



## XIII SIMPÓSIO DE RECURSOS HÍDRICOS DO NORDESTE

### **DIFUSÃO DE DADOS ORBITAIS NOS ESTUDOS DE MUDANÇAS NA COBERTURA DO SOLO NO SEMIÁRIDO BRASILEIRO**

*John Elton de Brito Leite Cunha<sup>1</sup>; Iana Alexandra Alves Rufino<sup>1</sup>; Carlos de Oliveira Galvão<sup>1</sup>; Thiago Emmanuel Pereira<sup>2</sup>; Francisco Vilar Brasileiro<sup>2</sup>; Esdras Vidal Pereira<sup>2</sup>*

**RESUMO** – A hidrologia e os recursos hídricos demandam o monitoramento do uso e cobertura do solo, relacionando-os ao impacto do clima e da ação humana. No entanto, muito frequentemente os dados necessários a tal monitoramento e análises decorrentes são de escalas espaciais que não podem ser coletados integralmente em campo. Dados e técnicas de sensoriamento remoto atendem a essas necessidades, pois incluem detecção de mudanças naturais e de origem humana em escalas que vão desde paisagens locais a continentais. Este artigo apresenta uma iniciativa intercontinental: o projeto *EUBrazil Cloud Connect*, desenvolvido entre parceiros brasileiros e europeus. O objetivo principal é fornecer uma infraestrutura para difusão dos dados obtidos por sensores orbitais e ferramentas para análise multitemporal de tendências de mudanças na superfície terrestre. Para tanto, são utilizadas ferramentas de computação em nuvem para o processamento do algoritmo SEBAL (*Surface Energy Balance Algorithm for Land*) para uma série superior a 30 anos de imagens orbitais para todo semiárido brasileiro e serviços *Web* para visualização, análise e disponibilização para tomadores de decisão e comunidade científica.

**ABSTRACT** – Hydrology and water resources demand monitoring land use and cover, related to the impacts of climate and human action. However, very often data for such monitoring and sequent analysis are from spatial scales that cannot be fully collected by field survey. Remote sensing techniques and data are suitable to those needs, since include land use/land cover changes detection in different scales (from local to continental landscapes). This paper presents an intercontinental initiative: the *EUBrazil Cloud Connect* project, developed by European and Brazilian partners. The main goal is to provide a cloud computing infrastructure to use tools for multi-temporal analysis and trend analysis of huge remote sensing databases to understand the main current drivers of land use changes. SEBAL (*Surface Energy Balance Algorithm for Land*) algorithm has been processed for a long time series (more than 30 years of satellite images) covering the whole Brazilian semi-arid area. Web services for visualization, analysis and deployment for decision makers and researchers are used.

**Palavras-chave:** processamento de imagens, *cloud computing*, análise multitemporal, semiárido, vegetação.

1) Universidade Federal de Campina Grande, Centro de Tecnologia e Recursos Naturais, john.e.cunha@gmail.com, iana.alexandra@ufcg.edu.br; carlos.galvao@ufcg.edu.br

2) Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática, thiago.manel@gmail.com; fubica@computação.edu.br

## 1- INTRODUÇÃO

Nos últimos anos têm sido observadas perdas de importantes serviços ecossistêmicos, resultantes do efeito antropogênico (Foley *et al.*, 2005; MEA, 2005; D'Odorico *et al.*, 2013), representando uma séria ameaça para a subsistência sustentável, incluindo os recursos hídricos. No entanto, a variedade de aplicações ambientais e de recursos hídricos exige dados de escalas espaciais que não podem ser coletados integralmente em campo. Dados e técnicas de sensoriamento remoto atendem a essas necessidades, pois incluem detecção de mudanças naturais e de origem humana em escalas que vão desde paisagens locais a continentais (Wang *et al.*, 2010; Pereira *et al.*, 2013).

As raízes do Sensoriamento Remoto Orbital (SRO) encontram-se nas disciplinas de geografia e engenharia (Pettrorelli *et al.*, 2014), sendo necessário na maioria das aplicações que seus usuários conheçam os fundamentos da cartografia, Sistemas de Informação Geográfica (SIG), ferramentas computacionais e física da atmosfera.

As pesquisas ambientais utilizando SRO estão provendo dados em grandes escalas, mas estão tendo dificuldades em encontrar arquiteturas de *software* para acomodar o grande volume de dados (Li *et al.*, 2010). As plataformas *Amazon's Elastic Compute Cloud*, *Google App Engine* e *Microsoft's Windows Azure* estão entre as recentes tecnologias que permitem maior capacidade computacional para análise e processamento de grandes volumes de dados (*Big Data*) ambientais advindos do SRO (Wang *et al.*, 2015; Ma *et al.*, 2015; Yang *et al.*, 2015).

No entanto, há desafios na conversão dos métodos e algoritmos utilizados tradicionalmente pela comunidade de usuários de SRO: algumas análises visuais e interativas, realizadas pelo operador, precisam ser substituídas por procedimentos computacionais automáticos; em muitos casos, há a necessidade de trabalhar com frações da área de estudo; automatizar os processos de coleta de dados em banco de dados, muitas vezes heterogêneo e inconsistente; e a necessidade de implementação de procedimentos computacionais para o gerenciamento dos processos, desde a coleta de dados à geração do produto final.

Neste sentido, este artigo apresenta as iniciativas do projeto *EU Brazil Cloud Connect* para difusão dos dados obtidos por sensores orbitais e ferramentas para análise multitemporal e análise de tendências de mudanças na superfície terrestre. Para tanto, são utilizadas ferramentas de computação em nuvem (*cloud computing*) para o processamento do Algoritmo SEBAL (*Surface Energy Balance Algorithm for Land*) para uma série superior a 30 anos de imagens orbitais para todo semiárido brasileiro e serviços Web para visualização, análise e disponibilização dos produtos para tomadores de decisão e comunidade científica.

## 2- METODOLOGIA DO TRABALHO

O sistema desenvolvido no âmbito do projeto *European Union Brazil Cloud Connect* prevê processamentos que ocorrem em segundo plano e em tempo real. Os processamentos que ocorrem em segundo plano utilizam o sistema *Blowout* para aquisição de dados climáticos e imagens de satélite e execução dos algoritmos *Fmask - Function of Mask* (Zhu e Woodcock, 2012) e *SEBAL* (Bastiaanssen, 2000). Os processamentos em tempo real são realizados no *BioClimate Scientific Gateway* (BSG) são destinados aos usuários finais, permitindo análise de tendências, geração de mapas e extração de informações em diferentes formatos (CSV, GeoTIFF e PNG).

No projeto *EUBrazilCC* optou-se por utilizar imagens dos satélites LANDSAT 5, 7 e 8, que apresenta resolução espacial de 30 metros e temporal de 16 dias. O conjunto de imagens disponíveis para o semiárido reúne cerca de 42 mil imagens, divididas em 57 cenas. As imagens do satélite LANDSAT são obtidas diretamente do repositório de imagens na *National Aeronautics and Space Administration* (NASA).

As informações sobre o clima necessárias para a aplicação do algoritmo SEBAL são selecionadas com base na localização das imagens de satélite e horário da passagem do sensor na região de interesse. São necessárias também as informações sobre a velocidade e temperatura do ar (últimos 30 anos) e para as imagens disponíveis a cada nova passagem do satélite pelo mesmo local. Os dados climáticos utilizados nesta aplicação são obtidos no Instituto Nacional de Meteorologia (INMET).

Para detecção das nuvens e sombra de nuvens nas imagens LANDSAT é utilizado o algoritmo *Fmask*. Em testes com imagens de referência globalmente distribuídas, foi verificado que a acurácia do método foi de 96,4%, sendo esse algoritmo uma boa opção para identificação de nuvens em imagens LANDSAT (Zhu e Woodcock, 2012).

Tradicionalmente, a aplicação do algoritmo SEBAL envolve a interação com o usuário para identificação dos pixels de referência, esse procedimento é conhecido como o CIMEC - *Calibration Using Inverse Modeling at Extreme Conditions* e é usado pelos algoritmos SEBAL e METRIC-*Mapping EvapoTranspiration at high Resolution with Internalized Calibration* (ALLEN *et al.*, 2011). A automatização do processo de escolha dos pixels de referência foi realizada com base em Allen *et al.* (2013), algumas adaptações foram incorporadas para melhor representação e adequação aos aspectos locais do semiárido brasileiro.

### 3.1- Arquitetura do Sistema *Blowout*

O sistema *Blowout* usado nesse estudo usa recursos de computação de uma federação de provedores de IaaS (*Infrastructure as a Service*) para execução dos algoritmos *Fmask* e SEBAL. Essa federação é formada pela agregação dos recursos ociosos disponíveis em cada provedor membro da

federação. Provedores IaaS geralmente mantêm uma fração de seus recursos computacionais ociosos para lidar com variações da demanda por seus recursos bem como para tratar falhas temporárias na infraestrutura. Do ponto de vista de um provedor, juntar-se à uma federação é bastante conveniente por uma razão simples: é possível ter acesso a um conjunto de recursos computacionais maior do que os recursos locais.

Nesta seção, descrevemos em linhas gerais a arquitetura do sistema *Blowout*. Essa arquitetura, ilustrada na Figura 1, tem seis componentes principais: *submission service*, *task catalog*, *crawler*, *scheduler*, *worker node* e *fetcher*. Todos esses componentes são implantados em máquinas virtuais providas em recursos da federação. Além desses componentes, o *Blowout* utiliza o *middleware* Fogbow (Barros *et al.*, 2015) para alocar e acessar os recursos federados.

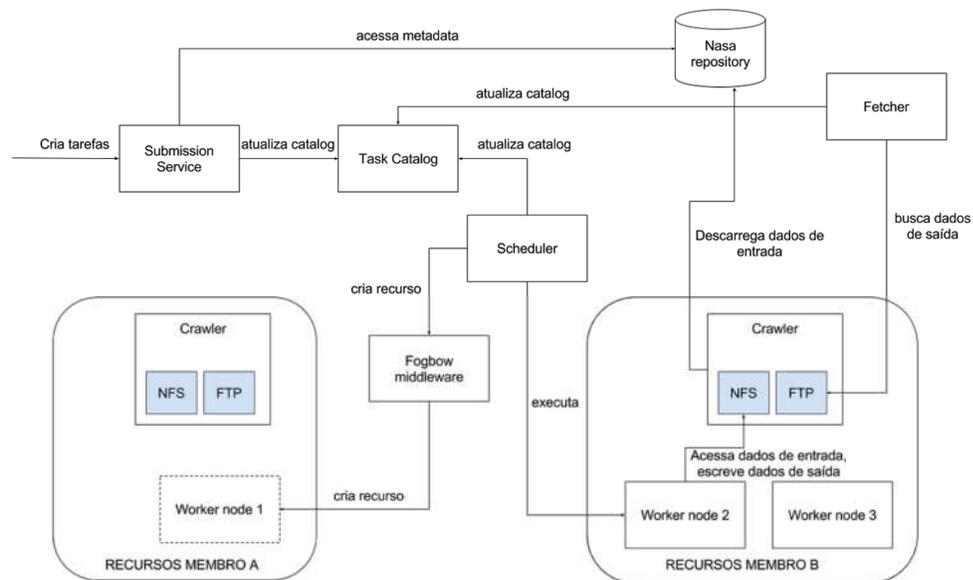


Figura 1 – Sistema *Blowout*

O componente *submission service* atende requisições dos usuários do *Blowout* de um membro da federação. Entre as requisições atendidas estão a criação de novas unidades de trabalho, o monitoramento da execução dessas unidades de trabalho e o expurgo de dados já processados (para os quais não se tem mais interesse).

O componente *task catalog* mantém registros que descrevem as unidades de trabalho em processamento em uma instalação *Blowout*. Cada unidade corresponde a uma imagem LANDSAT. Novos registros são adicionados ao *task catalog* como resultado de uma requisição de criação atendida pelo *submission service*; a requisição considera datas que correspondem ao momento de criação das imagens LANDSAT.

O componente *crawler* é responsável por obter as imagens LANDSAT da base de dados NASA relacionadas com as unidades de trabalho registradas. Após obter uma imagem, o *crawler* interage

com o componente *task catalog* para indicar sua obtenção no registro correspondente. As imagens obtidas são compartilhadas com os demais componentes do *Blowout* através do sistema de arquivos distribuídos NFS (Pawlowski *et al.*, 1994). Além das imagens LANDSAT, que são os dados de entrada usados pelo SEBAL, o componente *crawler* armazenará também os dados resultantes da execução do SEBAL. O acesso para esses dados é disponibilizado através de um serviço FTP em execução no *crawler*.

O componente *scheduler* é responsável por definir o *worker node* que processará uma determinada imagem obtida pelo *crawler*. É também função do *scheduler* interagir como o middleware Fogbow para alocar máquinas virtuais nas quais serão implantados novos componentes *worker nodes*. Isso pode ser feito tanto em resposta para falhas em recursos que tenham sido alocados antes quanto para aumentar o conjunto de *worker nodes* disponível; por exemplo, para aumentar a taxa de processamento de imagens. Ainda, o *scheduler* monitora a execução dos algoritmos *Fmask* e SEBAL no *worker node*. Uma vez terminada a execução (após os dados de saída terem sido copiados para o *crawler*), o *scheduler* indica ao *task catalog* para indicar que o registro relacionado indique que a imagem já foi processada.

O componente *fetcher* é responsável por agregar o resultado do processo dos algoritmos *Fmask* e SEBAL, após identificar no *task catalog* as unidades de trabalho que já foram processadas. Os dados de saída do algoritmo são coletados dos componentes *crawler* correspondentes através do serviço FTP disponibilizado por esses. Após obter esses dados de saída, o componente *fetcher* também modifica os registros correspondentes no *task catalog* indicando que os resultados foram agregados. Essa indicação permite que o *crawler* remova os dados que mantinha relacionados com a imagem processada, uma vez que o *fetcher* os copiou.

### 3.2 -BioClimate Scientific Gateway

O *BioClimate Scientific Gateway* (BSG) proporciona ao usuário uma interface de análise de alto nível, permitindo o acesso aos dados, análise e visualização através de múltiplas fontes de dados heterogêneos, expondo uma visão integrada. Além disso, suporta várias características, tais como séries temporais e análise estatística, análise dos dados e comparação entre as séries de diferentes variáveis.

Para execução das tarefas em tempo real, o sistema utiliza a técnica de *Parallel Data Analysis Service* (PDAS) e permite agendar as tarefas de acordo com os recursos disponíveis e ainda tem elasticidade para explorar os recursos de nuvens subjacentes. O trabalho desenvolvido por Elia *et al.*, (2016) apresenta maiores detalhes sobre as funcionalidades do sistema *BioClimate Scientific Gateway*.

No BSG estarão disponíveis nove variáveis obtidas pela aplicação do algoritmo SEBAL: *Normalized Different Vegetation Index* (NDVI), *Leaf Area Index* (LAI), *Enhanced Vegetation Index* (EVI), Albedo de superfície ( $\alpha$ ), Temperatura de superfície (TS), Saldo de Radiação (Rn), Fluxo de calor no Solo (G), Fração evapotranspirativa (EF) e Evapotranspiração de 24 horas (ET<sub>24h</sub>). Essas variáveis poderão ser analisadas conjuntamente com variáveis: precipitação e temperatura do ar disponibilizadas pelo *Climate Research Unit* (CRU), distribuição de ocorrências de plantas obtidos a partir de modelagem de nicho ecológico e métricas florestais obtidas a partir de imagens LiDAR 3D.

### 3- RESULTADOS PRELIMINARES

A Figura 2 apresenta as cenas que são submetidas ao processamento do algoritmo SEBAL usando o sistema *Blowout* desenvolvido no âmbito do projeto EUBrazilCC e a área escolhida para as análises usando o BSG. Para apresentação dos resultados preliminares foram escolhidas duas variáveis biofísicas obtidas pela aplicação do algoritmo SEBAL: *Enhanced Vegetation Index* (EVI) e Saldo da radiação (Rn), obtidos para a cena de orbita 215 e ponto 65. O período de análise é compreendido entre os anos de 1984 a 1995, reunindo 32 imagens.

O índice de vegetação EVI é sensível à variação da estrutura do dossel, redução das influências atmosféricas e do solo no monitoramento da vegetação. Com o EVI é possível monitorar a cobertura do solo e, assim, observar o padrão fenológico da vegetação, além de sua alteração em função do regime pluviométrico ou oriundos da intervenção humana.



Figura 2 – Detalhe do recorte utilizado para aplicação no *BioClimate Scientific Gateway*

Para testes, utilizou-se o *Bounding Box* indicado na Figura 2, observa-se uma tendência de decréscimo no índice de vegetação (EVI) durante o período analisado (Figura 3). Para esse período,

também foi observado uma tendência de redução das precipitações, seguindo comportamento semelhante ao EVI. Essa condição, evidencia a influência dos totais precipitados sobre a dinâmica da cobertura do solo na região de Caatinga. No entanto, o período de análise é relativamente pequeno, 12 anos, podendo ser necessário uma série temporal maior para distinguir se os efeitos da dinâmica da cobertura do solo são de origem antrópica ou resultantes da variabilidade climática.

Para o período de disponibilidade de dados, a análise indica que as mudanças ocorridas são motivadas pela variabilidade do clima (chuva), devido ao comportamento semelhante da tendência de precipitação e EVI.

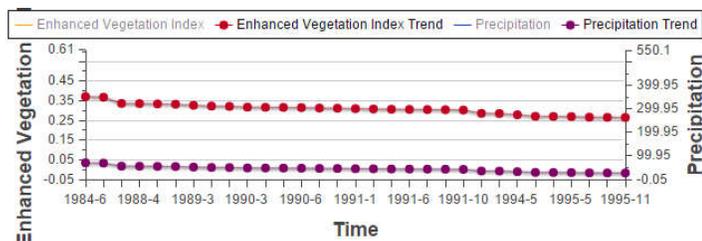


Figura 3 – Análise de tendência com EVI e precipitação

Ao analisar temporalmente os 32 momentos (Figura 4), observa-se, novamente, a influência das chuvas na cobertura do solo. Quando há aumento da disponibilidade hídrica, há uma resposta positiva da vegetação e vice versa. Entre os anos de 1989 e 1991, e 1992 a 1994 houve sequências de períodos com baixa precipitação, resultando nos menores valores de EVI. Isso não quer dizer necessariamente que houve perda de cobertura do solo, pois ao retorno das chuvas no ano de 1995, observa-se um novo aumento do EVI. Contudo, quando se compara os maiores valores da série, observa-se que eles estão presentes no início do período estudado. Esse efeito pode ser resultante da ação antrópica ou da persistência de anos secos que podem provocar a perda de alguns indivíduos no período analisado.

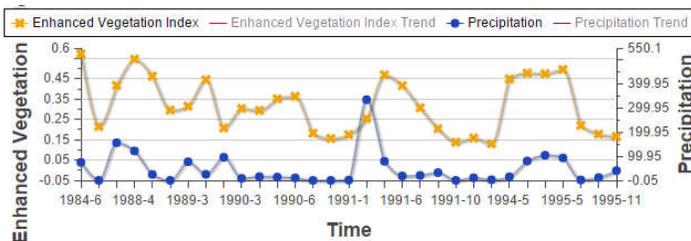


Figura 4 – Série temporal de EVI e precipitação

O Rn e a Temperatura do ar (Tar) sofrem influência da disponibilidade de energia na superfície (Figura 5). O Sol é o grande responsável por fornecer radiação para superfície da Terra, como essa disponibilidade não é uniforme durante o ano, observa-se o comportamento cíclico destas variáveis.

Observa-se que Rn e Tar apresentam comportamento semelhante, aumentos no Rn são acompanhados da Tar e vice e versa.

No entanto, ao analisar a tendência (Figura 6), observa-se um aumento da temperatura do ar e redução do Rn. A tendência observada, pode ser um indicativo de redução da cobertura vegetal, provocando um efeito inverso ao comportamento natural. O Rn é composto por radiação de onda curta (Solar) e longa (Temperatura). As reduções na cobertura vegetal, aumentam a refletividade da superfície, resultando em menor quantidade de radiação de onda curta e consequente, redução no Rn. Por sua vez, a redução da cobertura vegetal pode contribuir para aumento de temperatura de superfície e consequente aumento da temperatura do ar, conforme observado nos gráficos de tendência (Figura 6). Dessa forma, a análise integrada de diferentes produtos ambientais permite uma melhor compreensão da dinâmica da cobertura do solo e de suas respostas às mudanças.

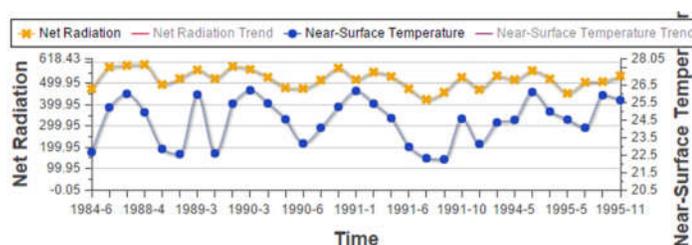


Figura 5 – Série temporal de Rn e temperatura do ar

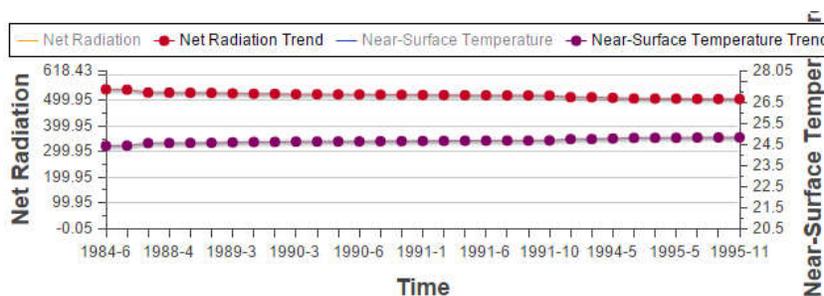


Figura 6 – Análise de tendência de Rn e temperatura do ar

## 5- CONCLUSÕES

Ferramentas e algoritmos para processamento de grandes massas de dados de imagens de satélite através de computação em nuvem permitem expandir a capacidade de análise de dados de mudança de uso e cobertura do solo atualmente disponível às comunidades de recursos hídricos e meio ambiente.

A análise de uma série de imagens (32) da forma como foi realizada, é um trabalho penoso do ponto de vista computacional e operacional, devido às limitações de processamento e de tempo. Neste sentido, a integração de análises estatísticas e análises cruzadas do BSG e a capacidade de processamento do algoritmo SEBAL em séries ainda mais longas no Sistema *Blowout*, representam um avanço significativo nos estudos de mudanças do uso e cobertura do solo no semiárido brasileiro.

## AGRADECIMENTOS

Este projeto é resultante do Edital MCTI/CNPq 13/2012 – Programa de Cooperação Brasil – União Europeia na Área de Tecnologias da Informação e Comunicação – TIC.

## 6- BIBLIOGRAFIA

- ALLEN, R. G.; PEREIRA, L. S.; HOWELL, T. A.; JENSEN, M. E. (2011) “*Evapotranspiration information reporting: I. Factors governing measurement accuracy*”. *Agricultural Water Management*, 98, pp.899–920.
- ALLEN, R.; BURNETT, B.; KRAMBER, W.; HUNTINGTON, J.; KJAERGAARD, J.; KILIC, A.; TREZZA, R. (2013) “*Automated calibration of the METRIC-Landsat evapotranspiration process*”. *Journal of the American Water Resources Association*, 49(3), pp. 563-576.
- BARROS, A.; BRASILEIRO, F.; FARIAS, G.; GERMANO, F.; NÓBREGA, M.; RIBEIRO, A.; SILVA, I.; TEXEIRA, L. (2015) “*Using Fogbow to federate private clouds*” in XXXIII Brazilian Symposium on Networks and Distributed Systems – Tools track. Vitória-ES, Maio-2015.
- BASTIAANSEN, W. G. M. (2000) “*SEBAL - based sensible and latent heat fluxes in the irrigated Gediz Basin, Turkey*”. *Journal of Hydrology*, 229, pp. 87-100.
- D’ODORICO, P.; BHATTACHAN, A.; DAVIS, K. F.; RAVI, S.; RUNYAN, C. W. (2013). “*Global desertification: drivers and feedbacks*”. *Advances in Water Resources*, 51, 326-344.
- ELIA, D.; NUZZO, A.; NASSISI, P.; FIORE, S.; BLANQUER, I.; BRASILEIRO, F. V.; RUFINO, I.; SEIJMONSBERGEN, A.; ANDERS, N.; GALVÃO, C.; CUNHA, J.; SOUSA-BAENA, M.; CANHOS, V.; ALOISIO, G. (2016) “*A Science Gateway for Biodiversity and Climate Change Research*” in 8th International Workshop on Science Gateways (IWSG 2016), June-2016.
- Fogbow project, <http://fogbowcloud.org/>
- FOLEY, J. A.; DEFRIES, R.; ASNER, G. P.; BARFORD, C.; BONAN, G.; CARPENTER, S. R.; HELKOWSKI, J. H. (2005). “*Global consequences of land use*”. *Science*, 309, 5734, pp. 570-574.
- LI, J.; HUMPHREY, M.; AGARWAL, D.; JACKSON, K.; VAN INGEN, C.; RYU, Y. (2010). “*E-Science in the cloud: a Modis satellite data reprojection and reduction pipeline in the windows azure platform.*” *Parallel & Distributed Processing (IPDPS) IEEE International Symposium*. pp.1-10.
- LIMA, R. D. C. C.; CAVALCANTE, A. D. M. B.; PEREZ, A. M. P. A. M.; PEREZ-MARIN, M. (2011). “*Desertificação e mudanças climáticas no semiárido brasileiro*”. Campina Grande-PB: Instituto Nacional do Semiárido, 209p.
- MA, YAN; WU, H.; WANG, L.; HUANG, B.; RANJAN, R.; ZOMAYA, A.; JIE, W. (2015) “*Remote sensing big data computing: Challenges and opportunities*”. *Future Generation Computer Systems*, 51, pp. 47-60.
- MEA. Millennium Ecosystem Assessment. (2005) “*Ecosystems and Human Well-Being: Desertification Synthesis*”. Washington DC: World Resource Institute.
- PAWLOWSKI, B.; JUSZCZAK, C.; STAUBACH, P.; SMITH, C.; LEBEL, D.; HITZ, D. (1994). “*NFS Version 3 Design and Implementation*”. USENIX.
- PEREIRA, H. M.; FERRIER, S.; WALTERS, M.; GELLER, G. N.; JONGMAN, R. H. G.; SCHOLES, R. J.; COOPS, N. C. (2013). “*Essential biodiversity variables*”. *Science*, 339(6117), pp.277-278.
- PETTORELLI, N.; SAFI, K.; TURNER, W. (2014). “*Satellite remote sensing, biodiversity research and conservation of the future*”. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369 (1643) pp. 20130190.

- R CORE TEAM. (2016). “*A language and environment for statistical computing*”. R Foundation for Statistical Computing, Vienna, Austria.
- RSTUDIO TEAM. (2016) “*RStudio: Integrated Development for R*”. RStudio, Inc, Boston.
- WANG, K.; FRANKLIN, S. E.; GUO, X.; CATTET, M. (2010). “*Remote sensing of ecology, biodiversity and conservation: a review from the perspective of remote sensing specialists*”. *Sensors*, 10(11), pp. 9647-9667.
- WANG, X. Z.; ZHANGA, H. M.; ZHAO, J. H.; LIN, Q. H.; ZHOU, Y. C.; LI, J. (2015). “*An Interactive Web-Based Analysis Framework for Remote Sensing Cloud Computing*”. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1, pp. 43-50.
- YANG, C.; SUN, M.; LIU, K.; HUANG, Q.; LI, Z.; GUI, Z.; LOSTRITTO, P.HAOWEI. (2015). “*Contemporary computing technologies for processing big spatiotemporal data*”. In: *Space-Time Integration in Geography and GIScience*. Springer Netherlands, pp. 327-351.
- ZHU, Z.; WOODCOCK, C.E. (2012) “*Object-based cloud and cloud shadow detection in Landsat imagery*”. *Remote Sensing of Environment*, 118, pp. 83-94.