# Revisão Bibliográfica

## Atividade agrícola no Brasil

O setor econômico considerado como mais importante do país é o setor primário, composto pela agricultura, pecuária e extrativas. O setor agrícola, além de fornecer matérias primas e alimento, essenciais para a sobrevivência humana, é responsável também por fontes de energias renováveis e responsáveis como parcela considerável em combustível, seja etanol ou biodiesel (CONAB 2014).

O Brasil está entre as dez maiores economias do mundo, é o segundo maior fornecedor mundial de alimentos e produtos agrícolas e, segundo a OECD/FAO, está prestes a se tornar o principal fornecedor, para atender à demanda mundial adicional. O crescimento da oferta deverá ser impulsionado por melhorias contínuas na produtividade, com o rendimento das culturas mais altas, conversão de pastagens em lavouras e a produção de gado mais intensa. Segundo projeções da OECD/FAO (2015) o Brasil deverá se tornar o maior exportador agropecuário do mundo em 2024 (Figura 1).

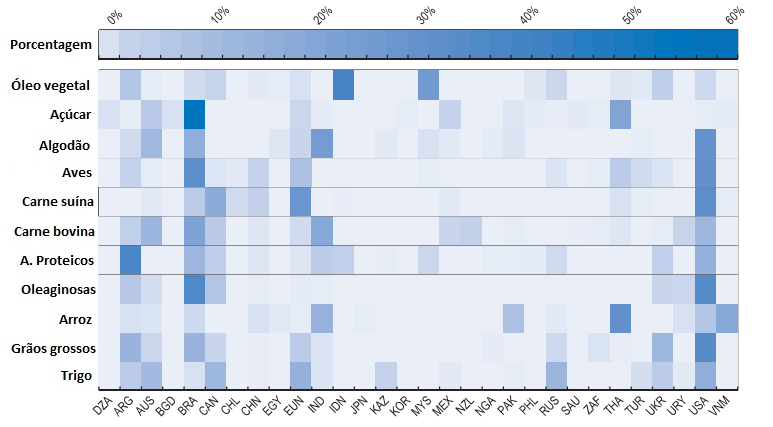


Figura 1 - Concentração das exportações por commodities, projeção 2024 - Adaptado de OECD/FAO, 2015.

A agricultura brasileira vem apresentando, nos últimos anos, importantes transformações em sua estrutura. Ao longo destas transformações, o governo vem sustentando uma ampla parcela por meio de uma política direcionada ao desenvolvimento tecnológico. Esta política pode ser facilmente compreendida quando se analisa a composição do agronegócio brasileiro (34% do PIB), distribuído entre agropecuária (30%), insumos agropecuários (6%), agroindústria (31%) e distribuição (33%) - (Rocha et al., 2014).

## Modelos de estimativa de produtividade agrícola

O principal objetivo do manejo da cultura agrícola em qualquer país é garantir recursos alimentares para a sua população. Vários fatores podem causar a escassez de alimentos, resultando em fome ou no desencadeamento de crises econômicas. A predição da produção antes de o tempo de colheita envolvendo grandes regiões é, dessa forma, uma questão importante para muitos países, sendo de suma importância no o norteamento de políticas agrícolas, no seu gerenciamento e planejamento, podendo ser realizada diretamente, por meio de amostras coletadas antes da colheita, período de maturação de acordo com calendário, ou indiretamente pela utilização de ferramentas, métodos e técnicas de previsão como modelos. No entanto, existe uma grande variedade de modelos e um número de problemas associados a eles (Delécolle et al., 1992; Soria-Ruiz et all, 2004).

Os estudos que buscam relacionar as interações entre o clima e a produtividade são geralmente desenvolvidos com o uso de modelos que procuram quantificar os efeitos das variações do clima sobre o comportamento vegetal (Robertson, 1983).

Acock & Acock (1991) e Delécolle et al. (1992) definem os estudos do efeito do clima na produtividade em três tipos de modelos: "empírico-estatístico" ou estatístico, "matemático-mecanicista" ou de crescimento e conceitual ou “semi-empírico”.

Os modelos estatísticos utilizam relações empíricas onde os fatores ambientais são inseridos em equações de regressão para predizer a produtividade.

Segundo Camargo et al. (1994) o modelo "empírico-estatístico", descreve as relações entre as variáveis, sem considerar os processos, apresentando muitas restrições para a extrapolação de resultados.

O modelo conceitual, o mais complexo, requer grande número de informações dos processos físicos e biológicos envolvidos, como o modelo de simulação de crescimento, maturação e produtividade em cítricos, proposto por Bem Mechlia e Carrol (1989).

Já o modelo matemático-mecanicista descreve matematicamente a casualidade ou a relação entre os processos envolvidos no ciclo da cultura, possuindo menores restrições em relação a extrapolação de resultados do que o empírico-estatístico.

### Balanço Hídrico do Solo - Método de Esgotamento FAO (BHFAO)

A determinação do balanço do solo é fundamental para o entendimento da relação solo-planta e permite determinar o nível de água no solo na zona de raízes em uma base temporal, diária, decendial ou mensal (STEDUTO et al., 2012). Por meio dele é possível determinar se a cultura está em estresse hídrico ou não e, dessa forma, se seu desenvolvimento será ou não prejudicado, tendo até, em casos extremos, perda completa da safra.

Para a determinação do balanço hídrico, dados agrometeorológicos, dados da cultura e também dados do solo são utilizados, tendo como entradas precipitação, irrigação e ascensão capilar e como saídas evapotranspiração, escoamento e percolação profunda. Assim, o balanço hídrico é a variação da água armazenada ou do esgotamento dessa água (Richetti, 2014).

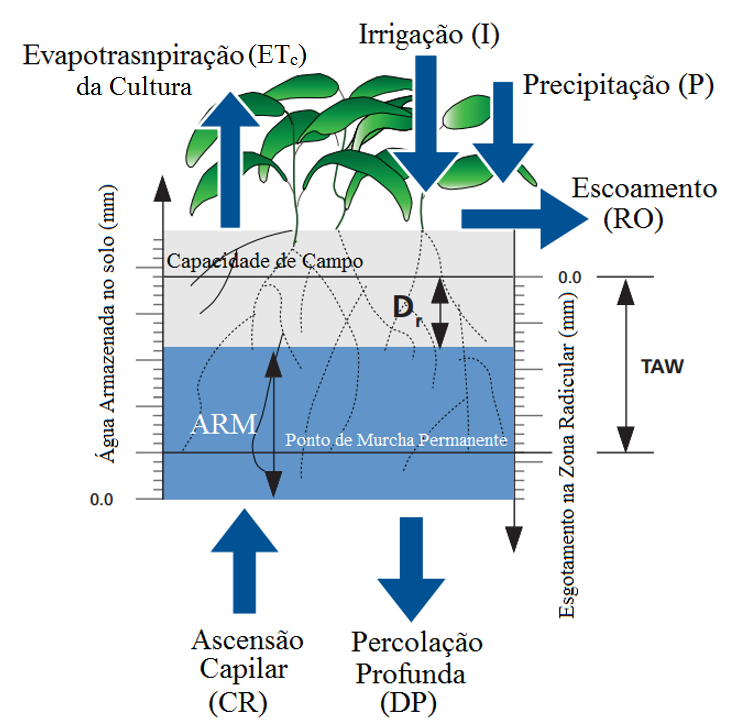


Figura 2 Balanço hídrico do Solo. (Adaptado de Steduto et al., 2012).

Dentre as formas de determinar essa necessidade hídrica da cultura, neste trabalho apresenta-se para validação do software o cálculo do balanço hídrico apresentado por Allen et al. (1998), doravante BHFAO.

O método de esgotamento da FAO (BHFAO) vem sido aplicado em estudos que utilizam de dados agrometeorológicos para estimativas de produtividade, Richetti et al. (2014) gerou dados do balanço hídrico, BHFAO, do solo para a cultura da soja para o oeste do Paraná, e Rojas (2007) utilizou-se do BHFAO para determinação de estresse hídrico e *ETa* para a cultura do milho no Quênia.

O balanço hídrico é apresentado pela FAO (1998) conforme a Equação 1.

|  |  |  |
| --- | --- | --- |
|  | 𝑩𝑯𝒓,𝒊 = 𝑩𝑯𝒓,𝒊−𝟏 + (𝑷 − 𝑹𝑶)𝒊 + 𝑰𝒊 − 𝑪𝑹𝒊 − 𝑬𝑻𝒄,𝒊 − 𝑫𝑷𝒊 | () |

em que: BHr,i é o balanço hídrico na zona radicular no final do tempo i (mm); BHr, i-1 é a quantidade de água no período anterior i-1 (mm); Pi é a precipitação no tempo i (mm); ROi é o escoamento superficial no tempo i (mm); Ii é a irrigação no tempo i (mm); CRi é a ascensão capilar do solo no tempo i (mm); ETC,i é a evapotranspiração da cultura no tempo i (mm); DPi é a perda de água devido a percolação profunda no tempo i (mm).

Os valores de ascensão capilar (*CR*), percolação profunda (*DP*), escoamento superficial (*RO*) e irrigação (*I*) são desconsiderados neste trabalho. Os primeiros, ascensão capilar, percolação profunda e escoamento superficial, pela impossibilidade de determinar tais valores e a irrigação, pois as áreas não possuem sistema de irrigação.

Para determinar-se da evapotranspiração da cultura (ETc) utiliza-se os valores de Kc, dependente da cultura e os valores de evapotranspiração de referência (ET0).

O total de água disponível no solo (*TAW*) é a capacidade de armazenamento de água do solo multiplicada pela profundidade desejada, ou seja, a capacidade de água disponível (CAD) multiplicada pela profundidade efetiva das raízes da cultura. É um fator limitante ao BHr,i (0 >= BHr,i <= TAW), ou seja o quantia de água (BHr) no momento i não pode ser menor que 0 e nem maior do que a TAW do momento i, qualquer água armazenada além dessa medida ficará impossibilitada de ser útil para a planta, portanto é desconsiderada (Allen et al., 1998).

Com esses dados calcula-se o esgotamento na zona radicular no final do tempo, desta forma tendo todos os parâmetros necessários para a realização do cálculo do balanço hídrico para a cultura da soja é o valor oposto do esgotamento na zona radicular no final do tempo, determinando assim o BHFAO para o período determinado. Maiores detalhes sobre o balanço hídrico FAO estarão presentes na seção métodos.

### Calculo de produtividade – Modelo FAO

O modelo utilizado pela FAO propõe uma equação simples em que a redução relativa da produtividade está relacionada com a correspondente redução relativa da evapotranspiração (ALLEN et al., 1998). Especificamente, a resposta da produtividade é expressa pela Equação 2:

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

em que:

*Yx* e *Ya* são os valores respectivos de produtividade potencial e produtividade atingível (t ha-1);

*ETc* e *ETa* são os valores máximo e atual de evapotranspiração (mm);

*Ky* é o fator de produtivida de que representa a sensibilidade da produtividade da cultura a estresses hídricos (adimensional).

O fator *Ky* captura a essência das complexas relações entre a produtividade de uma cultura e sua necessidade de água, em que diversos processos biológicos, químicos e físicos estão envolvidos (STEDUTO et al, 2012), e é o fator de sensibilidade ao estresse hídrico da cultura. São específicos para cada cultura e variam de acordo com o estádio da cultura, sendo que:

* *Ky* > 1: a cultura é muito sensível ao déficit hídrico com reduções de produtividade proporcionais ao estresse hídrico sofrido;
* *Ky* < 1: a cultura é tolerante ao déficit hídrico, recuperando-se parcialmente de estresses hídricos sofridos, sendo, assim, apresentam reduções de produtividade menores proporcionalmente em relação a redução de água disponível;
* *Ky* = 1:a produtividade é diretamente proporcional ao déficit hídrico.

Para se determinar o valor máximo de produtividade, também chamado de produção bruta de matéria seca da cultura padrão (*Yx*, kg ha-1) utiliza-se o método de Kassam (1977), que se baseia em níveis de radiação, temperatura e dias nublados e limpos. Determina-se a produtividade potencial pelo somatório das produtividades potenciais brutas (*PPR*, kg ha-1) para a cultura (Equação 3).

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

A produtividade potencial bruta para a cultura é determinada pela produtividade potencial bruta para cultura padrão (*PPBp*, kg ha-1, Equação 4) e o índice de colheita (*Cc*, adimensional). Este índice de colheita é responsável pela calibração do modelo na região em estudo.

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

O valor de *PPBp* é a produtividade potencial bruta para dias nublados (*PPBn*, Equação 5) somada à produtividade potencial bruta para dias limpos (*PPBc*, Equação 6)

|  |  |  |
| --- | --- | --- |
|  |  | (5) |
|  |  | (6) |

As produtividades potenciais são dependentes das correções de temperatura, para dias nublados (*ctn*, Equação 7) e para dias limpos (*ctc*, Equação 8), e da radiação no topo da atmosfera (*Q0*, Equação 9).

|  |  |  |
| --- | --- | --- |
|  |  | (7) |
|  |  | (8) |
|  |  | (9) |

em que:

*Q0* é a radiação no topo da atmosfera (*MJ m-2*);

*DR* é a distância relativa Sol-Terra (UA, Equação 10);

*lat* é a latitude do local (°);

*hn* é o ângulo horário (°, Equação 11);

é a declinação solar (°, Equação 12);

|  |  |  |
| --- | --- | --- |
|  |  | (10) |
|  |  | (11) |
|  |  | (12) |

em que:

*DJ* é o dia Juliano.

### Determinação das datas da cultura

Para determinação das datas de semeadura, máximo desenvolvimento vegetativo e colheita utilizando sensoriamento remoto é necessário analisar os dados de forma temporal. Ao se analisar perfis temporais de IV como o EVI é possível identificar as regiões de pico e os pontos de inflexão correspondentes as áreas onde encontram-se as datas da cultura (GRZEGOZEWSKI et al., 2013; BECKER et al., 2016), como exemplificado na Figura 2.

Ponto de inflexão

Ponto de pico

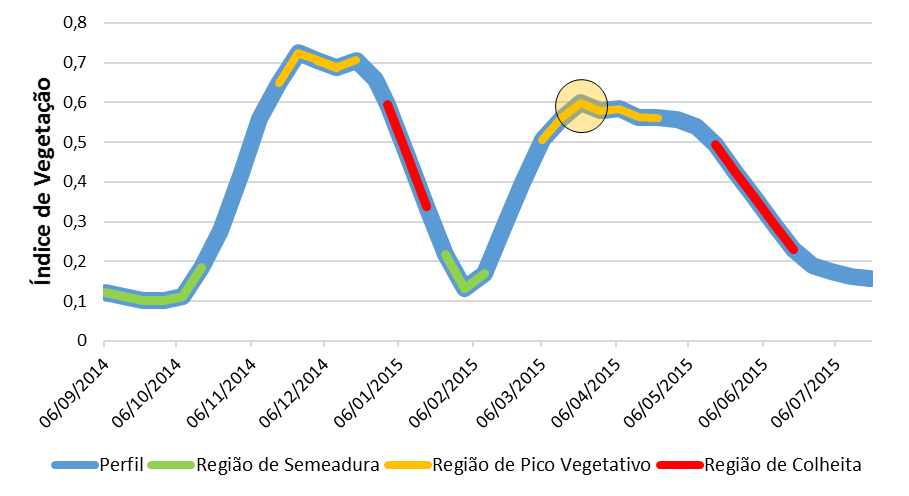


Figura 2 Perfil temporal do índice de vegetação EVI de culturas agrícolas.

Segundo Becker et al. (2016) as datas de semeadura, pico vegetativo e colheita podem ser identificadas com base nos pontos de pico e inflexão no perfil EVI, identificados as datas dos pontos de pico e inflexão é feita uma regressão, estimando as datas de semeadura e colheita utilizando os pontos de inflexão e pico. Dessa forma criasse um ajuste percentual em relação a diferença delta tempo entre o ponto de inflexão e o pico no perfil que é somado ao ponto de inflexão (Eq. 1)

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

em que:

= data da cultura a ser estimada (data de semeadura ou colheita);

= data do ponto de inflexão;

= data do ponto de pico;

= coeficiente de ajuste percentual;

## Sensoriamento remoto orbital

Relacionar as variáveis agrometeorologicas com a produtividade de uma cultura não é tarefa fácil (Araújo et al., 2011). Segundo Rudorff & Batista (1990), parte das questões que dificultam o estudo entre essa relação podem ser solucionadas ao se incorporar informações espectrais de culturas agrícolas no modelo, estas informações podem ser obtidas em imagens de satélites por sensoriamento remoto e são expressas como índice de vegetação. Resultados parecidos foram encontrados em Johann (2012) que realizou estimativas de produtividade de soja, obtidas por meio dos modelos espectrais e modelos agro-espectrais, observando que o desempenho dos modelos melhora com a inserção de dados agrometeorológicos.

### Sensor Modis

A técnica de sensoriamento remoto mostra-se eficiente no monitoramento da cobertura vegetal, evidenciando as variações temporais, detecção de mudanças, dinâmica sazonal e fenologia (YU et al., 2003). O sensor Moderate Resolution Imaging Spectroradiometer (MODIS) é um espectrorradiômetro orbital a bordo das plataformas (satélites) TERRA e AQUA, lançadas, respectivamente, em 1999 e 2002. Com 36 bandas, abrangendo as radiações do visível e do infravermelho, este sensor permite a obtenção de dados relacionados à vegetação, cobertura terrestre, nuvens e aerossóis, incidência de fogo, neve e cobertura de gelo no mar. As imagens desse sensor permitem o monitoramento da cobertura global quase em tempo real, possuindo as seguintes resoluções: espacial (250, 500 e 1.000 m), espectral (545 a 14.385 nm) e temporal (de 1 a 2 dias) (JUSTICE; TOWNSHEND, 2002; COUTO JUNIOR *et al.*, 2011).

Índices de vegetação podem ser obtidos a partir de dados do sensor MODIS, disponibilizados gratuitamente na forma composições de imagens de 16 dias, com resolução espacial de 250m. Entre os diversos produtos MODIS são disponibilizados dois índices de vegetação: o Índice de Vegetação por Diferença Normalizada (Normalized Difference Vegetation Index – NDVI) e o Índice de Realce da Vegetação (Enhanced Vegetation Index - EVI) (JUNGES, ALVES e FONTANA, 2007).

Os índices de vegetação são baseados em valores digitais de intensidade, para avaliar a biomassa ou vigor vegetativo. São formados por várias combinações de valores espectrais que são somados, divididos ou multiplicados de um modo a produzir um valor único que indica a quantidade biomassa ou o vigor vegetativo dentro de um pixel. Altos valores do IV identificam pixels cobertos por proporções substanciais de vegetação saudável. Algumas relações entre bandas têm sido definidas aplicando os conhecimentos de comportamento espectral da vegetação viva. No caso da vegetação viva essa estratégia pode ser especialmente efetiva por causa a relação inversa entre a refletância nas regiões espectrais do vermelho e do infravermelho (CAMPBELL & WYNNE, 2011).

Os índices de vegetação minimizam as interferências relativas às variações provenientes do ângulo solar e dos efeitos atmosféricos, observados em dados multitemporais. Enquanto o NDVI é sensível à clorofila, o EVI é mais sensível às variações na resposta estrutural do dossel, incluindo o índice de área foliar (LAI), a fisionomia da planta e a arquitetura do dossel (COUTO JUNIOR et al., 2011, HUETE et al., 2002)

O EVI foi desenvolvido para otimizar o sinal de resposta da vegetação, melhorando a sensibilidade em regiões com maiores densidades de biomassa, além de proporcionar o monitoramento da vegetação através de uma ligação do sinal de fundo do dossel e da redução das influências atmosféricas (HUETE et al. 1999). O EVI pode ser calculado por meio da Equação 2.

|  |  |  |
| --- | --- | --- |
|  |  | (13) |

em que, segundo Justice et al. (1998):

= reflectância no Infravermelho Próximo;

= reflectância no Vermelho;

= reflectância no Azul;

= Coeficiente de correção dos efeitos atmosféricos para o vermelho (6);

= Coeficiente de correção dos efeitos atmosféricos para o azul (7,5);

= Fator de correção para a interferência do solo (1);

= Fator de ganho (2,5).

O NDVI é um bom estimador de biomassa, sendo o índice mais utilizado em pesquisas que são relacionadas à dinâmica da cobertura vegetal, tais como o acompanhamento e a avaliação de rendimento de culturas na agricultura de precisão (PONTES, 2005; JUNGES *et al*., 2007). O NDVI, proposto por Rouse *et* *al*., (1973), é dado pela Equação 2:

|  |  |  |
| --- | --- | --- |
|  |  | (14) |

em que:

= reflectância no infravermelho próximo;

= reflectância no vermelho.

Embora numericamente os valores do NDVI possam variar entre -1 e 1, a vegetação está associada aos valores positivos. Materiais que refletem mais intensamente na porção do vermelho em comparação com o infravermelho próximo (nuvens, água e neve) apresentam NDVI negativo (PONTES, 2005).

### Modelo ECMWF

Grande parte dos estudos realizados obtém informações de dados de superfície proveniente de estações meteorológicas. Porém, os órgãos responsáveis pela coleta e organização destes dados, não os disponibilizam aos usuários em tempo real, numa quantidade (séries históricas) e formato que possam ser utilizados para análises espaciais (JOHANN, 2012). Uma fonte alternativa para obtenção de informações de dados agrometeorológicos tem sido o modelo ECMWF (*European Center for Medium-Range Weather Forecast*, 2010) que apresenta dados globais e gratuitos.

O ECMWF (*European Centre for Medium-Range Weather Forecasts*) é o Centro de Previsão do Tempo de Médio alcance Europeu. Trata-se de uma organização intergovernamental com 34 países, e com base em Reading, no oeste de Londres – Reino Unido (WOODS, 2006).

O ECMWF apresenta variáveis agrometeorológicas importantes para o estudo e aplicação de modelos de estimativa agrícola, as variáveis disponíveis no sistema são: temperatura média (ºC), temperatura máxima (ºC), temperatura mínima (ºC), precipitação total (somatório, mm), evapotranspiração total sobre a água (somatório, mm), evapotranspiração total sobre o solo (somatório, mm), evapotranspiração de Penman-Monteith (somatório, mm), radiação global total (somatório, kJ.m-2), média de profundidade da neve (cm), profundidade mínima de neve (cm), profundidade máxima de neve (cm), balanço hídrico climatológico (somatório, mm), velocidade média do vento (média, m.s-1) e pressão de vapor água médio (média, hPa). Além dos campos adicionais, como Latitude, Longitude, Ano, Mês, Decêndio e data da aquisição dos dados (ECMWF, 2015).

Os dados são coletados por meio de um sistema que engloba dados de estações meteorológicas espalhadas pelo mundo, radares meteorológicos, satélites, entre outras fontes. Estes dados são coletados a cada 6 horas. Os dados globais, em uma resolução espacial de um grau de longitude e latitude são obtidos, processados e depois organizados em forma de uma grade de 25 km (0,25º), sendo disponibilizados gratuitamente no *website* do JRC (*Joint Research Center*) em formato Shape. Os dados desde 1989 estão disponíveis no modelo ERA INTERIN, que é modelo de reanálise em escala reduzida, isto é, com a grade de 0,25º no lugar da antiga grade de um grau (JRC, 2015).

## Softwares e geotecnologias

Para desenvolvimento das ferramentas e estruturas de software propostas, uma série de tecnologias foram e serão estudas, visando garantir agilidade nos processos de construção, maleabilidade nas estruturas, re-usabilidade de códigos e facilidade de utilização e de manutenção.

### Geospatial Data Abstraction

A Geospatial Data Abstraction (GDAL) consiste em uma biblioteca construída na linguagem C++ para tradução de formatos de dados geográficos, (imagens, *raster*) e dados vetoriais (*shapefiles*) geoespaciais/georeferenciados. Abstraindo os mais diferentes formatos de arquivos matriciais *e* vetoriais possibilitando trabalhar com quaisquer tipos da mesma forma. Além disso a GDAL fornece uma série de algoritmos de processamento em linha de comando, como quatro tipos de algoritmo para interpolação, algoritmos de proximidade, para composição de *rasters* e outros (WARMERDAM, 2008).

A GDAL é distribuída pela Open Source Geospatial Foundation, sob a licença X/MIT estilo *Open Source*. Possui distribuições disponíveis para as linguagens: Perl; Python; VB6 Bindings; Java; C# / .Net; Ruby; R.

As vantagens de se utilizar esta biblioteca em *Geografic Information Systems* (SIG’s) são a capacidade de abstração dos variados tipos e por ela fornecer um grande desempenho computacional.

### Linguagem Python

Python é uma linguagem de alto nível, não tipada, orientada a objetos e livre para ser utilizada em fins comerciais. A linguagem Python foi concebida nos anos de 1980, mas só foi implementada em 1989 por Guido van Rossum nos Países Baixos, sua versão 2.0 criada em 2000 foi amplamente distribuída e sua utilização vem sendo também associada a processamento de imagens para mais diversos ramos da ciência, justamente por suas características de ser orientada a objetos, não ter tipagem e permitir sobrecarga de operadores (VENNERS, 2003). A utilização da linguagem Python para SIG’s não é recente, hoje ela é utilizada pra SIG’s mundialmente conhecidos como ArcGis e QGIS.

Hoje ela está na versão 3.0, lançada em 2008. Modificações importantes foram feitas ao longo de sua evolução, estas tornaram a linguagem limpa, de fácil aprendizado e de muito bom desempenho, apesar disso suas versões 2.6 e 2.7 ainda são as utilizadas na maioria dos SIG’s atuais por questões de compatibilidade e estabilidade.

Para este trabalho a versão Python utilizada será a 2.7 por fornecer maior compatibilidade com diversas distribuições de bibliotecas interessantes ao escopo do software.

### Orientação a objetos

A metodologia de programação orientada a objetos (OO) consiste em definir e implementar hierarquias abstratas de tipos de dados. Tipos abstratos são definidos por dados e procedimentos associados para manipular esses dados (Braz et al., 2001).

A OO é um paradigma, cada tipo de dado é decorado com características e comportamentos próprios que consistem em atributos e procedimentos/funções. Estes podem herdar características por meio de uma hierarquia, dadas as particularidades e limitações de cada linguagem. A OOP é diferentemente implementada em cada linguagem, porém segue sempre os mesmos princípios conceituais.

A modelagem orientada a objetos pode ser definida como uma técnica de projeto de software. Para essa técnica admite-se que se esteja interessado principalmente na clareza e organização do projeto, tendo por base uma representação clara e eficiente da aplicação, facilitando ao máximo o desenvolvimento e a manutenção do software (Agostini, 2002; RUMBAUGH et al., 1994).

### Unified Modeling Language

O Undefied Modeling Language (UML) é um dos métodos de modelagem mais utilizados no desenvolvimento de software. Ele inclui vários pontos de vista e diagramas para diferentes fins e uso. O UML também é utilizado para formar modelos conceptuais de vários tipos de objetivos (Li et al, 2009).

A UML é uma linguagem de modelagem gráfica de propósito geral. Os objetivos da UML incluem (Li et al, 2009):

* Especificar, visualizar, construir e documentar os artefatos de um sistema de software;
* Entender, desenhar, pesquisar, configurar, manter e controlar informações sobre os sistemas de software;
* Ser usado com todos os métodos de evolução, ciclos de vida, domínios de aplicativos e mídia.

### Padrões de projeto

Os padrões de projeto (*Design Patterns*) tornam mais fácil reutilizar padrões bem-sucedidos e arquiteturas. Utilizando técnicas comprovadas e conhecidas como padrões de projeto a aplicação se torna mais acessível aos novos desenvolvedores ao sistema. Os padrões de projeto podem até mesmo melhorar a documentação e manutenção de sistemas existentes, fornecendo uma definição explícita de classe e objeto interações e sua intenção subjacente (Gamma et al., 1995).

O movimento ao redor de padrões de projeto ganhou popularidade com o livro Design Patterns: Elements of Reusable Object-Oriented Software, publicado em 1995. Os autores desse livro, Erich Gamma, Richard Helm, Ralph Johnson e John Vlissides, são conhecidos como a "Gangue dos Quatro" (Gang of Four) ou simplesmente "GoF".

Os padrões "GoF" são organizados em 3 famílias:

* Padrões de criação: relacionados à criação de objetos;
* Padrões estruturais: tratam das associações entre classes e objetos;
* Padrões comportamentais: tratam das interações e divisões de responsabilidades entre as classes ou objetos.