FINAL REPORT

INTERNATIONAL WORKSHOP ON MATHEMATICS OF CLIMATE CHANGE AND NATURAL DISASTERS

Mathematics Threat Hazard

02 september 2017

List of Contents

1. Executive summary

This workshop brought together an extremely diverse set of world leading researchers, mathematicians, computer scientists, meteorolgists, statisticians, machine learning professionals, computational fluid dyanmics modellers, geodesists, engineers, biologists, geotechnicians, oceanographers and physicists were all represented.

Bringing these groups of people together is of paramount importance for the adaptation, mitigation and resilience of problems arising in climate change that require wide fields of expertise, a truly multidisciplinary effort is needed.

The workshop was organised into four working groups:

1. Modeling, Computer Architecture and Physical Process
2. Checking/Verification Time Series
3. Data Science & Machine Learning
4. Disaster Risk Reduction

Within each group, the dynamics allowed for the creation of a set of international collaborative projects between researchers to be taken forward. The resulting projects included extremely broad application areas, such as, model and data reduction through deep learning approaches, extreme event identification and uncertainty quantification through adjoint and Monte-Carlo methods, but also targetted projects to specifc scenarios, like the spread of disease through epidemiology, long range interactions between oceans and atmosphere and the predictability of landslides.

This resulted in a diverse set of high impact projects for which we plan to seek funding through appropriate research support agencies. This will allow for the fruition of the ideas initially developed here. Already applications are being developed resulting from these collaborations and further international knowledge exchange partnerships are being planned.

2. Workshop Schedule

AUGUST 29TH, 2017 (TUESDAY)

MORNING

08:30 Registration

09:30 Opening ceremony

10:00 Talk 1 - Haroldo Campos Velho-INPE/BR

11:00 Talk 2 - Jean Ometto-INPE/BR

12:00 Lunch

AFTERNOON

13:30 Talk 3 - Jair Koiller-SBMAC/BR and Leonardo Santos-Cemaden/BR

14:00 Talk 4 - Tristan Pryer-University of Reading/UK

15:00 Talk 5 - Luiz Fagundes-Edinburg/UK

15:30-16:00 Coffee-break

16:00 Talk 6 - Fangxin Fang-Imperial College/UK and Jeff Gomes-University of Aberdeen/UK

16:30 Talk 7 - Luciana Londe-Cemaden/BR

17:00 Talk 8 - Nicolas Rubido-Universidad de la República/UY

18:00 First day closing ceremony

19:00 Welcome Dinner (Ema Palace Hotel)

AUGUST 30TH, 2017 (WEDNESDAY)

MORNING

09:00-09:30 Working-groups formation

09:30-12:00 Working-groups Coordinator: explanation about the work dynamics

Participants will introduce themselves, indicating his/her institution, expertise, and current research.

12:00-13:30 Lunch

AFTERNOON

13:30-15:30 Working-groups: defining scientific questions

15:30-16:00 Coffee-break

16:00-17:00 Talk 9 – Marcelo Barreiro- Universidad de la República (UY)

17:00-18:00 Networking

AUGUST 31ST, 2017 (THURSDAY)

MORNING

09:00-12:00 Working-groups: discussion about data and methods

12:00-13:30 Lunch

AFTERNOON

13:30-15:30 Group report preparation

15:30-16:00 Coffee-break

16:00-17:00 Talk 10 - Tristan Pryer-University of Reading/UK

(Funding opportunities for Brazilians in the UK)

17h-18h Networking

SEPTEMBER 01ST, 2017 (FRIDAY)

MORNING

09:00-11:00 Working-groups: presentation preparation

11:00-11:30 Group-1: Report presentation (Plenary meeting)

11:30-12:00 Group-2: Report presentation (Plenary meeting)

12:00-13:30 Lunch

AFTERNOON

13:30-14:00 Group-3: Report presentation (Plenary meeting)

14:00-14:30 Group-4: Report presentation (Plenary meeting)

14:30-15:00 Group-5: Report presentation (Plenary meeting)

15:00-15:30 Coffee-break

15:30-17:00 Plenary discussion

SEPTEMBER 02ND, 2017 (SATURDAY)

09:00-10:30 Workshop final report

10:30-11:30 Letter addressed to support agencies, indicating priority scientific topics

11:30-12:00 Closing ceremony

12:00-15:00 Closing Lunch (Brazilian Barbecue)

3. Group discussion and future Projects

3.1 Working Group I - Modeling & Computer Architecture/Physical Process

Group: Tristan Martin Pryer (Moderator), Marcelo Barreiro (Moderator), Cassiano Bortolozzo, Jemima Tabeart, Luís Marcelo de Mattos Zeri, André Lanfer Marquez, Celso Von Randow, James Jackaman, Jean Ometto, Rodrigo S. Costa, Rong Zhang, Oliver Sutton, and James Targett.

3.1.1 Hydrological modelling.

A hydrological model is a simplification of a real world system to aid in the prediction and management of water resources.

The siltation of reservoirs is a consideration of paramount importance for regions that suffer from large amounts of erosion. This is particularly important for the maintenance and life span of the reservoir itself, allowing for the uninterrupted supply of water to both hydropower and, indeed, the population.

In the energy context, water is an extremely important resource. Most of electric matrix in Brazil comes from hydropower (68,1% in 2016 according to EPE), and it has a conflict evolved: the use of water to human consumption, agricultural uses and energy production. Know how are the real situation of the resource is crucial for decision makers.

Current practice in Cemaden is to use a Probability Distributed Model (PDM) [Moore, 2007] to forecast run-off/stream-flow. It is used due to the simplicity of the model, less demand on data and the computational time. This is necessary because of the limited operational time. Hydrological processes change in time, effected by many processes, for example, the change in land-use. The PDM model does not include distributional information on land-use, soil properties and other information such as rainfall, transpiration, etc., this is completely insufficient for hydrological simulation.

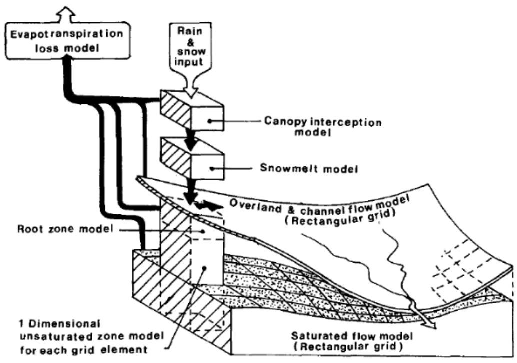

FIGURE 1: THE STRUCTURE OF THE SHETRAN NUMERICAL HYDROLOGICAL MODEL. SOURCE [ABBOTT ET. AL. 1986].

Another model is the SHETRAN package [Ewen et. al. 2000], this is a computational tool that uses a finite difference (Preissman) to approximate the solution of PDEs. This system of PDEs incorporates more information than the PDM model, including detailed surface, subsurface and channel flow processes along with detailed land use and soil information. It is currently not used for operational forecasting due to the high computational cost of the model, this is not feasible due to the limited operational time.

A particular challenge is how to reduce the complexity of the PDE system to allow for efficient approximation in short time. Adaptivity is crucial in the success of algorithms for geophysical multiscale problems with exactly this in mind. There are many notions of adaptivity, mesh refinement being one of the more popular, but local polynomial enrichment, movement of the nodes, are also possible, but the main goals are the balancing of computational time with accuracy of the numerical model.

References:

Ewen, J., Parkin, G., & O'Connell, P. E. (2000). SHETRAN: distributed river basin flow and transport modeling system. *Journal of hydrologic engineering*, *5*(3), 250-258.

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'connell, P. E., & Rasmussen, J. (1986). An introduction to the European Hydrological System—Systeme Hydrologique Europeen,“SHE”, 2: Structure of a physically-based, distributed modelling system. *Journal of hydrology*, *87*(1-2), 61-77.

R. J. Moore. (2007). The PDM rainfall-runoff model. Hydrology and Earth System Sciences Discussions, European Geosciences Union, 11 (1), pp.483-499.

3.1.2 Developing predictive theories for storm tracks' response to external forcing.

In the extratropical latitudes extreme meteorological events, like strong rainfall, heat waves and dry spells are related to the position and intensity of the jet streams and storm tracks as well as the occurrence of atmospheric blocking. In the context of climate change due to anthropogenic forcing, it has been shown that regional climate changes strongly depend on these dynamical aspects of climate. Moreover, the regions that do not show a robust response to radiative forcing have been tied to dynamical aspects of climate change (see Figure 1, Shepherd 2014) and have not changed significantly between IPCC AR4 and 5 reports. The energetic point of view that prioritizes global atmospheric warming and thermodynamic/radiative feedbacks makes sense to determine the global atmospheric response to greenhouse gases, but there is a need to include changes in atmospheric circulation to fully understand regional climate projections and reduce its uncertainties (Deser et al 2012, Sheperd 2014). Less uncertainty would produce more reliable projections to inform climate adaptation.

Thus, in order to reduce uncertainty in regional climate change projections it is extremely important to address the sensitivity of storm tracks to external forcing. In recent years it has been shown that the variability of storm tracks on seasonal and longer time scales depend not only on internal atmospheric dynamics, but also on the surface oceanic conditions (Chang et al 2014) and polar sea ice loss (Barnes and Screen 2015). Thus, it is a coupled ocean-atmosphere-ice problem.

On the other hand, climate dynamics is often studied using statistical methods like Principal Component Analysis, which has lead to “statistical modes of variability” thinking. This is a convenient way to condense the atmospheric behaviour but does not allow to develop a predictive theory for atmospheric dynamics, particularly for the behaviour of the response of storm tracks, blocks and jet streams to external forcing.

Given the complexity of the problem, there is a need to tackle it using a hierarchy of climate models of increased complexity. CMIP5 are very complex models and is very difficult to diagnose the physical processes involved as well as to determine if a particular response was expected.

To proceed, the first step is to develop process-based metrics based on the known theory, like one characterizing eddy-mean flow interaction, to diagnose the behaviour. Then use the hierarchy of models to develop an understanding and predictive theory that would tell what to expect about changes in jets and storm tracks under antropogenic forcing.

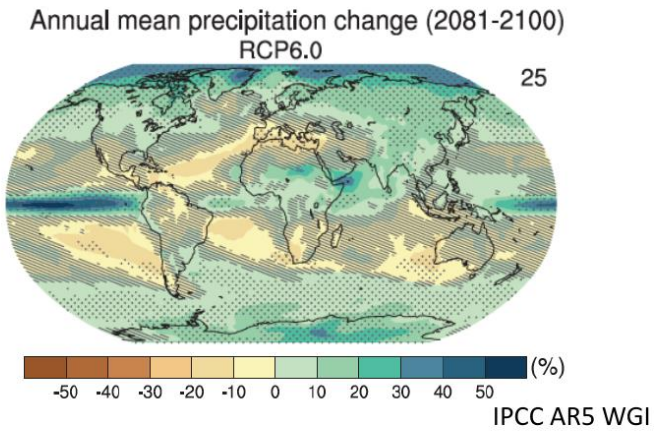

FIGURE 2: ANNUAL MEAN PRECIPITATION CHANGE FOR THE END OF THE 21ST CENTURY FOR AN ENSEMBLE OF CMIP5 CLIMATE MODELS. STIPPLED (ROBUST CHANGES); HATCHED (CHANGES SMALLER THAN NATURAL VARIABILITY); NO MARKS (INCONSISTENT MODEL RESPONSE).

Data-driven research can also help if it is connected to physical understanding. For example, stratify the data according to different physical/dynamical regimes and characterize each situation using tools from complex networks or machine learning. Do it in a controlled environment, as for example in reduced-complexity climate models and then try to find the associated signatures in more complex models at the same time determining how the inclusion of added physics modify the characteristics found before. A step-by-step process would then help understand how the dynamical regime would change under external forcing (e.g. radiative forcing).

References:

Shepherd, T.G. (2014) Atmospheric circulation as a source of uncertainty in climate change projections. Nature Geoscience, 7. pp. 703-708.

n precipitatiFrom Knutti and Sedlacek (2013).

Deser, C., A. S. Phillips, V. Bourdette, and H. Teng, 2012: Uncertainty in climate change projections: The role of internal variability. Climate Dyn., 38, 527-546.

Barnes, Elizabeth A. and James Screen, 2015: The impact of Arctic warming on the midlatitude jetstream: Can it? Has it? Will it?. WIREs Climate Change, 6.

3.1.3 Complexity and the cost of large scale simulations

Simulating large scale physical systems, such as atmospheric and ocean models, can be inherently very expensive. Recently, the sheer scale of these simulations is beginning to reach a limit dictated by the power consumption of the data centres in which the computations occur. In order to continue to improve the scale and accuracy of these simulations, it is therefore necessary to turn to new approaches to improve the efficiency of the methods in use. Broadly speaking, there are five design choices within these simulations which contribute to their cost: the choice of underlying physical model, the type of numerical method used, the resolution of the mesh, the precision of the arithmetic used in the computation, and the choice of computer hardware.

Four of these, the chosen model, the class of numerical method, its resolution, and the arithmetic precision all contribute to the error of a simulation — for a certain level of accuracy to be achieved, these must all be chosen to provide high enough accuracy. However, where any of these are set to a higher level of accuracy than is necessary, the computation will take longer and use more resources. The choice of what hardware the computation is run on will also affect the time, but also the power usage (Figure 3). For instance, a high performance computing cluster built around general purpose CPUs is highly flexible, but consumes a significant amount of energy per computation. At the other end of the spectrum, specialist hardware such as Field Programmable Gate Arrays (FPGAs) are often better suited to different methods than CPUs and use dramatically less power, but require very specialist programming.

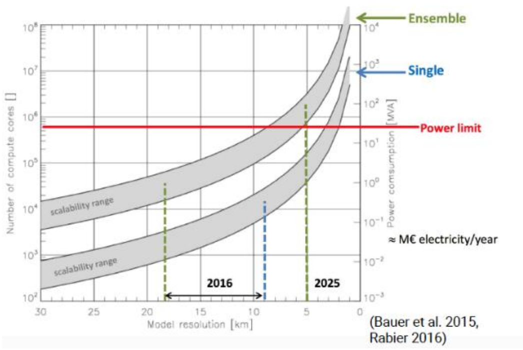

FIGURE 3: PREDICTED POWER USAGE ON LARGE SCALE CLIMATE MODELS.

In recent years, there have been several key advances which could particularly help decrease the cost associated with large simulations while maintaining similar levels of accuracy, or, equivalently, allow for larger scale or more accurate simulations for a similar cost. For instance, modern numerical techniques are often able to help here, by incorporating features such as tightly focussing computational effort to resolve features which required it. Automatic approaches for this include coarsening and refining their computational mesh or in response to estimates of the local quality of the simulation, or switching locally between high and low fidelity models, incorporating physical phenomena occurring at many different scales only where they are needed. Such techniques can retain the accuracy of a numerical scheme often without increasing the size of the simulation.

Coupled with this, the ECMWF have recently shown that usual 'double precision' arithmetic is often not required to ensure the accuracy of their simulations, observing that single precision was enough. Switching to reduced precision arithmetic, can offer sizable performance gains. Since the simulation error produced by this loss of arithmetic precision is less significant than the other sources of discretisation error, the quality of the outputs would not be noticeably affected. To take full advantage of this flexibility, however, requires specialised hardware, since conventional CPUs are generally restricted to either single and double precision. GPUs often additionally support half precision, while FPGAs have far greater flexibility as it is possible to choose nearly any level of precision.

Significant challenges remain before the efficiency gains offered by techniques such as these can be properly exploited by in the large scale simulations where they could be exploited. For instance, many of the numerical techniques still require further development before they can be applied to the complex models used in applications such as weather forecasting and climate modelling, and the precise savings possible in such models remain to be seen. Similarly, the impacts and potential savings of using reduced precision arithmetic remain to be investigated. In particular, it will be important to build a theoretical understanding of the behaviour of numerical methods in a reduced precision environment, and how fundamental properties such as conservation of mass or energy are affected.

References

Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, *525*(7567), 47-55.

N.P. Wedi, P. Bauer, W. Deconinck, M. Diamantakis, M. Hamrud, C. Kühnlein, S. Malardel, K. Mogensen, G. Mozdzynski, P.K. Smolarkiewicz. (2015). The modelling infrastructure of the Integrated Forecasting System: Recent advances and future challenges.

3.1.4 Correlated Observation Errors

Higher resolution of a numerical weather prediction system permits improved global and regional forecasts, and is key for natural disaster prediction. In order to obtain high resolution weather forecasts, more densely distributed observations are required. Currently up to 80% of observations are discarded ('thinned') to try and minimise the effect of correlated observation errors. Accounting for these error correlations would reduce the requirement to thin and means that better use of observations can be made. More observations allow a better understanding of small scale and local features and an improved ability to predict them [RBC15].

Correlated observation errors occur for three main reasons [JBB +]: errors in transformation of indirect variables, a difference in scale between the model grid and observations, and artificial correlations induced by errors in the model. Examples of transformation errors are seen for satellite and radar observations. Satellites measure radiances, whereas standard model variables are temperature, humidity etc. To compare model variables with observations a highly non-linear radiative transfer method is required. In order to calculate this transform in operational time (hours), approximate transfer techniques are needed. These approximations introduce correlated error in the observations.

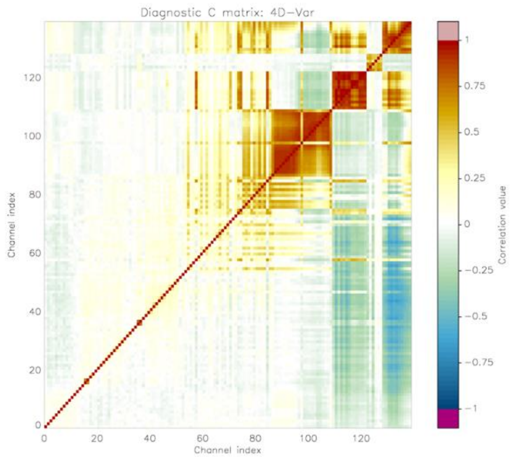

FIGURE 4: A DIAGNOSED CORRELATED MATRIX FOR THE IASI OBSERVING INSTRUMENT USING THE DESROZIERS DIAGNOSTIC [SDNEC14].

The introduction of correlated observation errors is very recent [WBE14], and is an active area of research both academically and operationally. Using correlated observation errors is more expensive computationally as a full matrix inversion is required. This, combined with a lack of knowledge about the structure of the correlations, has led most weather centres to use uncorrelated observation errors until relatively recently. One technique, commonly referred to as the Desroziers diagnostic, is almost exclusively used to diagnose observation error correlations. This technique relies on accurate knowledge of the errors in a corresponding model run, which is not necessarily true. Additional mathematical methodologies would increase confidence in the correlations.

The following problems have been highlighted as areas of potential interest:

- The development of alternative methods to determine observation error correlations. This could take the form of analytical methods arising from statistics, or operationally driven methods. One example of the latter is the use of conventional direct observations alongside satellite observations to extract correlation information.
- The identification and quantification of sources of error in models, with the aim of understanding how these introduce observation error correlations. With respect to satellite observations, this could be an analytic description of how correlations are expected to be introduced via an approximate transfer technique. Error attribution requires accurate knowledge of instrument calibration as well as good physical intuition.
- Temporal error correlations are currently not well understood. They will become increasingly important as knowledge of aforementioned forms of error correlations mature.

References

[DBCP05] G. Desroziers, L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation, background and analysis-error statistics in observation space. Quarterly Journal of the Royal Meteorological Society, 131(613):3385–3396, 2005.

[JBB +] T. Janjić, N. Bormann, M. Bocquet, J. A. Carton, S. E. Cohn, S. L. Dance, S. N. Losa, N. K. Nichols, R. Potthast, J. A. Waller, and P. Weston. On the representation error in data assimilation. Quarterly Journal of the Royal Meteorological Society. QJ-16-0326.R2.

[RBC15] Sabrina Rainwater, Craig H. Bishop, and William F. Campbell. The benefits of correlated observation errors for small scales. Quarterly Journal of the Royal Meteorological Society, 141(693):3439–3445, 2015.

[WBE14] P. P. Weston, W. Bell, and J. R. Eyre. Accounting for correlated error in the assimilation of high-resolution sounder data. Quarterly Journal of the Royal Meteorological Society, 140(685):2420–2429, 2014.

[SDNEC14] Stewart, L. M., Dance, S. L., Nichols, N. K., Eyre, J. R. and Cameron, J., Estimating interchannel observation-error correlations for IASI radiance data in the Met Office system. Quarterly Journal of the Royal Meteorological Society, 140:1236–1244, 2014.

3.1.5 Landslides Prediction

The monitoring of mass movement systems is one of the objectives of CEMADEN's natural hazards alert system, since many cities in Brazil (and in the world) suffer from the problem of mass movements in populated areas. Although many events are triggered by rain, the distribution of geological structures and layers has a key role in the landslide event. For this reason CEMADEN is implementing a system integrated to the SALVAR system for the monitoring of the underground water content in several critical areas in Brazil. The equipment corresponds to a pluviometer station coupled to sensors installed in wells that measure the humidity every 0.5m (Figure 5). The humidity values are acquired every hour in times without rain and every 10 minutes during the rainy period, with the sampling change done automatically. Some stations are already operating in Campos do Jordão for the last 2 years. During the drilling of the well, where the sensors go, a distribution of the geological layers is already obtained at that point. In addition, geophysical acquisitions are planned with the Resistivity and Inductive Electromagnetic methods in the station areas to determine the 2D variation of subsurface structures.

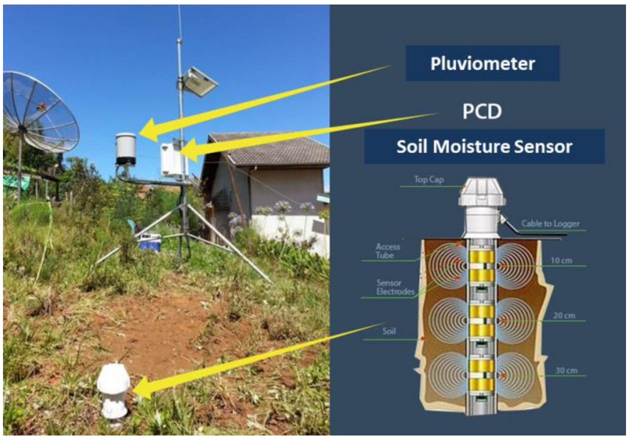

FIGURE 5: SOIL MOISTURE SENSOR INSTALLED IN CAMPOS DO JORDÃO – SP.

The initial use of the stations is to have a constant monitoring in the Situation Room, with the average values of humidity, critical values of humidity that produce a rupture and the current value of humidity. With an automatic warning when moisture values change rapidly. This system is still being developed and is not yet operational at CEMADEN. Although the system allows instantaneous visualization of the humidity in different layers in real time, its use for alert is restricted, once the alert will be in very short notice. In this way, a project is proposed to associate the rain forecast and the current situation of the geological layer's humidity to predict a possible rupture situation due the stability deterioration caused by the accumulation of water in the underground.

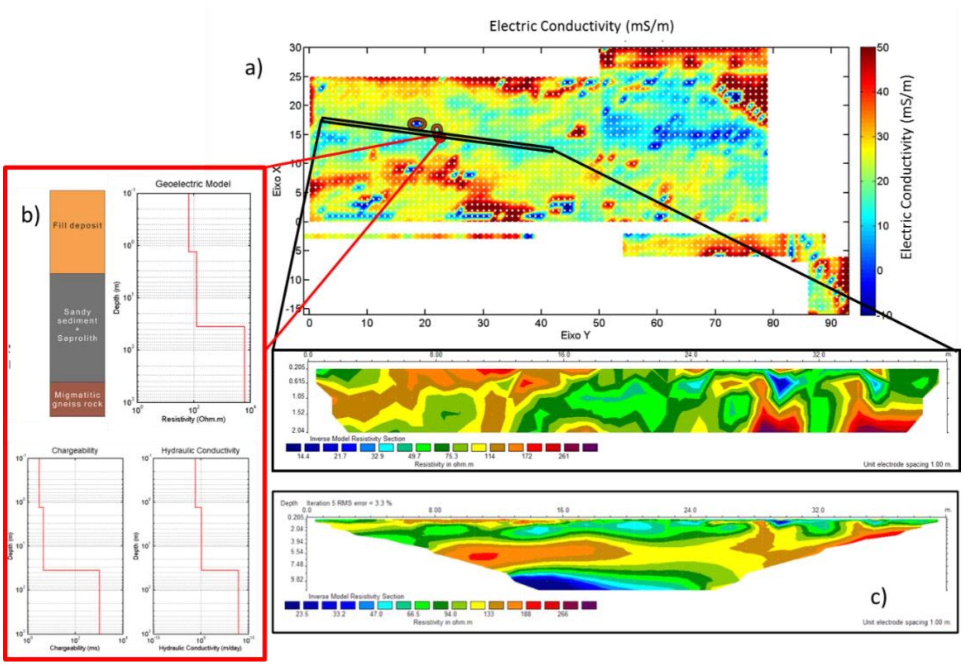

FIGURE 6: GEOPHYSICAL RESULTS IN CAMPOS DO JORDÃO – SP. IN A) THE RESULTS WITH INDUCTIVE ELETROMAGNETIC METHOD, IN B) THE RESULTS WITH VERTICAL ELECTRICAL SOUNDING AND C) THE RESULTS WITH ELECTRICAL RESISTIVITY TOMOGRAPHY.

The proposed multidisciplinary project has as main objectives:

1. Understand how geotechnical parameters (for example, hydraulic conductivity, porosity, permeability, chemical composition) can be associated with geophysical parameters (for example, electrical resistivity, magnetic susceptibility).
2. Determine how the water content in the layers is related to the amount and intensity of rainfall.
3. Predict how the expected rainfall will change the soil humidity.
4. Determine if with the expected humidity will have the occurrence of landslides.

The challenges of the proposed project are:

1. There are some equations and procedures to associate geotechnical parameters to geophysical parameters, but not many are addressed to the problem of landslides.
2. To understand the relationship between rain intensity and the water content of the layers, a numerical modeling process is required.
3. The prediction of how future rainfall will influence the humidity of the layers should be made in conjunction with future weather forecasting.
4. The modeling of how future rainfall will influence the hydrogeodynamic balance of the areas should be done automatically or semi-automatically.
5. In CEMADEN, the softwares that models the hydrogeodynamic balance of the slopes are GeoSlope and GeoStudio. That way the whole process is done as a black box and cannot be changed for the proposed goals.

3.2 Working Group II - Checking/ Verification & Time Series

Group: Nicolás Rubido Ober (Moderator), Alan James P. Calheiros (Moderator), Alessandra Corsi, Luana Albertani Pampuch, Luciano Xavier, Luis Ricardo Lage Rodrigues, Sheila Brito, and Tasmin Louise Symons

3.2.1 Forecasting severe storms

Collaborators: A. Calheiros, A. Corsi, N. Rubido, and T. Symons

Recents studies have shown that the number of extreme precipitation events have increased in recent decades [1,2] and, according to climate models, this condition will get worse in the coming years [3,4]. However, the current methods used to predict these severe events are inaccurate. Nowadays, the nowcasting of storms is essential in mitigating the effect (including life losses) of the natural disasters they can trigger, such as landslides or flooding. Nowcasting combines an ensemble of non-adaptive techniques that can make short range weather predictions (~6 hours). Consequently, our main goal is to develop adaptive nowcasting techniques for severe storm prediction that can go beyond the former time scale. In order to achieve this goal, we plan to use Compressive Sensing methods -- recovering a signal from sparse data -- [5-8] on different data platforms.

The data to develop this forecasting system are based on cloud top information from weather satellites (radiance from different channel measurements), weather precipitation radars (from polarimetric variables), and numerical weather models (re-analysis and prediction). It is worth noting that many of these channels and variables are irrelevant to the forecasting of storms, or have strong inter-dependencies. In order to keep the most relevant and independent measurements, a machine learning technique is therefore needed to separate them. Moreover, we are interested in forecasting storms in particular locations, such as areas at high risk of landslides or flooding, thus, the spatial characteristics of the data need to be discarded.

In order to tackle our forecasting goal, our approach is twofold and is explained in what follows. On the one hand, we want to "storm-chase" -- track the clouds -- and use the most relevant time-series in order to develop a forecasting tool. This approach is analogous to a Lagrangian view-point of the cloud system. The forecasting tool will be derived by using Compressive Sensing on the corresponding data, namely, we will derive a model for the evolution of the cloud system, fitting previous recordings and extrapolating to future states. Consequently, we will be able to predict the cloud/rain systems evolution and perform parameter changes to assess possible tipping points. On the other hand, we want to construct a severe storm early-warning indicator for locations that have experienced disastrous landslides or flooding in the past. This approach is complementary to the former, being similar to an Eulerian view-point of the cloud system. The forecasting tool in this case will again be an application of Compressive Sensing, this time to define a pre-convective situation (which occurs a few hours before the appearance of the first convective cells over a specific area) to the time-series of weather variables coming from numerical models in the chosen locations and local information, such as rain gauges, precipitation from radar, and soil humidity.

Our expected outcomes are the following: we will develop two parameter-free adaptive algorithms, one for severe storm prediction and another one for early-warnings of potentially imminent landslide or flooding in risk areas. Both algorithms will work in real time using live feeds of cloud/precipitation/weather data and local information.

References:

[1] MARENGO, J.A.; RUSTICUCCI, M.; PENALBA, O.; RENOM, M.; An intercomparison of observed and simulated extreme rainfall and temperature events during the last half of the twentieth century: part 2: historical trends. Climatic Change, v. 98, n. 3-4, p. 509-529, Oct. 2010.

[2] SILVA DIAS, M.A.F.; DIAS, J.; CARVALHO, L.M.V.; FREITAS, E.D.; SILVA DIAS, P.L. Changes in extreme daily rainfall for São Paulo, Brazil. Climatic Change, v. 116, n. 3-4, p. 705 – 722, feb. 2013.

[3] RUSTICUCCI, M.; JOSÉ MARENGO, J.A.; PENALBA, O.; RENOM, M. An intercomparison of model-simulated in extreme rainfall and temperature events during the last half of the twentieth century. Part 1: mean values and variability. Climatic Change, v. 98, n. 3-4 , p. 493-508, Feb. 2010.

[4] SILLMANN, J.; KHARIN, V.V.; ZWIERS, F.W.; ZHANG, X.; BRONAUGH, D. Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. J. Geophys. Res. Atmos., v. 118, p. 2473–2493, Mar. 2013..

[5] E. J. Candès, J. Romberg, and T. Tao, *IEEE Trans. Info. Theor.* **52**(2), 489-509 (2006).

[6] D. L. Donoho, *IEEE Trans. Info. Theor.* **52**(4), 1289-1306 (2006).

[7] Wen-Xu Wang, et al., Phys. Rev. Lett. **106**(15), 154101 (2011).

[8] W.-X. Wang, Y.-C. Lai, and C. Grebogi, *Phys. Rep.* **644**, 1-76 (2016).

Data specifics:

The storm tracking is based on the Forecast and Tracking the Evolution of Cloud Clusters (ForTraCC) algorithm, VILA et al. (2008). The system allows us to apply different input data, as brightness temperature from infrared channels in geostationary satellites or radar reflectivity as well. Several sources of information about the cloud tops and rain in the clouds will be analysed. The following list shows each variable for all data platforms.

- Satellite GOES-10,12,13 (GOES-16 the new platform):
  - Raw data 2D: reflectance and brightness temperature from different channels: 0.65µm, 3.9µm, 6.7µm, 10.7µm, and 12µm (GOES-16 -);
  - Spatial resolution: GOES-10,12,13 is 1x1km (shorter wavelength) and 4x4 km over (longer wavelength). GOES-16 is 0.5x0.5km and 1x1km;
  - Temporal resolution: every 30 minutes (GOES-16 every 15 minutes);
  - Estimated variables: Cloud top height, phase, pressure and temperature, total precipitable water, lightning detection, cloud type, size expansion, etc.
- RADAR:

    - Raw data 3D (lat,lon,level): reflectivity (horizontal and vertical wave), differential reflectivity, copolar correlation coefficient and Specific Differential Phase;
    - Spatial resolution: 2x2km;
    - Temporal resolution: 10 minute;
    - Estimated variables: Hydrometeor classification and severe signature.
- Landslide and Flood inventory from IPT
- Numerical Weather Models:
    - WRF, ETA, and BRAMS numerical models - 3D (lat,lon,level) Temperature, humidity, pressure, instabilities index, wind, etc.;
    - Spatial resolution: 5 km;
    - Temporal resolution: 3 hours (operational) and 1 hours (intensive operational period - SOS CHUVA project field experiment).

3.2.2 Assessment of droughts impacts and forecasting on Paraíba do Sul basin

Collaborators: S.S.B. Brito, L. R. Lage-Rodrigues, L.A. Pampuch, N. Rubido, T. Symons, and L. Xavier

This project aims to assess drought predictability and impact at different climatological time-scales in the Paraíba do Sul basin. This region is located between the two most populous and wealthiest Brazilian cities (São Paulo and Rio de Janeiro), hence, it is a very important basin for water supply and energy production, making drought forecasting for this basin an extremely important issue. In general, drought duration, severity, and frequency are commonly estimated from the SPI, SPEI, and PDSI indices, which come from observational data. In this project, we will assess the seasonal-drought predictability by using the North America Multimodel Ensemble (NMME) forecast systems. Particularly, our main focus is in making a forecast quality-assessments that include deterministic and probabilistic verification-measures. Namely, we plan to critically and quantitatively contrast the NMME forecast outputs with the observations. In addition, we will evaluate the capacity of the Coupled Model Intercomparison Project Phase 5 (CMIP5) forecast systems to represent droughts in present time. Future scenarios, duration changes, drought severity and frequency will also be evaluated. Complementarily, the impacts of past drought events on energy production and water supply in this basin will be studied as well as the impacts due to the projected changes on drought patterns. The outcomes of this project can be extended to other regions and applied by institutions, such as CEMADEN, CPTEC, INMET, ONS, ANA, and others dealing with drought forecasting, in order to support decision-makers in reducing the potential negative drought impacts.

Main Goals:

- To compare indices for determination of drought periods (SPEI, SPI, PDSI): drought duration, drought frequency, drought severity

- Verification of forecasting at seasonal time scale and quantification of sampling uncertainties
- Assessment of drought representation with climate system forecast (CMIP5) in the present period and changes of drought events on future climate scenarios.
- Verification of drought impacts in economic sectors, for example, energy production and water supply

Data:

- Variables: precipitation, temperature, evapotranspiration, stream flow
- Observation: GPCP, INMET, ANA, INPE, CEMADEN, ONS
- Models (forecast/hindcast): NMME (North America Multimodel Ensemble; Kirtman et al., 2014)
- Models (projections/simulations): CMIP5 and Regional Climate Models (RegCM and ETA)
- Streamflow: basin committee (AGEVASP, ANA)

Methodology:

- Index Calculations: SPI, PDSI, SPEI on different time scales (McKee et al., 1993; Palmer, 1965; Vicente-Serrano et al., 2010)
- Quantification of drought: duration, frequency, severity (Spinoni, et al. 2014; Spinoni, et al. 2015)
- Forecast quality assessment using several verification measures, such as Pearson and Spearman correlation coefficient, Brier skill score, continuous ranked probability skill score, and ROC curves (Hersbach, 2000; Stephenson et al., 2008; Wilks, 2006; Jolliffe and Stephenson, 2012)
- Quantify the sampling uncertainty using non-parametric bootstrap method (Mason, 2008; Jolliffe and Stephenson, 2012)
- Assessment of models projections: CMIP5 and Regional Climate Models - drought representation in historical period and quantification of changes for future periods in different representative concentration pathways (RCP4.5 and RCP8.5)

References

Hersbach, H. (2000) Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems, Weather and Forecasting, 15, 559-570.

Jolliffe, I. T., D. B. Stephenson (2012) Forecast Verification: A Practitioner's Guide in Atmospheric Science, Second Edition, John Wiley & Sons, Ltd, Chichester.

Kirtman, B. P. et al. (2014) The North American Multimodel Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction, Bulletin of the American Meteorological Society, 95, 585-601

Mason, S. J. (2008) Understanding forecast verification statistics, Meteorological Applications, 15, 31-40.

McKee, T.B.N., J. Doesken, and J. Kleist, 1993: The relationship of drought frecuency and duration to time scales. *Eight Conf. On Applied Climatology*. Anaheim, CA, Amer. Meteor. Soc. 179-184.

Nobre, C.A., Marengo, J.A., Seluchi, M.E., Cuartas, L.A. and Alves, L.M. (2016) Some Characteristics and Impacts of the Drought and Water Crisis in Southeastern Brazil during 2014 and 2015. *Journal of Water Resource and Protection*, 8, 252-262.

Palmer, W.C., 1965: *Meteorological droughts*. U.S. Department of Commerce Weather Bureau Research Paper 45, 58 pp.

Spinoni J, Naumann G, Carrão H, Barbosa P, Vogt J. 2014. World drought frequency, duration, and severity for 1951-2010. *International Journal of Climatology* **34:** 2792–2804. 10.1002/joc.3875.

Spinoni J, Naumann G, Vogt J, Barbosa P. 2015. European drought climatologies and trends based on a multi-indicator approach. *Global Planetary Change* **127:** 50–57. 10.1016/j.gloplacha.2015.01.012.

Vicente-Serrano S.M., Santiago Beguería, Juan I. López-Moreno, (2010) A Multi-scalar drought index sensitive to global warming: The Standardized Precipitation Evapotranspiration Index - SPEI. Journal of Climate 23: 1696-1718.

Wilks, D. (2006) Statistical methods in the atmospheric sciences, Second edition, International Geophysics Series, 59, Elsevier, Oxford

Paraíba do Sul :

Pacheco, F. S., Miranda, M., Pezzi, L. P., Assireu, A., Marinho, M. M., Malafaia, M., Reis, A., Sales, M., Correia, G., Domingos, P., Iwama, A., Rudorff, C., Oliva, P. and Ometto, J. P. (2017), Water quality longitudinal profile of the Paraíba do Sul River, Brazil during an extreme drought event. Limnology and Oceanography, doi:10.1002/lno.10586.

Kelman, J. (2015), Water Supply to the Two Largest Brazilian Metropolitan Regions, Aquatic Procedia, Volume 5, Pages 13-21, http://dx.doi.org/10.1016/j.aqpro.2015.10.004.

SMAP

Kwon, H., Souza Filho, F., Block, P., Sun, L., Lall, U. and Reis, D. (2012). Uncertainty assessment of hydrologic and climate forecast models in Northeastern Brazil. Hydrological Processes, volume 26, issue 25, 3875–3885. doi: 10.1002/hyp.8433.

3.3 Working Group III - Data Science & Machine Learning

Group: Fangxin Fang (Moderator), Haroldo F. de Campos Velho (Moderator), Albert Sánchez, Alejandro Cholaquidis Noblia, Christopher Castro, Eugenio S. Almeida, Ivo K. Koga, Jefferson L. M. A. Gomes, Lucas Massaroppe, Maha H. Kouri, Tiago Jose de Carvalho, Vinicius V. de Melo.

Climate simulation analysis: approach by Data Science

Introduction.

On the 29th of August, the city of São José dos Campos - Brazil held the International Workshop on Mathematics of Climate Change and Natural Disasters. Several researchers from different institutions collaborated in a joint effort to address academic and scientific problems that exist in the contemporary society and science regarding Climate Change and Natural Disasters. Many of them can make use of Machine learning and Data Science techniques.

In this report we identified some key points regarding Initial Conditions/Data Assimilation, Data Dimension Reduction, Urban Air Pollution and Extreme Events.

These are individually addressed in the following document, divided by four major sections. Section one discuss key issues of the current Data Assimilation techniques, summarizing its current obstacles, open questions and possible methodological approaches. Section two regards the Data dimensionality reduction, as the amount of actual data generated in a daily basis grows bigger, Big-Data approaches and dimensionality reduction techniques are proposed to address it. Sections three and four regards the air pollution that is mainly driven by urban areas and extreme events that experience surges in their frequency, forced by climate changes. In Section four, we investigate the application of machine learning approaches to improve dynamical model outputs in order to detect extreme events.

Preparing the initial condition

3.3.1 Machine Learning for Data Assimilation

(Fangxin Fang / Maha Kaouri)

Key questions:

Can we improve the efficiency (CPU) and accuracy of the current methodology/replace the current methods used for modelling atmospheric-ocean dynamics?

Can we provide new method tools and data science technology which provide real-time measurements and simulations to aid decisions that are needed to maintain a healthy environment?

Scientific background:

The key disadvantage of the current Data Assimilation methods is implementing and maintaining the adjoint – as soon as the model is changed, the adjoint must be updated. Due to the model complexity, solving such problems can be time consuming and computationally expensive.

A widely used method is Ensemble Kalman Filter which has unrealistic mathematical and statistical assumptions – linearity, multivariate normality, stationary state-transition functions (Evensen, 2009). Furthermore, confidence in the dynamical atmospheric model impacts the quality of the solution of the Kalman Filter. More general approaches based on particle filters are already under investigation, where the Gaussian assumptions are relaxed. However, such filters have also a high computational effort (Snyder et al., 2008).

Meteorological data is high dimensional. There are number of existing methods used to reduce the dimensional size of the data, for example, using super-observations (e.g. reducing the data by taking the average). Therefore, we propose a new approach to reduce the dimension by using Machine Learning – see Section 2.

In summary, there are two main problems affecting Data Assimilation CPU: Data and Model size and complexity. Both will benefit from Machine Learning and Singular Value Decomposition - Proper Orthogonal Decomposition (SVD-POD), in the next section, we will outline these.

Methodology:

SVD-POD or Deep (machine) Learning can be used to reduce the computational effort on several aspects: on the dimensional size of the original model (Xiao et al., 2017, Fang et al., 2014) and the dataset by identifying the most important data, or as a new method for data assimilation (Cintra et al., 2016; Cintra and Campos Velho, 2014; Härter and Campos Velho, 2008). Applying this new technology to data assimilation and optimization methods will reduce the CPU time (Fang et al., 2009, Chen et al., 2011, Fang et al., 2017).

We will propose a new Reduced Order Model (ROM) with domain decomposition methods and Deep Learning (Wang et al., 2017). The Deep Learning methods will be used to construct a set of hypersurfaces representing the reduced fluid dynamic system (including linear and nonlinear fluid dynamics). We will form the ROM basis functions, but we will do this within each subdomain of a solution domain decomposed into subdomains in order to resolve the systems energy more efficiently than global basis function methods. In each subdomain, an adaptive number of POD expansion size will be determined. It also means that a ROM can be built up region by region without solving the full model across the whole solution domain.

Applications: academic impact and on the society

Having a very rapid ROM compatibility developed here will be nothing short of revolutionary for a large number of disciplines not least of all pollutant flow based disciplines.

For all application areas, having a rapid ROM could potentially make tractable computationally demanding predictive problems: uncertainty analysis, data assimilation, optimizing monitoring/sensors/instrumentation/observations, optimal experimental design and control. ROM's may be used in place of highly simplified (e.g. single column models or Gaussian plume dispersion modelling for air pollution) whilst maintaining the complexity of 3D models.

Optimal design and sensitivity analysis can be used to identify the key parameters which are polluting the source (Fang et al., 2017). Therefore, we will be able to guide the decision makers about the pollutant sources so they can prevent/combat/mitigate this by identifying where the pollutant is coming from.

The Adjoint method is the best method for uncertainty quantification (Cacuci et al., 2005). However, it is computationally demanding. Instead of this, we propose using Machine Learning and ROM to get a sensitivity analysis in order to find the uncertainty parameters (not the initial conditions alone).

We can use rapid response modelling to use a reduced model so that the CPU time is reduced. Therefore, we can generate the prediction in seconds, enabling us to link with the data and provide real-time predictions. This enables us to inform the public sooner, reducing injuries and fatalities. The decision makers and public will benefit from the rapid new generation of atmosphere modelling, efficient pollutant transport and processing as well as the rapid assessment of environmental impact.

3.3.2 Data dimension reduction

(Eugenio Almeida / Ivo Koga / Lucas Massaroppe / Vinicius Melo)

During recent years, Climate Science and Natural Disasters researchers have to deal with an exponential growth in large and heterogeneous databases to conduct their research. For instance, the number of meteorological and environmental satellites have increased, with new sensors onboard and more spectral bands.

In (big) data science, there is a need of preprocessing due to data redundancy making it difficult to extract relevant data. These redundancies must be filtered out in order for techniques to be effective.

In the next section we introduce some possible techniques of dimension reduction.

Dimension reduction techniques

To induce a model to explain a phenomenon, a model building technique uses a dataset of observations to develop a hypothesis that explains a phenomenon. Each observation contains one or more variables (features). As the number of features increase, finding a high-quality model may become exponentially harder. This is a well-known problem called "curse of dimensionality".

To deal with it, many researchers have been suggesting techniques to perform dimensionality reduction (Xiao et al., 2017, Fang et al., 2014, Ruivo et al., 2015). The most popular approaches are feature subset selection and feature extraction.

The rationale behind feature subset selection is that not all features may be useful to describe or explain the phenomena, meaning that the dataset may be composed of relevant and irrelevant features. There are several techniques for reducing the data dimension without losing information.

There are basically two approaches: filter and wrapper. In the filter approach, a technique analyses the data and tries to identify the relevance of a feature according to the expected output (supervised learning) or not (unsupervised learning). The wrapper approach uses a technique to select a subset of features and then induces a model on them.

In a classification problem, one may check if the Probability Density Function (PDF) of a feature can be used to discriminate the classes in the dataset. This means that each class has a different PDF according to a specific feature. On the other hand, if there is overlapping, then such feature is not useful for classification. As one may notice, in this approach, features are analysed individually. However, they may be complementary, meaning they should be evaluated together. In this case, one could use the wrapper approach. Using this approach, the model is then evaluated to provide a quality measurement for that subset. After that, a new subset can be generated and evaluated. An iterative process is employed to find the best feature subset that optimizes a specified quality criterion. As one may notice, the wrapper approach is more time consuming as, for each feature subset, one has to induce a model and evaluate it. Also, the best feature subset is tied to the model building technique, meaning that a different technique could have a different best feature subset because different techniques use different ways of selecting features based on an importance criterion.

Some techniques automatically perform feature subset selection during the model building process. Algorithms that employ regularization can eliminate features by weighting them with zero weights (Zou, H., & Hastie, T., 2005). Decision-trees (Quinlan, 1986) commonly use some sort of entropy measure to select the currently most discriminative feature to use as a decision node; thus, not all features may be necessary to build a tree. In the Random Forest algorithm (Breiman, 2001), a number of trees is built upon randomly selected feature subsets. This way, one can calculate a feature importance based on how much they are used in the forest.

In feature extraction, the idea is to build a new set of features from the original feature set. One may optimize the weights applied to each feature in order to create a new one. Notice that all original features are used in this process. This process can be repeated to generate other features. Examples of techniques are projection methods such as Empirical Orthogonal Functions (EOF) was proposed by Lorenz (Lorenz, 1956), also known as Principal Component Analysis (PCA), Kernel PCA (Schölkopf, 1998), Random Projection (Bingham and Mannila, 2001) Projection Pursuit (Huber, 1985). Statistical analysis methods can combine with artificial intelligence procedures to reduce the data dimension: *p*-value approach and decision trees (Ruivo et al., 2015b).

An alternative is to use Artificial Neural Networks (ANN) (Haykin, 2009) and Deep Learning. In this approach, the ANN gets the original data as input and must generate the same data as output. However, the hidden layer has the desired number of features to be extracted. Thus, the ANN must condense the information into the desired number of features and be able to recover the original data from the reduced feature set. After trained, the ANN is cut to the hidden layer, so it outputs just the reduced data. Self-Organizing Map (SOM) (Sheridan, 2011) is an ANN employed for data reduction. SOM can be used as a clustering technique. The idea is to map the original data into a low-dimensional grid.

Another approach for feature extraction is to combine some of the original features (not necessarily all of them) into new features. Such approach is also known as feature construction or generation. If done manually, it is also known as feature engineering. The combination can be nonlinear and any kind of transformation can be used. For instance, one can apply arithmetic, geometric, or conditional operations on the original features to obtain new features. Examples of techniques are evolutionary algorithms such as Genetic Programming (Sotto et al., 2016), Grammatical Evolution (Miquilini et al., 2016) and, more recently, Kaizen Programming (de Melo and Banzhaf, 2016).

Our focus is on machine-learning related methods, like SOM, Kernel PCA and Kaizen Programming, as well as the canonical techniques, such as EOF (also called PCA or SVD), as pre-processing step.

Application examples:

- ANN for downscaling medium-range ensemble forecasts and probabilistic prediction of local precipitation in Japan (Ohba et al., 2016; Valverde et al., 2014, 2006, 2005),
- Assessing the forecast skill of eight North American Multi Model Ensemble (NMME) models by improving their skill using Bayesian updating (BU) (Zhang et al, 2017).
- Comparison of two different methods (PCA and SOM) of teleconnection pattern recognition (Rousi et al., 2015).

- Evaluation of Machine Learning tools using historical flood data collected in the State of Iowa, the United States and associated hydrometeorological variables from 1948 to 2010. (Yahui et al., 2015).
- Principal Component Analysis (PCA) in combination with two post-processing techniques for the prediction of wind power produced over Sicily, and of solar irradiance measured by Oklahoma Mesonet measurements' network (Frederica et al., 2016).
- Usage of genetic programming on ensemble models (Dufek, 2017).
- A GIS-based multi-criteria statistical methodology developed to quantify hazard potential and to map flood characteristics (Arpita, et al., 2016)

3.3.3 Urban Air Pollution

(David Franca / Alejandro Cholaquidis / Jeff Gomes)

Motivation: Pollution at urban street canyons have strong impact on the health of communities. Assessment and prediction of pollutant dispersion at street level are not simple problems that can be solved with operational methods as they are strongly dependent on meteorological data (pluviometric, temperature, wind velocity/direction etc), traffic information (Garcia *et al.*, 2011), online and in-situ monitoring etc. Sustainable cities rely on *smart* use of available data from sensors (fixed and mobile) at different parts of the city and optimal use of such data through integration with detailed and operational models. Thus the aim is to produce a ***V**irtual City **A**ir **Po**llution Fast **R**esponse Model (VAPOR)* to help policy-makers, health and safety authorities, traffic-controllers and rapid response teams to manage city pollution and/or to mitigate its impact on the general population, conceptualized in Figure 7.

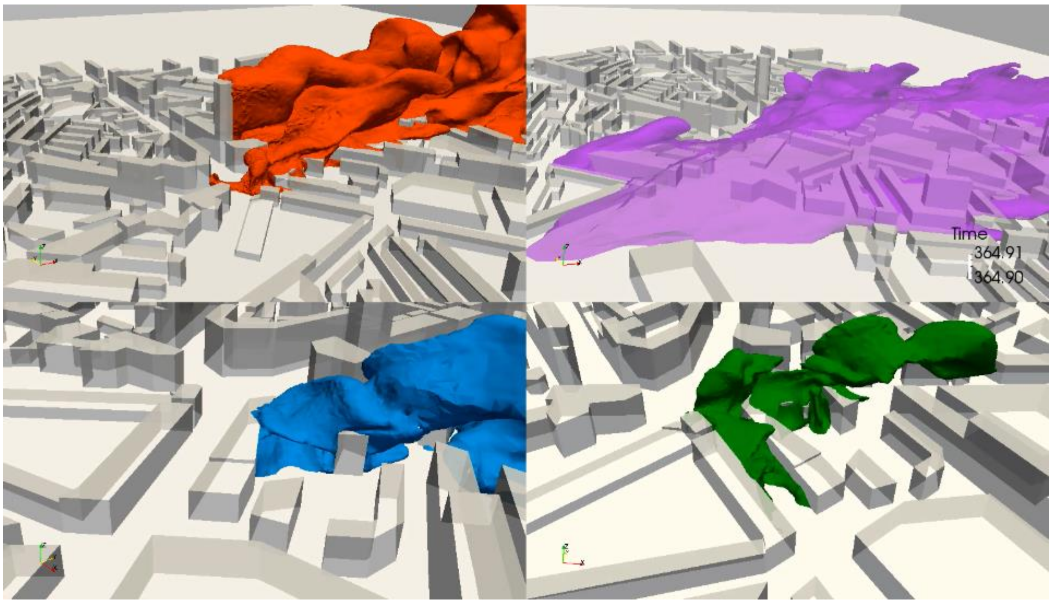

FIGURE 7: CONCEPTUAL ILLUSTRATION OF THE VAPOR MODEL FOR PASSIVE-TRACER RELEASE AND DISPERSION IN CITY STREET-CANYONS. ADAPTED FROM PAVLIDIS ET AL. (2010).

Science Questions: Prediction of pollution dispersion in urban street canyons is a complex inter-disciplinary topic which involves:

- *coupled detailed models* of pollutant flows through street canyons (i.e., turbulent CFD models), traffic models (different pollutant sources: vehicles, industries, people movement, and other considerations), and meteorological models (Pavlidis *et al.*, 2010);
- *movement of people* in the urban landscape, which may be affected by pollution. It can be predicted using a number of anthropology-based methods (e.g., mobile methods, Elliot *et al.*, 2017);
- *detailed simulations:* they can be computationally very expensive as the number of potential degrees of freedom increases (contaminants + velocity + pressure + temperature + chemistry + radiation + density); thus, the reduction of problem dimensionality is a potential strategy to make such models running in a feasible time-scale;
- *data produced by models:* they need to be continuously updated with existing data from city sensors (data assimilation);
- *data management* from sensors and models (Stingone et al., 2017; Keller, 2014; Xi et al., 2015; Kalapanidas & Avouris, 1999).

During our discussion we identified two key science questions that we should address:

  - How to best *classify and identify* potential pollutant sources (spatial and time distribution) from fixed and mobile data sensors using Machine Learning technologies (ML, hybrid methods, e.g., ANN + decision trees);
  - How to use this information to *mitigate the impact* of pollutants/contaminants on street canyons during rush hours or extreme events (e.g., terrorist attack, accidental release of chemicals etc). Rapid response models (i.e., based on reduced-order models) through coupled models and ML may be an efficient way to help predict pollutants/contaminants pathways and concentrations in the urban canopy at such events.

3.3.4 Addressing Extreme Events by Machine Learning

(Alber Sánchez, Christopher Cunningham, Thiago Carvalho)

Reasoning

Currently, records are showing the increasing frequency of severe weather (IPPC, 2014). Extreme weather events are known to trigger natural disasters and this is a warning call to scientists to refine their forecasting tools to help government agencies to better deploy their resources (Editorial NM (2017), IPCC (2014)).

Not only is it well known that extreme events will become more frequent in future decades (IPCC, 2014), but also there is an increasing demand on operational prediction and applications communities for forecasts that fill the gap between medium-range weather (up to two weeks) and long-range or seasonal (3-6 months).

Current modeling is limited by the spatio-temporal resolution level required by society under a changing climate. Modeling the sub-seasonal level is better suited to tackle the challenges that a changing environment poses to humankind.

Despite the fact that climate projections and sub-seasonal predictions produce large amounts of data, the signal of extreme events is frequently hidden in the midst of the dataset. The projections fail to account for the nonlinear interactions among the components of the climate system.

Specifically, we would like to address the following research question: *How can Machine Learning enhance dynamic model outputs to improve the recognition of extreme events, considering the sub-seasonal and climate change time scales?*

Extreme events

Extreme weather events deeply affect society, since they can trigger socio-environmental disasters. During the period 1995-2015, the majority (90%) of disasters have been caused by floods, storms, heatwaves and other weather-related events (Wahlstrom and Guha-Sapir, 2015).

Due to the chaotic nature of the atmosphere, small imperfections in models' initial conditions and parameters cause huge differences in the predictions as the forecast horizon becomes longer. In other words, very similar initial conditions can make a model diverge, increasing the prediction uncertainty. In order to overcome this issue, scientists use ensemble predictions to quantify their uncertainty. Hence, predictions beyond a week can only be probabilistic.

Prediction of extreme events is addressed differently for climate change and sub-seasonal time scales. On the latter case, the goal is to obtain a better anticipation of dry spells, wet spells, and heat weaves. This anticipation is on a probabilistic sense, i.e., to predict more reliable chances for the forthcoming dry/wet spell, and heat wave. In the climate change scale, we are interested in shifts in probability distribution of extreme events (Coumou and Rahmstorf, 2012, Sippel and Otto, 2014) during relevant seasons in future decades. For instance, the increase in frequency of extreme wet spells during the monsoon season.

Machine Learning Methods Applied to Extreme Events Detection

For extreme events detection, one of the goals is to improve numerical modeling *a posteriori* in order to determine the precursors of extreme subseasonal weather events. To address this goal, it is possible to adopt Machine Learning approaches, which have been used to solve similar problems.

Chauhan and Vig (2015) propose the use of Long-Short Term Memory (LSTM) networks for an anomaly detection approach towards analyzing Electrocardiogram (ECG) signals. Addressing a different problem that also involves time-series, Filonov et al. (2016) adopt an approach based on an LSTM neural network to monitor and detect faults in industrial multivariate time series data. The authors validate their approach applying it into a part of a real gas-oil plant, where the system was responsible by fault-detections. Ahmad and Purdy (2016) address the problem of detecting anomalies in streaming data (a kind of time-series data problem) based on an on-line sequence memory algorithm called Hierarchical Temporal Memory (HTM).

Similar to previously mentioned approaches, when analyzing time-series data produced by environmental models, the presence or absence of extreme events is a difficult task to perform. This project intends to investigate the possibility of designing and using Machine Learning solutions (e.g., Deep Learning, decision trees, random forests) to address the problem of detecting extreme events in environmental data generated by different models.

But Machine Learning methods not only need to be precise but also fast. And their speed is a function of the amount of data processed.

The number of Earth observation satellites is increasing. As time passes, satellite technology gets smaller and cheaper and more countries are putting their own sensors into orbit. Accordingly, we need more efficient tools (hardware and software) in order to process terabytes of data (Belward, 2015, Gottfried, 2004).

Regarding hardware, the main approaches are High Performance Computing (HPC) and Grid computing (Berman, 2003, Dowd, 1993, Pordes, 2007). Traditionally, HPC is prefered in fields such as meteorology and Grid computing in physics where institutions such as CERN leads the development of new Grid technologies. However, which of those fits better machine learning for detecting extreme event detection is still an open question.

Regarding data, the principal approaches are Map Reduce and array databases. In particular, array databases seem to better fit our analysis requirements because they join the experience of almost 50 years of relational databases to a well-known mathematical abstraction, the array. For example, atmospheric data is collected by static sensors which build long time series. This data is interpolated or simulated into regular dynamic grids which are simple to represent as multidimensional arrays, that is, arrays of 3 spatial dimensions and one temporal. Handling these array as linear (array) algebra is a well known and understood field on mathematics (Baumann 1999, Camara 2014). The same way relational databases are based on relational calculus, array databases are founded in array algebra. This sound foundation on mathematics allow users of array databases to express complex analysis queries in terms of abstract functions. For example, the array database SciDB is known to efficiently handle petabytes of data. It eases data management of large datasets but more importantly, SciDB can re-use R code because it separates data management from modeling through the use of high level programming languages such as MATLAB, R, Python, or an array functional language (Stonebraker (2009), Stonebraker (2013)).

REFERENCES

Ahmad S. and Purdy S. (2016) Real-time anomaly detection for streaming analytics. CoRR abs/1607.02480, URL http://arxiv.org/abs/1607.02480.

Nandi, Arpita, et al. "Flood hazard mapping in Jamaica using principal component analysis and logistic regression." *Environmental Earth Sciences* 75.6 (2016): 465.

Baumann P. (1999) A Database Array Algebra for Spatio-Temporal Data and Beyond, Proc. 4th Int. Work. Next Gener. Inf. Technol. Syst., no. Mdd, pp. 76–93.

Berman, Fran, Geoffrey Fox, and Anthony JG Hey, eds. Grid computing: making the global infrastructure a reality. Vol. 2. John Wiley and sons, 2003.

Bingham E., Mannila H. (2001) Random projection in dimensionality reduction: Applications to image and text data. KDD-2001: Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: Association for Computing Machinery. pp. 245–250. doi:10.1145/502512.502546.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Caccuci D.G., Ionescu-Bujor M., Navon I.M. (2005) *Sensitivity and uncertainty analysis, volume II: applications to large-scale systems*. Vol. 2. CRC Press.

Camara G., Egenhofer M.J., Ferreira K., Andrade P., Queiroz G., Sanchez A., Jones J.,. Vinhas L. (2014) Fields as a Generic Data Type for Big Spatial Data. In Geographic Information Science, vol. 8728, Springer Verlag, 159–172.

Chauhan S., Vig L. (2015) Anomaly detection in ecg time signals via deep long short-term memory networks. In: 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA), pp 1–7.

Chen X, Navon I.M, Fang F. (2011) dual weighted trust-region adaptive POD 4D-Var applied to a Finite-Element shallow-water Equations Model, *International Journal for Numerical Methods in Fluids*, 65, 520-541.

Cintra R. S., Campos Velho H. F. , Cocke S. (2016): Tracking the model: Data assimilation by artificial neural network. *International Joint Conference on Neural Networks* (IJCNN), Vancouver, Canada (USA: IEEE Computational Intelligence Society), vol. 1, 403-410.

Cintra, R. S., Campos Velho, H. F. (2014) Data Assimilation by Artificial Neural Networks for an Atmospheric General Circulation Model: Conventional Observation. arXiv:1407.4360 / https://arxiv.org/abs/1407.4360

Coumou, Dim, and Stefan Rahmstorf. 2012. "A Decade of Weather Extremes." Nature Climate Change, no. March 2012. doi:10.1038/nclimate1452.

Davò F. Alessandrini S., Sperati S., Monache L.D., Airoldi D., Vespucci M.T (2016) Post-processing techniques and principal component analysis for regional wind power and solar irradiance forecasting. Solar Energy 134:327-338.

Di D., Ding W., Mu Y., Small D., Islam S. Chang N-B. (2015) Developing Machine Learning tools for long-lead heavy prediction with multi-sensor data. Networking, Sensing and Control (ICNSC), IEEE 12th International Conference.

Dowd, Kevin. High performance computing. O'Reilly, 1993.

Dufek A., Augisto D.A., Dias P.L.S., Barbosa H.J.C. (2017) Application of Evolutionary Computation on Ensemble Forecast of Quantitative Precipitation. *Computers & Geosciences* 106, 139-149.

Editorial-Nature (2017) Extreme weather events are the new normal. Nature pp 54–62, DOI doi:10.1038/548499b.

Elliot A., Norum R., Salazar N.B. (2017) Methodologies of Mobility: Ethnography and Experiment. Hb Publisher.

Evensen G. (2009) *Data assimilation: the ensemble Kalman filter*, Springer Science & Business Media.

Fang F., Pain C.C., Navon I.M., Xiao D. (2017) An efficient goal based reduced order model approach for targeted adaptive observations, International Journal for Numerical Methods in Fluids, 83, 263-275.

Fang, F., Pain, C.C., Navon, I.M., Piggott, M.D., Gorman, G.J., Allison, P.A., Farrell, P.E. Goddard, A.J.H. A POD reduced order 4D-Var adaptive mesh ocean modelling approach. *International Journal for Numerical Methods in Fluids*. 60(7), (2009), 709-732.

Fang, F. et al. Reduced order modelling of an unstructured mesh air pollution model and application in 2D/3D urban street canyons. **Atmospheric Environment**, v. 96, p. 96-106, 2014.

Filonov P., Lavrentyev A., Vorontsov A. (2016) Multivariate industrial time series with cyber-attack simulation: Fault detection using an lstm-based predictive data model. CoRR abs/1612.06676, URL http://arxiv.org/abs/1612.06676.

Garcia X., Pavlidis D., Gorman G.J., Gomes J.L.M.A., Piggott M.D., Aristodemou E., Mindel J., Latham J-P, Pain C.C., ApSimon H. (2011) A two-phase adaptive finite element method for solid-fluid coupling in complex geometries. *International Journal for Numerical Methods in Fluids*, 66:82-96.

Grotjahn, Richard, et al. (2016) "North American extreme temperature events and related large scale meteorological patterns: a review of statistical methods, dynamics, modeling, and trends." *Climate Dynamics* 46(3-4), 1151-1184.

Gottfried, K. "Small satellites–a tool for earth observation." Invited Paper: Commission IV, Institute of Photogrammetry and GeoInformation, University of Hannover, Hannover (2004).

Härter, F. P., Campos Velho, H. F. (2008) New approach to applying neural network in nonlinear dynamic model. *Applied Mathematical Modelling* , 32(12), 2621-2633.

Haykin, S. S., et al. *Neural networks and learning machines*. Vol. 3. Upper Saddle River, NJ, USA:: Pearson, 2009.

Huber, P.J. "Projection pursuit." *The annals of Statistics* (1985): 435-475.

IPCC (2014) Climate Change 2014 Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change pp 1–112, DOI 10.1017/CBO9781107415324, arXiv:1011.1669v3.

Kalapanidas E., Avouris N.M. (1999) Applying Machine Learning techniques in air quality prediction, Proceedings of the ECCAI Advanced Course on Artificial Intelligence.

Keller R.H. (2014) A Machine Learning model of Manhattan air pollution at high spatial resolution, PhD Thesis, MIT.

Li X., Peng L., Hu Y., Shao J., Chi T. (2016) Deep Learning Architecture for air quality predictions. Environmental Science and Pollution Research 23:22408-22417.

Lorenz, E. N., 1956: Empirical orthogonal functions and statistical weather prediction. Technical report, Department of Meteorology, MIT, science Report 1. — 1963: Deterministic nonperiodic flow. J. Atmos. Sci., 20, 130–141.

Miquilini P., Barros R.C., de Melo V.V., Basgalupp, M.P. (2016). Enhancing discrimination power with genetic feature construction: A grammatical evolution approach. In Evolutionary Computation (CEC), 2016 IEEE Congress.

de Melo V.V.,Banzhaf W. (2016) Kaizen Programming for Feature Construction for Classification. In Genetic Programming Theory and Practice XIII (pp. 39-57). Springer International Publishing.Ohba, M., Kadokura S., Nohara D., Toyoda Y. (2016) Rainfall downscaling of weekly ensemble forecasts using self-organising maps. Tellus A: Dynamic Meteorology and Oceanography 68:29293.

Pavlidis D., Gorman G.J., Gomes J.L.M.A., Pain C.C., ApSimon H. (2010) Synthetic-eddy method for urban atmospheric flow modelling. *Boundary-Layer Meteorology* 136:285-299.

Pavlidis D., Gomes J.L.M.A, Gorman G.J., Pain C.C., Aristodemou E., ApSimon E., Robins A.G. (2010) Novel Mesh Adaptive LES Simulations for Multi-Scale Atmospheric Flows: Application to the Urban Environment. In *Air Pollution Modeling and its Application XX.* NATO Science for Peace and Security Series C: Environmental Security, Springer Netherland, pp. 63-68.

Pordes R., Petravick B., Kramer B., Olson D., Livny M., Roy A., Avery P., Blackburn K., Wenaus T., Würthwein F. (2007) The open science grid." Journal of Physics: Conference Series. 78. No. 1. IOP Publishing.

Quinlan, J. R. (1986). Induction of decision trees. Machine learning, 1(1), 81-106.

Rousi E., Anagnostopoulou C., Tolika K., Maheras P. (2015) Representing teleconnection patterns over Europe: A comparison of SOM and PCA methods. Atmospheric Research 152:123-137.Schölkopf, Bernhard, Alexander Smola, and Klaus-Robert Müller. "Nonlinear component analysis as a kernel eigenvalue problem." *Neural computation*10.5 (1998): 1299-1319.

Ruivo H. M., Ramos, F. M., Campos Velho, H. F., Sampaio, G. (2015) Study of Extreme Brazilian Meteorological Events. In: Christian Constanda, Andreas Kirsch. (Org.), 539-550, Birkhäuse, 1ed.New York (USA).

Ruivo, H. M., Campos Velho, H. F., Ramos, F. M., Freitas, S. R. (2015) Analysis of extreme events using a data mining approach. In: 12th International Conference on Applied Computing, 113-120, Maynooth (Ireland).

Sheridan, Scott C., and Cameron C. Lee. "The self-organizing map in synoptic climatological research." *Progress in Physical Geography* 35.1 (2011): 109-119.

Sippel, Sebastian, and F. E L Otto. 2014. “Beyond Climatological Extremes - Assessing How the Odds of Hydrometeorological Extreme Events in South-East Europe Change in a Warming Climate.” Climatic Change 125 (3–4): 381–98. doi:10.1007/s10584-014-1153-9.

Snyder C., Bengtsson T., Bickel P., Anderson J. (2008) Obstacles to high-dimensional particle filtering, Monthly Weather Review 136:4629-4640.

Sotto L.F., Coelho R.C., de Melo, V.V. (2016). Classification of Cardiac Arrhythmia by Random Forests with Features Constructed by Kaizen Programming with Linear Genetic Programming. In Proceedings of the 2016 on Genetic and Evolutionary Computation Conference (pp. 813-820). ACM.

Stingone J.A., Pandey O.P., Claudio L., Pandey G. (2017) Using Machine Learning to identify air pollution exposure profiles associated with early cognitive skills among U.S. children, Environmental Pollution 230:730-740.

Stonebraker M., Becla J., DeWitt D.J., Lim Kt., Maier D., Ratzesberger O., Zdonik S.B. (2009) Requirements for Science Data Bases and SciDB. In: {CIDR} 2009 Fourth Biennial Conference on Innovative Data Systems Research, Asilomar, CA, USA, January 4-7, 2009, Online Proceedings

Stonebraker M., Brown P., Zhang D., Becla J. (2013) Scidb: A database management system for applications with complex analytics. Computing in Science & Engineering 15(3):54–62.

Valverde M.C., Araujo E., Campos Velho H.F. (2014) Neural network and fuzzy logic statistical downscaling of atmospheric circulation-type specific weather pattern for rainfall forecasting. *Applied Soft Computing* 22:681-694.

Valverde-Ramirez M.C., Ferreira N.J., Campos Velho H.F. (2006) Linear and Nonlinear Statistical Downscaling for Rainfall Forecasting over Southeastern Brazil, *Weather and Forecasting*, 21:969-989.

Valverde-Ramirez M.C., Campos Velho H.F., Ferreira N.J. (2005) Artificial Neural Network Technique for Precipitation Forecasts Applied to the São Paulo Region. *Journal of Hydrology*, 301:146-162.

Xi X., Wei Z., Xiaoguang R., Yijie W., Xinxing B., Wenjun Y., Jin W. (2015) A comprehensive evaluation of air pollution prediction improvement by a Machine Learning method, IEEE International Conference on Service Operations and Logistics, and Informatics.

Xiao D., Fang F., Pain C.C., Navon I.M. (2017) A parameterized non-intrusive reduced order model and error analysis for general time-dependent nonlinear partial differential equations and its applications, Computer Methods in Applied Mechanics and Engineering, 317:868-889.

Wang Z., Xiao D., Fang F., Govindan R., Pain C.C., Guo Y. (2017) Model identification of reduced order fluid dynamics systems using deep learning, International Journal for Numerical Methods in Fluids, DOI: 10.1002/fld.4416.

Zhang W., Villarini G., Slater L., Vecchi G.A., Allen Bradley (2017) Improved ENSO Forecasting using Bayesian Updating and the North American Multi Model Ensemble (NMME). *Journal of Climate, DOI:10.1175/JCLI-D-17-0073.1.*

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67(2), 301-320.

3.4 Working Group VI - Disaster Risk Reduction

Group: Luciana R Londe, Leonardo Santos, Francesco Ferrulli, Matthew Brown, Luiz Carvalho, Andres Sosa, Jair Koiller, Guillermo Obregón Párraga, Walter Mendes Filho, Viviana Aguilar Muñoz, Alice Nardoni Marteli and Selma Santos

The risk of disasters can be understood through indices based on spatial-temporal models, with the difficult objective to preview their scales of magnitude and intensities. For floods, for instance, it is possible to adjust prediction systems using observational data, such as rainfall, soil moisture, vegetation and geology. We can also compare the cost/benefit of preventive measures with the economic and social costs for remediation of impacts. One of our challenges is to link models on a global scale and those on local scales. Infectious diseases often happen in sequence of an impact (floods, landslides). Innovative research must be done in a way that links mathematics and the different themes related to the study of disasters, to develop improved risk models.

3.4.1 Modelling rainfall with applications to flood risk and food insecurity

Floods are an important factor threatening food security around the globe. We aim at developing an effective precipitation estimator that could be useful in assessing the risk of flooding at a global scale.

Flooding is arguably the weather-related hazard that is most widespread around the globe. It is an important factor affecting living condition in terms of health and economic stability. It heavily influences the whole agricultural sector as well as it is responsible for epidemic outbreak and shortages in drinking water.

Atmospheric events and in particular exceptional precipitations are believed to be the leading causes for floods. So far, little attention in literature has been given to the interaction between the amount of precipitation and the moisture level of the soil.

The development of a global indicator of flood risk when combined with other climatic variables would help gauge the overall vulnerability to floods in across the globe.

Furthermore, an improved indicator could find use in larger models such as of food insecurity Richardson (2017) developed at the Met office in UK.

The main challenges in this project are (i) combining the data at different scales (ii) developing optimal weighting schemes for the precipitation series and (iii) developing and validating an index for floods at the appropriate scale.

Data

- Daily rainfall: observations and model output
- Vegetation: MODIS Land Cover;
- Disaster databases (for calibration): Emergency Events Database - EM-DAT (of the Centre for Research on Epidemiology of Disasters – CRED) ; NatCatSERVICE Natural catastrophe statistics online

Methods

In order to account for incomplete drainage of rainfall, we need an autoregressive model that relates rainfall at day i, p_i , with rainfall in previous days. The idea is to use a smoothed estimator of rainfall, P_i such that:

$$P_i = \sum_{j=0}^{364} w_{ij} \, p_{i-j} \quad (1)$$

The formulation in (1) suggests an autoregressive structure for the corrected (smoothed) precipitation estimator. One methodological challenge is to choose a weighting scheme (w_{ij}) that allows for the correct representation of the effective quantity of retained rainfall at any given day. We propose a first stage of exploratory analysis with the use of semi-variograms to compare spatial correlation at several time points and also the application of autoregressive models to understand the temporal autocorrelation structure.

At a later stage in the project, we intend to calibrate our index against flood occurrence data using artificial neural networks (ANNs). ANNs are a robust and computationally affordable technique that has been very successful in many applications in several areas including climate, hydrology and epidemiology.

Expected results

We expect to have an informative global flood index at grid point resolution 0.5º .The major implication of working on such resolution is that it enables future projections by mean of climate model simulations. The precipitation index that we want to develop will be a cumulative index which will play the role of a proxy for a soil moisture indicator.

References

Kwon (2007), Hyun-Han, Upmanu Lall, and Abedalrazq F. Khalil. "Stochastic simulation model for nonstationary time series using an autoregressive wavelet decomposition: Applications to rainfall and temperature." Water Resources Research 43.5 (2007).

Richardson (2017) Richardson, K. J., Lewis, K., Krishnamurthy, P. K., Kent, C., Wiltshire. A. and Hanlon, H. Food security outcomes under a changing climate: impacts of mitigation, adaptation and development on vulnerability to food insecurity. Submitted to Climatic Change

Sigrist (2011), Fabio, Hans R. Künsch, and Werner A. Stahel. "An autoregressive spatio-temporal precipitation model." Procedia Environmental Sciences 3 (2011): 2-7.

SREX (2012) Intergovernmental Panel on Climate Change. . Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaption. Cambridge University Press.

Westra (2014), Westra S., Fowler, H. J., Evans, J. P., Alexander, L. V., Berg, P., Johnson, F., ... & Roberts, N. M. . Future changes to the intensity and frequency of short-duration extreme rainfall. Reviews of Geophysics, 52(3), 522-555.

Collaborators: Leonardo Santos, Francesco Ferrulli, Matthew Brown, Luiz Carvalho, Andres Sosa, Jair Koiller, Guillermo Obregón Párraga, Walter Mendes Filho, Viviana Aguilar Muñoz, Sheila Brito and Alice Nardoni Marteli.

3.4.2 Leptospirosis phylodynamics in an outbreak after a flooding event

Natural disasters can often create favourable conditions for the onset of infectious diseases outbreaks (e.g. cholera, leptospirosis and dengue). New integrative approaches are needed in order to explore the interplay between the impact of floods, landslides and tornadoes on infectious diseases reservoirs. We aim to recover a phylogenetic tree from a leptospirosis outbreak after a flood in the Amazon region, and to incorporate hydrological measurements in a descriptive model of Leptospirosis epidemics.

Introduction

Some Brazilian states are severely affected by floods (EM-DAT, 2014, TOMINAGA et al., 2009). As consequences for public health, there is frequently a raise in the number of cases of diarrheas, hepatitis-a, hepatitis-e and leptospirosis - a disease caused by *Leptospira* bacteria. These microorganisms are scattered in environment through rodent's urine and transmission occurs through contact between skin and contaminated water, soils, vegetation (such as sugarcane) and mud. Leptospirosis in urban places is often associated to sanitation characteristics, especially concerning the accumulation of garbage and contact with sewages. (KO et al, 1999; BARCELLOS & SABROZA, 2000). As floodwaters can contain trash and sewage, the scattering of *Leptospira* bacteria is often raised.

Leptospirosis impacts many sectors with expressive socioeconomic importance, because it raises the financial cost for hospitalizations.

Most authors stress the synergy of factors triggering leptospirosis cases: poor sanitation services, mud and garbage accumulation, large populations of mice, environment and housing characteristics, low socioeconomic levels and work activities (Barcellos and Sabroza, 2000; Costa et al., 2001; Oliveira et al., 2012; Pelissari et al., 2011). Although many factors relate to leptospirosis transmission, there are particularities: post-flood situations may favor the appearance (or increasing) of outbreaks of this disease (Londe et al., 2016).

Scientific question

Can we associate hydrological predictors with the spread and maintenance of Leptospirosis as revealed by phylodynamics?

Data and Method

- Amazon region: the flooding time-scale can last many months
- Leptospirosis data: data concerning leptospirosis cases in Brazil will be retrieved from "DataSUS" website (www.datasus.gov.br/).
- Rainfall: precipitation data will be downloaded from INMET - Brazilian Institute of Meteorology - website.
- Phylogenetic data: complete genome sequences from environmental and clinical samples.
- Hydrological data: Based in a Digital Elevation Model (DEM, from SRTM project) it is possible to determine the Height Above Nearest Drainage (HAND) matrix.

Expected Results

A statistically principled reconstruction of the temporal and spatial dynamics of Leptospirosis and an assessment of the relative importance of several hydrological predictors on disease spread and maintenance.

Researchers

PI: Luciana Londe (Cemaden), Luiz Carvalho (Edinburgh), Leonardo Santos (Cemaden), Alice Marteli (Cemaden), Jair Koiller (INMETRO)

Collaborators: Francesco Ferruli, Matthew Brown, Andrés Sosa, Guilhermo Obregón, Walter Mendes Filho, Viviana Aguilar Muñoz

REFERENCES

EM-DAT, 2014 EM-DAT (2014) The OFDA/CRED International Disaster Database. http://www.em-dat.net/. Visitado em 30 jan. 2014

TOMINAGA et al., 2009 Tominaga LK, Santoro J, Amaral R (2009) Desastres naturais: conhecer para prevenir. Instituto Geológico, São Paulo.

Ko AI et al (1999) Urban epidemic of severe leptospirosis in Brazil. Salvador leptospirosis study group. Lancet 354:820

BARCELLOS & SABROZA, 2000 Barcellos C, Sabroza P (2000) Socio-environmental determinants of the leptospirosis outbreak of 1996 in western Rio de Janeiro: a geographical approach. Int J Health Search 10:301–313.

Costa et al., 2001 Costa E et al (2001) Formas graves de leptospirose: aspectos clínicos, demográficos e ambientais. Rev Soc Bras Med Trop 34(3):261–267

DataSUS (2016) Banco de dados do Sistema Único de Saúde. Accessed Sep 2017.

Oliveira et al., 2012 Oliveira TVS, Marinho DP, Neto CC, Kligerman DC (2012) Climate variables, living conditions and the health of the population: leptospirosis in the city of Rio de Janeiro from 1996 to 2009. Cienc Saud Col 17:1569–1576

Pelissari DM, Maia-Elkhoury ANS, Arsky MD (2011) Revisão sistemática dos fatores associados a` leptospirose no Brasil, 2000–2009/Systematic review of factors associated to leptospirosis in Brazil, 2000–2009. Epidemiol serv saude 20(4):565–574

Londe, L.., da Conceição, R.S., Bernardes, T. et al. Nat Hazards (2016) 84: 1419. https://doi.org/10.1007/s11069-016-2493-8

3.4.3 Soil Moisture - interaction with the group from Cemaden (Matt, Sheila, Chris)

Abstract

The Tropical Applications of Meteorology using SATellite and ground-based observations (TAMSAT) group at the University of Reading have developed a prototype decision-support tool for early warnings of drought - TAMSAT-ALERT (TAMSAT-AgricuLtural dEcision suppoRT). Currently, TAMSAT-ALERT relies upon accurate diagnosis of the current soil moisture state to make accurate sub-seasonal to seasonal assessments of drought risk in certain African countries. CEMADEN uses local data from farmers/children to drive predictive crop/weather models on sub-seasonal time scales in semi-arid regions. By combining the two methods and sharing data it should be possible to improve drought risk assessments for both Africa and Brazil.

The question we are asking is: Can soil moisture persistence, locally collected data and predictive models be combined to improve drought risk estimates?

The question TAMSAT-ALERT addresses is *Given the current state of the land surface and the historical climatology, what is the likelihood of agricultural drought?*

Specifically, it uses observed meteorological data - such as that from satellites, reanalysis datasets or local weather stations, and a climatology of the area and a land surface model to determine the current state of the land surface - e.g. the soil moisture. From this state the climatological data is used to project forward what the soil moisture could do in the future, and a distribution of possible soil moisture averages for a certain time period (such as the growing season) is produced. As the land surface takes time to respond to rainfall, and agricultural drought is a function of past and future soil moisture, it is possible to predict an agricultural drought in this way. For more details see http://onlinelibrary.wiley.com/doi/10.1002/wea.3033/full

The exact details of how the soil moisture persistence drought risk estimates of TAMSAT will be integrated with the drought risk products produced by CEMADEN depends on the details of the two individual products. A meeting is being set up on the 4th September to discuss said details.

Currently, the TAMSAT-ALERT system has only been trialled in a few African countries. However the method is quite general, in that it can be applied to various meteorological and land-surface parameters. It can also be used in countries in or near the tropics that have one or more rainy/growing seasons in a 365-day span with little additional configuration.

The TAMSAT-ALERT method relies upon the accuracy of land surface models to produce accurate drought risk assessments. The accuracy of these models is largely dependent on the accuracy of the soil properties and crop parameters used in the model. This information could be obtained using the framework that CEMADEN have set up to collect local data via apps used by the communities most affected by drought/flood.

The expertise at CEMADEN in using crop models for the purposes of drought risk assessments is something that TAMSAT could make use of, as this is an obvious next step for the group/product.

It is hoped that the expertise at TAMSAT in exploiting the soil moisture memory will also prove useful in helping CEMADEN improve their drought/flood risk assessment products.

Ultimately, both CEMADEN and TAMSAT’s risk assessment products should be improved, with some of the CEMADEN techniques applied/trialled in conjunction with TAMSAT-ALERT in Africa, and TAMSAT techniques applied/trialled in Brazil in a similar way.

3.4.4 Flooding damage estimation using urban mobility data - including economic impacts

Collaborators: Leonardo, Luciana, Andres, Viviana, Guillermo

Question: How to assess the different economic risks in floods?

To be able to understand and evaluate the impacts of disasters, it is necessary to study their economics consequences, locally and globally.

The aim of the work is to analyze the spatial and temporal distribution of flood events in a particular risk zone, observing the social, economic and financial impacts that entail this type of event. The method can be extrapolated to different events such as drought, and floods in different areas where they happen.

This type of analysis is an approach to the problem for evaluating options to mitigate the potential economic losses and, in case of not being able to mitigate the risks, to obtain the necessary capacity to adapt to the new situation. Besides it is important to generate capacity of transmission to the inhabitants of the highest risk regions, because the capacity to adapt varies considerably among regions, countries, and socioeconomic groups and will vary over time.

Based on an origin and destination data from the metropolitan region of Rio de Janeiro and using a Google Maps based script, Santos et al. (2015) estimated the amount of people directly and indirectly affected in their mobility in a potential flood episode. In cities as Rio de Janeiro and São Paulo, for example, there is an economic impact related to floods that could be studied to understand where the most vulnerable regions are. A Complex Network approach could be applied to this task.

The social-economic problem behind is the trade-off that exists between Climate change and economic globalization accused by all countries. In the phrase: "*economic development at the cost of ...*" the idea is that the challenges of climate change and globalization occur simultaneously and this leads to "games in the sense of Nash's theory" where there are winners and losers in the system. In the actuality, the Economic globalization develops a set of processes whereby production and consumption activities shift from the local or national scale to the global scale with new technologies that often lead to a negative situation towards the environment, such as gas emissions and water pollution. The policy makers should aim to allow technical progress but punish through taxes or prohibitions (and roughly) those who do not collaborate in the sustainable development of the environment. The aim is to analyze "the game" between growth and climate change in order to obtain some optimum to try to target.

Our aim is to develop a mathematical model, in order to evaluate the properties of risks in various social, economic and financial factors. Some mathematical and statistical techniques can be used, such as analysis of time series, modeling via partial differential equations, modeling of stochastic processes, utility function.

As expected results, we should understand the implications that floods can trigger in different areas. It is supposed that the study uses data from different databases such as hydrologic, topography, soil, vegetation. To analyze the social structure, data on the social-economic situation would be used to add the household level data to the model, and to analyze the economic part, data on relevant macroeconomic variables that are influenced by these catastrophic events should be used.

One possible application in terms of the financial aspect of the countries is to examine the effects of extreme events on agricultural trade. Economically, it is necessary for countries to evaluate the effects on yields, commodity prices, and imports and exports.

A clear example is the economy of developing economies, based on agriculture and the sale of commodities abroad, where this type of events has significant repercussions in the gross domestic product.

3.4.5 An early warning indicator for arboviroses in urban environment

Abstract

Arboviroses (dengue, yellow fever, zika, etc) are major health threats in Brazil and other developing tropical countries. Widespread epidemics occasionally overwhelm the healthcare system. Early warning systems give policy makers and health authorities time to prepare.

The final goal of the project would be to provide a complementary tool to early warning systems already in use (https://info.dengue.mat.br/).

Introduction

Arboviroses are human diseases caused by viruses transmitted by arthropods, in special mosquitoes. Arboviroses are a major health threat in tropical developing countries, causing deaths and losses in productivity due to illness-related work absenteeism.

During the summer, transmission rates increase dramatically, leading to widespread epidemics that overwhelm the healthcare system. In such a context, an early warning system offers policy makers and health authorities a tool to aid the best allocation of resources and personnel.

Much effort has been made over the years to model and understand these diseases, but only recently the role of urban mobility in disease spread and maintenance has been realised.

We aim at capitalising on an extensive body of work based on ordinary differential equations while at the same time incorporating data on human mobility data in order to construct an structured meta-population epidemic model.

Methods and Data

As a first stab, we propose a minimal deterministic meta-population model using EDOs (see eg. Bayley, 1975).

$$\begin{cases} \dot{S}_{h,i} = & \Lambda_{h,i} - T_{h,i}(S_{h,i}, I_v) - \mu_{h,i}\, S_{h,i} \\ \dot{I}_{h,i} = & T_{h,i}(S_{h,i}, I_v) - \gamma_{h,i}\, I_{h,i} - \mu_{h,i}\, I_{h,i} \\ \dot{R}_{h,i} = & \gamma_{h,i}\, I_{h,i} - \mu_{h,i}\, R_{h,i} \\ \dot{S}_{v,i} = & \Lambda_{v,i} - T_{v,i}(I_h, S_{v,i}) - \mu_{v,i}\, S_{v,i} \\ \dot{I}_{v,i} = & T_{v,i}(I_h, S_{v,i}) - \mu_{v,i}\, I_{v,i}, \end{cases}$$

The right hand side of the equations for $\dot{I}^h_j$ are evaluated with the current available data. ${I}^h_j(t)$ is the number, at time t, of infected humans living in district j. When the rate changes from negative to positive the red flag is raised. Since there is about a week of incubation period for the disease to manifest after the infection, this flag could help the policy-makers to plan health measures. Moreover, by solving the ODEs forward, one could make a prediction for the intensity of the disease on the various districts (updated daily as current data is assimilated).

Questions:

i) How to assimilate the available data? The number of infected humans can be estimated from notifications, and of infected mosquitoes by household indices such as Breteau. Estimating the susceptible fractions of human and mosquitoes is the major problem; we propose using the technique presented in the Workshop by Nicolas Rubido (compressive sensing), and the method of observers, currently being developed by Max Souza and A. Iggidr. Both estimate unknown parameters and non accessible data by sophisticated mathematical techniques.

Starting from the multi-scaled dengue system derived in [M.O. Souza, Multiscale analysis for a vector-borne epidemic model. J. Math. Biol., 68 (5), 1269–1293, 2014], the authors construct a pair of observers to estimate the dynamics of the disease. The nature of both the observers and the multi-scaled system allows to estimate both the number of Susceptible and Recovered hosts, as well as to provide information on the vector population, using only infected population data.

ii) A related question: disease notifications are assigned to the residence areas; however, many of the infections may have happened in other nodes (city sections) of the graph formed by the city districts (a mobility network - Santos et al., 2009). Can the inverse problem be addressed, namely finding the amount of real infections occurring on a given place? This can help authorities to decide where to apply vector control measures.

Expected results

Give to policy makers a computable threshold (updated weakly) indicating that an outbreak may occur at a given district of a city.

References

Bailey, N.J.T. (1975) The Mathematical Theory of Infectious Diseases and Its Application. 2nd Edition, Macmillan Publishers.

IGGIDR, A. ; KOILLER, J. ; PENNA, M.L.F. ; SALLET, G. ; SILVA, M.A. ; SOUZA, M.O. . Vector borne diseases on an urban environment: The effects of heterogeneity and human circulation. Ecological Complexity, v. 30, p. 76-90, 2017.

IGGIDR, ABERRAHMAN ; SALLET, GAUTHIER ; Souza, Max O. . On the dynamics of a class of multi-group models for vector-borne diseases. Journal of Mathematical Analysis and Applications (Print), v. 441, p. 723-743, 2016.

Santos, L. B. L.; Costa, M. C.; Pinho, S. T. R.; Andrade, R. F. S.; Barreto, F. R.; Teixeira, M. G.; Barreto, M. L. (2009): Periodic forcing in a three-level cellular automata model for a vector-transmitted disease. Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics (Print), v. 80, p. 016102

https://hal.inria.fr/hal-00839351/document

Onset of a vector borne disease due to human circulation - uniform, local and network reproduction ratios

http://www.dengue.mat.br/Anais2014.pdf An observer for a multi-scaled dengue. This is an announcement of work by Max Souza and A. Iggidr. This work is for an aggregated system. They are now extending their work to the metapopulation system.

Researchers

PI: Jair Koiller, Luiz Carvalho

Collaborators: Luciana R. Londe

Additional Questions

In this section we lay out some broader scientific questions that would be worth investigating in the long term.

Question 0: Which environmental and populational cues lead to toxin production by cyanobacteria?

Background: Cyanobacteria are unicellular organisms that exist in high numbers in large bodies of water, including those used in urban water supplies. Changes in environmental conditions sometimes prompt the bacteria to produce toxins that are harmful to several forms of life, from fish to humans. The drivers of this toxin production are poorly understood and mathematical modelling could be employed to describe and understand this phenomenon. Quorum sensing modelling has been a topic of increasing interest to mathematicians. It also offers the opportunity to applying techniques of machine learning combining data from bio-molecular testing with those of remote sensing, as well as meteorological information (winds, rain, temperature).

As this understanding evolves, a more ambitious project that the emerging consortium could work on, but that would require a more substantial funding and maturation time is creating a protocol for early warning of water supply toxicity outbreak. Predicting algae bloom events, specially if toxins are about to be released has great public health importance.

For this, two distinct operational lines exist nowadays:

Remote sensing via satellite: these exploit changes in the electromagnetic spectrum due to the presence of biota. INPE has the capability to introduce this service.

In situ, low cost microchips for real time detection of bio-molecular events have been used as early warning indicators .

References:

K. Montgomery, J. C. Charlesworth, R. LeBard, P.T. Visscher and B. P. Burns, Quorum Sensing in Extreme Environments, Life 2013, 3, 131-148; doi:10.3390/life3010131

Miller, M.; Bassler, B. Quorum Sensing in Bacteria. Ann. Rev. Microbiol. 2001, 55, 165199.

A. Ouellette, S.W Wilhelm, Toxic cyanobacteria: the evolving molecular toolbox, Front Ecol Environ 1(7): 359–366, 2003.

M.K. Rogalus, M. C.Watzin. Evaluation of sampling and screening techniques for tiered monitoring of toxic cyanobacteria in lakes, Harmful Algae, 7: 4, 504-514, 2008.

M. W. Matthews, D. Odermatt, Improved algorithm for routine monitoring of cyanobacteria and eutrophication in inland and near-coastal waters, Remote Sensing of Environment, 156, 374-382, 2015.

M. W. Matthews, Eutrophication and cyanobacterial blooms in South African inland waters: 10 years of MERIS observations, Remote Sensing of Environment, 155, 161-177, 2014.

D. J. Grimes, T. E. Ford, R. Colwell, C. Baker-Austin, J. Martinez-Urtaza, A. Subramaniam, D. G. Capone,

Viewing Marine Bacteria, Their Activity and Response to Environmental Drivers from Orbit Satellite Remote Sensing of Bacteria, Microb Ecol. 67(3): 489–500, 2014. doi: 10.1007/s00248-013-0363-4

S. Saha, A. Majumder, M. Hasanuzzaman, A. Ekbal, Bio-molecular event extraction using Support Vector Machine, Third International Conference on Advanced Computing (ICoAC), 2011 DOI: 10.1109/ICoAC.2011.6165192

4. Participants

Name	Organization	e-mail
Alan James P. Calheiros	Instituto Nacional de Pesquisas Espaciais (INPE)	alan.calheiros@inpe.br
Alice N. Martelli	CEMADEN	alicenmart2@gmail.com
Ana Carolina Xavier	IAC/Campinas	Anacarolinaf.xavier@gmail.com
Alejandro Cholaquidis Noblia	Univ. de la Republica	acholaquidis@cmat.edu.uy
Alessandra Corsi	IPT	accorsi@ipt.br
Alber Hammerson Sanchez	INPE	alber.ipia@inpe.br
André Lanfer Marquez	INPE	andre.lanfer@cptec.inpe.br
Andrés Sosa	Univ. de la Republica	asosa@cmat.edu.uy
Cassiano Antonio Bortolozo	CEMADEN	cassianoab@gmail.com
Celson Von-Randow	INPE	celso.vonrandow@inpe.br
Christopher Cunningham	CEMADEN	christopher.castro@cemaden.gov.br
David Guimaraes	CEMADEN	dvdgmf@gmail.com
Eugênio Sper de Almeida	INPE	eugenio.almeida@cptec.inpe.br
Fangxin Fang	Imperial College London	f.fang@imperial.ac.uk
Fernanda Aguiar	CEMADEN	fernanda.aguiar@cemaden.gov.br
Francesco Ferrulli	Univ. of Reading	f.ferrulli14@imperial.ac.uk
Guillermo Oswaldo Obregón Párraga	INPE	guillermo.obregon@inpe.br
Haroldo F. de Campos Velho	INPE	haroldo.camposvelho@inpe.br
Ivo Kenji Koga	INPE	ivo.koga@gmail.com
Jair Koiller	INMetro	jairkoiller@gmail.com
James Lain Jackaman	Univ. of Reading	james.jackaman@pgr.reading.ac.uk
James Targett	Imperial College London	james.targett10@imperial.ac.uk

Jean P. H.B. Ometto	INPE	jeanometto@gmail.com
Jefferson Luis Melo de Almeida Gomes	Universsity of Aberdeen	jefferson.gomes@abdn.ac.uk
Jemima Maple Tabeart	Univ. of Reading	jemima.m.tabeart@reading.ac.uk
Leonardo Bacelar Lima Santos	CEMADEN	santoslbl@gmail.com
Luana Albertani Pampuch	UNESP	luana.pampuch@ict.unesp.br
Lucas Massaroppe	USP	lucasmassaroppe@usp.br
Luciana de Resende Londe	CEMADEN	luciana.londe@cemaden.gov.br
Luciano Xavier	CEPEL	xavier@cepel.br
Luis Ricardo Lage Rodrigues	CCST	rodrigues.luisricardo@gmail.com
Luiz Max Carvalho	University of Edinburgh	lm.carvalho@ed.ac.uk
Maha Hussein Kaouri	Univ. of Reading	m.h.kaouri@pgr.reading.ac.uk
Marcelo Barreiro	Univ. de la Republica	mar.bar.par@gmail.com
Marcelo Zeri	CEMADEN	marcelo.zeri@cemaden.gov.br
Matthew James Brown	Univ. of Reading	matthew.brown@physics.ox.ac.uk
Nicolás Rubido Ober	Univ. de la Republica	nrubido@fisica.edu.uy
Oliver James Sutton	Univ. of Reading	o.sutton@reading.ac.uk
Rodrigo S. Costa	INPE	rodrigo.costa@inpe.br
Rong Zhang	CEMADEN	rong.zhang@cemaden.gov.br
Selma R.S. Santos	CEMADEN	Selma.santos@cemaden.gov.br
Sheila Santana de Barros Brito	CEMADEN	sheilasbarros@gmail.com
Tasmin Louise Symons	Imperial College London	tls111@ic.ac.uk
Tiago Jose de Carvalho	IFSP	tiagojc@ifsp.edu.br
Tristan Pryer	Univ. of Reading	t.pryer@reading.ac.uk
Vinícius Veloso de Melo	UNIFESP	vinicius.melo@unifesp.br
Viviana Aguilar Muñoz	CEMADEN	viviana.munoz@cemaden.gov.br
Walter M. Mendes Filho	ITA	walter.mendesfilho@gmail.com

5. Talks

Mathematics of Climate Climate change and natural disasters

HAROLDO CAMPOS VELHO

In his talk, a description of challenges associated to the weather, climate and disasters were addressed. Model calibration is one topic cited. Firstly, this issue was shown as a multi-objective problem. Secondly, the process requires a sensitivity analysis, classifying into different groups the impact of those parameters in the model. In addition, the sensitivity analysis depends on the time-scale of prediction period. Finer model resolution implies higher computational cost. The tendency is to use hybrid computer machines. The companies are offering different types of co-processors to be linked with the CPU: GP-GPU (Nvidia, AMD), FPGA (Xilinx, Altera), MIC (Xeon Phi Intel). Some worked examples with different resolutions were shown. Data science application was also commented. A method combining statistical analysis (p-value) and artificial intelligence scheme (Decision Tree) was applied to identify extreme events: deep drought in the Amazon region, and intense rainfall events in the South of Brazil.

Research Group “Mathematics and Disaster Risk Reduction”

Towards International collaborations in Natural Disasters prevention: personal experiences and contacts

LEONARDO BACELAR LIMA SANTOS

JAIR KOILLER

Leonardo presented the research group “Mathematics and Disaster Risk Reduction”, held in the Brazilian Applied Mathematic Society (SBMAC). “We are now thinking about internationalization”, with the group "Mathematics of Planet Earth", from Imperial College and University of Reading. The workshop was the first step in this direction. Jair Koiller described some of the results of the arbovirus diseases network he participated in, presented his contacts related to early warning and prevention, specially the connection with MPECDT, and mentioned some themes he would be interested to collaborate with. Besides the promotors of the workshop (ICL/Reading Math for Planet Earth, Cemaden/INPE, Universidad de la Republica) he suggested the following Institutions that could be invited for a network: Inmetro, IMPA, FGV, Fiocruz, UFRJ, IPRJ (Institutions in Rio de Janeiro), and PTI (Itaipu Technological Park) in Iguassu, Paraná. Internationally, his main contacts are INRIA (France), Caltech (USA), IST (Instituto Superior Tecnico. Lisbon) and the working group “Statistics Without Borders”, from the American Statistical Association.

Adaptivity in geophysical flows

TRISTAN PRYER

In Meteorological simulations, as computational capacity grows, one typically invests in increased resolution, more ensemble runs, addition of PDEs governing new physical processes. This has a massive impact on the computational complexity. Indeed, a new problem arising in simulations is the power cost of the total ensemble runs. The next generation of dynamical cores will require considerable innovation for them to be viable. Adaptivity is crucial in the success of algorithms for geophysical multiscale problems. One of the novelties introduced in this talk is the idea of 'model adaptivity', the automatic switching between models of different complexity in real time as and when required.

Evolutionary Epidemiology

LUIZ MAX FAGUNDES DE CARVALHO

Evolutionary Epidemiology (phylodynamics) is an emerging field of scientific enquiry that combines phylogenetics (genomic data) with epidemiological, environmental and socio-economical data in order to understand the driving factors of pathogen evolution, spread, virulence and maintenance. In my talk I motivate the use of phylodynamic methods with two examples on Ebola virus spread and virulence, respectively. I show that approaches that combine several sources of data in a principled way can give deep insight into the driving factors behind disease dynamics. I also discuss the methodological and data acquisition challenges that we face in the 21st century. I argue that we need better, more realistic models as well as more (and better data). I also argue that addressing the methodological challenges is the most pressing problem due to the exponential growth in data availability over the past decade.

Next Generation Multiscale Adaptive Mesh Atmospheric Modelling, Rapid Response and Data Assimilation

JEFF GOMES

FANGXIN FANG

A multiscale adaptive mesh fluid model (Fluidity) for general multi-physical problems has been presented. The innovative and novel features of this model include anisotropic adaptive mesh technology library, user-friendly GUI (Graphical User Interface) and a Python interface for model setup (initial and boundary conditions and diagnostic test fields). The model has been successfully applied to simulate natural disaster problems, e.g., air pollution in urban street canyons, atmospheric cyclones, Tsunami events, dust storms during Asian Monsoon and urban flooding. Predictive modelling methods (combination of data analysis and management and flow simulation) embedded in Fluidity were also presented to: (a) introduce uncertainty analysis; (c) demonstrate improvement of accuracy of numerical predictions and; (c) reduce computational costs. Among the methods, data assimilation, reduced order modelling with deep learning and optimized sensor locations have been further discussed here.

Brazilian Centre for Monitoring and Early Warnings of Natural Disasters

LUCIANA LONDE

In Brazil hurricanes and earthquakes are not common. On the other hand, droughts, floods, flash floods and landslides are frequent and may cause many problems. The flash floods are more common in the South and Southeast regions, wild fires in the north and west central, droughts in the Southeast and Northeast, landslides in the South and Southeast and floods basically all over the country. The highest mortality rates due to disasters in Brazil occur in the southeast region. The droughs in the brazilian semi-arid are responsible for severe long term impacts to the local people. The landslides in Rio de Janeiro state in 2011, which caused over 900 fatalities, motivated the creation of Cemaden, the Brazilian center for monitoring and early warnings of natural disasters. It is a center located in São José dos Campos, which works 24 hours every day, 7 days a week, for a permanent monitoring of hazards in the country. Disasters are signs of failures - failures of preparedness, response, and recovery. Most often they are failures of long-term development and risk reduction planning. They grow on underlying societal challenges such as inequality or poverty, termed "root causes" and "unsafe conditions". We need investigations that assume that the goal is to increase development sustainability, not only to reduce losses and damages from poor development in the past, although this is already part of the equation. We need research that works from the positive side of the action in an innovative way. The final goal is a sustainable and fair development and not the management of risks and disasters, built on an unfair and bewildering development model.

Compressive Sensing in Non-linear Dynamics

NICOLÁS RUBIDO

Compressive Sensing (CS) is a rather novel method to efficiently acquire and reconstruct signals from solving indeterminate linear systems of equations. CS requires sparsity, namely, few data points suffice. Hence, its results are counter-intuitive, since the sampling requirement apparently violates the well-known minimum sampling frequency established by the Shannon-Nyquist theorem. Despite its linear nature and apparently odd requirements, its application goes far beyond linear systems. In this talk, I review some applications suitable for time-series analysis. Particularly, in the cases when the objective is to predict the trajectory of a chaotic system from few measurements or when the intention is to model the equations of motion from time-series measurements to predict catastrophes.

Interacting Networks in Climate

MARCELO BARREIRO

The talk titled "Interacting networks in climate" presented the use of complex networks for the study of several climate phenomena. A brief description of the methodology was first introduced, followed by three examples that cover the different climatic time scales. In the first example we showed how climate networks can improve our understanding of global teleconnections on intraseasonal to interannual time scales. In the second example the use of a directionality measure to construct the climate network allowed to study the impact of the equatorial Pacific on other regions of the world, as well as to uncover the extratropical atmospheric dynamics that dominate on different time scales. In the last example, we constructed a network with nodes including the climate modes that dominate the tropical oceans and rainfall over southeastern South America (SESA) to study the evolution of their collective behavior during the 20th century and how it will change in the future under a scenario of anthropogenic forcing. We found that the collective influence of the oceans in rainfall in SESA varies on interannual and interdecadal time scales and that the response to radiative forcing is nonlinear.

Earth System Science Center

JEAN PIERRE OMETTO

The presentation explored the different lines of research carried out by the Center for Earth System Science, of the National Institute for Space Research, from environmental monitoring networks to climate change modeling and future scenarios. The goal was to expose the audience to the opportunities for collaborations and to drawn a proposal among the researcher's team from the Institutions represented at the workshop.