work in multidisciplinary teams. The PICSA case study presents essential questions around communication:

- How do you communicate climate risks to non-technical audiences?
- What graphs are appropriate to show climate variability?
- How would you discuss these graphs in a workshop to farmers?
- What makes a graph readable?

These questions have been addressed in the implementation of PICSA and statisticians, with the appropriate experience, can make valuable contributions. These questions illustrate issues that can facilitate students gaining valuable skills in communication and presentation of data and information inherent in data and in results of statistical analyses.

The PICSA case study also introduces skills that usually are acquired through considering the design and collection of additional complementary data around the routinely collected climate data.

- How do you design a survey to understand farmers' perceptions of climate change?
- Should a survey or individual case studies be used, or should both frameworks be used?
- How do you design questions on climate change that are not biased?
- How do you evaluate the success of a project that works with farmers?

Thus, the examples lead into more areas of interest to students that introduce other forms of data and important general skills around how to collect and process such data.

Section 4 illustrated climate data and their relative merits for innovations in teaching statistics: varied sets of large, real, and complex data in a context rich with questions. Here, we describe some of the important skills that are necessary to make progress on the questions and that can be acquired through exploration of these data sets.

When reasonably large data sets are used, descriptive statistics becomes a more fascinating subject. With daily historical climate data for a single station over thirty years, there are over 10,000 observations. These include numerical variables, some normally distributed but some with other distributions, such as rainfall, which has many zeros, or sunshine hours, which is bounded below by zero and above by the day length. As mentioned in Section 2, this leads naturally to consider the structure of the data, e.g., wet/dry days or none/some/totally sunny days, an important general data skill that is intuitive in climate data. There are also natural categorical variables, such as year and month, which lead to examples of multi-level data through calculating monthly or annual summaries. Here, we are not considering a conventional multi-level analysis, but simply giving users the experience of handling multiple-linked rectangles of data. That is important for almost any type of analysis, e.g., a simple survey with individual and household levels. With data from multiple stations, a spatial element is added. This leads naturally to simple mapping, yet another topic that is rarely included as part of descriptive statistics courses.

Describing and exploring the primary (often daily) data before moving to the summary data to be analysed is an important skill for statisticians. One reason for keeping access to the daily data, while analysing monthly summaries, is for quality-control purposes. If a monthly summary seems odd, there are statistical analyses to check whether this is a real outlier. Being able to look in more detail at the source of a problem, i.e., back to the daily records, is usually highly effective. Often, in teaching, the examples start with semi-processed data and move straight to the method of analysis. Starting with the primary data leads into discussions on another important set of skills for statisticians, namely the abilities to clean, tidy, and manipulate data.

With access to relatively large datasets from different sources, as part of statistics education, skills to organise data efficiently becomes a natural part of the training. This may be through work with spreadsheet software or include programming. Examples of routinely collected climate data often provide a gold mine for data cleaning exercises. The historical data is still often recorded on paper sheets and later entered into a computer, allowing mistakes to be introduced. Some are easy to detect and correct with a spreadsheet, such as a rainfall of 54.3 mm being recorded as 543 mm. Others require more sophisticated methods for detection, such as unrealistic day-to-day variations in maximum and minimum temperatures. These cleaning processes can be explored from school level to higher university levels, as different degrees of details of correction methods.

The concept of tidy data is also relevant (see Wickham, 2014), where the shape of the data may not be in a natural structure of observations as rows and variables as columns. In Wickham, the most untidy example is a climate data set, which needs several transformations to get the data into a tidy format. This is less untidy than some we have had to process!

Going beyond a single source of structured data, such as adding satellite data to station data, implies that students need skills to merge data from different file formats, with different structures. The distinction between statistics and computer science becomes blurred and this introduces useful discussions around the data science movement and the potential advantages and disadvantages of different data sources. We do not claim that students must always start with messy data, but they should be exposed to examples in their training if they are to be useful in the jobs of the future that require skills to work with increasingly large and complex data.

R-Instat “was designed to be simple-to-use while encouraging good statistical practice. It also had an extra customized menu specifically for the analyses of historical climatic data” (Stern, 2017, p. 4). Hence it permits users to concentrate on the objectives, the data, and the analyses rather than on mastering the tool. However, it is not an essential tool. It is simply a front end to R. Hence, an alternative is to use R directly. And many analyses can be done using only the general facilities in R-Instat for organising and analysing the data, rather than the special climate menu. These general facilities are like those in any statistics package.

The extension of skills proposed in this section has similarities to those proposed for data science with some differences. The most notable, that is central to many promoting data science, is programming. The omission here does not reflect its lack of importance; rather, it indicates that programming and data skills can be acquired largely independently. Programming remains essential for advanced analyses, particularly methods such as machine learning, and the importance of programming skills in the education of the next generation cannot be overstated. However, good data skills can be acquired without students being expert coders, and this opens the door to making statistical education more accessible.

Given the rapid growth of data in all aspects of society, the skill of working with data needs to be accessible to a broad audience. Strong data skills are needed for students across many disciplines, from health to environment, politics to business, and everything in between. Hence, these data skills are needed for service teaching as well as the training of specialists.

In African contexts, separating data from programming skills is important as many school students do not yet get exposed to either. We support the inclusion of both in future school curricula but currently such skills are often not introduced properly before tertiary education. The need for data skills warrants an inclusive approach focussing on skills with low entry points; this need ties back to the flattening-prerequisites discussion in Section 2 and leads to an approach, which cuts across academic levels.

The emphasis on softer skills, realistic exercises, and experiential learning does not undermine the potential for learning also the more complex details of the methods. On the contrary, such an approach could embed that within a broader problem-solving skill set. There have been recent efforts to develop problem-solving and project-based introductory statistics courses with a focus on data skills, independent of the choice of statistical software by the lecturers and the students (Parsons, Stern, & Stern, 2019; Dierker et al., 2018). This problem-solving approach focuses on data skills and is almost independent of mathematical background or programming skills. The same problems can be presented to all students with those having better mathematical background or programming skills able to access additional tools and methods.

The case studies presented in Section 3 have a meaningful context within the area of climate. Indeed, for statisticians to fully engage in the research questions within the case studies, they would require a combination of statistical understanding, together with some understanding of the domain itself, plus the softer skills mentioned above. If the training of the statisticians of tomorrow includes a problem-solving approach that combines statistical learning with an appreciation for the need to engage in different subject areas, this could lead to developing statistical specialists who are able to play a more central role in transdisciplinary projects.

6. CONCLUSIONS

We have presented a vision of how future statisticians could be trained by engaging them across disciplines and all academic levels in meaningful experiential learning with contextually relevant questions, substantive data, and practical skills. We have used case studies on climate issues to illustrate how such experiences could be achieved. Our vision is that working with data should be seen as a basic skill rather than exclusively for specialists. A statistical specialist's role shifts from being a consultant, brought in when needed, to that of a full collaborator.

In Cobb (2015), a strong case is made for radical reform in the undergraduate training of statisticians. This paper has shown that the suggested framework resonates strongly with the goals of the African Data Initiative (ADI). The vision applies to all level of students and also to those needing service courses in statistics, including professional staff that need in-service training.

Our ideas have been shaped by projects on climate issues and we think that this context is not only suitable to communicate our challenges and criteria for innovative courses in statistics. The context of climate forms also a suitable area to build resources that could support the suggested kind of reform of statistics education.

The context has been presented in the form of case studies that are supported by real and large open datasets; a range of skills that we think are the goals of this intended reform, has been illustrated; our elaborations should highlight how these case studies can enrich student learning. The goal for statistics teaching embraces also in-service training for the staff of National Meteorological Services (NMS). As the main custodians of the historical climate data for their country, the staff could benefit from stronger statistical skills to make full use of their data.

Possibly the most surprising insight we have gained from this study is the liberation of statistical problem-solving from both its mathematical foundation and programming skills. We do not want to underestimate the importance of either; on the contrary, we actively support the strengthening of all three areas in education. However, recognising that one is not a prerequisite of another enables a change in the order, in which the skills can be taught and acquired. Internationally there is a precedent, with New Zealand including statistical thinking as a central component of both primary and secondary education (Forbes, 2014). This realisation means that we can worry less about the sequential order and instead think about creating positive interactions between these three key areas of statistics.

This work is part of the African Data Initiative (ADI), a collaboration to support improvements in African statistics education through providing open tools and resources together with initiatives to assist those who are trying to instigate positive change on the continent. ADI is currently transitioning from a focus on software development into a second phase of assembling a broad set of resources that can facilitate change in African institutions. We have outlined a route for future ADI efforts both in terms of the focus on statistical problem solving and the use of climate as a key area of application.

The last 50 years have seen considerable change in statistical methods and reform in statistics training. Much has been driven from outside Africa and many courses on the continent have not yet been affected. ADI is trying to change this process by embedding cutting-edge initiatives in Africa from the start.

The ideas exposed in this paper may lead to theoretical thinking on curricular reforms as they show how such reform can be developed into initiatives for low-resource environments that are adapted to the challenges faced by trainers across Africa. This approach can support potential international partners to join ADI by providing a possible research framework to underpin collaborations and communicate the global relevance of the outcomes. The core of our approach remains valid for innovations in statistic education in general, to prepare the future generation of statisticians with a wider repertoire of skills and views.

ACKNOWLEDGEMENTS

We would like to thank the referees whose detailed comments and suggestions helped significantly to improve this paper. We also wish to thank the various NMS who we have collaborated with over a number of years and who have provided data freely for the teaching and learning of statistics.

REFERENCES

Amrhein, V., Greenland, S., & McShane, B. (2019). Retire statistical significance. *Nature, 567*, 305–307. [Online: www.nature.com/articles/d41586-019-00857-9]

Australian Government (n.d.). *Climate statistics for Australian locations*. Bureau of Meteorology.
[Online: www.bom.gov.au/climate/averages/tables/cw_066062.shtml]

Bilgin, A., & Petocz, P. (2013). Students' experience of becoming a statistical consultant. In *Proceedings of the Joint IASE / IAOS Satellite Conference on "Statistics Education for Progress"* (pp. 1–6). The Hague, The Netherlands: International Statistical Institute und

International Association for Statistics Education. [Online: iase-web.org/Conference_Proceedings.php?p=2013_Stats_Education_for_Progress]

Borovcnik, M. (2018, May). The potential of statistical consulting for statistics education. *Plenary talk at the 15th International Students Statistics Colloquium*, Istanbul, 4–5 May. [Online: doi.org/10.13140/RG.2.2.35255.01447]

Carslaw, D. C., & Ropkins, K. (2012). *openair* – An R package for air quality data analysis. Environmental Modelling & Software, 27–28, 52–61.

Clarkson, G., Dorward, P., Osbahr, H., Torgbor, F., & Kankam-Boadu, I. (2019). An investigation of the effects of PICSA on smallholder farmers' decision-making and livelihoods when implemented at large scale – The case of Northern Ghana. *Climate Services, 14*, 1–14.

Climsoft (n.d.). *Free open source climate data management system*. Open-Source Community project. London: Met Office, Department for Business, Energy and Industrial Strategy of the United Kingdom of Great Britain and Northern Ireland. [Online: www.climsoft.org/]

Cobb, G. W. (2015). Mere renovation is too little too late: We need to rethink our undergraduate curriculum from the ground up. *The American Statistician, 69*(4), 266–282.

Cooper, P. J., Dimes, J., Rao, K. P., & Shapiro, B., Shiferaw, B., & Twomlow, S. (2008). Coping better with current climatic variability in the rain-fed farming systems of sub-Saharan Africa: An essential first step in adapting to future climate change? *Agriculture, Ecosystems and Environment, 126*(1–2), 24–35.

Copernicus (2018, June 14). *ERA5 hourly data on single levels from 1979 to present*. [Online: doi.org/10.24381/cds.adbb2d47]

Davenport, T. H., & Patil, D. J. (2012). Data scientist: The sexiest job of the 21st century. *Harvard Business Review, 90*(10), 70–76.

Dierker, L., Evia, J. R., Singer-Freeman, K., Woods, K., Zupkus, J., Arnholt, A., . . . Rose, J. (2018). Project-based learning in introductory statistics: Comparing course experiences and predicting positive outcomes for students from diverse educational settings. *International Journal of Educational Technology and Learning, 3*(2), 52–64.

EUMETSAT (n.d.). *Satellite application facility on climate monitoring*. CM SAF Web User Interface. Darmstadt: European Organisation for the Exploitation of Meteorological Satellites. [Online: wui.cmsaf.eu/safira/action/viewHome]

Fletcher, L. (2014). Statistical consultation as part of statistics education. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics* (pp. 1–4). Voorburg, The Netherlands: International Statistical Institute and IASE. [Online: iase-web.org/Conference_Proceedings.php?p=ICOTS_9_2014]

Forbes, S. (2014). The Coming of age of statistics education in New Zealand, and its influence internationally. *Journal of Statistics Education, 22*(2), 1–19. [Online: doi.org/10.1080/10691898.2014.11889699]

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report*. American Statistical Association. [Online: www.amstat.org/asa/education/Guidelines-for-Assessment-and-Instruction-in-Statistics-Education-Reports.aspx]

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., . . . Michaelsen, J. (2015). The climate hazards infrared precipitation with stations – A new environmental record for monitoring extremes. *Scientific Data, 2*: 150066. [Online: www.nature.com/articles/sdata201566]

Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG) (2014). *A world that counts: Mobilising the data revolution for sustainable development*. Report prepared for the UN Secretary General by the Independent Expert Advisory Group on a Data Revolution for Sustainable Development. [Online: www.undatarevolution.org/report/]

ISLP (n.d.). *International Statistical Literacy Project*. ILSP Poster Competition [Online: iase-web.org/islp/]

Jersky, B. (2002). Statistical consulting with undergraduates – A community outreach approach. In *Proceedings of the Sixth International Conference on Teaching Statistics* (pp. 1–5). Voorburg, The Netherlands: International Statistical Institute and IASE. [Online: iase-web.org/Conference_Proceedings.php?p=ICOTS_6_2002]

Maidment, R. I., Grimes, D., Black, E., Tarnavsky, E., Young, M., Greatrex, H., . . . Alcántara, E. M. (2017). A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa. *Scientific Data, 4*: 170063.

Ministry of Education (2012). Achievement objectives / Mathematics and statistics / The New Zealand Curriculum. NZ Curriculum Online. [Online: nzcurriculum.tki.org.nz/The-New-Zealand-Curriculum/Mathematics-and-statistics/Achievement-objectives]

Musyoka, J. (2015). Western Kenya climatic information. African Maths Initiative Dataverse. Harvard Dataverse. [Online: doi.org/10.7910/DVN/28472]

Parsons, D., Stern, R., & Stern, D. (2019). A problem-solving course in statistics for mathematical science students. In S. Budgett (Ed.), Decision Making Based on Data. Proceedings of the IASE Satellite Conference (pp. 1–6). The Hague, The Netherlands: International Statistical Institute and International Association for Statistics Education. [Online: iase-web.org/Conference_Proceedings.php?p=2019_Decision_Making_Based_on_Data]

Pearson Education (2018). *A level mathematics*. Pearson Edexcel Level 3 Advanced GCE in Mathematics. London: Pearson Education. [Online: qualifications.pearson.com/en/qualifications/edexcel-a-levels/mathematics-2017.html]

Rao, K. C., Ndegwa, W., Kizito, K., & Oyoo, A. (2011). Climate variability and change: Farmer perceptions and understanding of intra-seasonal variability in rainfall and associated risk in semi-arid Kenya. Experimental Agriculture, 47(2), 267–291. [Online: doi.org/10.1017/S0014479710000918]

Royal Statistical Society (2019). *Stats careers – Career stage – School (11–16)*. London: Royal Statistical Society. [Online: www.statslife.org.uk/stats-careers-11-16]

Schulz, J., Albert, P., Behr, H. -D., & Caprion, D. (2009). *Operational climate monitoring from space: The EUMETSAT satellite application facility on climate monitoring* (CM-SAF). Atmospheric Chemistry and Physics, 1687-1709.

Solar Cookers International (n.d.). [Online: www.solarcookers.org/]

Stern, D. (2014). Reflections on using technology to teach statistics in Kenya. In K. Makar, B. de Sousa, & R. Gould (Eds.), Sustainability in statistics education. *Proceedings of the Ninth International Conference on Teaching Statistics* (pp. 1–6). Voorburg, The Netherlands: International Statistical Institute and IASE. [Online: iase-web.org/Conference_Proceedings.php?p=ICOTS_9_2014]

Stern, D. (2017). Seeding the African Data Initiative. In A. Molnar (Ed.), *Decision Making Based on Data. Proceedings of the IASE Satellite Conference* (pp. 1–5). The Hague, The Netherlands: International Statistical Institute und International Association for Statistics Education. [Online: iase-web.org/Conference_Proceedings.php?p=2017_Teaching_Statistics_in_a_Data_Rich_World]

Stern, D., Ongati, O. N., Agure, J. O., & Ogange, B. (2010). Incremental modernization of statistics teaching and curriculum at Maseno University, Kenya. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eight International Conference on Teaching Statistics* (pp. 1–6). Voorburg, The Netherlands: International Statistical Institute and IASE. [Online: iase-web.org/Conference_Proceedings.php?p=ICOTS_8_2010]

Stern, R. D., & Cooper, P. J. (2011). Assessing climate risk and climate change using rainfall data – A case study from Zambia. *Experimental Agriculture, 47*(2), 241–266. [Online: doi.org/10.1017/S0014479711000081]

University of Reading (2019). *PICSA – Participatory Integrated Climate Services for Agriculture*. [Online: research.reading.ac.uk/picsa/]

Wickham, H. (2014). Tidy data. *Journal of Statistical Software, 59*(10), 1–23.
[Online: www.jstatsoft.org/index.php/jss/article/view/v059i10/v59i10.pdf]
Witt, G. (2014). Using data from climate science to teach introductory statistics. *Journal of Statistics Education, 21*(1), 1–23. [Online: doi.org/10.1080/10691898.2013.11889667]

DAVID STERN
Idems International
15 Warwick Road
RG2 7AX Reading
UK