

Relatório do projeto de plot de gráficos do TIMED Doppler Interferometer para São João do Cariri, PB

O projeto criado se encontra disponível no link: <https://github.com/KateMorf/skiymet-tidi-plot>

Objetivo:

O objetivo do projeto é plotar os ventos médios meridionais e zonais para o instrumento TIDI (**TIMED Doppler Interferometer**) para a cidade de São João do Cariri, PB.

Período de análise:

O período de análise é janeiro de 2005 a dezembro de 2006.

Obtenção dos dados:

Os dados foram obtidos a partir da base disponibilizada pela Universidade de Michigan que é responsável por manter o projeto. A base foi baixada via protocolo FTP pelo link: <ftp://tidi.engin.umich.edu/tidi/vector/>

O projeto skiyet-tidi-plot se divide em duas partes, uma responsável por ler os arquivos .VEC e convertê-los em um .csv pré-processado. Essa parte pode ser observada no jupyter notebook `preparing_data_to_plot.ipynb`. A segunda parte diz respeito a leitura dos arquivos pré-processados e execução e plot dos gráficos dos ventos médios meridionais e zonais. Essa parte é evidenciada pelo jupyter notebook `tidi_plot.ipynb`.

Primeira parte – preparando os dados para o plot

Inicialmente obtemos os dados das coordenadas de São João Cariri, PB e filtramos apenas os dados que estão na latitude e longitude da região.

Foram considerados para a análise somente os dados da versão D010_R01, que é uma das versões do algoritmo de ajuste espectral que possui mais dias mapeados **[Obs.: perguntar para a orientadora se isso é correto]**

A seguir são apresentados os exemplos marcados são de arquivos com o sufixo D010_R01 que foram considerados para a realização do trabalho:

```

TIDI_PB_2005001_P0100_S0200_D007_R01.VEC
TIDI_PB_2005002_P0100_S0200_D007_R01.VEC
TIDI_PB_2005003_P0100_S0302_D010_R01.VEC
TIDI_PB_2005004_P0100_S0450_D011_R01.VEC
TIDI_PB_2005006_P0100_S0200_D007_R01.VEC
TIDI_PB_2005007_P0100_S0302_D010_R01.VEC
TIDI_PB_2005008_P0100_S0450_D011_R01.VEC
TIDI_PB_2005010_P0100_S0200_D007_R01.VEC
TIDI_PB_2005011_P0100_S0302_D010_R01.VEC
TIDI_PB_2005012_P0100_S0450_D011_R01.VEC
TIDI_PB_2005014_P0100_S0200_D007_R01.VEC

```

Os campos considerados para a análise foram:

- **'ascending': [Obs.: não sei o que significa, perguntar para a orientadora, mas a informação não foi utilizada para o cálculo]**
- **'alt_retrieved':** altitude registrada
- **'ut_date':** marcação do ano e dia (formato: yyyydd, ex: 2005001 (primeiro dia de 2005, não há separação do mês, ele conta de 1 a 365 dias do ano))
- **'UT_time':** horário do registro convertido em horas
- **'lat':** latitude
- **'lon':** longitude
- **'data_ok': [não sei o que significa]**
- **'u':** velocidade zonal [seria bom perguntar a orientadora se isso é correto]
- **'var_u':** variância da velocidade zonal [seria bom perguntar a orientadora se isso é correto]
- **'v':** velocidade meridional [seria bom perguntar a orientadora se isso é correto]
- **'var_v':** variância da velocidade meridional [seria bom perguntar a orientadora se isso é correto]

Depois disso salvamos os dados das medições diárias dos anos de 2005 a 2006 para csv. A seguir um exemplo de arquivo csv:

na1ts	nvce	ascending	alt_retrieved	ut_date	UT_time	lat	lon	data_ok	u	var_u	v	var_v
4	448	b'F'	80.0	b'2005001'	7.584722222222222	-13.928479	315.54437	b'T'				
4	964	b'T'	80.0	b'2005001'	16.3225	-10.331198	322.62607	b'T'	-12.295723	559.0333	-78.536606	459.26843
4	965	b'T'	80.0	b'2005001'	16.358611111111111	-3.0482023	324.57245	b'T'	-27.740772	874.9862	-90.49597	538.84644
4	1044	b'T'	80.0	b'2005001'	17.977222222222224	-11.78036	331.24957	b'T'				
4	1045	b'T'	80.0	b'2005001'	18.013055555555557	-4.329226	332.63715	b'T'				
5	244	b'F'	82.5	b'2005001'	4.2375	-0.630256	328.00845	b'T'				
5	245	b'F'	82.5	b'2005001'	4.273055555555556	-8.046699	329.51636	b'T'				
5	246	b'F'	82.5	b'2005001'	4.308888888888889	-15.527763	330.80066	b'T'				
5	448	b'F'	82.5	b'2005001'	7.584722222222222	-13.928479	315.54437	b'T'				
5	964	b'T'	82.5	b'2005001'	16.3225	-10.331198	322.62607	b'T'	-18.252918	506.04907	-51.82612	447.36417
5	965	b'T'	82.5	b'2005001'	16.358611111111111	-3.0482023	324.57245	b'T'	-37.583786	767.43604	-80.42148	475.36847
5	1044	b'T'	82.5	b'2005001'	17.977222222222224	-11.78036	331.24957	b'T'	-13.946257	353.0026	-8.672956	458.59717
5	1045	b'T'	82.5	b'2005001'	18.013055555555557	-4.329226	332.63715	b'T'	-12.211397	593.9559	-7.388207	666.62286
6	244	b'F'	85.0	b'2005001'	4.2375	-0.630256	328.00845	b'T'				
6	245	b'F'	85.0	b'2005001'	4.273055555555556	-8.046699	329.51636	b'T'				
6	246	b'F'	85.0	b'2005001'	4.308888888888889	-15.527763	330.80066	b'T'				
6	448	b'F'	85.0	b'2005001'	7.584722222222222	-13.928479	315.54437	b'T'				

Apesar do arquivo já estar um formato de apresentação compreensível, é preciso converter esses dados para um formato mais acessível para realizar o plot dos gráficos. O formato utilizado possui os seguintes atributos:

- **tempo decorrido:** marcação da hora da marcação
- **marcação da velocidade(u/v) para cada altitude de 70km a 100km para o tempo decorrido marcado**

- **marcação do desvio padrão (raiz do var_u e var_v) para cada altitude de 70km a 100km para o tempo decorrido marcado.**

Em outras palavras, dividimos o arquivo .csv em dois arquivos .txt com os dados das altitudes de 70km a 100 km, um com as velocidades meridionais(v) e o outro arquivo com as velocidades zonais(u). Exemplos dos dois arquivos são apresentados abaixo:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
tempo_decorrido	70.0km	72.5km	75.0km	77.5km	80.0km	82.5km	85.0km	87.5km	90.0km	92.5km	95.0km	97.5km	100.0km	std_70.0km	std_72.5km	std_75.0km	std_77.5km	std_80.0km
4.2375								-29.443619	-91.668465	-58.01477	-31.099606	-6.627021	16.407536					
4.273055555555556																		
4.308988888888889																		
7.584722222222222																		
16.3225					-12.295723	-18.252918	-12.7073145	-15.544464	-21.688604	-13.053576	30.25003	55.833044	28.889711					23.643885044
16.358611111111111					-27.740772	-37.583786	-34.80297	-54.69547	-36.029278	-18.078068	-16.80974	-23.97808	-67.941286					29.580165652
17.977222222222224						-13.946257	10.700877	11.786410	18.649472	6.321148	-11.248838	-7.031294	0.13635635					

As imagens abaixo representam o resultado da separação dos dados zonais e meridionais do arquivo csv. Está marcado para cada arquivo o ano, mês e dia do ano. Ex.: 200501_001 (2005 é o ano, 01 é o mês de janeiro e _001 é o primeiro dia do ano, podendo variar de 1 a 365).

```

▼ dados_zonal
  ~.lock.dados_zonal_200501_001.txt# U
  dados_zonal_200501_001.txt M
  dados_zonal_200501_002.txt M
  dados_zonal_200501_003.txt M
  dados_zonal_200501_004.txt M
  dados_zonal_200501_005.txt M
  dados_zonal_200501_006.txt M
  dados_zonal_200501_007.txt M
  dados_zonal_200501_008.txt M
  dados_zonal_200501_009.txt M
  dados_zonal_200501_010.txt M
  dados_zonal_200501_011.txt M
  dados_zonal_200501_012.txt M

```

```

▼ dados_meridional
  dados_meridional_200501_001.txt
  dados_meridional_200501_002.txt
  dados_meridional_200501_003.txt
  dados_meridional_200501_004.txt
  dados_meridional_200501_005.txt
  dados_meridional_200501_006.txt
  dados_meridional_200501_007.txt
  dados_meridional_200501_008.txt
  dados_meridional_200501_009.txt
  dados_meridional_200501_010.txt
  dados_meridional_200501_011.txt
  dados_meridional_200501_012.txt
  dados_meridional_200501_013.txt
  dados_meridional_200501_014.txt
  dados_meridional_200501_015.txt

```

Segunda parte – plotando os dados

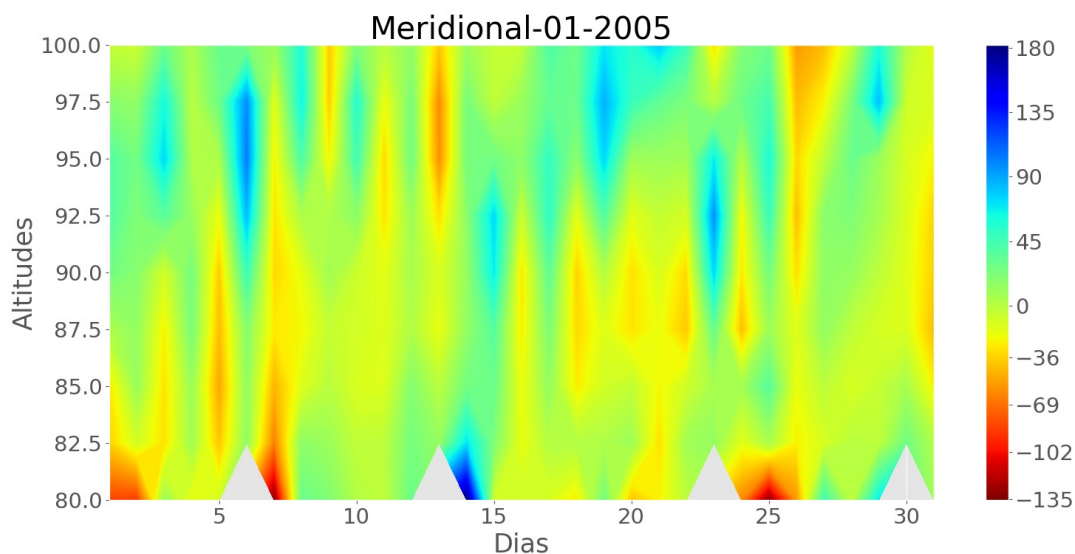
Antes de começarmos a plotagem dos gráficos fazemos um último pré-processamento nos dados. Como pôde ser visto nos dados anteriores, para um mesmo dia, temos várias medições. O nosso objetivo é plotar um gráfico de contorno representando o comportamento médio dos ventos meridionais e zonais. Para isso convertemos os arquivos dos dias que representam um mês em um único arquivo que representa todo o mês. Ex.: para o mês de janeiro temos 31 arquivos que representam cada um dia do mês de janeiro. O que fazemos aqui é calcular a média das velocidades de cada arquivo e salvamos em um único arquivo as médias das medições de cada dia do mês de janeiro. Abaixo temos um exemplo de arquivo de média:

You, há 3 semanas 1 author (You)										
1	tempo	80	82.5	85	87.5	90	92.5	95	97.5	100 dia
2	11.134444	-84.516288	-37.077191	-29.181674	3.617037	27.323490	38.639939	38.184168	17.465549	-5.056134 001
3	10.204833	-88.016905	-10.356659	11.352427	13.532518	17.962179	21.346305	25.940079	14.404904	-7.691143 002
4	10.186806	10.661430	-31.390949	-31.417763	-25.918412	-11.070759	35.731642	74.042580	61.411755	28.630177 003
5	7.730243	-20.461229	8.525717	10.817472	23.654989	25.056604	11.831165	3.119202	-1.331916	3.729986 004
6	9.857623	-10.134791	-40.760205	-51.577233	-44.353715	-39.192305	-21.282287	3.305036	27.452075	30.195835 005
7	8.090486	NaN	14.268007	14.411061	18.768271	47.118648	82.853950	102.547112	98.082140	22.268653 006
8	9.738413	-135.540117	-65.636459	-46.625375	-28.703849	-34.152752	-30.427284	-26.511411	-18.116641	9.467645 007
9	9.835938	29.719336	16.418852	-14.747579	-23.244248	-18.869264	3.507708	40.698946	64.720828	58.571542 008
10	8.906000	20.666161	10.303945	-0.346234	-5.410332	8.333408	2.729219	-25.433571	-39.967194	-39.806151 009
11	9.455432	1.491302	-3.427210	-12.930375	-9.600188	-7.721404	16.315602	45.348337	56.331283	21.306685 010
12	7.739653	-2.794852	-1.691392	-15.144296	-14.969339	-17.615390	-30.811165	-34.773666	-21.487849	-1.788876 011
13	8.558426	29.356441	24.796352	19.012756	2.758629	5.666323	9.877600	27.069588	23.808958	14.798421 012
14	6.791771	NaN	12.005069	-1.458478	-19.642186	-13.782025	-32.026837	-60.172556	-61.496475	-40.750456 013
15	7.051343	181.969400	70.741948	25.107543	13.477065	12.056382	15.746405	24.240451	24.013969	10.334883 014
16	8.324931	-10.539773	20.901491	29.416004	37.132590	67.043553	73.992495	27.194958	-3.534441	-4.592565 015
17	6.895895	-8.974996	-19.465507	-21.167451	-27.830577	-30.056682	-9.134929	15.079756	10.940274	-4.865149 016
18	8.897569	-12.978514	1.524888	13.523699	21.985692	33.730605	52.749538	50.621386	32.165432	40.288179 017
19	7.349877	-27.601914	2.192948	-27.859140	-34.152023	-36.469003	-18.562735	15.546868	29.225492	27.582961 018
20	6.818306	25.349837	3.358541	-9.729787	-15.306275	1.054611	38.691619	74.890265	86.506081	70.108917 019
21	5.587083	-41.701940	9.880286	-3.893396	-30.064957	-31.143019	-16.853374	8.774770	48.082479	55.448050 020
22	4.943889	-25.749078	-31.719065	-23.738989	-19.792372	-13.231890	-4.468979	10.299000	32.465295	77.884343 021
23	7.025035	-1.568279	8.491584	-8.779289	-39.305421	-33.471553	-11.933694	7.480755	22.229586	47.487512 022
24	4.704167	NaN	12.873030	5.722042	30.607006	74.501959	95.120025	67.587730	-0.474170	-31.424670 023
25	12.880741	-61.740891	-17.275709	6.780081	-43.021825	-29.774317	-23.396458	-0.328429	28.369523	19.796520 024
26	15.551111	-133.601659	4.755945	36.169026	10.433566	25.875724	48.492528	59.802214	44.106826	28.264793 025
27	7.941230	-57.241005	-26.589686	-19.871556	-16.197882	-32.553435	-43.812410	-35.166379	-44.242844	-54.934804 026
28	8.954472	45.954470	-11.998618	-1.769807	13.564885	15.124328	17.172918	-3.527917	-25.094480	-43.513259 027
29	12.184846	-5.421025	0.269290	-12.797764	-2.336961	12.375108	21.681465	29.758209	21.172956	-6.068835 028
30	13.261080	68.987663	-2.692327	-5.354926	-11.829416	-2.191387	6.260382	10.255178	81.910624	59.415169 029
31	5.831167	NaN	26.593859	6.006748	-13.974649	-15.356619	-13.345063	-12.483039	-9.166069	1.071901 030

Essa imagem é um exemplo do arquivo das médias do mês de janeiro para os dados meridionais, existe um equivalente do mês de janeiro para os dados zonais. Fizemos também mais duas alterações, removemos os dados de desvio padrão e consideramos apenas as altitudes a partir de 80 km.

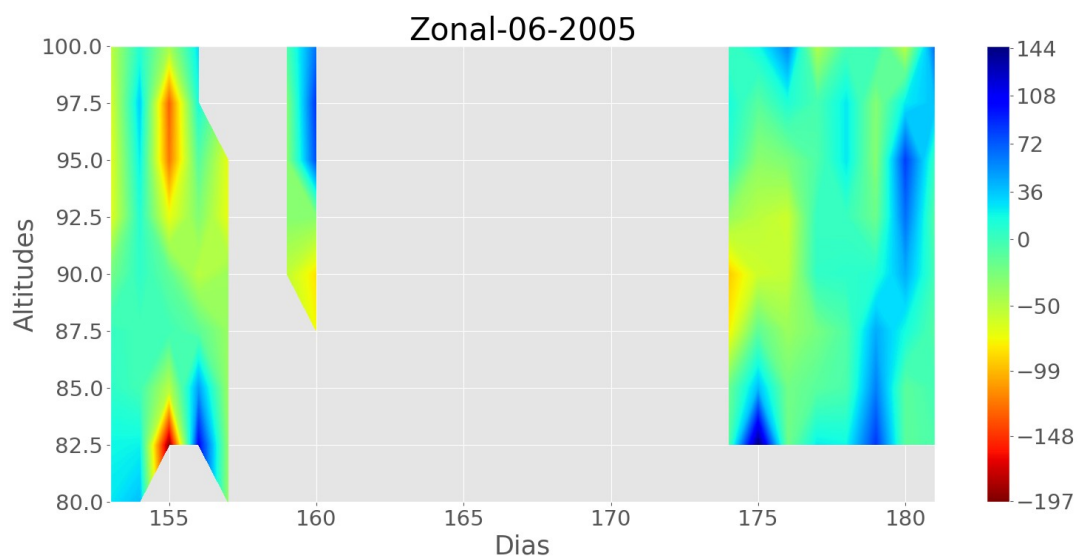
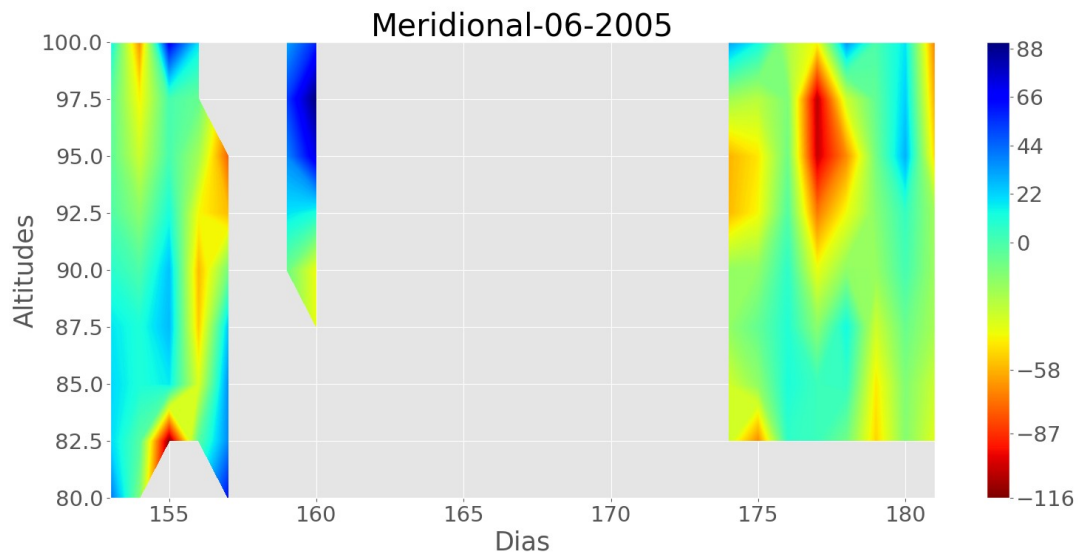
Após isso, fazemos a plotagem dos gráficos de contornos meridionais e zonais para cada mês dos anos de 2005 e 2006. Utilizamos como paleta de cores o jet_r.

A seguir temos um exemplo de plot:



Observações: Existem alguns plots um pouco estranhos, seria interessante perguntar a orientadora se esses plots estão corretos e como analisá-los. Vou apresentar alguns exemplos abaixo:

Tanto para os dados Meridionais e Zonais o mês de junho de 2005 tem vários dias sem registros.



Para o mês de janeiro de 2006, tanto para os dados Zonais e Meridionais temos alguns valores de velocidade muito distoantes (possíveis outliers) que podem estar causando esse problema.

