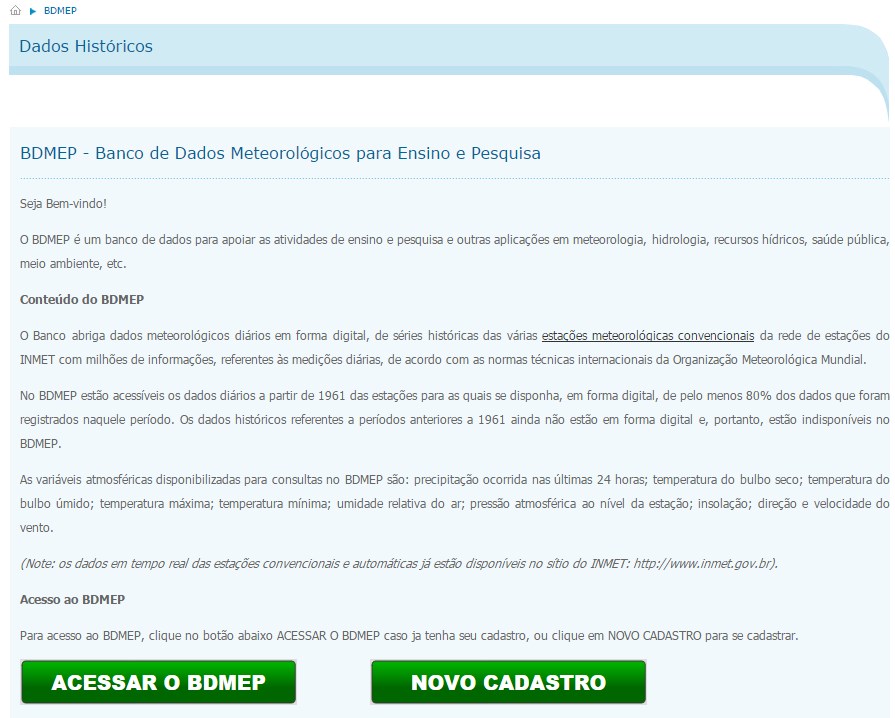
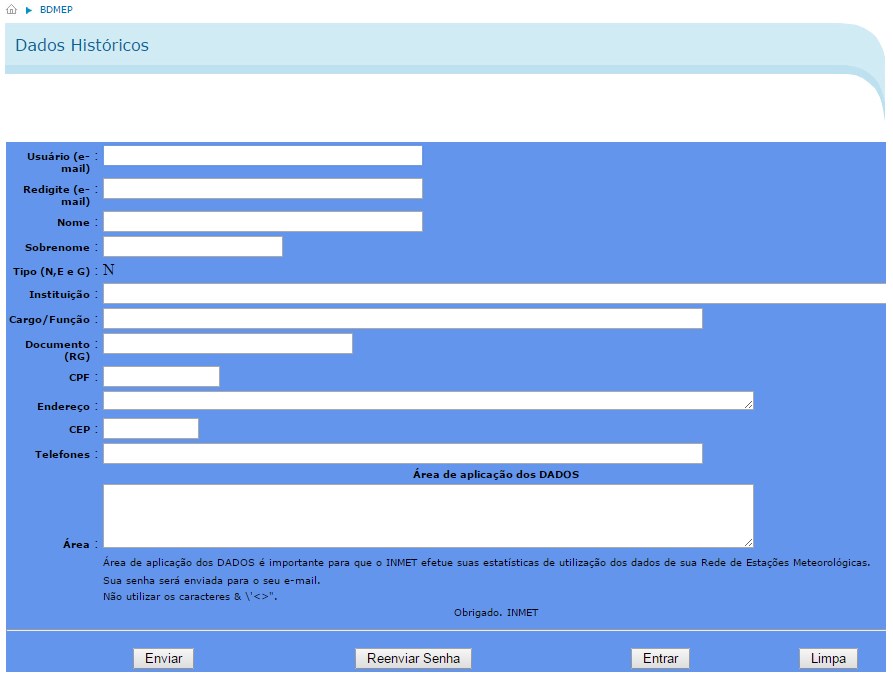
**Recuperação de Dados Meteorológicos do BDMEP**

O BDMEP oferece uma grande massa de dados meteorológicos que podem ser consultados gratuitamente por estudantes e pesquisadores. Para ter acesso a esses dados é necessário acessar o site do BDMEP:

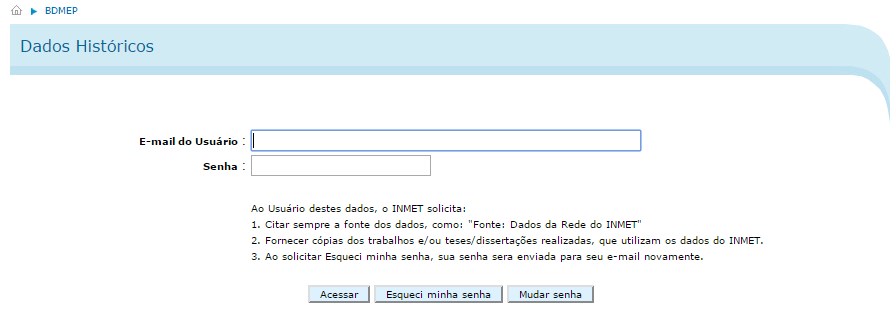
<http://www.inmet.gov.br/portal/index.php?r=bdmep/bdmep>



Em seguida realizamos um novo cadastro aonde são fornecidos dados pessoais e aonde serão aplicados os dados fornecidos pelo BDMEP.

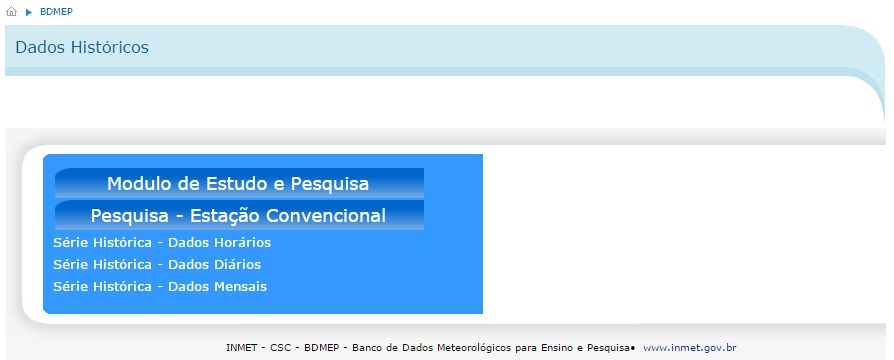


Depois de ter completado o cadastro no BDMEP será enviado para seu email as informações de acesso para que possamos usufruir dos dados de estações meteorológicas espalhadas por todo o país.



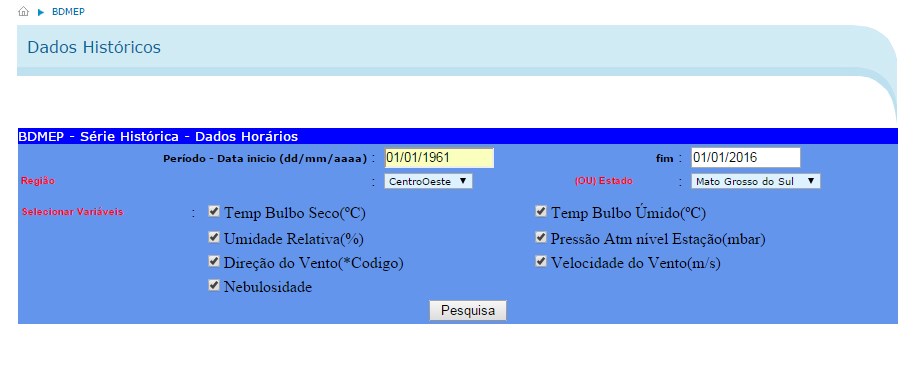
O BDMEP disponibiliza uma série histórica de dados que podem ser consultadas em três escalas: Horaria, Diária e Mensal.

Cada uma dessas escalas fornecem uma quantidade de variáveis do tempo sendo a mais completa a mensal por ser medida em um espaço maior de tempo. Depois de realizar o login com suas credencias recebidas por e-mail, é disponibilizada as opções de pesquisa.



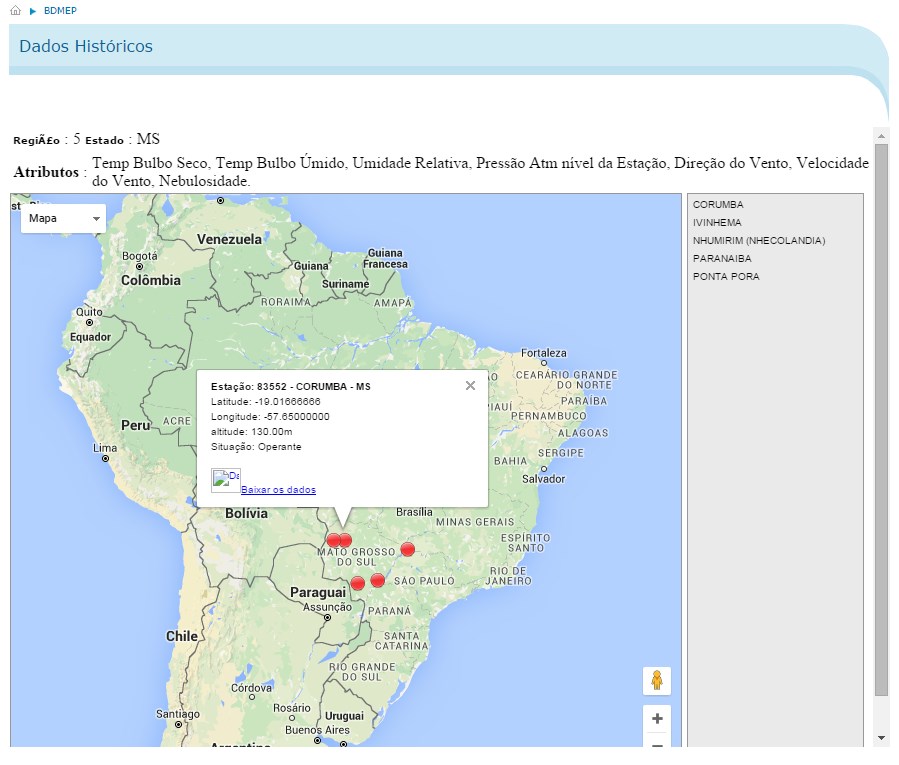
Ao clicar em alguma das opções apresentadas é apresentado um formulário, aonde podemos escolher em qual período de tempo será realizada a pesquisa. Além do período de tempo, também podem ser informadas quais variáveis do tempo serão retornadas.

Outro filtro interessante fornecido é possibilidade de escolher a região ou o estado em que serão buscados os dados.

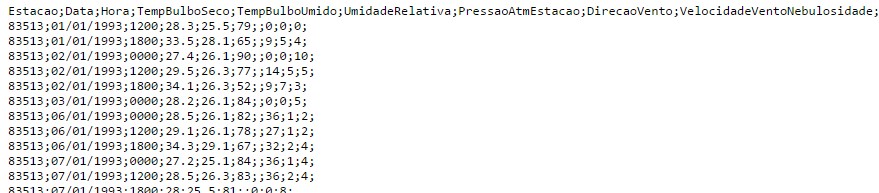


No protótipo desenvolvido nessa monografia serão abordados somente dados do estado do Mato Grosso do Sul, como especificado na imagem acima.

Ao clicar em pesquisa é apresentada uma página onde existe um mapa informando a localização das plataformas meteorológicas. Clicando em cima de um ponto vermelho são apresentadas informações de latitude, longitude e altimetria. Logo abaixo desses dados existe um link aonde pode-se baixar os resultados da pesquisa para aquela plataforma.



Clicando no link “Baixar os dados” será disponibilizado um arquivo txt onde a primeira linha representa os nomes das variáveis do tempo e as posteriores os dados em questão.

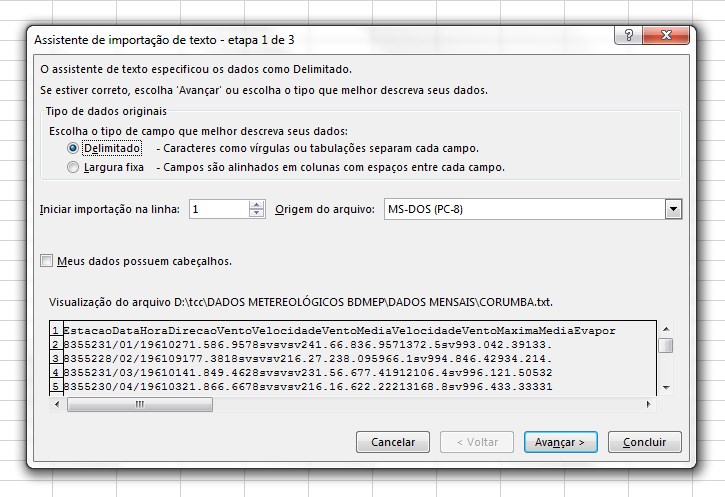


A próxima etapa será realizar a importação desses dados em uma planilha Excel para que possa ser manipulada mais facilmente pelo software que realizará o carregamento desses dados no MongoDB.

**Transformando os arquivos TXT obtidos do BDMEP para planilhas Excel**

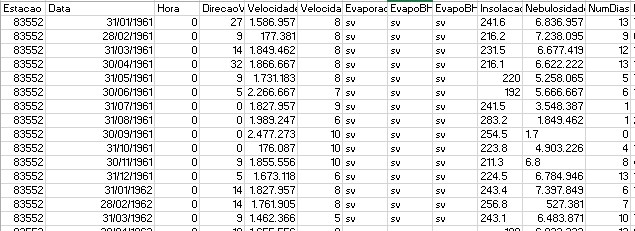
Como pode ser visto no tópico anterior, os dados são fornecidos em uma estrutura aonde as colunas são separadas pelo caractere “;”. Com essa estrutura os dados podem ser importados facilmente e transformados em uma planilha Excel.

O próprio Excel fornece uma ferramenta para realizar essa importação, para executar a ferramenta basta abrir o arquivo TXT.



Ao avançar para a próxima aba é possível escolher como delimitador o caractere “;” e abaixo já é possível ver como ficara a planilha após o processamento.

Em algumas linhas é possível notar que existem campos vazios, para esse projeto foi definido que todos os campos aonde não há valor serão substituídos pela palavra “sv” que significa “Sem Valor”.



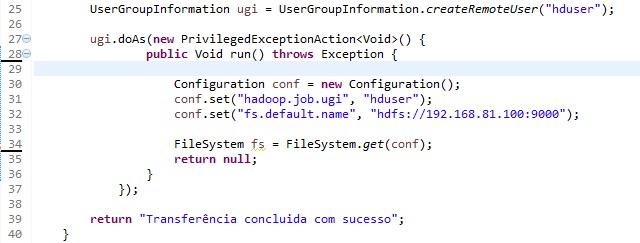
Acima podemos visualizar como fica a planilha ao terminar o processamento do Excel.

**Manipulando Arquivos no HDFS**

O Hadoop Distributed File System é o sistema de arquivos distribuído nativo do Hadoop, que é capaz de armazenar e transmitir grandes massas de dados em maquinas de baixo custo.

Para manipular os arquivos existentes no HDFS é usada a linha de comando no servidor onde o Hadoop está rodando. Para facilitar testes implementados nesse projeto foi desenvolvido um software utilizando a linguagem Java para executar ações no HDFS.

Para que a aplicação tenha acesso ao HDFS que está rodando em uma máquina virtual externa é necessário fazer algumas configurações. O Hadoop disponibiliza uma biblioteca chamada Hadoop Core. Veja um exemplo na imagem a baixo:



Com a biblioteca do hadoop no projeto é possível criar uma instancia do FileSystem que é usada para fazer qualquer operação