



UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE ENGENHARIA GEOGRÁFICA, GEOFÍSICA E ENERGIA

MOS – *Model Output Statistics* – Aplicação a  
Previsões MM5 de curto prazo em Portugal Continental

Raquel Alexandra Pratas Marujo

Mestrado em Ciências Geofísicas

Especialização em Meteorologia

2009



UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE ENGENHARIA GEOGRÁFICA, GEOFÍSICA E ENERGIA

MOS – *Model Output Statistics* – Aplicação a  
Previsões MM5 de curto prazo em Portugal Continental

Raquel Alexandra Pratas Marujo

Mestrado em Ciências Geofísicas

Especialização em Meteorologia

Trabalho de Projecto orientado pelo Prof. Doutor Carlos Pires (FCUL)

2009

## Abstract

The need to predict the weather accurately is one of the oldest dreams of mankind. In recent decades there has been significant progress in weather forecasting models. However, the models have systematic errors that may be corrected with the application of linear statistical models, from the accumulated statistics of a particular model prediction. These corrective models have the generic name of MOS (Model Output Statistics). The linear version of the MOS is based on the use of multilinear regressions. In this study, there were made predictions MOS within 24 hours, for the variables of temperature and horizontal components of the wind, using the 61 stations of the network of Portuguese Institute of Meteorology. The model equations were obtained with a stepwise regression or step-by-step, up to five predictors and assessed by cross-validation. The pre-selected predictors were forecasts of MM5 model (The Fifth-Generation NCAR / Penn State Mesoscale Model (MM5)) and observations available at the release of the forecast. The results show an improvement in the mean square error of 60% in the temperature forecasting and 80% in the horizontal components of the wind forecasting, related to the results of the forecast, in 24 hours, by the numerical model MM5. In terms of forecasts of 48 and 72 hours, the application of the model improves MOS of 70% of the MM5 forecasts for the three variables. These results motivate the operational use of this method and its development for other predictive variables.

**Key-words:** statistical forecasting models, numerical models, forecasts, temperature, horizontal components of wind.

## Resumo

A necessidade de prever o tempo com exactidão é um dos sonhos mais antigos da humanidade. Nas últimas décadas houve um progresso significativo nos modelos de previsão meteorológica. Contudo, os modelos apresentam erros sistemáticos que podem ser corrigidos com a aplicação de modelos estatísticos lineares ou não, a partir das estatísticas acumuladas de um certo modelo de previsão. Esses modelos correctivos a posteriori têm o nome genérico de MOS (Model Output Statistics). A versão linear do MOS tem como base o uso de regressões multilíneas. Neste trabalho, realizaram-se previsões MOS com prazo de 24 horas, para as variáveis da temperatura e componentes horizontais do vento, nas 61 estações da rede do Instituto de Meteorologia Português. As equações do modelo foram obtidas com uma regressão *stepwise* ou passo-a-passo, até cinco predictores e aferidas em validação cruzada. Os predictores pré-seleccionados foram previsões do modelo MM5 (The Fifth-Generation NCAR / Penn State Mesoscale Model (MM5)) e observações, disponíveis no lançamento da previsão. Os resultados mostram uma melhoria do erro médio quadrático de 60% na previsão da temperatura e 80% na previsão das componentes horizontais do vento, relativamente aos resultados da previsão, a 24 horas, do modelo numérico MM5. Para prazos de previsões de 48 e 72 horas, a aplicação do modelo MOS melhora na ordem dos 70% as previsões do MM5, para as três variáveis. Estes resultados motivam a utilização operacional deste método e o seu desenvolvimento para outras variáveis de previsão.

**Palavras-chave:** modelos de previsão estatísticos, modelos numéricos, previsões, temperatura, componentes horizontais do vento

# Índice

Abstract .....	iii
Resumo.....	iii
Agradecimentos.....	vi
1. Introdução.....	1
2. Modelo de Mesoescala – MM5.....	3
2.1 Breve descrição do modelo numérico .....	3
2.2 Dados de previsão MM5 .....	3
2.3 Dados de observação.....	4
2.4 Análise do output .....	7
2.4.1 Medidas da análise da qualidade do modelo MM5.....	7
2.4.2 Avaliação da qualidade das simulações do modelo MM5 .....	8
3. Modelo Estatístico -MOS.....	11
3.1 Breve descrição do modelo .....	11
3.1.1 Regressão linear simples .....	11
3.1.2 Regressão linear múltipla.....	12
3.2 Metodologia da construção do modelo estatístico .....	13
3.2.1 Selecção de predictores .....	13
3.2.2 Validação cruzada .....	15
3.3 Análise do output .....	16
3.3.1 Medidas da qualidade do modelo estatístico.....	16
4. Apresentação dos resultados do modelo estatístico .....	17
4.1 Resultados dos MOS .....	19
4.1.1 Caso A .....	19
4.1.2 Caso B .....	20
4.1.3 Caso C .....	23
4.1.4 Caso D .....	25
4.1.5 Caso E .....	27

4.2	Discussão dos resultados do modelo.....	29
4.2.1	Comparação dos resultados MOS .....	31
4.2.2	Coeficientes da equação de regressão .....	32
4.2.3	Comparação entre o MM5 e o MOS .....	37
5.	Breve apresentação e discussão dos resultados do MOS para diferentes prazos de previsão .....	40
6.	Conclusões .....	44
7.	Referências.....	45

## **Agradecimentos**

Venho por este meio expressar os meus mais sinceros agradecimentos:

*Ao meu orientador Professor Doutor Carlos Pires pelo apoio, disponibilidade, interesse e empenho que demonstrou na realização deste trabalho, bem como pelas sugestões e críticas que contribuíram para a valorização deste trabalho.*

*Aos meus colegas da FCUL e do LNEG pela amizade, interesse e ajuda demonstrado, em especial a Paulo Costa, Emanuel Dutra, Ricardo Tomé, Sónia Domingos, Liliana Madeira, Nuno Silva, doutora Ana Estanqueiro e Teresa Simões pela disponibilidade, paciência e apoio ao longo da elaboração da tese.*

*À minha família, em especial aos meus pais e irmã, pela oportunidade de realizar mais um sonho.*

*A todos a minha gratidão pelo apoio demonstrado.*

## 1. Introdução

A previsão do tempo com precisão é um dos sonhos mais antigos da humanidade. Durante séculos a previsão do tempo estava ligada à mitologia, superstição e folclore, e era realizada por altos padres, médicos feiticeiros ou bruxos. Foi só nos últimos cem anos que as melhorias na meteorologia operacional têm sido ligadas aos avanços científicos.

A base da matemática meteorológica ou previsão numérica do tempo foi formalizada por Vilhelm Bjerkness em 1919. Ele sugeriu que a previsão do tempo fosse vista como um problema de condições iniciais. A primeira pessoa a tentar a simulação numérica do movimento da atmosfera foi Lewis Fry Richardson em 1922 no seu trabalho: '*Weather Prediction by Numerical Process*'. Este trabalho foi recebido calorosamente nos círculos académicos, mas não foi considerado na altura como sendo um possível método de previsão por requer um número elevado de condições iniciais, exigindo uma enorme carga de trabalho computacional para realizar uma previsão de 24 horas (Tomé, 2004).

Nas últimas décadas houve um progresso significativo no campo da previsão do tempo, graças em grande parte às facilidades proporcionadas pela tecnologia moderna, especialmente na área da electrónica e de sistemas de comunicação de dados (Cressman, 1981). A evolução dos computadores permitiu o aperfeiçoamento dos modelos de previsão numérica da atmosfera, culminando durante a década de 60 do século XX com os modelos de equações primitivas. A simulação numérica do movimento atmosférico foi, então, recebida como sendo uma valiosa aproximação para a previsão do tempo. Hoje em dia os modelos numéricos são usados como a principal ferramenta para a previsão do tempo a curto prazo.

O *output* das simulações numéricas são previsões dos campos das variáveis meteorológicas em diferentes níveis numa grelha específica. Nestas previsões existem erros sistemáticos que não dependem somente das falhas de parametrização física dos modelos dinâmicos, mas também da topografia local, da resolução do modelo e dos erros de interpolação das estações para os pontos de grelha do modelo entre outras causas. Para interpretar estes erros e caracterizá-los existe a possibilidade de implementar um modelo estatístico que use como variáveis de entrada as simulações.

Os modelos estatísticos lineares de correcção da previsão de tempo são, em geral, equações de regressão linear múltipla. Segundo Carter et al. (1989), no final da década de 80 do séc. XX, mais de 110.000 dessas equações estavam a ser processadas, nos EUA, para previsões de elementos meteorológicos.

O MOS (*Model output statistics*) é um modelo estatístico, considerado como um suporte principal para a previsão do tempo nos dias de hoje. A origem desta técnica estatística foi introduzida, nos anos 60 e 70 do século XX por Glahn, Lowry e Klein, para pós-processamento dos dados de saída das simulações numéricas. De um modo geral, os resultados dos modelos numéricos representam melhor os padrões da atmosfera em altitude sendo mais débeis quando tentam prever as variações locais à superfície. Por outro lado, os modelos estatísticos puros conseguem representar a previsão do tempo local, mas perdem qualidade para além de seis horas de previsão. O MOS é uma técnica que combina estes dois métodos, usando as previsões numéricas com base na física da atmosfera para prever os padrões climáticos de grande escala e, as equações de regressão estatística resultantes dos pós-processamento para esclarecer detalhes do tempo à superfície. A precisão dos resultados é geralmente muito melhor do que utilizar apenas um modelo estatístico ou numérico puro.

Este método, MOS, foi provado como sendo bem sucedido para modelos de previsão global (Glanh e Lowry 1972; Klein e Glahn 1974; Lemcke e Krusinga 1988; Dallavalle et al. 2004; Taylor e Leslie 2005), assim como para modelos de previsão regionais (Sokol 2003; Termonia e Deckmyn 2007; Cheng e Steenburgh 2007).

A deficiência principal deste método estatístico é necessitar de um elevado número de dados disponíveis das saídas de uma mesma versão de modelo numérico para permitir uma melhor *performance* das equações de previsão. Os modelos numéricos são continuamente modificados

tornando as relações estatísticas obtidas, com as saídas do modelo anterior, não representativas, ou seja em geral os modelos permanecem em geral menos tempo que o necessário para obter estatísticas fiáveis para construir um modelo MOS. No entanto, alguns estudos mostram que pequenas alterações nos modelos numéricos não degradam substancialmente as equações do MOS [Erickson et al. (1991)].

Há no entanto uma versão que permite a adaptação do MOS a alterações das versões do modelo dinâmico (UMOS – Updateable MOS) (Wilson e Valée, 2003).

Dado o sucesso alcançado, o MOS tem sido aplicado a diversas variáveis, tais como: temperatura mínima e máxima, probabilidade de ocorrência de precipitação, temperatura do ponto de orvalho, direcção e velocidade do vento, ocorrência de tempestades (severas), altura da base

das nuvens, visibilidade, temperatura horária, nevoeiro, entre outras [Jacks et al. (1989); Wilks (1990); Esterle (1992); Vislocky e Fritsch (1995); Vislocky e Fritsch (1997)]; e está sendo continuamente implementado operacionalmente em vários centros de meteorologia (Esterle (1992); Vasiliev (1991); Francis et al. (1982); Lemcke e Kruizinga (1988); Conte et al. (1980); Tapp et al. (1986); Lu (1991); Azcarraga e Ballester (1991); Brunet et al. (1988); Carter et al. (1989)).

Existem várias técnicas de MOS para além das técnicas de regressão multilinear. Assim existem técnicas baseadas em redes neuronais (neural networks NN) (Yuaval e Hsieh, 2003).

#### Objectivos e organização da dissertação

Neste trabalho pretende-se aplicar o MOS para caracterizar os erros espacio-temporais das previsões de modelos atmosféricos de área limitada. O MOS vai ser aplicado a um ano de previsões MM5, disponíveis em períodos de 30 em 30 min até às 72 horas de previsão, em confronto com as observações da rede do Instituto de Meteorologia (IM). As variáveis que vão ser estudadas vão ser apenas um subconjunto das variáveis disponibilizadas pelo modelo: componentes do vento (u,v) e temperatura T à superfície. O modelo estatístico tem como finalidade obter previsões estatísticas para as 24h, utilizando como predictores observações e previsões do modelo numérico com diferentes tempos de previsão (24h, 48h e 72h) e aplicado a diferentes casos de estudo (variação do número e tipo de predictores do MOS).

O trabalho está organizado em seis capítulos. No capítulo 1 é feita uma breve introdução, bem como a apresentação dos objectivos e organização do trabalho. No capítulo 2 é apresentada uma breve descrição do modelo numérico MM5, apresentação da base de dados e análise da qualidade do output do modelo através da metodologia apresentada. No capítulo 3 é descrito o modelo estatístico - MOS, assim como, a metodologia associada e a aplicar aos vários casos de estudo. O capítulo 4, expõe os resultados obtidos da previsão de 24 horas, bem como a comparação dos resultados entre os modelos MOS e MM5. O capítulo 5 apresenta uma breve exposição dos resultados obtidos para os prazos de previsão de 48 e 72 horas. Finalmente, no capítulo 6, apresentam-se as conclusões obtidas.

## 2. Modelo de Mesoescala – MM5

Um modelo de previsão numérica do tempo é um modelo integrado numericamente desenvolvido para simular e prever o comportamento da atmosfera. Este modelo faz em geral recurso a meios informáticos poderosos em termos de memória e velocidade de processador. Este modelo resolve um sistema complexo de equações matemáticas baseadas em leis físicas de modo a prever o estado futuro da atmosfera partindo de condições iniciais específicas. As equações matemáticas, são as formas mais completas e complexas das equações dinâmicas e termodinâmicas. Em adição, processos físicos como a radiação, nuvens, precipitação e trocas à superfície são representados no modelo usando esquemas de parameterizações.

Os modelos de previsão numérica do tempo podem ser classificados segundo a sua escala espacial dentro da qual a previsão é feita:

*Global*, com resolução da ordem de 200km, visa identificar o comportamento geral da atmosfera sobre uma área extensa, estes modelos são apenas capazes de identificar fenómenos meteorológicos de larga escala (ditos sinópticos);

*Regional ou local*, com resolução espacial de alguns metros a cerca de 50km, visa identificar com maior detalhe o comportamento da atmosfera sobre uma região específica, sendo capazes de identificar fenómenos meteorológicos de pequena escala (ditos de mesoscala).

### 2.1 Breve descrição do modelo numérico

O modelo MM5 (The Fifth-Generation NCAR / Penn State Mesoscale Model (MM5)) é um modelo numérico de mesoscala utilizado para simular padrões atmosféricos à escala regional. O código do modelo é de acesso livre e continua sendo aperfeiçoado por diversos utilizadores em universidades e institutos de investigação. O modelo utiliza coordenadas sigma, que acompanham o terreno, o que o torna capaz de ser utilizado operacionalmente para fins de previsão atmosférica (Costa, 2004).

Um tutorial do modelo pode encontrar-se em:

<http://www.mmm.ucar.edu/mm5/>

### 2.2 Dados de previsão MM5

Para este trabalho foram disponibilizadas para o ano completo de 2005, previsões diárias das componentes zonal (u) e meridional (v) do vento e da temperatura T com um horizonte temporal (prazo de previsão) até 72 horas, oriundas do modelo MM5, instalado na Faculdade de Ciências da Universidade de Lisboa a correr operacionalmente com uma malha de resolução espacial de 6km x 6km. As previsões são fornecidas regularmente no página da internet (<http://www.weather.ul.pt/>). Foram processadas, com intervalos de tempo de 30 em 30 minutos, séries anuais de vento e temperatura interpoladas bilinearmente para os pontos das estações da rede do Instituto de Meteorologia. O output do valor médio de cada variável seleccionada, do modelo MM5 para 24 horas de previsão, está representado nos mapas da figura 2.1.

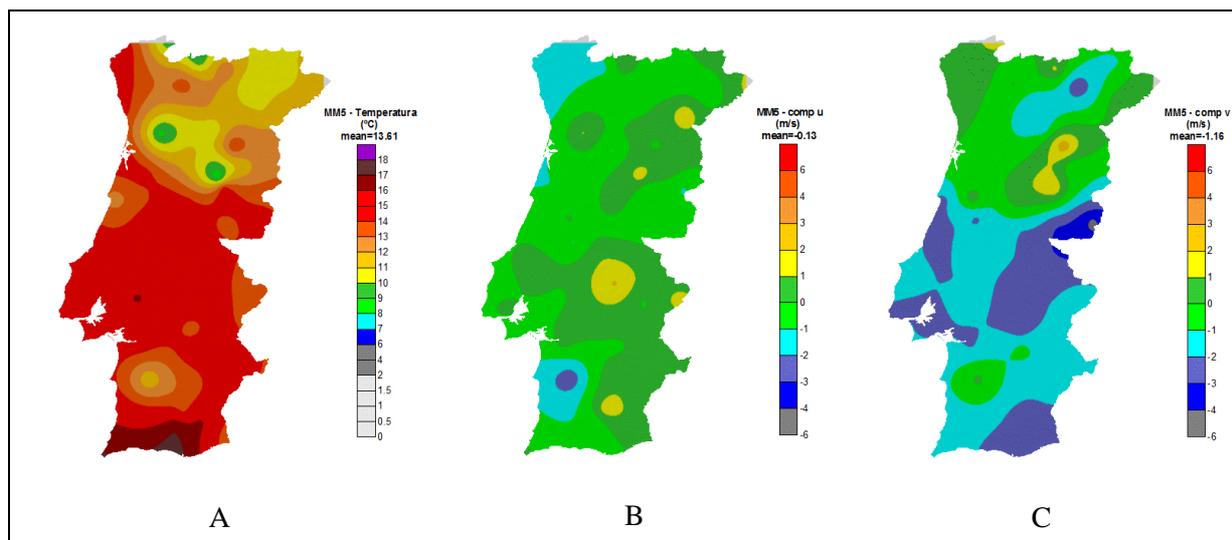


Figura 2.1 – Média das previsões numéricas MM5 para as 0h com prazo de 24 horas, referente ao ano de 2005. A) Temperatura; B) componente u do vento; C) componente v do vento.

### 2.3 Dados de observação

Neste estudo usaram-se as observações de 61 estações meteorológicas da rede do Instituto de Meteorologia (IM), cuja localização pode ser observada na figura 2.2 e tabela 2.1.

Tabela 2.1- Características das estações meteorológicas da rede do Instituto de Meteorologia.

Nº	Nome da Estação	Latitude (E)	Longitude (N)	Altitude (m)
531	C. CARVOEIRO / FAROL	-9.4073778	39.360772	32
533	SAGRES / QUARTEL DA MARINHA	-8.95	37	25
535	LISBOA/GEOFISICO	-9.149652	38.719093	77
545	PORTO/P.RUBRAS	-8.681367	41.2335	70
548	COIMBRA/CERNACHE AERO.	-8.468541	40.15711	171
554	FARO/AEROPORTO	-7.971964	37.016578	8
558	EVORA/C.C. AERO.	-7.888029	38.536419	245
560	VISEU/C.C. AERO.	-7.895987	40.714943	644
562	BEJA	-7.867582	38.025728	246
567	VILA REAL/C.C.	-7.717101	41.274187	562
568	PENHAS DOURADAS	-7.558658	40.411366	1380
570	CASTELO BRANCO/C.C.	-7.478924	39.839441	386
575	BRAGANCA/C.C.	-6.7428389	41.803889	690
577	ODEMIRA/S. TEOTONIO	-8.7541	37.525	119
579	LISBOA/GAGO COUTINHO	-9.128303	38.766621	104
604	VILA NOVA DE CERVEIRA/AERO.	-8.676167	41.973084	18
606	LAMAS DE MOURO/PORTO RIB.	-8.183333	42.033333	880
611	MONTALEGRE	-7.783333	41.816667	1005
616	CHAVES / AERO.	-7.466667	41.716667	360
619	CABRIL/ S. LOURENCO	-8.016667	41.716667	585
632	MIRANDELA	-7.966667	41.533333	250
633	MACEDO DE CAVALEIROS / IZE.	-7.2	41.516667	702
635	MIRANDA DO DOURO	-6.783333	41.566667	693
637	MOGADOURO	-6.283333	41.516667	644
654	MONCORVO	-6.733333	41.333333	600
663	MOIMENTA DA BEIRA	-7.016667	41.183333	715
666	TRANCOSO/BANDARRA	-7.6	40.983333	850

Tabela 2.1 (cont.) - Características das estações meteorológicas da rede do Instituto de Meteorologia.

Nº	Nome da Estação	Latitude (E)	Longitude (N)	Altitude (m)
668	AROUCA	-7.366667	40.783333	340
683	GUARDA	-8.25	40.933333	1020
686	PAMPILHOSA DA SERRA	-7.266667	40.533333	890
687	COVILHA/AERO.	-7.916667	40.133333	414
697	LOUSA / AERODROMO	-7.4823	40.264291	195
702	AVEIRO/UNIVERSIDADE	-8.233333	40.133333	15
705	ANADIA/EST. VITIVINICOLA	-8.6595833	40.635297	45
713	FIGUEIRA DA FOZ / VILA VER	-8.439816	40.438751	9
716	ANSIAO (DEP. DE AGUA)	-8.80596	40.139769	405
724	TOMAR / VALDONAS	-8.416667	39.9	75
726	ALCOBACA / EST. FRUTIC. VI	-8.366667	39.6	38
734	SANTAREM / FONTE BOA	-8.966667	39.516667	73
739	TORRES VEDRAS / DOIS PORTOS	-8.733333	39.2	110
744	CORUCHE / ESTACAO DE REGADIO	-9.179	39.043894	25
765	CABO RASO / FAROL	-8.533333	38.95	25
766	BARREIRO / LAVRADIO	-9.485503	38.709035	6
767	ALMADA / PRAIA DA RAINHA	-9.047661	38.674487	10
776	ALCACER DO SAL / BARROSINH	-9.212815	38.617118	29
783	ALVALADE / EST. REGADIO	-8.483333	38.366667	61
800	SABUGAL / MARTIM REI	-8.4	37.95	858
803	ZEBREIRA	-7.033333	40.25	375
806	PROENCA-A-NOVA / PISTA MOI	-7.016667	39.866667	379
812	ALVEGA	-7.866667	39.733333	51
824	AVIS / BENAVIDA (ESCOLA AB)	-8.05	39.466667	150
835	ELVAS / EST. MELHORAMENTO	-7.866667	39.1	208
837	ESTREMOZ / TECHOCAS	-7.13976	38.8908	366
848	PORTEL / ORIOLA	-7.516667	38.866667	205
850	AMARELEJA	-7.866667	38.316667	192
863	MERTOLA / VALE FORMOSO	-7.216667	38.216667	190
864	CASTRO VERDE / NEVES CORVO	-7.55	37.75	255
865	ALCOUTIM / MARTINLONGO	-7.9721139	37.576558	290
867	CASTRO MARIM / RES. NAC. D	-7.766667	37.433333	5
868	ALMODOVAR / CERRO NEGRO	-7.45	37.216667	400
878	PORTIMAO / AERODROMO	-8.083333	37.4	14

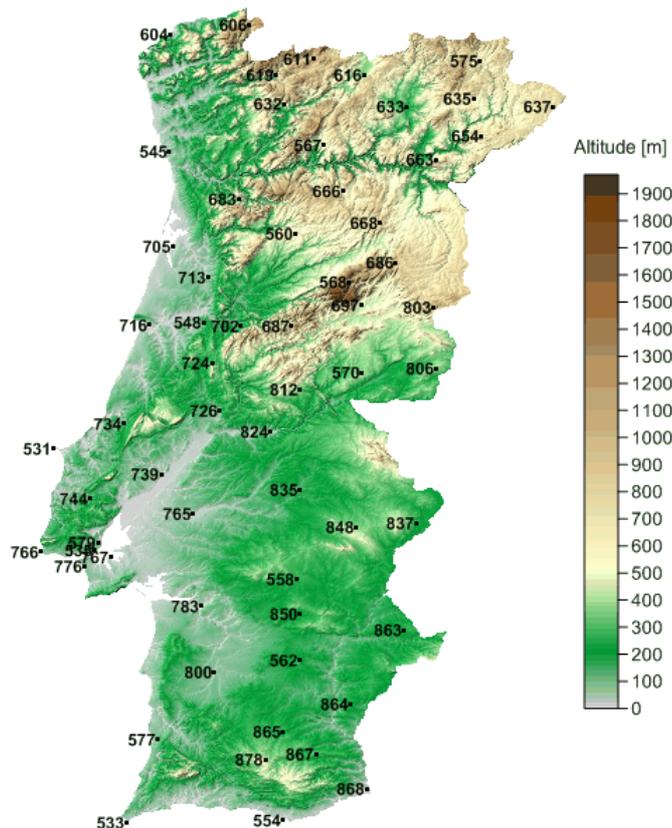


Figura 2.2 – Localização das estações da rede do Instituto de Meteorologia.

A base de dados das observações foi disponibilizada em intervalos de 10 minutos para o ano de 2005. Para o estudo do MOS foram construídas séries, com espaçamento temporal de 30 minutos, das duas componentes horizontais da velocidade do vento e da temperatura. No caso de verificação de falhas, foi efectuada uma média centrada no passo de tempo pretendido.

Na figura 2.3 estão apresentados os mapas médios anuais dos dados observados às 0h, para cada uma das variáveis disponíveis durante o ano de 2005.

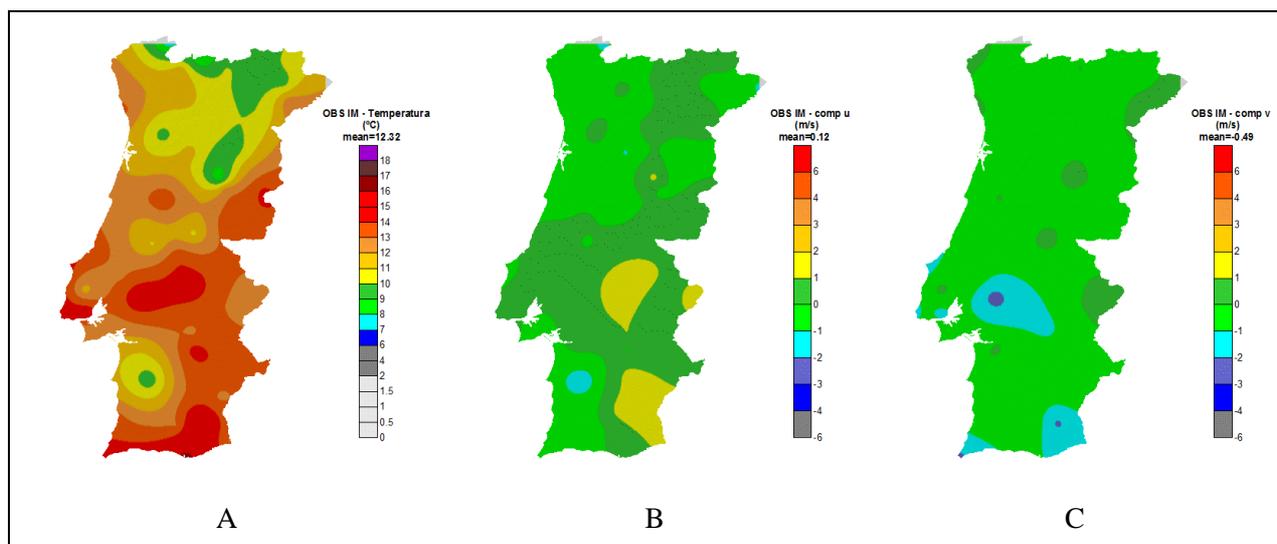


Figura 2.3 – Médias das Observações da rede IM das 0 horas referente ao ano de 2005. A) Temperatura; B) componente u do vento; C) componente v do vento.

## 2.4 Análise do output

Como o objectivo deste estudo é realizar um modelo de ajuste estatístico aplicado ao output do modelo numérico, para as variáveis da temperatura e das componentes horizontais do vento, o primeiro passo é analisar o comportamento do modelo numérico face às observações através de parâmetros estatísticos.

### 2.4.1 Medidas da análise da qualidade do modelo MM5

As medidas estatísticas como o viés, a raiz do erro médio quadrático e a correlação de Pearson, podem ser usados para quantificar a capacidade do modelo.

O viés (BE, Bias Error) mede a tendência do modelo para sobreavaliar ou subavaliar uma variável, e é definido matematicamente para uma dada variável  $x$ , por,

$$BE(x) = \frac{1}{N} \sum_{i=1}^N (x_i^p - x_i^o) \quad (2.1)$$

onde  $N$  é o número total de observações/previsões e o índices  $p$  e  $o$  indicam previsão e observação, respectivamente. Um valor positivo do viés indica uma predisposição do modelo em sobreavaliar uma variável particular. Reciprocamente, um valor negativo do viés implica uma subavaliação do modelo na avaliação de uma variável.

A raiz do erro médio quadrática (RMSE, Root Mean Square Error) é a raiz quadrada do erro médio quadrático (MSE, Mean Square Error), ou seja, a raiz quadrada da média das diferenças individuais quadráticas entre a previsão e as observações e é definido pela equação,

$$MSE(x) = \frac{1}{N} \sum_{i=1}^N (x_i^p - x_i^o)^2 \quad (2.2)$$

$$RMSE(x) = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i^p - x_i^o)^2} \quad (2.3)$$

esta medida retém as unidades físicas da variável prevista e mede o erro típico da previsão do modelo. Valores grandes do RMSE representam grandes erros nos campos previstos, e valores próximos de zero indicam uma previsão quase perfeita. Quadrando o termo da diferença, o RMSE tende a dar maior peso às grandes discrepâncias entre os campos observados e previstos.

O coeficiente de correlação de Pearson, também designado de coeficiente de correlação produto-momento ou simplesmente de  $r$  de Pearson mede o grau da correlação (e a direcção dessa correlação - se positiva ou negativa) entre duas variáveis. O coeficiente de correlação de Pearson calcula-se segundo a seguinte equação,

$$r = cor(x, y) = \frac{cov(x, y)}{\sqrt{var(x) var(y)}} = \frac{\sum_{i=1}^N (x_i^p - \bar{x}^p)(x_i^o - \bar{x}^o)}{\sqrt{\sum_{i=1}^N (x_i^p - \bar{x}^p)^2} \sqrt{\sum_{i=1}^N (x_i^o - \bar{x}^o)^2}} \quad (2.4)$$

onde a barra representa uma média sobre os casos observados e/ou de previsão. Este coeficiente, normalmente representado pela letra "r", assume apenas valores entre -1 e 1. Quando  $r=1$ , significa uma correlação perfeita positiva entre as duas variáveis. Se  $r = -1$ , significa uma correlação negativa perfeita entre as duas variáveis – isto é, se uma aumenta, a outra sempre diminui. No caso de  $r = 0$ , significa que as duas variáveis não dependem linearmente uma da outra. Quanto maior o valor de r, mais forte é a relação. Portanto, para  $r > 0$  ambas as variáveis crescem juntas, e para  $r < 0$  quando uma das variáveis diminui, a outra aumenta.

### 2.4.2 Avaliação da qualidade das simulações do modelo MM5

Na tabela 2.2 e nas figuras 2.4 a 2.6 estão apresentados os parâmetros estatísticos de avaliação da qualidade das simulações do MM5. Os resultados correspondem a um valor médio da temperatura (T), da componente zonal do vento u (u) e da componente meridional do vento v (v), para Portugal Continental, relativos ao ano de 2005 para uma previsão a 24 horas.

Tabela 2.2 – Parâmetros estatísticos de avaliação da qualidade da previsão do modelo MM5 para a temperatura e cada componente horizontal do vento.

<i>Parâmetros Estatísticos</i>	<i>Variáveis</i>		
	Temperatura	Componente zonal do vento - u	Componente meridional do vento - v
BE	1,29	-0,25	-0,67
r	0,92	0,49	0,51
RMSE	2,99	3,83	4,07
MSE	8,95	14,66	16,55

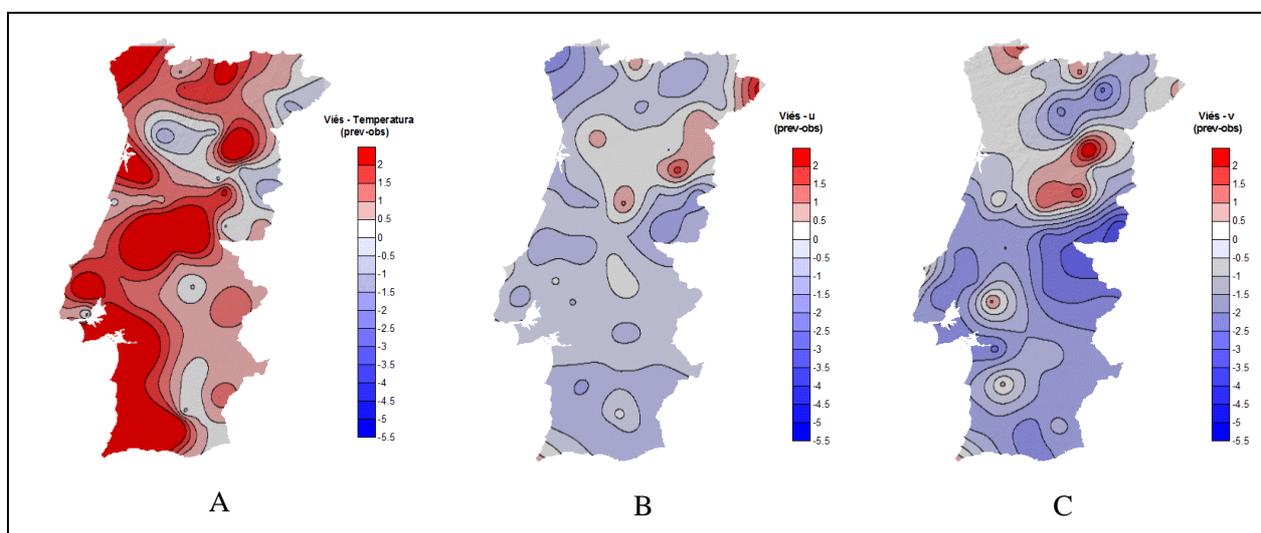


Figura 2.4 –Viés (BE) entre as previsões do MM5 e as observações da rede IM. A) Temperatura; B) componente u do vento; C) componente v do vento.

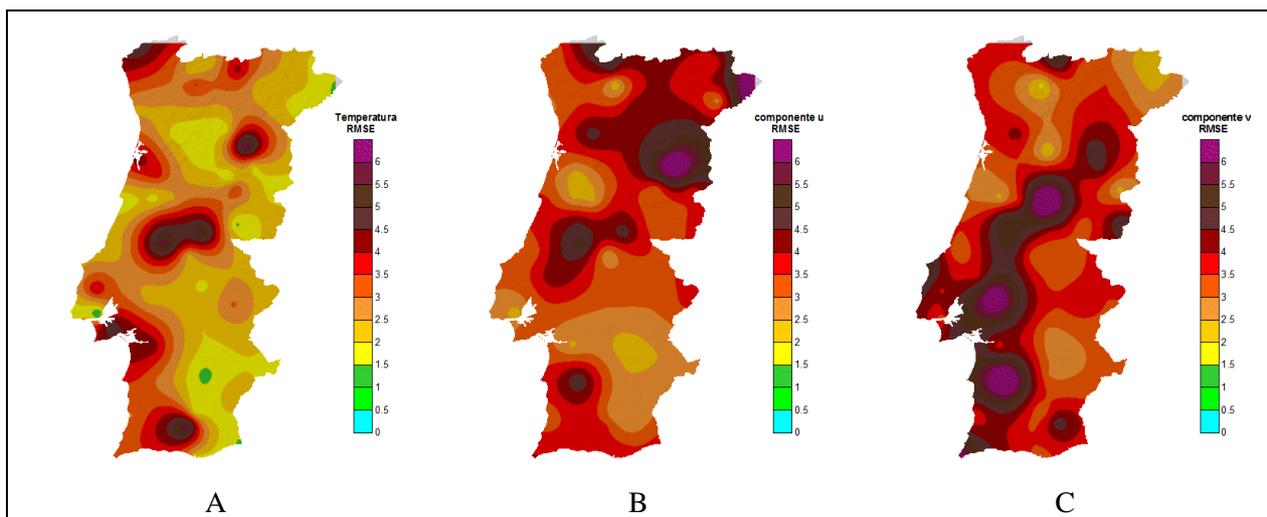


Figura 2.5 – Raiz do erro médio quadrático (RMSE) entre as previsões do MM5 e as observações da rede IM. A) Temperatura; B) componente u do vento; C) componente v do vento.

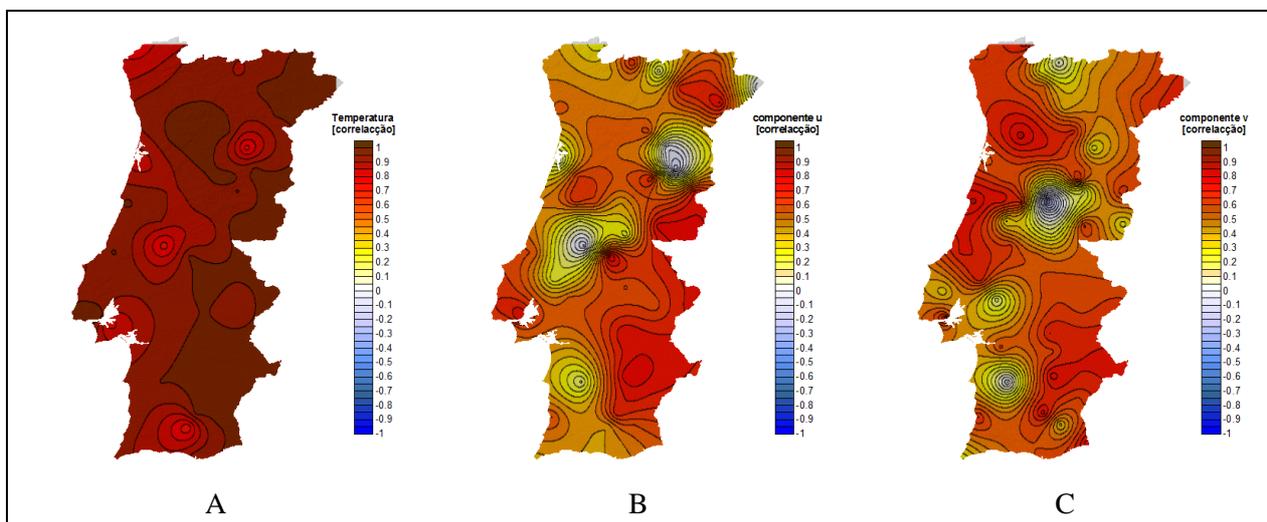


Figura 2.6 – Correlação (r) entre as previsões do MM5 e as observações da rede IM. A) Temperatura; B) componente u do vento; C) componente v do vento.

Em geral, os resultados demonstram que o modelo sobrestima os valores da temperatura, em especial quanto mais próximo da costa e subestima as componentes do vento, excepto na zona da Serra da Estrela. Contudo, o modelo vai sobreavaliar o vento nas zonas em altitude, principalmente na componente v do vento.

Na tabela 2.3 são apresentados os valores máximos e mínimos obtidos na análise dos parâmetros estatísticos RMSE e correlação.

Tabela 2.3 – Valores extremos do parâmetro RMSE e as correspondentes correlacões, para a temperatura e cada componente horizontal do vento.

Nº	Nome da Estação	RMSE			r		
		T	u	v	T	u	v
579	LISBOA/GAGO COUTINHO	<b>1.18</b>	2.34	4.11	0.97	0.75	0.55
726	ALCOBACA/ EST. FRUTIC. VI	<b>5.97</b>	5.32	5.28	0.72	-0.11	0.61
668	AROUCA	<b>6.27</b>	4.82	4.77	0.73	-0.04	0.27
637	MOGADOURO	1.44	<b>8.15</b>	2.73	0.98	-0.04	0.69
686	PAMPILHOSA DA SERRA	1.88	<b>7.78</b>	4.79	0.96	-0.19	0.48
632	MIRANDELA	3.58	<b>2.05</b>	<b>1.82</b>	0.93	0.40	0.51
687	COVILHA/AERO.	3.11	3.83	<b>8.28</b>	0.93	0.40	-0.35
765	CABO RASO/ FAROL	2.20	3.61	<b>7.27</b>	0.91	0.53	0.12
800	SABUGAL/ MARTIM REI	2.82	5.10	<b>7.82</b>	0.95	0.14	-0.06
878	PORTIMAO/ AERODROMO	<b>6.08</b>	3.61	3.41	0.71	0.40	0.59

Os valores máximos extremos (vermelho bold) indicam a estação onde o modelo consegue prever com menos exactidão, ou seja, as estações onde existem as maiores discrepâncias entre os campos observados e previstos. As estações nas quais o erro do modelo é significativo apresentam valores de correlacão baixo.

Estes resultados são a base de partida para a construção das equações de regressão do modelo estatístico, que visa minimizar os desvios médios e melhorar a previsão local dos parâmetros em estudo.

### 3. Modelo Estatístico -MOS

Os modelos de previsão estatísticos têm por base equações de regressões lineares múltiplas, onde a variável dependente é expressa em função das quantidades que contribuem para a sua variação. A aplicação desses métodos é descrita na literatura por Wilks (1995) e von Storch e Zwiers (1999).

#### 3.1 Breve descrição do modelo

O MOS é um exemplo de um modelo estatístico que estabelece relações entre as previsões do modelo numérico e as observações. O caso mais simplificado deste método é a equação da regressão linear simples, que estabelece a relação entre um único predictor e um predictando.

##### 3.1.1 Regressão linear simples

O método da regressão linear consiste em determinar a recta de regressão que melhor se ajusta aos dados  $(x,y)$  (onde  $x$  é a variável predictor e  $y$  é a variável predictando) minimizando a soma dos desvios quadráticos dos dados à recta. A equação de regressão é uma função de  $x$  dada pela expressão:

$$\hat{y}(x) = a + bx \quad (3.1)$$

onde  $\hat{y}$  é o valor previsto de  $y$  (observação - predictando) a partir de  $x$  (previsão do modelo numérico - predictor). Os erros da previsão do modelo de regressão são dados por:

$$e_i = \hat{y}_i(x_i) - y_i \quad (3.2)$$

Os coeficientes da equação de regressão  $a$  e  $b$  são obtidos através das seguintes relações:

$$b = \frac{\sum_{i=1}^n [(x_i - \bar{x})(y_i - \bar{y})]}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n (x_i)^2 - \left( \sum_{i=1}^n x_i \right)^2} \quad (3.3)$$

$$a = \bar{y} - b\bar{x} \quad (3.4)$$

onde  $\bar{x}$  e  $\bar{y}$  são os valores médios da previsão e da observação, respectivamente.

O coeficiente  $b$ , designado pelo declive da recta de regressão e dado pelo quociente entre a covariância de  $(x,y)$  e a variância de  $x$ , é proporcional ao coeficiente de correlação Pearson entre  $x$  e  $y$  (equação (2.4)).

A correlação é uma quantidade adimensional que funciona como um indicador da qualidade de previsão de  $x$ .

### 3.1.2 Regressão linear múltipla

Nos casos onde se considera  $k$  predictores, as equações de previsão são obtidas através de regressões lineares múltiplas. A equação de previsão vem:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_k x_k, \quad (3.5)$$

onde  $b_0$  é denominado a constante da regressão e os coeficientes da regressão são os termos  $b_i$  ( $i=1, \dots, k$ ), correspondentes aos  $k$  predictores lineares.

A equação (3.2) dos resíduos continua válida, se o valor previsto  $\hat{y}$  for função do vector dos predictores,  $x_i$ ,  $i=1, \dots, k$ .

Os  $k+1$  coeficientes da equação (3.5) são encontrados, minimizando a soma dos quadrados dos resíduos. Através da matriz dos dados  $[X]$ , o vector  $\hat{y}$  das previsões é dado por:

$$\hat{y} = [X]b \quad (3.6)$$

ou seja:

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} = \begin{bmatrix} 1 & x_{1,1} & x_{1,2} & \dots & x_{1,k} \\ 1 & x_{2,1} & x_{2,2} & \dots & x_{2,k} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{n,1} & x_{n,2} & \dots & x_{n,k} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix} \quad (3.7)$$

Onde  $\hat{y}$  é uma matriz ( $n \times 1$ ) (i.e., vector) de  $n$  observações do predictando,  $[X]$  é uma matriz ( $n \times k + 1$ ) contendo os valores dos predictores, e  $b^T = [b_1 \ b_2 \ \dots \ b_k]$  é um vector ( $k + 1 \times 1$ ) dos parâmetros da regressão.

As equações normais são obtidas pela multiplicação em cada lado na equação (3.6) por  $[X]^T$ :

$$[X]^T y = [X]^T [X]b \quad (3.8)$$

$$\begin{bmatrix} \sum y \\ \sum x_1 y \\ \vdots \\ \sum x_k y \end{bmatrix} = \begin{bmatrix} n & \sum x_1 & \dots & \sum x_k \\ \sum x_1 & \sum x_1^2 & \dots & \sum x_1 x_k \\ \vdots & \vdots & \dots & \vdots \\ \sum x_k & \sum x_k x_1 & \dots & \sum x_k^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix} \quad (3.9)$$

A matriz  $[X]^T[X]$  tem dimensão  $(k+1 \times k+1)$  e, cada lado da equação (3.8), tem dimensão  $(k+1 \times 1)$ .

### 3.2 Metodologia da construção do modelo estatístico

Depois de uma análise cuidadosa, da qualidade do modelo numérico e preparação da base de dados, que irá ser aplicada na construção da equação de regressão multilinear, o passo seguinte é seleccionar os possíveis predictores e validar o modelo estatístico para cada estação e variável em estudo.

#### 3.2.1 Selecção de predictores

Os predictores para a elaboração da equação de regressão, devido às inúmeras associações possíveis disponíveis, foram pré-seleccionados (figura 3.1).

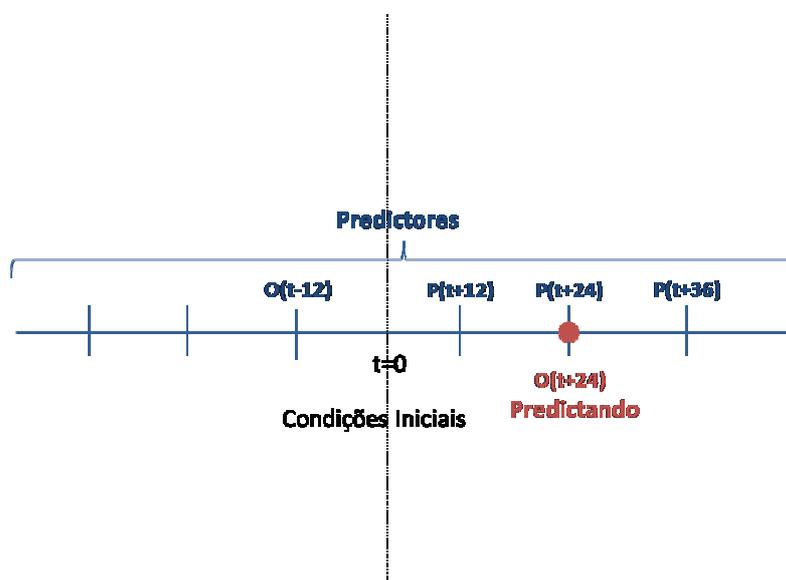


Figura 3.1 – Representação espacial e temporal dos predictores pré-seleccionados.

Na tabela 3.1 estão apresentados os predictores para os cinco casos de estudo realizados e na figura 3.1 está representada uma visualização temporal dos predictores escolhidos.

Tabela 3.1 – Predictores seleccionados para o MOS em t+24h e para cada caso de estudo

Casos	Predictores
A	P(t+24)
B	P(t+24) O(t)
C	P(t+24) O(t) O(t-12)
D	P(t+24) O(t) O(t-12) P(t+12)
E	P(t+24) O(t) O(t-12) P(t+12) P(t+36)

As 5 versões escolhem observações em t=0 ou anteriores e previsões lançadas em t=0 para diferentes prazos, mesmo que diferentes do instante para o qual se pretende executar o MOS (t+24h).

Duas metodologias de aplicação do modelo foram efectuadas, de forma a estudar a sensibilidade do modelo de regressão em relação aos predictores introduzidos. Na primeira, o modelo foi construído sem selecção de predictores, onde a ordem corresponde à descrita na tabela anterior. No segundo, foi realizado um método de selecção de predictores, baseado no método de regressão ‘forward stepwise’ (forward stepwise regression) (von Storch, Zwiers, 1999) que permite seleccionar de forma progressiva os predictores de acordo com um critério de variância explicada. A escolha dos predictores termina quando o diferencial de variância explicada por inclusão de novos predictores já não ser estatisticamente significativo.

### 3.2.1.1 Regressão *forward stepwise*

A filtragem da regressão consiste em seleccionar, de entre o conjunto de potenciais predictores, um subconjunto de predictores estatisticamente robustos. A metodologia da regressão stepwise (stepwise regression) consiste em seleccionar este sub-conjunto através da exclusão dos piores ou melhores predictores.

O método escolhido para esta selecção foi o *forward stepwise regression*, que consiste em adicionar, sucessivamente, o melhor predictor disponível, na elaboração de cada equação de regressão, em cada estação e variável de previsão. Este processo de selecção foi realizado através da seguinte estrutura:

- 1- No primeiro passo todos os  $k$  predictores ( $x_i$ ) são avaliados quanto à sua capacidade predictiva do predictando  $y$ . Onde o primeiro predictor ( $i_1$ ) é escolhido através da maximização da correlacção entre os predictores e a observação, através de:

$$r = \text{correlacção}(y, x_i) \rightarrow i_1 = \arg \max |r(y, x_i)| \quad (3.10)$$

- 2- Constrói-se o modelo de regressão teste e calculam-se os coeficientes  $a$  e  $b$  de regressão da previsão (através da equação (3.1)).
- 3- O segundo predictor ( $i_2$ ) é seleccionado através da correlacção máxima entre o residuo (3.3) da previsão anterior e os restantes predictores.

$$r = \text{correlacção}(e_1, x_i) \rightarrow i_2 = \arg \max |r(e_1, x_i)| \quad (3.11)$$

- 4- Contudo, em cada passo em que se tenham escolhido  $m-1$  predictores, haverá um aumento significativo da variância explicada no  $m$ -ésimo predictor escolhido se o teste:

$$\frac{SSR_m - SSR_{m-1}}{(SSE_m / (n - m))} \quad (3.12)$$

for superior ao quantil  $\alpha$  da distribuição  $F(1, n-m)$ . Na equação (3.12), o termo  $SSR$  é a soma dos quadrados da regressão, ou a soma dos quadrados das diferenças entre as previsões da regressão e a média de  $y$ , com  $m$  predictores, e  $SSE$  é a soma dos quadrados dos resíduos.

Caso não haja aumento significativo da variância explicada com a adição de um novo predictor, o predictor é excluído da equação de regressão.

Este processo continua com a escolha do terceiro melhor predictor, continuando até não existirem mais predictores disponíveis. Dado o número reduzido de predictores, não é aplicado o teste de significância da redução de variância explicada. O método aqui apresentado (método com selecção) difere da regressão multilinear no seguinte. Em vez de se resolver o sistema de equações dos coeficientes de regressão multilinear com todos os predictores, fazem-se regressões lineares dos resíduos das regressões anteriores até esgotar todos os predictores. A equação de regressão final difere da obtida com a regressão multilinear (método sem selecção). O método sem selecção garante um menor erro médio quadrático no período de calibração do modelo do que o método com selecção. No entanto o método com selecção, pelo menos quando o número de predictores é reduzido (neste caso no máximo de 5) e não há perigo de overfitting (excesso de predictores), pode produzir menores erros médios quadráticos quando avaliado em validação cruzada, tal como veremos neste trabalho.

### 3.2.2 Validação cruzada

A fase final da construção do modelo é a validação. O MOS será aferido em validação cruzada, onde as equações de regressão utilizadas para previsão meteorológica são testadas numa amostra de dados independentes da amostra de calibração do modelo estatístico. A validação cruzada é uma técnica de re-amostragem, os dados disponíveis são divididos repetidamente em subconjuntos de dados de desenvolvimento ou calibração do modelo estatístico e verificação. Tal procede-se, simplesmente avaliando a performance da equação de previsão numa parte dos dados, enquanto se usa o resto dos dados como amostra de calibração.

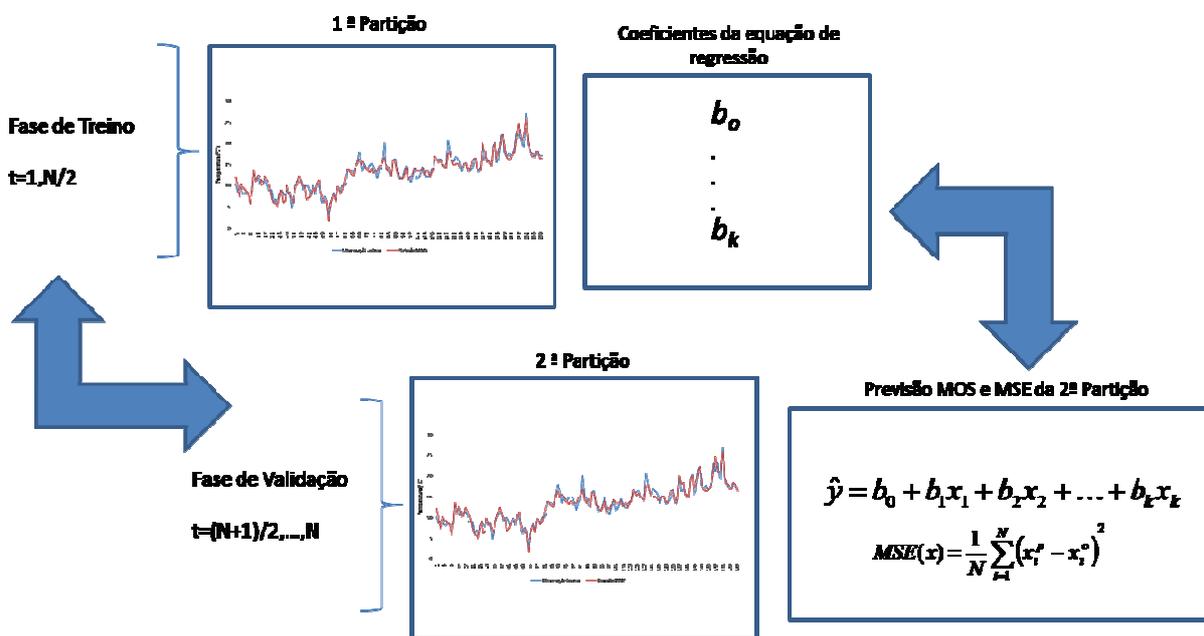


Figura 3.2 – Diagrama do método de validação aplicado (validação cruzada).

A figura 3.2 ilustra o método de validação aplicado. A validação cruzada é feita usando os conjuntos de dados de desenvolvimento do tamanho  $N/2$ , e os conjuntos de verificação contendo a restante observação do predictando, onde  $N$  é o número total de dados da série. O modelo de regressão é recalculado para cada uma das duas partições (fase de treino e fase de validação). Uma vez obtidas as previsões, calcula-se o erro médio quadrático de previsão [MSE – equação (2.2)] obtido pela média dos quadrados das diferenças entre observações e previsões. O processo de validação cruzada usa

todas as  $N$  observações do predictando para estimar o MSE, permitindo que cada observação seja tratada de cada vez, como um dado independente.

### 3.3 Análise do output

#### 3.3.1 Medidas da qualidade do modelo estatístico

Para além das medidas apresentadas na secção 2.4.1, o score (SC), também, irá ser uma medida adoptada para a avaliação da qualidade do modelo estatístico obtido para cada caso de estudo e variável.

O SC - MESS (Mean square error skill score) é a razão entre a diferença do erro médio quadrático do modelo numérico e do modelo estatístico, com o erro médio quadrático do modelo numérico e é definido pela equação,

$$SC = \frac{MSE_{(MM5)} - MSE_{(MOS)}}{MSE_{(MM5)}} \quad (3.12)$$

Se o valor máximo for igual a 1 a previsão do MOS é perfeita, enquanto que se o valor máximo for igual a 0 o modelo de previsão estatístico é equivalente à previsão numérica. Um valor negativo corresponde ao pior resultado, ou seja, que a previsão do modelo estatístico é pior que as previsões do modelo numérico.

#### 4. Apresentação dos resultados do modelo estatístico

Nesta secção vão ser apresentados e discutidos os resultados finais do modelo estatístico, bem como a avaliação estatística, para os cinco casos de estudo. Foram aplicadas as metodologias apresentadas na secção anterior (método com selecção e sem selecção).

Da análise da qualidade do modelo numérico, as previsões desinviadas do MM5, ou seja, obtidas por uma correcção aditiva simples:  $y_{\text{unbiased}} = y_{\text{MM5}} - \text{BE}$ , têm no período de avaliação do viés, o valor de  $\text{MSE}_{\text{unbiased}} = (\text{MSE}_{\text{MM5}} - \text{BE})^2$  e portanto o respectivo score (SC) vale  $(\text{BE})^2 / \text{MSE}_{\text{MM5}}$ . Recorrendo à tabela 2.2 e aos valores médios de BE e de  $\text{MSE}_{\text{MM5}}$  sobre todas as estações, obtemos os scores de 0.19, 0.004 e 0.027 para as previsões desinviadas, respectivamente da temperatura T, componente u e componente v do vento.

##### Correlacção dos predictores com o predictando

O primeiro passo para a obtenção do modelo de regressão é a análise da qualidade da relação entre os dados disponíveis, ou seja, a relação entre as observações das 0h e os diferentes predictores pré-seleccionados. A correlacção (r) é a medida estatística, apresentada na tabela 4.1, que irá servir para estabelecer esta relação. Assim, para cada estação IM e cada grandeza (u,v,T), destaca-se a cor, de entre os 5 predictores, o que exhibe maior correlacção com as observações respectivas, utilizando-se as cores vermelha, azul e verde, respectivamente para T,u e v.

Tabela 4.1 – Correlacção entre predictores e predictando para as variáveis em estudo (temperatura, componente zonal e meridional do vento).

	Predictores														
	P(t+24)			P(t+12)			P(t+36)			O(t)			O(t-12)		
	T	u	v	T	u	v	T	U	v	T	u	v	T	u	v
531	0,88	0,60	0,44	0,86	0,62	0,43	0,86	0,68	0,45	0,89	0,45	0,60	0,84	0,46	0,48
533	0,91	0,37	0,23	0,91	0,16	0,53	0,89	0,27	0,46	0,89	0,35	0,60	0,83	0,22	0,41
535	0,97	0,74	0,73	0,93	0,61	0,67	0,93	0,67	0,74	0,91	0,50	0,60	0,89	0,47	0,36
545	0,92	0,52	0,64	0,87	0,45	0,62	0,87	0,55	0,65	0,88	0,31	0,42	0,78	0,29	0,46
548	0,94	0,65	0,74	0,87	0,54	0,63	0,86	0,55	0,64	0,87	0,35	0,54	0,83	0,33	0,44
554	0,95	0,42	0,57	0,94	0,19	0,48	0,94	0,31	0,52	0,92	0,46	0,32	0,88	0,29	0,07
558	0,96	0,75	0,70	0,93	0,60	0,63	0,94	0,59	0,65	0,91	0,61	0,53	0,86	0,38	0,39
560	0,97	0,48	0,67	0,91	0,58	0,62	0,93	0,65	0,57	0,90	0,51	0,52	0,82	0,42	0,46
562	0,97	0,79	0,77	0,93	0,61	0,66	0,94	0,61	0,68	0,92	0,55	0,51	0,88	0,33	0,40
567	0,95	0,44	0,37	0,92	0,34	0,39	0,94	0,38	0,42	0,91	0,23	0,31	0,86	0,12	0,19
568	0,97	0,68	0,65	0,95	0,55	0,41	0,95	0,57	0,61	0,89	0,44	0,39	0,86	0,27	0,23
570	0,98	0,77	0,61	0,94	0,57	0,57	0,95	0,67	0,54	0,93	0,49	0,46	0,90	0,35	0,35
575	0,96	0,56	0,52	0,94	0,42	0,22	0,95	0,46	0,20	0,93	0,28	0,18	0,89	0,14	0,07
577	0,94	0,55	0,65	0,90	0,45	0,54	0,89	0,47	0,60	0,90	0,48	0,41	0,85	0,38	0,25
579	0,97	0,75	0,55	0,92	0,52	0,68	0,93	0,68	0,73	0,91	0,49	0,59	0,88	0,29	0,39
604	0,78	0,40	0,59	0,75	0,34	0,59	0,75	0,50	0,58	0,83	0,14	0,43	0,77	0,21	0,34
606	0,83	0,26	0,72	0,81	0,47	0,65	0,82	0,55	0,71	0,81	0,32	0,49	0,72	0,19	0,34
611	0,95	0,65	0,04	0,91	0,49	0,22	0,93	0,63	0,33	0,90	0,43	0,18	0,84	0,31	0,06
616	0,89	0,12	0,38	0,89	0,07	0,34	0,89	0,08	0,29	0,90	0,20	0,37	0,87	0,05	0,14
619	0,95	0,40	0,29	0,91	0,35	0,35	0,93	0,56	0,43	0,90	0,46	0,31	0,83	0,28	0,28
632	0,93	0,40	0,51	0,92	0,26	0,17	0,92	0,33	0,29	0,92	0,14	0,27	0,89	0,08	0,08
633	0,96	0,68	0,48	0,93	0,56	0,30	0,94	0,66	0,27	0,92	0,49	0,21	0,88	0,28	0,16
635	0,95	0,63	0,54	0,94	0,50	0,49	0,93	0,61	0,39	0,93	0,56	0,47	0,90	0,31	0,32
637	0,98	-0,04	0,69	0,94	0,55	0,62	0,96	0,59	0,65	0,91	0,51	0,46	0,87	0,30	0,30
654	0,97	0,76	0,64	0,94	0,64	0,56	0,96	0,69	0,55	0,92	0,49	0,35	0,89	0,34	0,24
663	0,95	0,49	0,50	0,93	0,34	0,35	0,94	0,36	0,46	0,93	0,32	0,18	0,88	0,28	0,07

Tabela 4.1(Cont.) – Correlação entre preditores e predictando para as variáveis em estudo (temperatura, componente zonal e meridional do vento).

	Predictores														
	P(t+24)			P(t+12)			P(t+36)			O(t)			O(t-12)		
	T	u	v	T	u	v	T	u	v	T	u	v	T	u	V
666	0,97	0,57	0,66	0,91	0,40	0,63	0,93	0,53	0,68	0,91	0,54	0,59	0,85	0,37	0,45
668	0,73	-0,04	0,27	0,74	0,41	0,30	0,73	0,36	0,19	0,76	0,16	0,17	0,73	0,12	0,04
683	0,97	0,54	0,82	0,93	0,45	0,57	0,95	0,46	0,70	0,92	0,45	0,55	0,86	0,25	0,33
686	0,96	-0,19	0,48	0,90	0,09	0,51	0,92	0,20	0,54	0,89	0,40	0,57	0,82	0,26	0,39
687	0,93	0,40	-0,35	0,92	0,45	0,13	0,93	0,51	0,30	0,91	0,29	0,43	0,89	0,21	0,27
697	0,89	0,48	0,19	0,88	0,37	0,08	0,88	0,18	0,17	0,90	0,19	0,04	0,86	0,23	0,04
702	0,88	0,58	0,66	0,84	0,49	0,61	0,83	0,50	0,63	0,87	0,31	0,44	0,79	0,37	0,38
705	0,82	0,08	0,52	0,80	0,20	0,49	0,77	0,37	0,49	0,79	0,19	0,40	0,78	0,09	0,28
713	0,86	0,60	0,55	0,78	0,47	0,53	0,77	0,51	0,48	0,79	0,20	0,32	0,74	0,29	0,28
716	0,96	0,47	0,81	0,89	0,63	0,66	0,90	0,69	0,73	0,87	0,47	0,59	0,83	0,38	0,38
724	0,91	0,20	0,29	0,87	0,18	0,58	0,88	0,21	0,59	0,88	0,59	0,57	0,83	0,36	0,28
726	0,72	-0,11	0,61	0,71	0,00	0,54	0,69	0,04	0,51	0,79	0,23	0,43	0,70	0,14	0,33
734	0,95	0,54	0,71	0,91	0,30	0,62	0,91	0,40	0,65	0,91	0,51	0,51	0,86	0,21	0,36
739	0,92	0,21	0,69	0,88	0,15	0,55	0,86	0,23	0,44	0,89	0,33	0,49	0,82	0,26	0,35
744	0,91	0,61	0,32	0,87	0,40	0,39	0,88	0,47	0,31	0,90	0,47	0,19	0,84	0,30	0,17
765	0,91	0,53	0,12	0,88	0,47	0,25	0,88	0,61	0,30	0,87	0,55	0,60	0,78	0,42	0,42
766	0,97	0,62	0,49	0,94	0,51	0,63	0,94	0,57	0,72	0,93	0,51	0,49	0,90	0,46	0,23
767	0,77	0,61	0,46	0,73	0,50	0,61	0,72	0,54	0,60	0,66	0,40	0,39	0,70	0,43	0,28
776	0,93	0,51	0,50	0,90	0,40	0,48	0,90	0,43	0,44	0,92	0,47	0,35	0,85	0,32	0,28
783	0,91	0,41	0,56	0,89	0,28	0,57	0,88	0,25	0,53	0,91	0,38	0,39	0,84	0,26	0,25
800	0,95	0,14	-0,06	0,93	0,22	0,39	0,93	0,29	0,56	0,93	0,35	0,57	0,89	0,18	0,46
803	0,97	0,71	0,62	0,93	0,62	0,53	0,95	0,73	0,58	0,93	0,53	0,39	0,89	0,33	0,23
806	0,95	0,76	0,27	0,91	0,52	0,35	0,93	0,63	0,39	0,90	0,50	0,17	0,87	0,33	0,09
812	0,85	0,22	0,21	0,86	0,33	0,18	0,85	0,41	0,23	0,87	0,46	0,15	0,82	0,31	0,03
824	0,96	0,81	0,53	0,94	0,54	0,56	0,94	0,71	0,50	0,94	0,51	0,24	0,91	0,39	0,16
835	0,97	0,54	0,59	0,95	0,41	0,54	0,96	0,34	0,57	0,93	0,28	0,33	0,91	0,23	0,23
837	0,96	0,71	0,62	0,94	0,52	0,52	0,94	0,60	0,65	0,93	0,49	0,34	0,89	0,32	0,15
848	0,93	0,58	0,57	0,92	0,37	0,54	0,92	0,53	0,48	0,92	0,37	0,30	0,87	0,20	0,25
850	0,97	0,71	0,59	0,95	0,56	0,59	0,96	0,60	0,64	0,95	0,55	0,39	0,91	0,30	0,26
863	0,97	0,71	0,72	0,95	0,57	0,61	0,96	0,56	0,68	0,95	0,51	0,45	0,90	0,29	0,28
864	0,96	0,68	0,59	0,92	0,64	0,55	0,93	0,57	0,63	0,93	0,48	0,42	0,87	0,28	0,27
865	0,97	0,69	0,77	0,94	0,51	0,58	0,95	0,47	0,70	0,94	0,47	0,49	0,89	0,27	0,31
867	0,94	0,55	0,33	0,94	0,50	0,47	0,95	0,39	0,57	0,94	0,50	0,42	0,89	0,34	0,04
868	0,97	0,54	0,77	0,92	0,38	0,61	0,94	0,45	0,67	0,90	0,54	0,50	0,86	0,28	0,37
878	0,71	0,40	0,59	0,85	0,21	0,44	0,84	0,25	0,45	0,89	0,40	0,35	0,80	0,21	0,04

Em geral, os resultados mostram que existe uma maior relação entre a observação e a previsão das 24h. Este resultado é mais persistente para a temperatura, e menos consistente para as componentes zonal e meridional do vento. Em geral e para além das previsões P(t+24), as observações da temperatura no instante de lançamento da previsão, O(t) e nas 12 horas precedentes, O(t-12) são igualmente bons preditores conforme se verifica pelas elevadas correlações, da ordem de ~0.85 a ~0.96. Tal, mostra que previsões por persistência amortecida da temperatura (regressão linear com observações antecedentes) são de razoável qualidade. Tem-se que, para a maioria das estações e para a temperatura em O(t+24h) (instante da validação), a sequência de preditores por ordem decrescente de qualidade são: P(t+24), (P(t+12) ou P(t+36)), O(t) e O(t-12). No que respeita às componentes (u,v) do vento, as correlações das previsões P(t+24) com as observações O(t+24) são da ordem de ~0.40 a ~0.70 com excepção de algumas estações francamente más (e.g. 668, 686, 689, 726). Para algumas estações, o melhor preditor do vento em t+24 é a previsão P(t+36). Tal poderá mostrar que no que diz respeito ao vento, a previsão MM5 está atrasada. Uma hipótese de explicação é a de que a velocidade de propagação dos sistemas sinópticos é subestimada no modelo MM5, juntamente com o efeito de certos efeitos locais. Apenas, num número ainda mais reduzido de estações, os melhores preditores do vento são P(t+12) ou O(t), apesar de as diferenças para P(t+24) não serem muito significativas.

## 4.1 Resultados dos MOS

Nas seguintes secções, vão ser apresentados os resultados dos cinco casos de estudo do modelo estatístico MOS. Para os últimos quatro casos, vão ser analisadas duas abordagens, de acordo com a metodologia aplicada à ordem dos predictores e apresentadas as respectivas avaliações estatísticas. Os resultados mapeados apresentam um valor médio de previsão estatística MOS de 24 horas e do score (SC), da temperatura (T), da componente zonal u do vento e da componente meridional v do vento para Portugal Continental, relativos ao ano de 2005. Os scores globais calculam-se pela mesma fórmula de score em cada estação com a excepção de que o MSE é acumulado sobre todas as estações, ou seja, para o cálculo faz-se a concatenação das séries para todas as estações avaliando-se as estatísticas com a essa série alongada. O mesmo se utiliza para as viés global, correlação global, MSE global e RMSE global.

### 4.1.1 Caso A

No primeiro caso de estudo, apenas foi pré-seleccionado um predictor (previsão das 24 horas do modelo MM5 – P(t+24)). Os resultados finais do modelo para cada variável estão apresentados na figura 4.1.

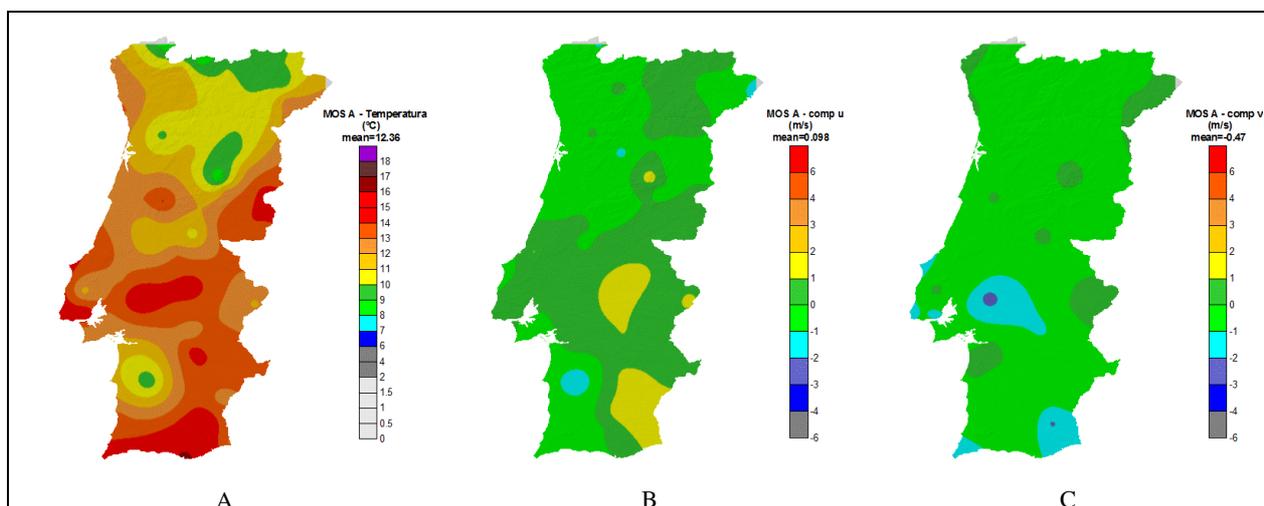


Figura 4.1 – Previsões de 24 horas do MOS com um predictor A) Temperatura; B) componente u do vento; C) componente v do vento.

#### 4.1.1.1 Avaliação da qualidade das previsões do MOS (caso A)

Na tabela 4.2 estão apresentados os parâmetros estatísticos de avaliação da qualidade dos resultados do modelo estatístico MOS, para o caso A. Os scores médios em validação cruzada são: 0.41, 0.79 e 0.83, respectivamente para T, u e v. Ao compará-los com os scores respectivos das previsões desinviesadas: 0.19, 0.004 e 0.027, verificamos que a afectação de um factor multiplicativo à previsão MM5 P(t+24), através da regressão linear simples, introduz uma elevada correcção. Esta correcção e o valor do score tem duas origens. A primeira provém do *downscaling* linear entre a previsão MM5 e a estação IM com correcções devidas a diferenças de altitude e correcções horizontais. A segunda correcção é de natureza intrínseca ao modelo, dependendo dos erros de representatividade do modelo associados à sua resolução espacial e temporal, e ainda aos erros puramente associados a incertezas das condições iniciais e fronteira (erros de previsão associados à predictabilidade).

Tabela 4.2 – Parâmetros estatísticos de avaliação da qualidade da previsão do modelo MOS – caso A para a temperatura e cada componente horizontal do vento.

Parâmetros Estatísticos	Variáveis		
	Temperatura	Componente zonal do vento – u	Componente meridional do vento - v
BE	0.04	-0.02	0.02
r	0.92	0.50	0.53
RMSE	2.30	1.77	1.69
MSE	5.31	3.14	2.87

Os mapas da figura 4.2 representam a análise da qualidade da previsão MOS relativamente à previsão MM5 (SC).

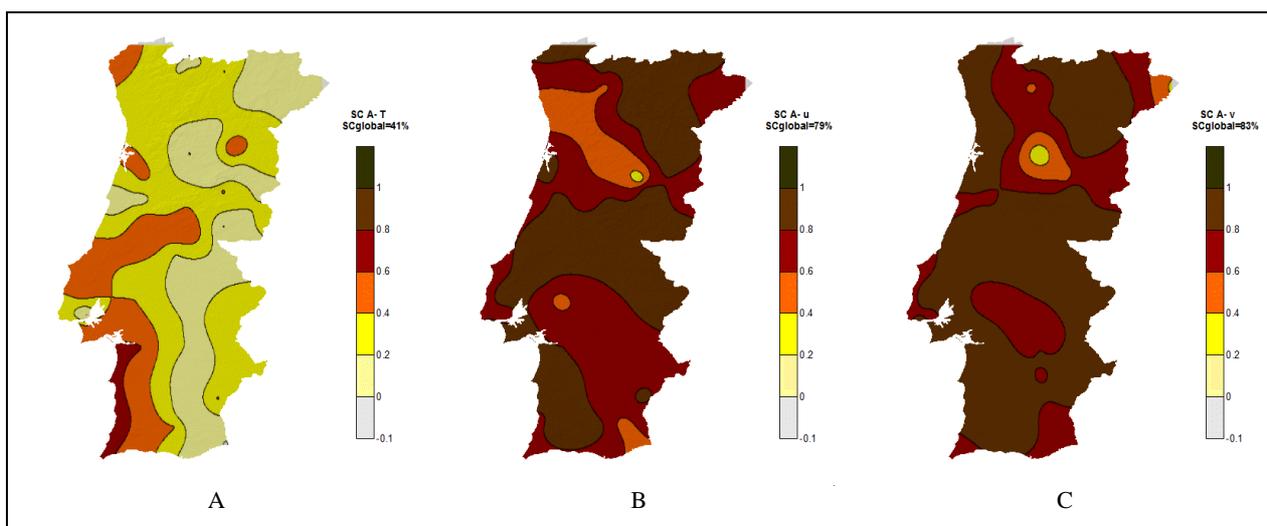


Figura 4.2 – Score (SC) entre as previsões do MOS – caso A e as previsões MM5. A) Temperatura; B) componente u do vento; C) componente v do vento.

#### 4.1.2 Caso B

Neste caso foi adicionado o predictor da observação efectuada para  $t=0$ , ou seja, a observação realizada na hora da simulação do modelo. A selecção dos dois predictores ( $P(t+24)$  e  $O(t)$ ), para as duas abordagens, seguiu a metodologia apresentada no capítulo anterior. Os resultados finais do modelo estatístico sem selecção e com selecção são apresentados na figuras 4.3 e 4.4, respectivamente.

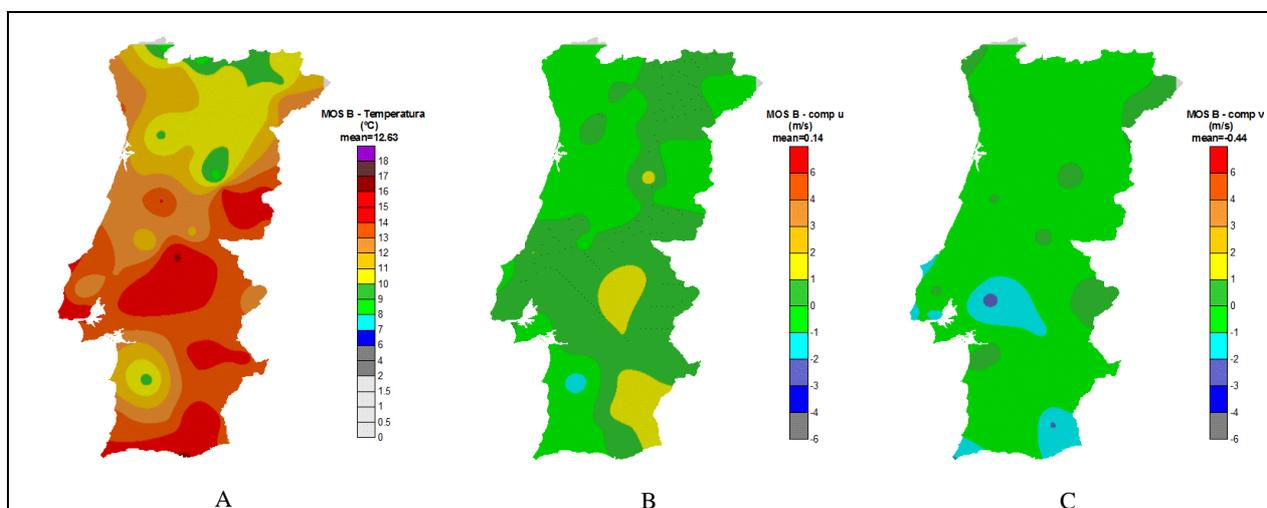


Figura 4.3 – Previsões de 24 horas do MOS – caso B com dois predictores A) Temperatura; B) componente u do vento; C) componente v do vento.

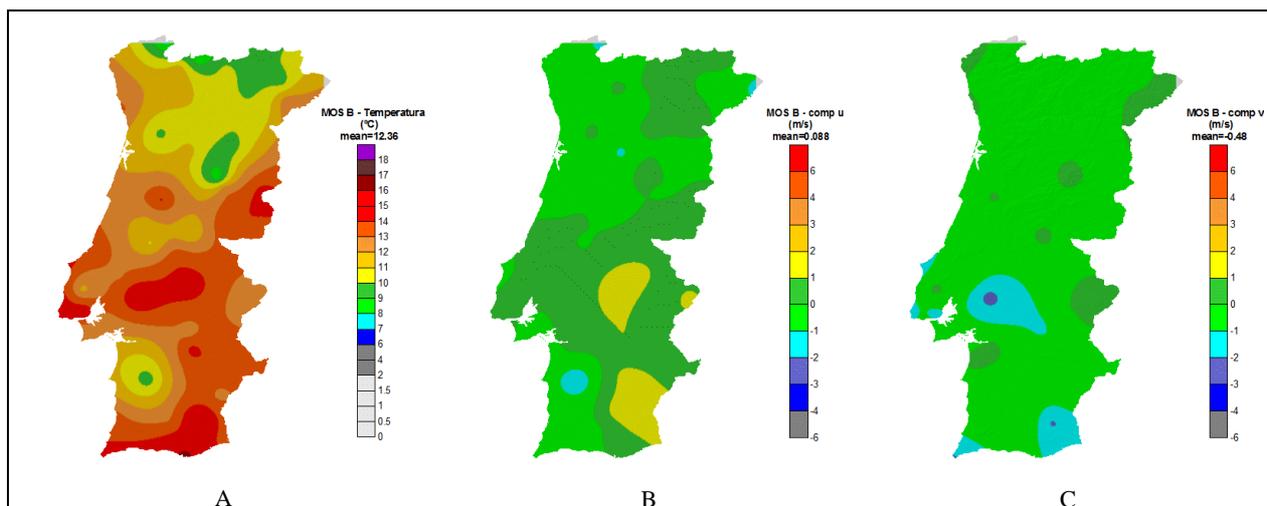


Figura 4.4 – Previsões de 24 horas do MOS – caso B com selecção de dois predictores. A) Temperatura; B) componente u do vento; C) componente v do vento.

#### 4.1.2.1 Avaliação da qualidade das previsões do MOS (caso B)

Os resultados da avaliação estatística, para ambos os casos, estão apresentados na tabela 4.3 e 4.4. As figuras 4.5 e 4.6 mostram os mapas do score do modelo MOS- caso B em relação à previsão MM5. Na regressão sem selecção, isto é na regressão com 2 predictores simultâneos ( $P(t+24)$  e  $O(t)$ ), obtiveram-se scores globais e valores de MSE que praticamente não diferem dos obtidos com  $P(t+24)$  (tabela 4.2). Quando se opta pelo método de regressão com selecção, isto é em que se faz regressão linear do resíduo resultante da previsão linear do melhor dos dois predictores, estação a estação, então obtém-se melhores resultados, isto é menores valores de MSE em validação cruzada. Tal evidencia que o segundo método, dito com selecção é mais eficiente, pelo menos para o tipo de predictores considerados. Estes resultados verificam-se especialmente para a temperatura em que se passa de um score de 0.41 para 0.56, ao optar por um método com selecção de predictores em vez de sem selecção. Quanto às componentes do vento as melhorias são marginais. De notar ainda que, o MOS da temperatura, para ambos os métodos é especialmente eficaz, nas regiões litorais a Sul do Tejo. Quanto ao vento, a distribuição regional do score é relativamente homogénea.

Tabela 4.3 – Parâmetros estatísticos de avaliação da qualidade da previsão do modelo MOS – caso B, sem selecção de predictores, para a temperatura e cada componente horizontal do vento.

<i>Parâmetros Estatísticos</i>	<i>Variáveis (sem selecção)</i>		
	Temperatura	Componente zonal do vento – u	Componente meridional do vento - v
BE	0.308	0.018	0.046
r	0.93	0.54	0.58
RMSE	2.29	1.75	1.63
MSE	5.25	3.05	2.67

Tabela 4.4 – Parâmetros estatísticos de avaliação da qualidade da previsão do modelo MOS – caso B, com selecção de dois predictores, para a temperatura e cada componente horizontal do vento.

Parâmetros Estatísticos	Variáveis (com selecção)		
	Temperatura	Componente zonal do vento – u	Componente meridional do vento - v
BE	0.039	-0.032	0.007
r	0.94	0.58	0.59
RMSE	1.98	1.66	1.56
MSE	3.92	2.77	2.44

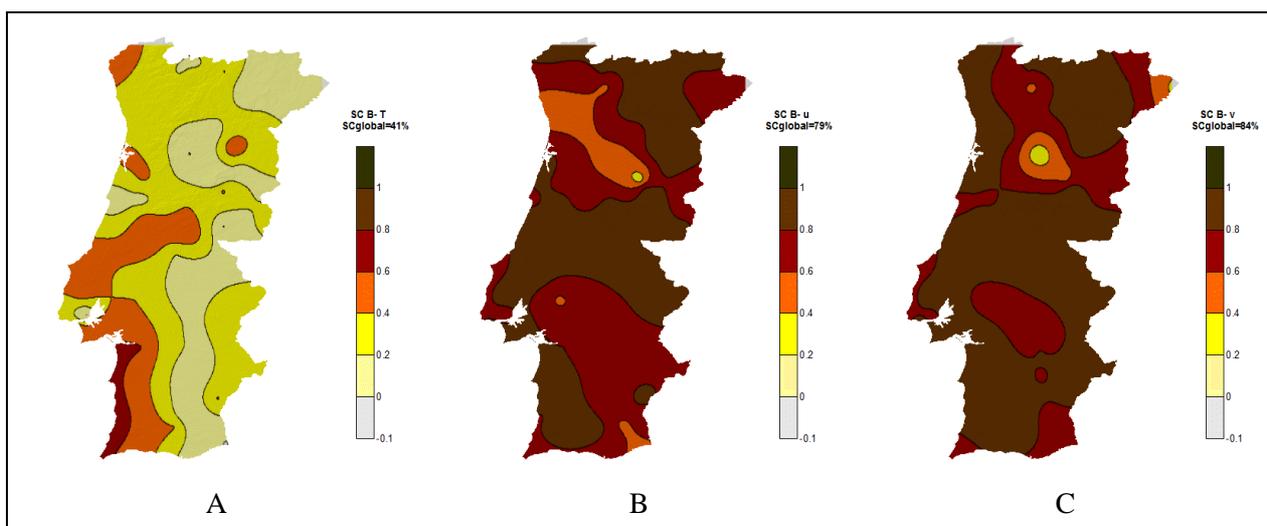


Figura 4.6 – Score (SC) entre as previsões do MOS- caso B, com dois predictores, e as previsões MM5. A) Temperatura; B) componente u do vento; C) componente v do vento.

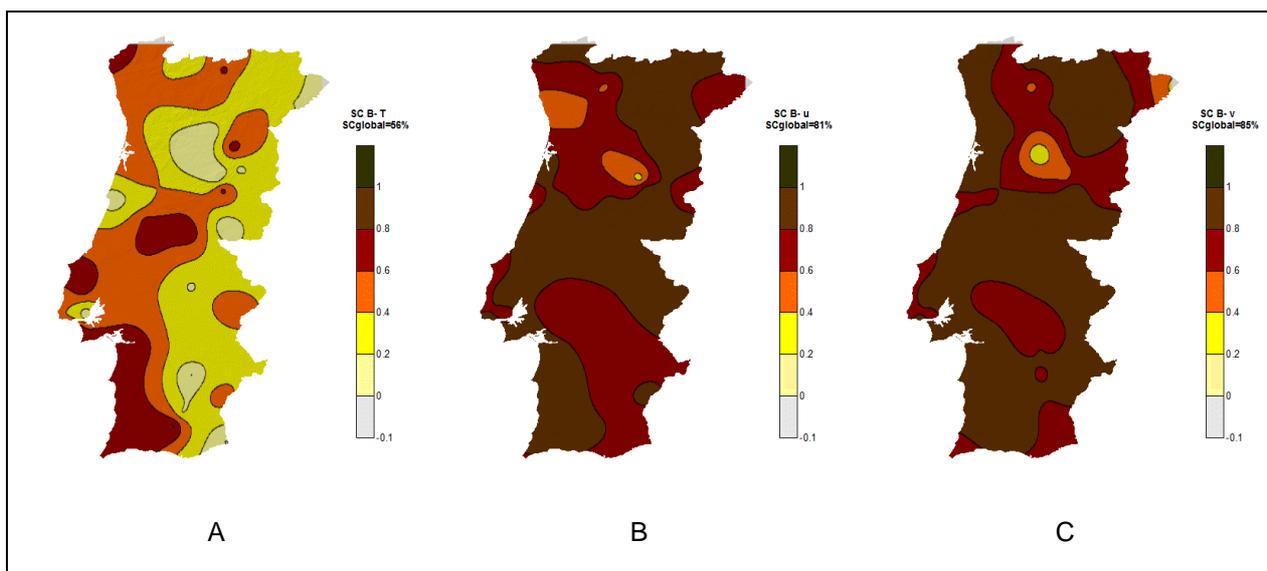


Figura 4.7 – Score (SC) entre as previsões do MOS – caso B, com selecção de dois predictores, e as previsões MM5. A) Temperatura; B) componente u do vento; C) componente v do vento.

### 4.1.3 Caso C

O modelo de regressão para o caso C foi elaborado com 3 predictores ( $P(t+24)$ ,  $O(t)$  e  $O(t-12)$ ). Relativamente ao caso B, foi adicionado como predictor, os dados da observação efectuada para 12 horas antes do lançamento da previsão. Os resultados finais para as duas metodologias estão apresentados nas figuras 4.8 e 4.9.

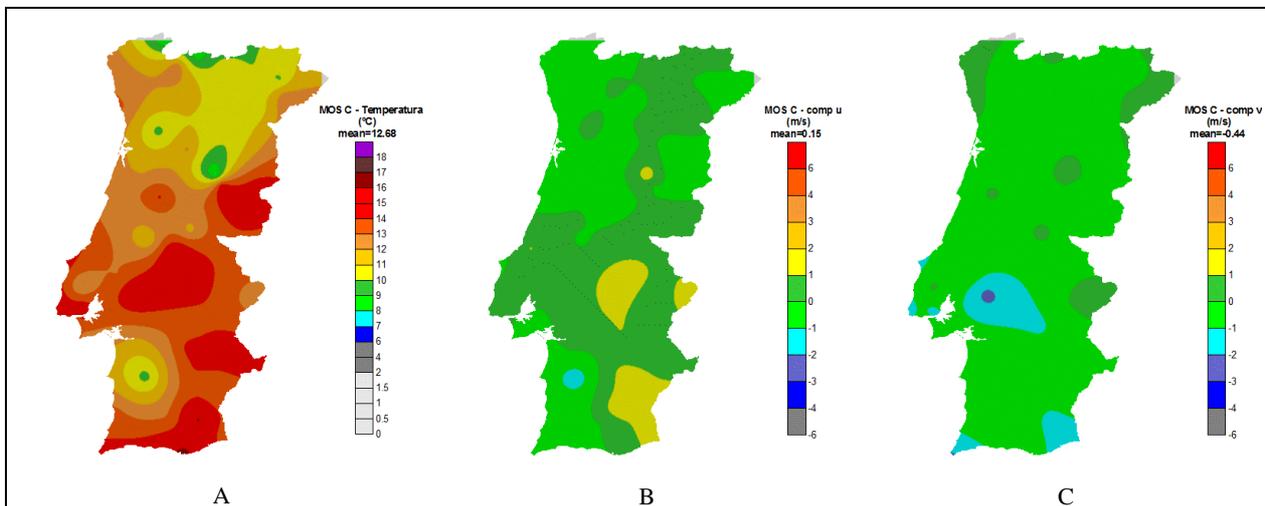


Figura 4.8 – Previsões de 24 horas do MOS- caso C com três predictores A) Temperatura; B) componente u do vento; C) componente v do vento.

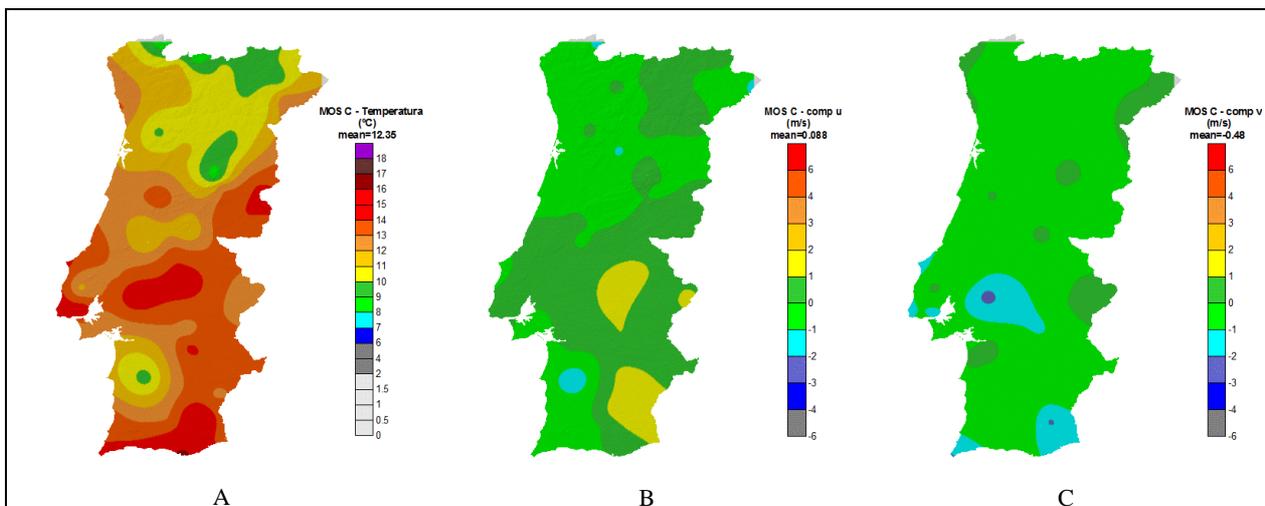


Figura 4.9 – Previsões de 24 horas do MOS – caso C com selecção de três predictores. A) Temperatura; B) componente u do vento; C) componente v do vento.

#### 4.1.3.1 Avaliação da qualidade das previsões do MOS (caso C)

Os resultados da avaliação estatística estão apresentados na tabelas 4.5 e 4.6. Os mapas das figuras 4.10 e 4.11 ilustram a medida estatística SC (score) para as três variáveis previstas pelo MOS – caso C. Os valores médios do score em validação cruzada não variaram muito devido à inclusão de um terceiro predictor em relação ao caso com dois predictores, quer pelo método sem selecção, quer pelo método com selecção. Desse modo o MOS linear fica praticamente estabilizado com a versão B (com selecção).

Tabela 4.5 – Parâmetros estatísticos de avaliação da qualidade da previsão do modelo MOS – caso C, sem selecção de preditores, para a temperatura e cada componente horizontal do vento.

<i>Parâmetros Estatísticos</i>	<i>Variáveis (sem selecção)</i>		
	Temperatura	Componente zonal do vento – u	Componente meridional do vento - v
BE	0.364	0.034	0.047
r	0.93	0.54	0.59
RMSE	2.29	1.75	1.61
MSE	5.23	3.06	2.60

Tabela 4.6 – Parâmetros estatísticos de avaliação da qualidade da previsão do modelo MOS – caso C, com selecção de três preditores, para a temperatura e cada componente horizontal do vento.

<i>Parâmetros Estatísticos</i>	<i>Variáveis (com selecção)</i>		
	Temperatura	Componente zonal do vento – u	Componente meridional do vento - v
BE	0,033	-0,03	0,01
r	0,94	0,58	0,59
RMSE	1,98	1,66	1,56
MSE	3,91	2,76	2,43

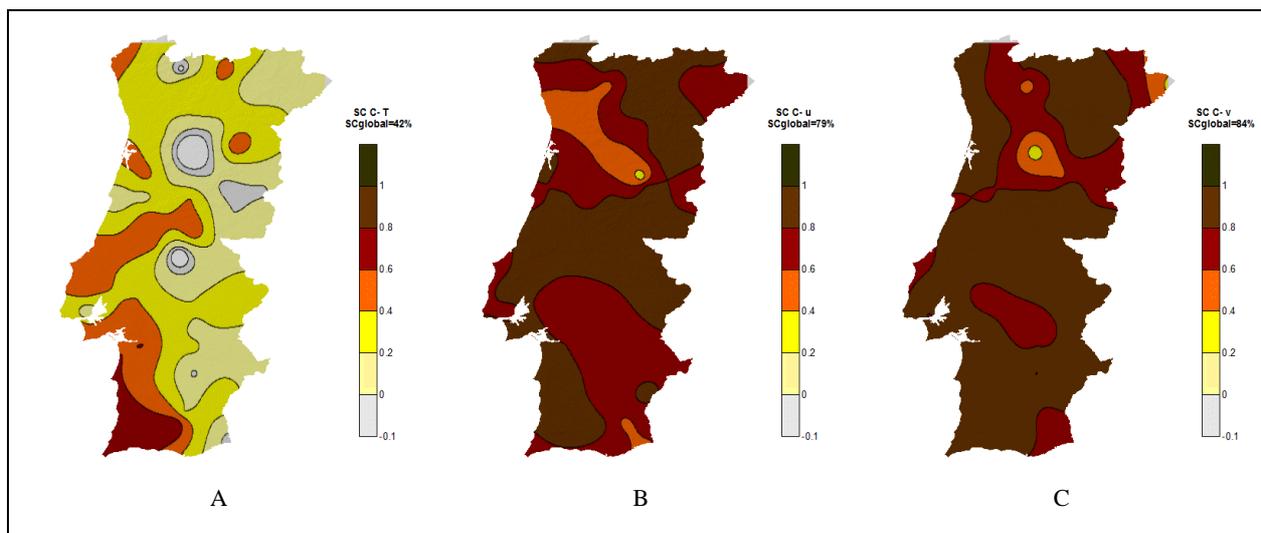


Figura 4.10 – Score (SC) entre as previsões do MOS- caso C, com três preditores, e as previsões MM5. A) Temperatura; B) componente u do vento; C) componente v do vento.

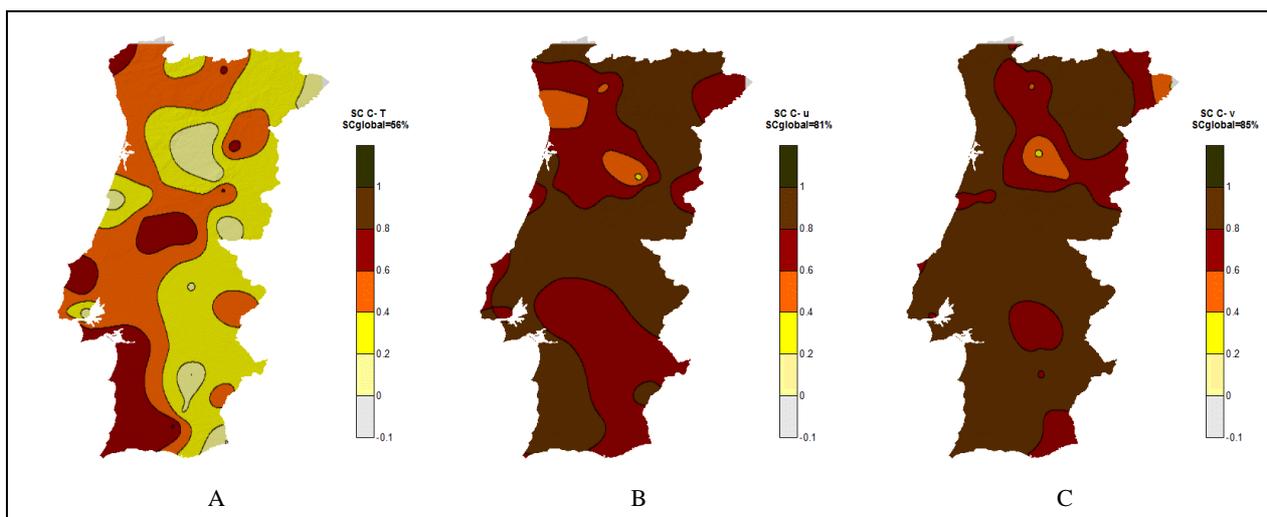


Figura 4.11 – Score (SC) entre as previsões do MOS – caso C, com selecção de três predictores, e as previsões MM5. A) Temperatura; B) componente u do vento; C) componente v do vento.

#### 4.1.4 Caso D

Para este caso, quatro predictores ( $P(t+24)$ ,  $O(t)$ ,  $O(t-12)$  e  $P(t+12)$ ) foram utilizados para a obtenção da equação do modelo estatístico. Os mapeamentos das previsões do MOS das médias anuais de cada varável estão apresentados na figura 4.12 e 4.13.

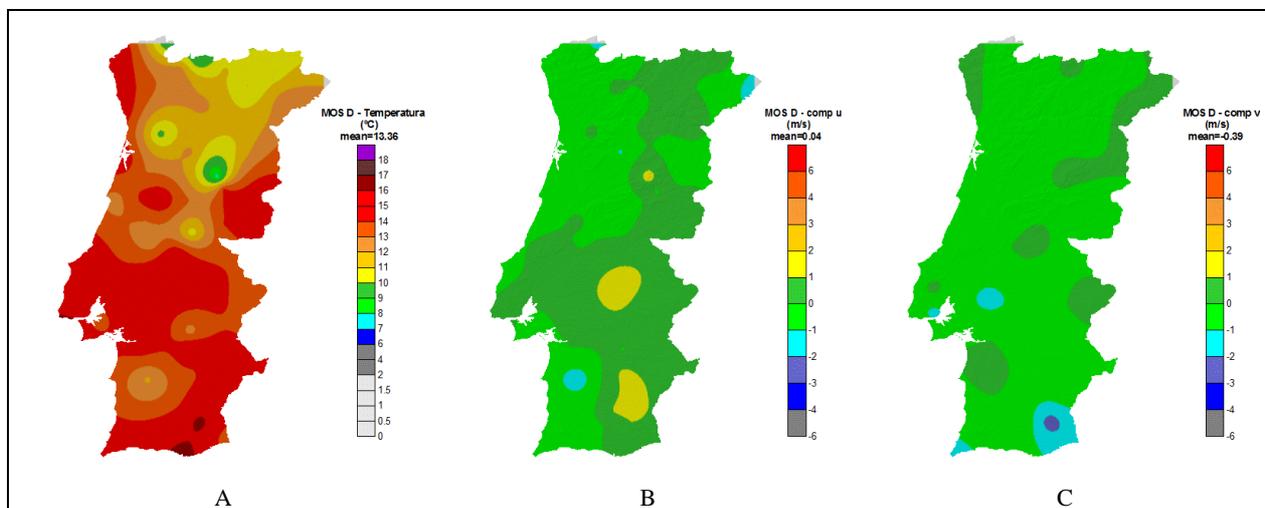


Figura 4.12 – Previsões de 24 horas do MOS- caso D com quatro predictores. A) Temperatura; B) componente u do vento; C) componente v do vento.

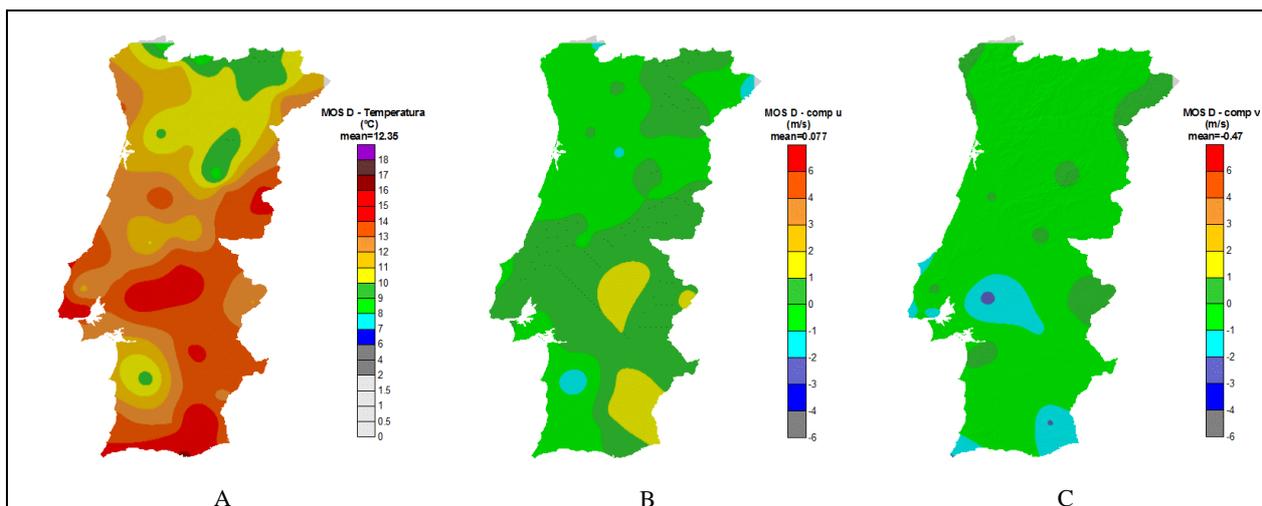


Figura 4.13 – Previsões de 24 horas do MOS- caso D com selecção de quatro predictores A) Temperatura; B) componente u do vento; C) componente v do vento.

#### 4.1.4.1 Avaliação da qualidade das previsões do MOS (caso D)

Os resultados da avaliação estatística estão apresentados na tabelas 4.7 e 4.8.

Os mapas das figuras 4.14 e 4.15 ilustram a medida estatística SC (score) para as quatro variáveis previstas pelo MOS – caso D. Os scores para este caso de 4 predictores, relativos ao método sem selecção baixaram bastante em relação ao caso C (3 predictores sem selecção). Por exemplo para a temperatura passou-se de um score de 0.41 para 0.24, o que mostra que regressão multilinear com 3 predictores deixa de ser eficaz. Os scores no método com selecção ficam praticamente inalterados em relação ao caso C, na verdade o quarto predictor tem já correlação praticamente nula com o resíduo da regressão linear (com selecção) obtida com 3 predictores.

Tabela 4.7 – Parâmetros estatísticos de avaliação da qualidade da previsão do modelo MOS – caso D, sem selecção de predictores, para a temperatura e cada componente horizontal do vento.

<i>Parâmetros Estatísticos</i>	<i>Variáveis (sem selecção)</i>		
	Temperatura	Componente zonal do vento – u	Componente meridional do vento - v
BE	1.04	-0.079	0.098
r	0.93	0.51	0.57
RMSE	2.68	1.85	1.78
MSE	6.78	3.43	3.17

Tabela 4.8 – Parâmetros estatísticos de avaliação da qualidade da previsão do modelo MOS – caso D, com selecção de quatro predictores, para a temperatura e cada componente horizontal do vento.

<i>Parâmetros Estatísticos</i>	<i>Variáveis (com selecção)</i>		
	Temperatura	Componente zonal do vento – u	Componente meridional do vento - v
BE	0,031	-0,042	0,017
r	0,94	0,59	0,61
RMSE	1,97	1,63	1,53
MSE	3,89	2,68	2,33

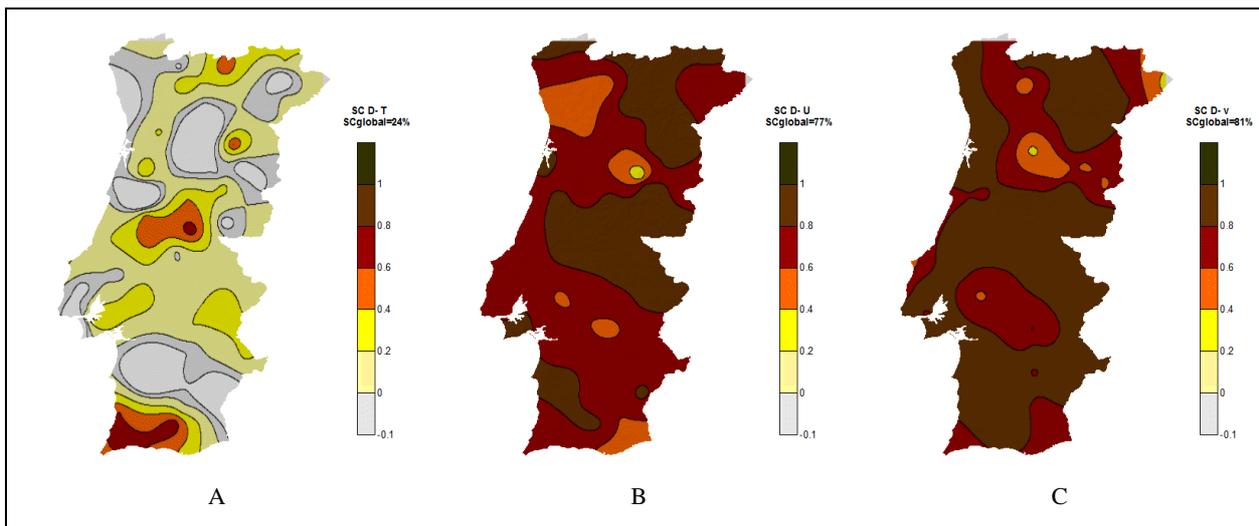


Figura 4.14 – Score (SC) entre as previsões do MOS- caso D, com quatro predictores, e as previsões MM5. A) Temperatura; B) componente u do vento; C) componente v do vento.

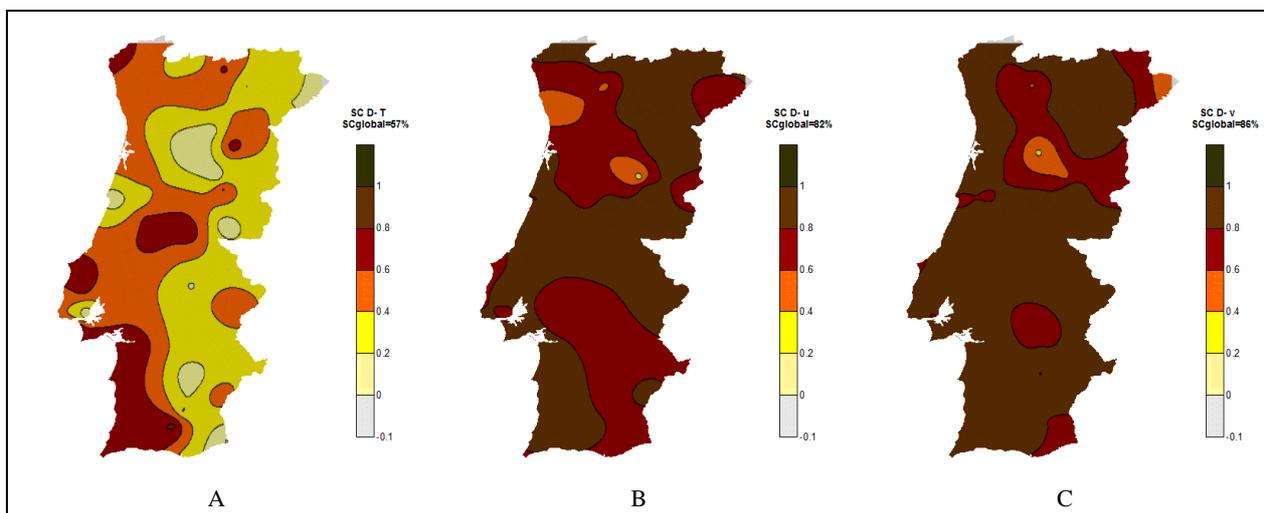


Figura 4.15 – Score (SC) entre as previsões do MOS – caso D, com selecção de quatro predictores, e as previsões MM5. A) Temperatura; B) componente u do vento; C) componente v do vento.

#### 4.1.5 Caso E

Neste caso foram utilizados todos os predictores pré-seleccionados ( $P(t+24)$ ,  $O(t)$ ,  $O(t-12)$ ,  $P(t+12)$  e  $P(t+36)$ ). Os resultados finais para as duas metodologias estão apresentados nas figuras 4.16 e 4.17, respectivamente.

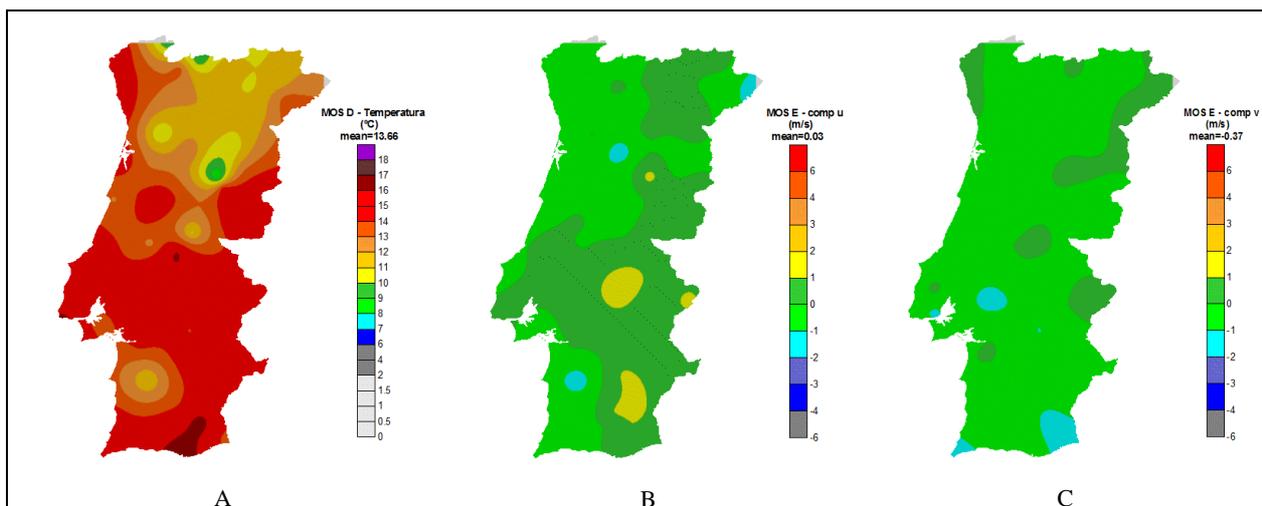


Figura 4.16 – Previsões de 24 horas do MOS- caso E com cinco preditores. A) Temperatura; B) componente u do vento; C) componente v do vento.

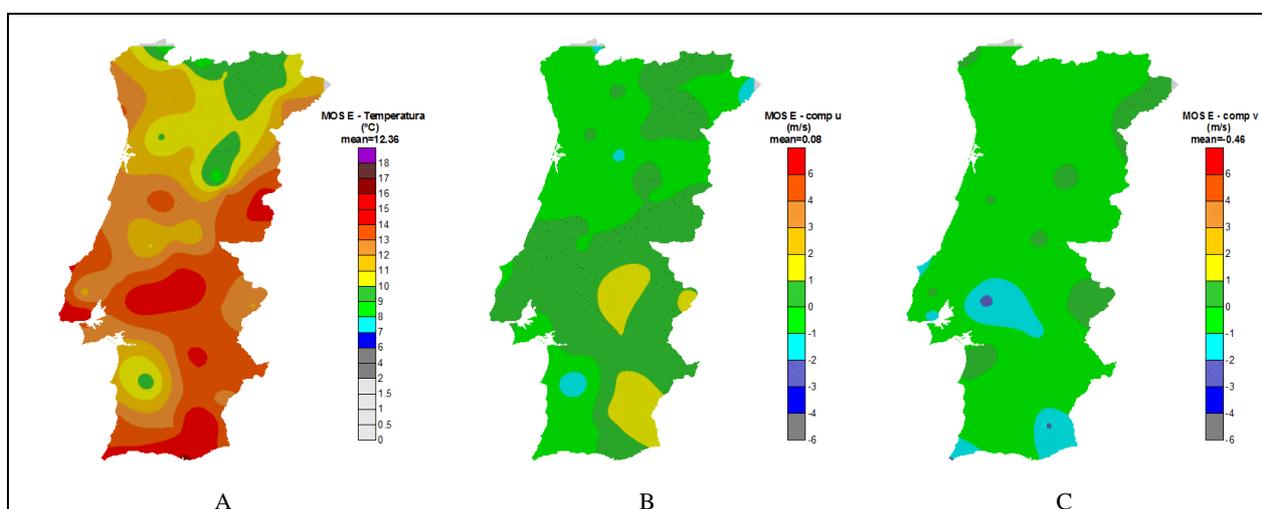


Figura 4.17 – Previsões de 24 horas do MOS- caso E com selecção de cinco preditores. A) Temperatura; B) componente u do vento; C) componente v do vento.

#### 4.1.5.1 Avaliação da qualidade das previsões do MOS (caso E)

Os resultados da avaliação estatística estão apresentados nas tabelas 4.9 e 4.10.

Tabela 4.9 – Parâmetros estatísticos de avaliação da qualidade da previsão do modelo MOS – caso E para a temperatura e cada componente horizontal do vento.

Parâmetros Estatísticos	Variáveis (sem selecção)		
	Temperatura	Componente zonal do vento – u	Componente meridional do vento – v
BE	1.34	-0.088	0.115
r	0.93	0.50	0.54
RMSE	2.74	1.89	1.80
MSE	7.51	3.57	3.25

Tabela 4.10 – Parâmetros estatísticos de avaliação da qualidade da previsão do modelo MOS – caso E, com selecção de cinco predictores, para a temperatura e cada componente horizontal do vento.

Parâmetros Estatísticos	Variáveis (com selecção)		
	Temperatura	Componente zonal do vento – u	Componente meridional do vento - v
BE	0,03	-0,03	0,026
r	0,94	0,62	0,64
RMSE	1,95	1,58	1,48
MSE	3,82	2,49	2,19

Os mapas das figuras 4.18 e 4.19 apresentam a medida estatística SC (score) para as três variáveis previstas pelo MOS – caso E. À semelhança do caso D, assiste-se à continuação da degradação do MOS com o método sem selecção e à inexistência de melhorias em relação ao método com selecção.

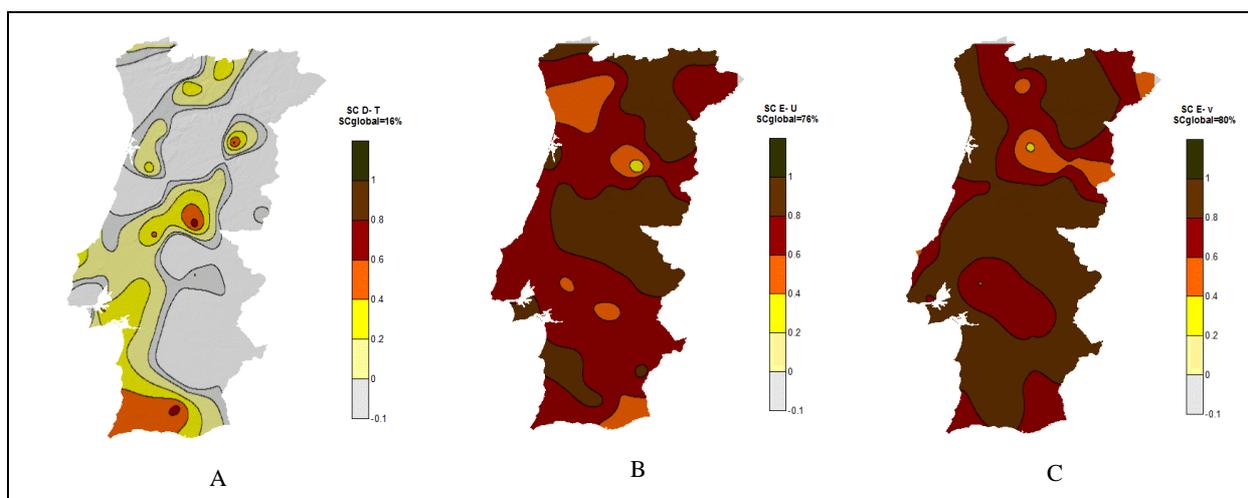


Figura 4.18 – Score (SC) entre as previsões do MOS- caso E, com cinco predictores, e as previsões MM5. A) Temperatura; B) componente u do vento; C) componente v do vento.

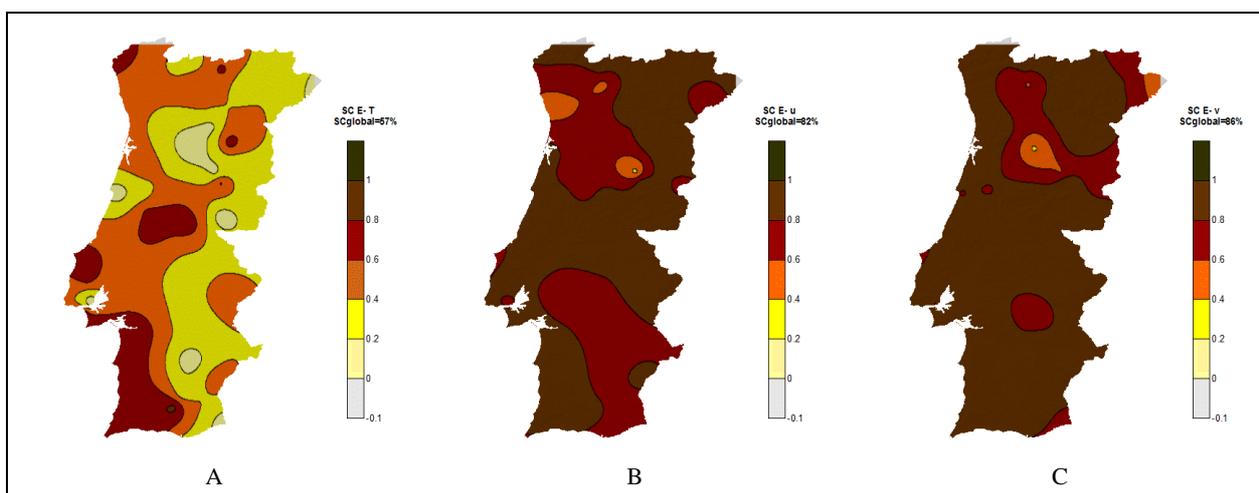


Figura 4.19 – Score (SC) entre as previsões do MOS – caso E, com selecção de cinco predictores, e as previsões MM5. A) Temperatura; B) componente u do vento; C) componente v do vento.

## 4.2 Discussão dos resultados do modelo

A construção das equações de regressão dos modelos estatísticos são metodologias sensíveis à selecção da ordem dos seus predictores. Na tabela 4.11 está uma síntese dos resultados apresentados na secção anterior.

Tabela 4.10 – Resumo dos parâmetros estatísticos de avaliação da qualidade da previsão dos modelos MOS e MM5, para a temperatura e cada componente horizontal do vento.

Casos	selecção	Variáveis	Parâmetros Estatísticos				
			BE	r	RMSE	MSE	SC [%]
MM5	----	T	1.29	0.92	2.99	8.95	----
		u	-0.25	0.49	3.83	14.66	
		v	-0.67	0.51	4.07	16.55	
A	não	T	<b>0.04</b>	<b>0.92</b>	<b>2.30</b>	<b>5.31</b>	<b>41</b>
		u	<b>-0.02</b>	<b>0.50</b>	<b>1.77</b>	<b>3.14</b>	<b>79</b>
		v	<b>0.02</b>	<b>0.53</b>	<b>1.69</b>	<b>2.87</b>	<b>83</b>
B	sim	T	0.04	<b>0.94</b>	<b>1.98</b>	<b>3.92</b>	<b>56</b>
		u	-0.03	<b>0.58</b>	<b>1.66</b>	<b>2.77</b>	<b>81</b>
		v	<b>0.007</b>	<b>0.59</b>	<b>1.56</b>	<b>2.44</b>	<b>85</b>
	não	T	0.308	<b>0.93</b>	<b>2.29</b>	<b>5.25</b>	<b>41</b>
		u	<b>0.018</b>	<b>0.54</b>	1.75	<b>3.05</b>	79
		v	0.046	<b>0.58</b>	<b>1.63</b>	<b>2.67</b>	<b>84</b>
C	sim	T	<b>0.033</b>	0.94	1.98	<b>3.91</b>	56
		u	-0.03	0.58	1.66	<b>2.76</b>	81
		v	0.01	0.59	1.56	<b>2.43</b>	85
	não	T	0.364	0.93	2.29	5.23	42
		u	0.034	0.54	1.75	3.06	79
		v	0.047	<b>0.59</b>	<b>1.61</b>	<b>2.60</b>	<b>84</b>
D	sim	T	<b>0.031</b>	0.94	1.97	<b>3.89</b>	<b>57</b>
		u	-0.042	<b>0.59</b>	<b>1.64</b>	<b>2.68</b>	<b>82</b>
		v	0.02	<b>0.61</b>	<b>1.53</b>	<b>2.33</b>	<b>86</b>
	não	T	1.04	0.93	2.60	6.78	24
		u	-0.079	0.51	1.85	3.43	77
		v	0.098	0.57	1.78	3.17	81
E	sim	T	0.03	<b>0.94</b>	<b>1.95</b>	<b>3.82</b>	<b>57</b>
		u	<b>-0.03</b>	<b>0.62</b>	<b>1.58</b>	<b>2.49</b>	<b>82</b>
		v	0.03	<b>0.64</b>	<b>1.48</b>	<b>2.19</b>	<b>86</b>
	não	T	1.34	0.93	2.74	7.51	16
		u	-0.09	0.50	1.89	3.57	76
		v	0.115	0.54	1.80	3.25	80

Em primeira análise, os melhores resultados obtidos são verificados no caso E com selecção de predictores. Estes resultados indicam uma melhoria na previsão das variáveis na ordem de: 60% na previsão da temperatura, e 85% na previsão das componentes horizontais do vento. Em geral, para todos os casos de estudo, a selecção de predictores, através da técnica da regressão forward stepwise apresenta sempre melhores resultados. Este método tem em consideração a minimização dos resíduos do modelo, seleccionando o predictor que melhor satisfaça esta condição.

Os resultados em geral, mostram uma diminuição do erro médio quadrático e do viés, e uma ligeira melhoria da correlacção, mais significativa nas componentes horizontais do vento no caso E (aumento de 10%).

Os resultados do modelo para os casos onde não se seleccionou a ordem dos predictores é menos satisfatória. Estes resultados mostram que a partir da adição do quarto e quinto predictor, o modelo deixa de apresentar melhoria nos resultados. Analisando a tabela 4.1, onde estão indicadas as correlacções de todos os predictores com os dados observados, verifica-se que apesar dos predictores P(t+12) e P(t+36) apresentarem valores altos de correlacção (sendo por vezes os melhores, principalmente para as variáveis do vento), a sua contribuição não é significamente positiva. Este facto, traduz a importância da ordem de selecção dos potenciais predictores a utilizar na construção das equações de regressão do modelo.

Uma segunda análise, permite-nos verificar que no caso A, onde o modelo de regressão apresenta um único predictor (a previsão das 24 horas –  $P(t+24h)$ ), a previsão melhora cerca de: 40% para a temperatura e 80% para as componentes do vento. Com a adição de um novo predictor na construção do modelo os valores das previsões melhoram cerca de 2 %, em relação ao caso A. Com a adição contínua e sucessiva dos predictores, a previsão do modelo sofre apenas ligeiras melhorias, inferiores ao verificado na introdução do segundo predictor – caso B. A introdução sucessiva de mais predictores permitem a refinação do modelo.

Será que a ligeira melhoria dos modelos de previsão com o aumento de número de predictores é significativa? Para melhor avaliar estas diferenças, e interpretar a importância da adição de predictores, vão ser apresentados alguns resultados de comparação entre os casos de estudo.

#### 4.2.1 Comparação dos resultados MOS

A figura 4.20 apresenta a diferença entre o erro médio quadrático (MSE) do caso A e do caso B. Esta avaliação vai permitir analisar as zonas onde a adição de predictores é mais significativa.

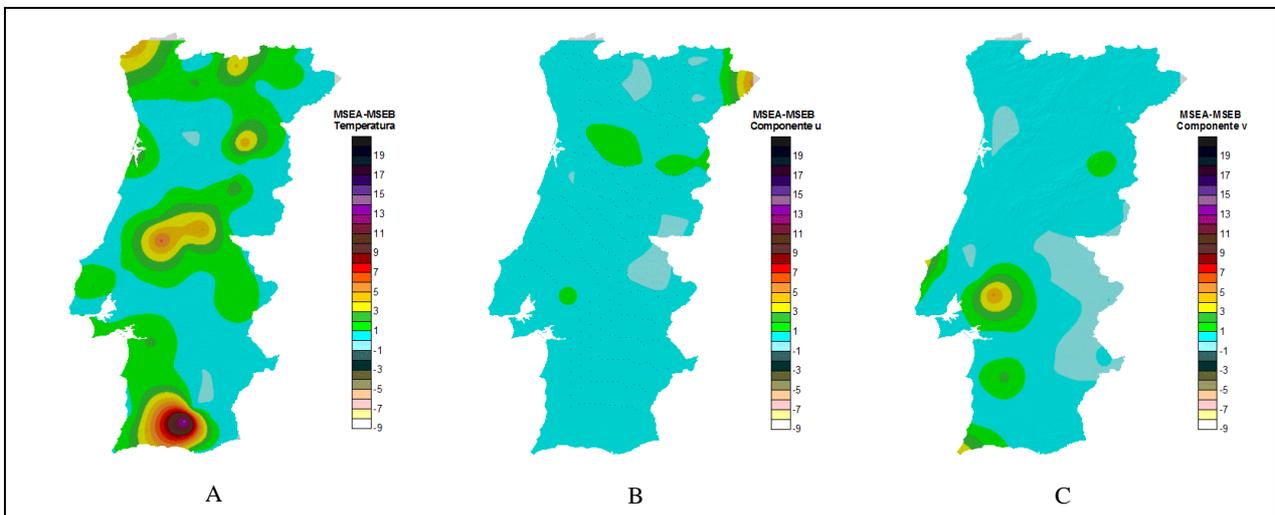


Figura 4.20 – Diferença entre o erro médio quadrático (MSE) dos casos de estudo A e B – com selecção de predictores. A) Temperatura; B) componente u do vento; C) componente v do vento.

Estes resultados comprovam a ligeira melhoria nas previsões anteriormente verificadas, resultantes da incrementação de predictores na equação de regressão do modelo. Os valores próximos de zero indicam que o modelo para o caso B não é significativo (indiferente), ou seja, não vai melhorar a previsão do modelo no caso A. No mapa da temperatura existem algumas estações onde esta diferença é significativa, valores da diferença superiores a zero, o que nos indica que para algumas estações a aplicação do modelo de regressão caso B é estatisticamente significativo ( exemplo: estação de Portimão). Para as componentes horizontais do vento, apenas poucas estações vão beneficiar com a adição de mais um predictor. Contudo, na figura 4.21 é apresentada a análise da incrementação de mais um predictor (MSEB-MSEC).

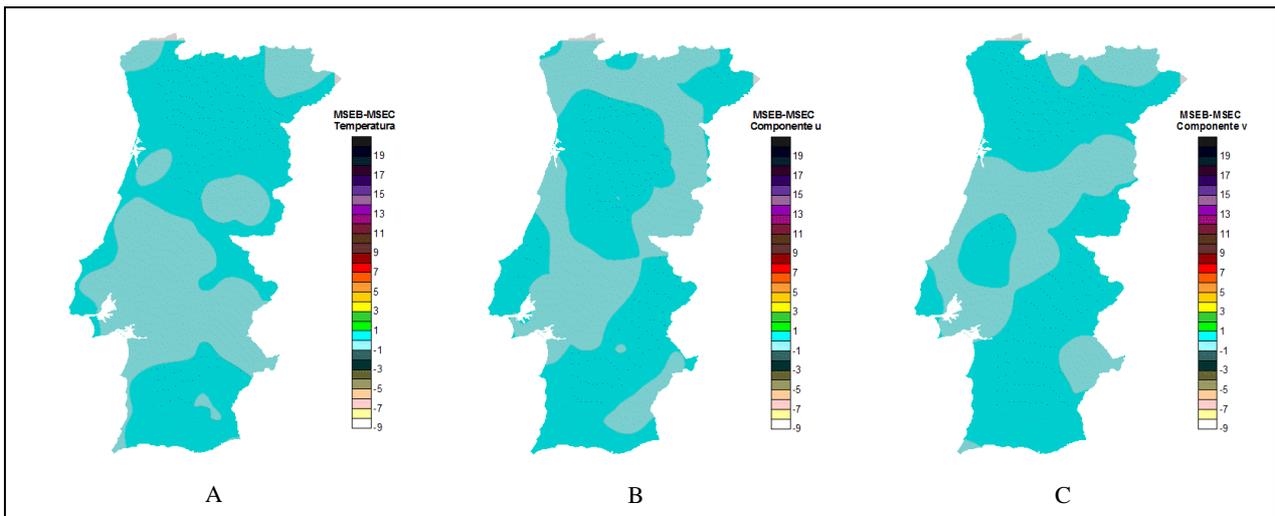


Figura 4.21 – Diferença entre o erro médio quadrático (MSE) dos casos de estudo B e C – com selecção de predictores. A) Temperatura; B) componente u do vento; C) componente v do vento.

Como esperado, o modelo caso C não apresenta melhorias significativas na redução do erro médio quadrático. Este resultado está de acordo com as conclusões apresentadas na secção anterior.

Contudo, os melhores resultados incluem este predictor. O modelo melhora 1% no caso D. Os melhores resultados da previsão estatística, da temperatura e das componentes horizontais do vento, são os apresentados no caso E. Apesar de último predictor ser também pouco significativo, existe uma diminuição do erro médio quadrático, uma melhoria da correlacção e, para a temperatura e componente zonal do vento, uma diminuição do viés.

Este estudo revela que, para cada variável e cada estação, existe um sub-conjunto de potenciais predictores, onde a metodologia aplicada na selecção da sua ordem é um método indispensável na construção de cada equação de regressão.

#### 4.2.2 Coeficientes da equação de regressão

As equações do modelo têm como base os coeficientes de regressão. Nas figuras 4.22 e 4.23 estão apresentados os valores de  $b_0$  e  $b_1$  para o primeiro caso de estudo (Caso A). O coeficiente  $b_0$  está relacionado com o viés BE e a média das observações  $\overline{x_{obs}}$  através da equação:

$$b_0 = -b_1(BE) + (1 - b_1)\overline{x_{obs}} \quad (4.1)$$

Note-se igualmente que o coeficiente  $b_1$  é o mesmo que seria obtido ao centrar a série do predictor, isto é ao subtrair-lhe a média, obtendo a série das anomalias. Desse modo, o coeficiente multiplicativo  $b_1$  é invariante a uma mudança puramente aditiva do sistema de unidades (ex. utilização de graus Kelvin em vez de centígrados na temperatura).

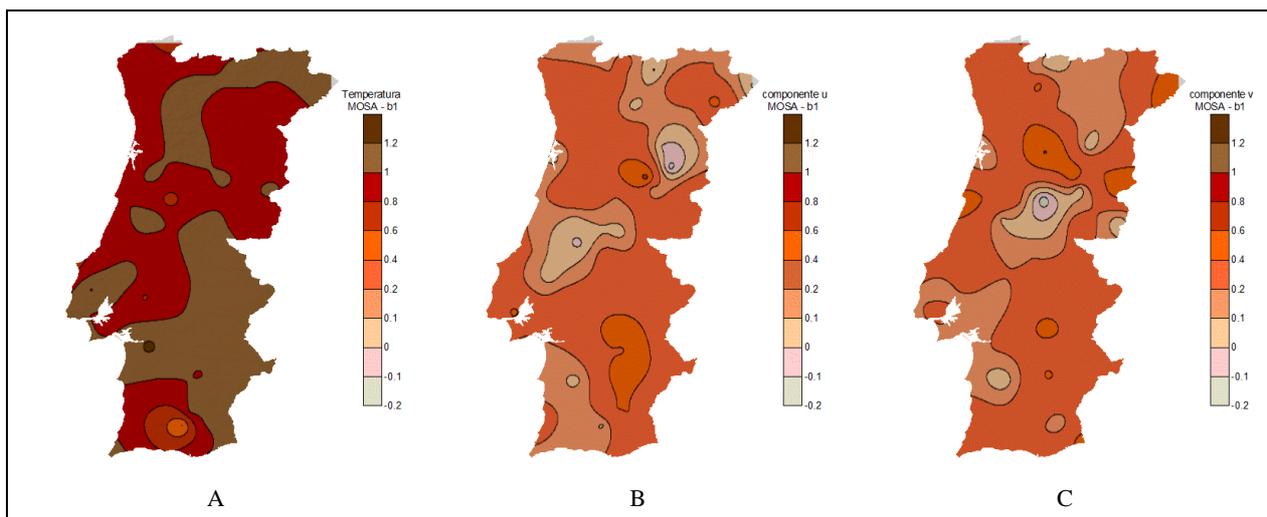


Figura 4.22 – Coeficiente de regressão  $b_1$ . A) Temperatura; B) componente u do vento; C) componente v do vento.

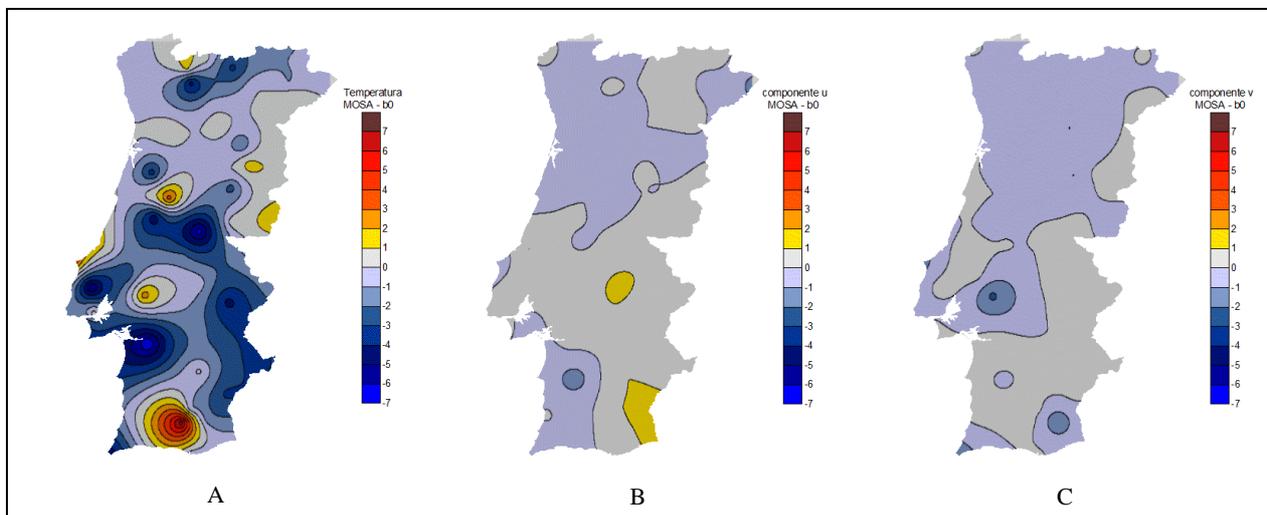


Figura 4.23 – Constante de regressão  $b_0$ . A) Temperatura; B) componente u do vento; C) componente v do vento.

Analisando os resultados verifica-se que o coeficiente  $b_1$  é proporcional à correlacção. Quanto maior a O coeficiente  $b_1$  exprime-se pela fórmula:

$$b_1 = \frac{cor(x_o, x_p)\sigma(x_o)}{\sigma(x_p)} \quad (4.2)$$

onde  $cor(x_o, x_p)$  é a correlacção entre o predictando  $x_o$  e o predictor  $x_p$  e  $\sigma(x_o)$  e  $\sigma(x_p)$  são os seus desvios padrão. Analisando os resultados verifica-se que o coeficiente  $b_1$  é proporcional à correlacção. Quanto maior a correlacção, da previsão do MM5 com as observações, maior o valor do coeficiente  $b_1$ .

Os coeficientes das equações de regressão pelo método com selecção, bem como a ordem dos predictores seleccionados, do modelo casoE, para a temperatura e componentes horizontais do vento, estão apresentados nas tabelas 4.11 a 4.13.

Verificámos anteriormente que o método com selecção praticamente estaciona em termos de score ao fim de escolhidos 3 predictores. Este facto é compatível com os resultados dos coeficientes  $b_4$  e  $b_5$ , relativos ao quarto e quinto predictores serem muito inferiores aos coeficientes  $b_1$ ,  $b_2$  e  $b_3$ .

Tabela 4.11 – Coeficientes de regressão do caso E, com selecção de 5 predictores, para a variável da temperatura.

Nº Estação	b0	np	b1	np	b2	Np	b3	np	b4	np	b5
531	0,7928028	5	0,1762292	1	0,2514222	4	0,4730408	3	0,1240265	2	-0,1196021
533	-3,66132	2	-0,0123351	4	0,3845835	1	0,7243338	3	0,1493616	5	-0,122306
535	0,0910649	2	-0,0232992	1	0,7814235	4	0,1950552	3	0,0209171	5	0,0104084
545	-0,1055791	2	-0,0006384	4	0,4580004	1	0,586934	3	0,0275647	5	-0,0947625
548	-0,0742786	2	-0,0932388	4	0,2681988	1	0,8219353	3	0,0434747	5	-0,0370634
554	-1,200228	2	0,1029377	1	0,2155486	4	0,3475769	3	0,3553662	5	-0,009585
558	-0,8956981	2	-0,0629423	1	0,7520561	3	0,2141066	4	0,1369984	5	-0,0290582
560	-0,4321701	2	-0,2569674	1	0,9462779	4	0,1370449	3	0,2364281	5	-0,0095649
562	0,191926	2	-0,0716235	1	0,884128	4	0,0584889	3	0,1207263	5	-0,0089937
567	-0,4538539	2	-0,191226	1	0,6597483	4	0,3136377	3	0,1931657	5	-0,0104739
568	-0,6126997	2	0,0632316	1	0,8030124	3	0,1108311	5	0,0805507	4	-0,0554084
570	0,3225559	2	-0,1772934	1	0,9856172	4	0,1429265	5	0,0335862	3	0,014839
575	-1,421612	2	0,0291607	1	0,4947751	4	0,256874	3	0,2123539	5	0,0105284
577	-0,9099014	2	-0,023182	4	0,3538138	1	0,6526798	3	0,0722792	5	-0,0796596
579	-0,5182192	2	-0,0972169	1	0,8379647	4	0,2319432	3	0,0687152	5	-0,0018334
604	-1,128476	5	-0,0234462	4	0,5450116	1	0,4957548	3	0,1959834	2	-0,2281923
606	0,5359835	2	-0,1599028	4	0,3973289	1	0,1740113	3	0,4514274	5	-0,0617093
611	0,0898912	2	-0,2308666	1	0,6239908	4	0,3284469	3	0,2514752	5	-0,020808
616	-2,751015	5	0,0287349	1	0,2761352	4	0,3648423	3	0,22841	2	0,0932803
619	0,9730072	2	-0,2017595	1	0,4784184	4	0,3599348	3	0,3484075	5	-0,0709928
632	-2,948929	2	-0,1896887	4	0,4508387	1	0,5276733	3	0,2787489	5	0,0311487
633	-0,8474485	2	-0,0317008	1	0,5581139	4	0,2887723	3	0,179872	5	-0,0180258
635	-0,8666345	2	-0,0171572	4	0,3687133	1	0,5808227	3	0,1700935	5	-0,0616114
637	-0,8415248	2	-0,0617158	1	0,8806336	3	0,188672	4	0,0493788	5	-0,006887
654	0,3147821	2	-0,142958	1	0,7528229	4	0,1641213	3	0,1528864	5	0,0578078
663	0,280875	2	-0,1598104	1	0,490865	4	0,4488783	3	0,1943118	5	-0,0209615
666	0,6892935	2	-0,3649186	1	0,8855704	4	0,3010989	3	0,1998451	5	-0,0191238
668	-0,4999695	5	0,1480691	4	0,4685546	1	0,2494121	2	-0,1010804	3	0,0559608
683	0,821781	2	-0,2358771	1	0,6990892	3	0,3913856	4	0,1968066	5	-0,0511852
686	1,304639	2	-0,2723694	1	0,8354614	4	0,3313745	3	0,1714823	5	-0,0838828
687	-0,5578436	2	-0,1844922	4	0,3129055	1	0,4687328	3	0,284098	5	0,046689
697	-2,850887	5	0,1095384	1	0,3293291	4	0,4040418	3	0,1119571	2	0,0783081
702	1,113375	2	-0,0598702	4	0,5621084	1	0,4532747	3	0,0095201	5	-0,0478711
705	-1,015215	2	0,1039884	4	0,4507715	1	0,403153	5	0,0249267	3	-0,0464781
713	-2,908501	2	-0,1416715	4	0,2791579	1	0,9358948	5	0,0997391	3	-0,077864
716	-0,2789731	2	-0,1465738	1	0,9074005	4	0,201897	3	0,0981713	5	-0,0276371
724	-3,406483	2	0,0096985	4	0,3330297	1	0,4552214	3	0,2242586	5	0,0084938
726	-4,335577	5	0,2451454	4	0,430059	1	0,1686362	2	0,119284	3	0,0508693
734	0,3850686	2	-0,0001782	4	0,2550122	1	0,6719296	3	0,0615363	5	-0,0591102
739	-2,101433	2	0,084473	4	0,3630461	1	0,6574993	3	-0,0290886	5	-0,0047173
744	-3,637485	2	-0,070448	4	0,4661779	1	0,5396396	3	0,161367	5	0,0133692
765	1,608014	2	0,0420645	4	0,2884847	1	0,5859585	3	-0,0505437	5	-0,0113869
766	-0,6519471	2	0,0420432	1	0,4624844	4	0,3636153	3	0,1730119	5	-0,0458949
767	-1,627142	2	0,0917285	4	0,2842791	1	0,4369988	5	0,0854245	3	0,0199614
776	-1,686823	2	-0,0544669	4	0,4822158	1	0,6577449	3	0,1073196	5	-0,1127498
783	-3,777315	2	-0,1453867	4	0,5127554	1	0,3908621	3	0,2641915	5	0,0677527
800	-1,117153	2	-0,1980916	4	0,4637654	1	0,5790228	3	0,2143159	5	-0,0292646
803	0,357147	2	-0,1634299	1	0,7587817	4	0,2398616	3	0,2713254	5	-0,0697498
806	1,16654	2	-0,2550707	1	0,6467388	4	0,2561534	3	0,2806816	5	-0,025119
812	-0,9980395	5	-0,1193518	4	0,4862609	1	0,3164958	2	0,2230248	3	0,0310679
824	-2,374336	2	0,0042527	1	0,3038107	4	0,3846802	3	0,2354993	5	0,0723965
835	-0,4423049	2	0,0479196	1	0,5533184	4	0,1693923	3	0,2684549	5	-0,0396678
837	-1,767776	2	0,0687741	1	0,5504942	4	0,3455226	3	0,1660588	5	-0,0781082
848	-2,876633	2	0,1966504	4	0,4561127	1	0,20377	3	0,257594	5	-0,0611638
850	-1,266188	2	0,1076514	4	0,3217899	3	0,1313078	1	0,5912205	5	-0,0995728
863	-2,993977	2	-0,0548241	1	0,4749793	4	0,2668977	3	0,3344967	5	0,0514732
864	-1,991753	2	-0,0386798	4	0,4582427	1	0,3893113	3	0,2817271	5	-0,0510721
865	-1,344357	2	0,0232336	1	0,7122019	4	0,1285635	3	0,1928904	5	-0,006422
867	-1,581524	4	0,4376301	3	0,3814289	1	0,1310391	2	0,1262492	5	-0,0655855
868	-0,5519456	2	-0,1510574	1	0,8062853	3	0,3160389	4	0,0801348	5	-0,0354688
878	-1,49053	5	-0,2212648	4	0,4731693	3	0,3736198	1	0,1721849	2	0,1763223

Tabela 4.12 – Coeficientes de regressão do caso E, com selecção de 5 predictores, para a variável da componente zonal do vento ( $u$ ).

Nº Estação	b0	np	b1	np	b2	np	b3	np	b4	np	b5
531	-1,184861	4	-0,1433118	3	0,299892	2	0,1376735	1	0,1290966	5	0,2149788
533	-0,4254164	2	0,0198296	4	0,2626088	1	0,0903416	3	0,0603385	5	0,0060625
535	0,1497891	3	0,019429	1	0,1017253	2	0,3070714	4	0,0639499	5	-0,0419032
545	-0,912408	3	0,0391919	4	0,2799064	2	0,1041878	1	0,0723071	5	0,0342371
548	-0,1893886	2	-0,0022819	1	0,2345736	5	0,0596102	3	0,0752579	4	0,0335234
554	0,3963452	5	0,0438307	1	0,3020011	4	0,2513106	3	-0,0213064	2	-0,0952641
558	0,6520802	2	0,0148096	1	0,3320442	4	0,3151183	3	0,1042993	5	-0,056256
560	-1,189237	4	0,221306	3	0,3582371	2	0,2069882	5	0,063481	1	-0,1036361
562	1,01057	2	0,1174024	1	0,3578969	4	0,112483	3	0,088929	5	-0,0869193
567	-0,2798101	2	0,016584	1	0,0228428	4	0,1281489	3	0,0309708	5	0,0055064
568	-0,1486743	2	0,1053824	1	0,3413926	4	0,2323157	3	0,1422098	5	-0,1335731
570	0,9163399	2	-0,0096909	1	0,2391851	3	0,1235983	4	0,0978184	5	0,0238264
575	0,5469115	2	0,0075631	1	0,1866452	4	0,0897497	3	0,0113882	5	-0,0666896
577	-0,0626391	2	0,0406545	1	0,1577589	4	0,3414325	3	0,0449077	5	0,0306559
579	0,3810627	2	-0,0494447	1	0,3117777	3	0,1632216	4	0,2099429	5	-0,0777816
604	-0,162047	4	-0,083184	3	0,1582166	1	0,0751295	5	0,0397945	2	-0,0077977
606	-1,158996	4	0,1154525	3	0,2195256	2	0,1328539	1	-0,0193631	5	-0,1274161
611	-0,4786475	2	0,0432071	3	0,194017	1	0,0439295	4	0,1282846	5	-0,0344842
616	0,2122873	5	0,0847303	4	0,0165119	1	-0,0194024	3	-0,013255	2	0,0197032
619	-0,5908129	4	0,2923091	3	0,1915636	1	0,0375178	5	-0,0117965	2	-0,0431293
632	0,1731049	2	0,0705215	1	0,0886105	4	0,1522895	3	0,2923878	5	0,0495669
633	0,0893998	2	0,016875	3	0,0928799	4	0,1965629	1	0,1587435	5	-0,0310045
635	-0,4869652	2	-0,0254962	3	0,2136766	4	0,458002	1	0,0287401	5	-0,0763669
637	-1,422063	4	0,3254645	3	0,4798777	2	0,1841052	5	-0,0914999	1	-0,0042474
654	-0,42222	2	0,0570623	3	0,1852461	1	0,3054041	4	0,0975639	5	-0,030196
663	0,2490181	2	-0,0458184	1	0,1898457	4	0,1810228	3	-0,0370905	5	0,041948
666	-0,0361294	2	0,0300844	1	0,1490399	4	0,3630149	3	0,0461253	5	-0,0377758
668	-0,2750355	1	-0,0587965	2	0,0826007	3	0,0558966	4	0,0022685	5	-0,0186435
683	-0,3430896	2	0,0917797	1	0,1476511	4	0,3327165	3	0,166037	5	-0,0787625
686	-0,4753627	5	0,0378326	4	0,3206714	3	0,1650634	2	0,1317077	1	-0,2040652
687	-0,5132706	4	0,0721511	3	0,1570882	2	0,088531	1	-0,0082961	5	-0,0184914
697	-0,0494907	2	-0,0056048	1	0,1652216	4	0,14176	5	0,0844775	3	-0,0474521
702	-0,4586169	1	0,0326894	2	0,2314918	5	0,149954	3	0,0423402	4	-0,0620364
705	-0,4509599	4	0,2037866	3	0,1278397	2	0,0034027	5	-0,0472948	1	0,0067952
713	-0,3260491	2	0,1358021	1	0,2160489	3	0,0022164	5	0,006675	4	-0,2223787
716	-0,2048663	4	0,131829	3	0,3069912	2	0,1441757	1	0,0337038	5	0,0336705
724	0,068971	5	0,0084018	4	0,7055736	3	0,0842484	2	0,0199513	1	-0,0561849
726	-0,126297	5	0,0200824	4	0,3549262	3	0,0086201	2	-0,0058274	1	-0,0151222
734	0,613049	2	-0,0874301	4	0,3313159	1	0,2264316	3	0,0090287	5	-0,0044005
739	0,0763145	5	0,0890377	4	0,2535945	3	0,0660033	2	0,0165134	1	-0,026901
744	0,3999747	2	-0,0585948	1	0,2255591	4	0,3036087	3	0,0443933	5	-0,0463927
765	-0,1360248	4	0,2084881	3	0,1704417	1	0,1122814	5	0,269387	2	0,0218092
766	0,2183436	2	-0,0520335	4	0,2879923	3	0,092691	1	0,2074104	5	0,0743478
767	-0,5435132	2	0,000439	3	0,057772	1	0,1743282	4	-0,049146	5	0,0971971
776	0,1527123	2	-0,0274009	4	0,4332467	1	0,1464044	3	0,0473755	5	0,0188229
783	0,1443044	2	-0,0245376	4	0,3278703	1	0,2677496	5	0,0232617	3	-0,0775445
800	-1,128038	5	-0,0282011	4	0,2583428	3	0,1294472	2	0,0276533	1	0,0153985
803	0,6645807	4	0,2211552	3	0,280994	1	0,1344482	2	0,0867402	5	-0,060362
806	0,3291703	2	-0,0140079	1	0,2348739	3	0,084211	4	0,2108338	5	-0,0577232
812	0,1973682	5	0,0197057	4	0,2987151	3	0,0622875	2	-0,0055988	1	-0,0051232
824	0,1794521	2	-0,0834033	1	0,2499261	3	0,0838311	5	0,059999	4	0,0663096
835	1,136992	2	-0,0327687	1	0,2560794	4	0,0633806	5	-0,0361876	3	-0,0322103
837	0,6369994	2	0,0088946	1	0,0863691	3	0,1649835	4	0,3178848	5	-0,0663609
848	0,2087312	2	0,0260803	3	0,1654492	1	0,1809093	4	0,1339295	5	-0,0678679
850	0,0204546	2	-0,0148678	1	0,243698	4	0,1563275	3	0,1960603	5	0,0129319
863	0,3267284	2	-0,0007747	1	0,2762567	4	0,249749	3	0,053629	5	0,0101558
864	0,9829361	2	0,1396338	1	0,1821991	4	0,1480676	3	0,0630318	5	-0,0333222
865	0,6816087	2	0,0361799	1	0,2983514	4	0,2482188	3	0,0359794	5	-0,029485
867	1,121607	2	0,0977887	4	0,2734849	1	0,1152776	3	0,1669176	5	0,0486552
868	0,5663279	5	-0,0106589	1	0,2376896	4	0,309309	3	0,1498829	2	-0,0271399
878	0,0944446	2	0,0174219	4	0,3359396	1	0,036273	5	0,0008935	3	0,0364823

Tabela 4.13 – Coeficientes de regressão do caso E, com selecção de 5 predictores, para a variável da componente meridional do vento (v).

Nº Estação	b0	np	b1	np	b2	np	b3	np	b4	np	b5
531	-0,6734533	5	-0,0049684	3	-0,0385048	4	0,5307243	1	0,2042408	2	0,0163843
533	-0,8833135	5	0,0024109	3	0,2371948	4	0,3996393	2	0,0852022	1	0,0344636
535	-0,4755698	4	0,2635178	3	0,2380448	1	0,1591748	2	0,1106792	5	-0,0627381
545	-0,2295198	4	-0,0399714	3	0,0675264	1	0,0886097	2	0,0875554	5	0,1056043
548	0,039825	2	0,0207812	1	0,1730039	4	0,3055552	3	0,1254864	5	-0,0396119
554	-0,2484771	2	0,0441496	1	0,2146419	3	0,2421846	4	0,0985447	5	-0,0030421
558	-0,2314544	2	0,0561786	1	0,2403432	4	0,1987189	3	0,2301136	5	0,0723111
560	-0,5412625	2	0,1049932	1	0,4308926	4	0,1925928	3	0,0269631	5	0,046484
562	0,0560894	2	0,0844689	1	0,2547576	3	0,126073	4	0,0605846	5	-0,0332449
567	-0,1886269	4	0,1154	3	0,1533011	1	0,0823499	2	0,0834505	5	-0,0818413
568	0,0363223	2	-0,0921872	3	0,2589727	1	0,2790804	4	0,2403692	5	-0,0574006
570	-0,1038184	2	0,1655022	1	0,1489127	3	0,1627063	4	0,0861769	5	0,0300969
575	0,1902126	2	0,0120599	1	0,2249344	4	0,05025	5	-0,0954699	3	-0,0003711
577	0,2977925	2	-0,0415517	3	0,0890102	1	0,2534369	4	0,2185819	5	-0,1168624
579	-0,4273367	4	0,2272933	3	0,3098753	1	0,0687809	2	0,0467872	5	-0,0294438
604	-0,1411195	1	0,0288228	4	0,1847519	3	0,1346071	2	0,1096799	5	-0,0396835
606	-0,7509317	2	0,1146401	3	0,1022436	1	0,2325424	4	0,1127065	5	-0,149495
611	-0,3453433	4	0,0600314	3	0,129	2	0,0843269	5	-0,056541	1	-0,0808431
616	-0,2410119	2	0,0752163	4	0,2738062	1	0,0604413	3	0,064748	5	-0,0161617
619	-0,5613621	4	0,06981539	3	0,126453	2	0,06227757	5	0,1133965	1	0,00425547
632	-0,2084213	2	-0,064094	1	0,3307153	4	0,16511	3	-0,0235407	5	-0,0305807
633	-0,2415416	2	0,0632244	1	0,1263354	4	0,0760215	5	-0,0057243	3	-0,0252085
635	-0,2270788	2	0,0866293	4	0,3238682	1	0,2006704	3	-0,013636	5	-0,0439296
637	-0,1062485	2	0,204281	3	0,2297492	1	0,2544341	4	0,0927873	5	-0,0443501
654	0,4534663	2	0,0941589	1	0,2508882	3	0,0234642	4	0,0450635	5	-0,0498486
663	0,0156334	2	0,0203329	1	0,10319	3	0,0667753	4	-0,0459495	5	-0,0535719
666	-0,0710686	4	0,3033412	3	0,2033708	1	0,1031519	2	0,0401195	5	-0,0468274
668	-0,3112751	1	-0,0012964	4	0,0801237	2	0,0616063	3	-0,0149157	5	-0,0346769
683	-0,5905707	2	0,0332262	1	0,3154134	3	0,0378396	4	0,003741	5	-0,1220682
686	-0,111024	5	-0,0303245	3	0,1901449	4	0,4316534	1	0,1757663	2	-0,032393
687	-0,5775534	5	-0,0722897	4	0,3321798	3	0,1156068	2	0,0191134	1	-0,1379321
697	-0,3268096	2	-0,0277871	1	0,0337014	3	0,0420068	4	-0,0193859	5	0,000985
702	-0,093371	2	-0,0118181	1	0,1711847	3	0,1011011	4	0,1691927	5	-0,0088222
705	-0,0281661	2	0,0471257	4	0,2700258	1	0,0479229	3	0,0171084	5	-0,0759041
713	-0,3883305	2	0,0400839	1	0,140182	4	0,0833305	3	0,0466415	5	-0,0056574
716	0,0808893	2	0,0369455	1	0,3173418	3	0,1244642	4	0,1019134	5	-0,0213314
724	-0,534732	4	0,43296	3	0,1021904	1	0,0095671	2	0,1744073	5	-0,1509991
726	0,0538832	2	0,0186899	1	0,0839597	4	0,1786508	3	0,0444565	5	-0,0344397
734	0,0283808	2	0,0541151	1	0,1919693	4	0,1854145	3	0,0793209	5	-0,0004697
739	0,3514226	2	0,0498474	1	0,2712868	4	0,110538	3	-0,0174086	5	-0,0393854
744	0,2564895	1	-0,0157876	2	0,0772606	4	0,1081492	3	0,080183	5	-0,0055141
765	-0,8631527	5	0,0421525	4	0,5306942	3	0,1731743	2	0,0445397	1	-0,0486859
766	-0,3145257	4	0,1797657	3	0,1754951	1	0,140991	2	0,1754748	5	-0,1662004
767	-0,0601369	1	0,0386344	2	0,1481694	3	0,0815312	4	0,2154366	5	-0,1333761
776	0,2002909	2	0,0713542	1	0,0920523	4	0,165174	3	-0,0017171	5	0,0014317
783	0,5240372	1	0,1001504	4	0,1376501	3	0,0480103	2	0,1290759	5	-0,0594931
800	-0,4627758	5	0,01801	3	0,3148192	4	0,3226756	2	0,0327871	1	0,0150689
803	0,1778437	2	0,1053855	3	0,2519062	1	0,2301148	4	0,1222286	5	-0,080553
806	-0,5634338	4	0,0777175	3	0,0827259	1	0,0077619	2	0,129086	5	-0,1165432
812	0,2861064	4	0,0276826	3	0,0614816	1	0,0214657	2	0,0061544	5	-0,0225946
824	-0,1981028	1	0,0213721	2	0,1897275	3	0,0648657	4	-0,0825485	5	-0,0304068
835	0,0018492	2	0,0818814	1	0,1450116	3	0,1429695	4	0,094267	5	0,0062761
837	0,4897448	4	0,1442152	3	0,1845235	1	0,0979041	2	0,0539941	5	-0,1311868
848	0,5133967	2	0,1416817	1	0,198263	3	0,0044309	4	-0,0206205	5	-0,037369
850	0,0475341	4	0,128805	3	0,2396955	1	0,0423588	2	0,1362864	5	-0,0880253
863	-0,0857298	2	0,0419038	1	0,1651662	3	0,1681508	4	0,0637451	5	-0,0173751
864	0,0803875	4	0,1293131	3	0,1518761	1	0,1456248	2	0,0422197	5	-0,0268511
865	0,0987334	2	-0,0043883	1	0,237863	3	0,1559244	4	0,2306549	5	-0,0763758
867	-1,408778	4	0,1432962	3	0,3516547	1	0,0426496	2	0,1938974	5	-0,1261603
868	-0,2376626	2	0,0290251	1	0,2901534	3	0,2284362	4	0,0703037	5	0,0086174
878	0,1940115	1	0,0456044	2	0,1605311	4	0,1091382	3	0,038256	5	-0,0763567

### 4.2.3 Comparação entre o MM5 e o MOS

Nesta secção, vão ser comparadas as medidas estatísticas do modelo de previsão numérico MM5 e o melhor modelo de previsão estatístico MOS obtido (caso E com selecção de predictores). Considerando os valores estatísticos extremos do modelo numérico, já apresentados na tabela 2.3, foram particularizados os estudos do MOS para as mesmas medidas estatísticas e estações. Nas tabelas 4.14 a 4.16 está apresentado um resumo das medidas estatísticas dos dois modelos para cada variável.

Tabela 2.3 – Comparação dos valores extremos do parâmetro RMSE, r e BE, obtidos pelo modelo numérico, com o modelo MOS, para a previsão da temperatura.

Nº	Nome da Estação	Temperatura (T)						SC [%]
		MM5			MOS			
		RMSE	r	BE	RMSE	r	BE	
579	LISBOA/GAGO COUTINHO	1.18	0.97	-0.04	1.14	0.97	0.03	8
726	ALCOBACA/ EST. FRUTIC. VI	5.97	0.72	4.25	3.45	0.82	0.01	67
668	AROUCA	6.27	0.73	4.55	3.36	0.82	0.13	71
637	MOGADOURO	1.44	0.98	-0.16	1.32	0.98	-0.06	17
686	PAMPILHOSA DA SERRA	1.88	0.96	-0.32	1.65	0.96	0.08	23
632	MIRANDELA	3.58	0.93	2.02	2.43	0.95	-0.003	54
687	COVILHA/AERO.	3.11	0.93	1.82	2.17	0.95	0.08	51
765	CABO RASO/ FAROL	2.20	0.91	1.13	1.57	0.92	0.03	49
800	SABUGAL/ MARTIM REI	2.82	0.95	1.83	1.79	0.97	0.055	60
878	PORTIMAO/ AERODROMO	6.08	0.71	3.90	2.33	0.92	0.06	85

Tabela 2.3 – Comparação dos valores extremos do parâmetro RMSE, r e BE, obtidos pelo modelo numérico, com o modelo MOS, para a previsão da componente zonal do vento.

Nº	Nome da Estação	Componente zonal do vento (u)						SC [%]
		MM5			MOS			
		RMSE	r	BE	RMSE	r	BE	
579	LISBOA/GAGO COUTINHO	2.34	0.75	-0.33	1.09	0.78	-0.044	78
726	ALCOBACA/ EST. FRUTIC. VI	5.32	-0.11	-0.94	0.69	0.27	-0.02	98
668	AROUCA	4.82	-0.04	-0.016	0.59	0.50	-0.06	98
637	MOGADOURO	8.15	-0.04	2.31	3.30	0.70	-0.03	84
686	PAMPILHOSA DA SERRA	7.78	-0.19	1.80	2.76	0.50	0.09	87
632	MIRANDELA	2.05	0.40	-0.25	1.41	0.40	0.02	53
687	COVILHA/AERO.	3.83	0.40	1.13	1.30	0.55	-0.13	88
765	CABO RASO/ FAROL	3.61	0.53	-0.55	2.13	0.68	-0.03	65
800	SABUGAL/ MARTIM REI	5.10	0.14	-1.14	1.66	0.41	-0.07	89
878	PORTIMAO/ AERODROMO	3.61	0.40	-0.85	0.88	0.47	-0.02	94

Tabela 2.3 – Comparação dos valores extremos do parâmetro RMSE, r e BE, obtidos pelo modelo numérico, com o modelo MOS, para a previsão da componente meridional do vento.

Nº	Nome da Estação	Componente meridional do vento (v)						SC [%]
		MM5			MOS			
		RMSE	r	BE	RMSE	r	BE	
579	LISBOA/GAGO COUTINHO	4.11	0.55	-1.33	1.41	0.79	0.01	88
726	ALCOBACA/ EST. FRUTIC. VI	5.28	0.61	-0.99	0.93	0.64	0.09	97
668	AROUCA	4.77	0.27	3.06	0.57	0.31	-0.04	99
637	MOGADOURO	2.73	0.69	0.69	2.06	0.73	0.05	43
686	PAMPILHOSA DA SERRA	4.79	0.48	-0.46	2.43	0.68	0.08	74
632	MIRANDELA	1.82	0.51	0.43	1.19	0.55	0.09	57
687	COVILHA/AERO.	8.28	-0.35	1.50	2.14	0.54	-0.03	93
765	CABO RASO/ FAROL	7.27	0.12	1.12	2.99	0.63	0.10	83
800	SABUGAL/ MARTIM REI	7.82	-0.06	0.66	1.95	0.69	0.14	94
878	PORTIMAO/ AERODROMO	3.41	0.59	-1.91	0.86	0.64	0.04	94

Após análise dos resultados das medidas estatísticas, para as estações onde o modelo numérico apresentava os melhores e piores resultados, verificou-se que a aplicação do modelo estatístico melhora, significativamente as piores previsões e ligeiramente as melhores. O melhor resultado observado é para a estação de Arouca, que apresenta um SC de 99% na previsão da componente meridional do vento (v). Para esta estação, o modelo de previsão MM5 sobrestima o valor observado na ordem dos 3.06 m/s, com um RMSE de 4.77m/s. O resultado do MOS, apesar da correlação continuar inferior a 50%, mostra que a previsão subestima 0.04 m/s o valor observado, produzindo um erro (RMSE) de 0.57 m/s. Na figura 4.24 estão representadas as observações e as respectivas previsões dos modelos, para a componente meridional do vento da estação de Arouca.

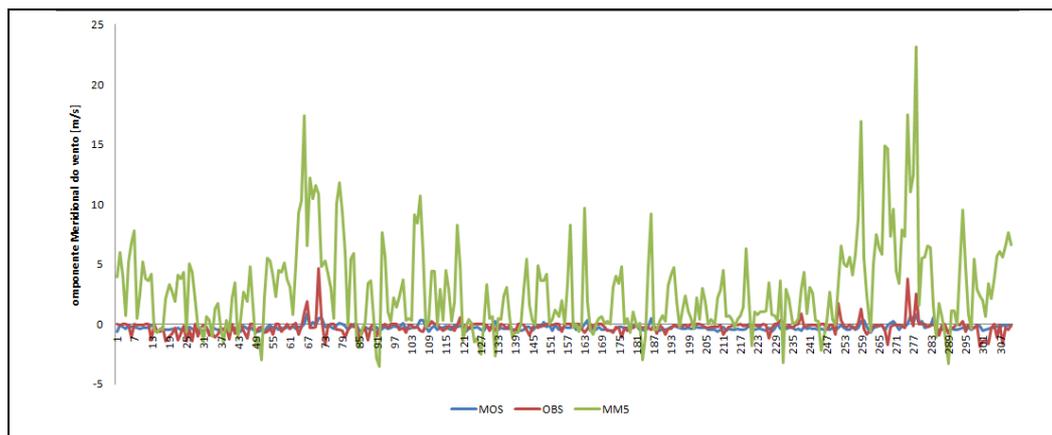


Figura 4.24 – Previsões MM5, MOS e observações da componente meridional do vento, da estação de Arouca.

Nas figuras seguintes estão apresentados os resultados do MOS, do MM5 e das observações para os casos em que o modelo MM5 apresenta os piores resultados. No caso da temperatura, na estação de Arouca, o MOS produziu uma previsão com um RMSE de 3.36 °C , melhorando a previsão do MM5 em 71% (figura 4.25).

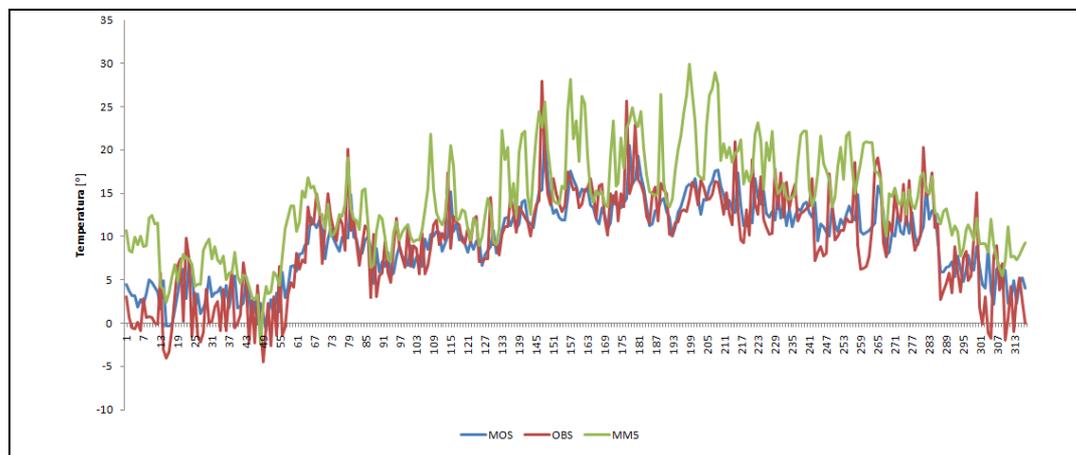


Figura 4.25 – Previsões MM5, MOS e observações da temperatura, da estação de Arouca.

O pior resultado da previsão MM5, no caso da componente zonal do vento, foi observado na estação do Mogadouro, com um RMSE de 8.15 m/s. A aplicação do MOS melhorou em 84% a previsão, diminuindo o RMSE para 3.30 m/s. Estes resultados, em confronto com as observações na estação, estão apresentados na figura 4.26. Na componente meridional do vento, o pior resultado registado foi na estação da Covilhã. O Mos melhorou cerca de 93% a previsão, diminuindo o RMSE de 8.28 m /s para 2.14m/s (figura 4.27).

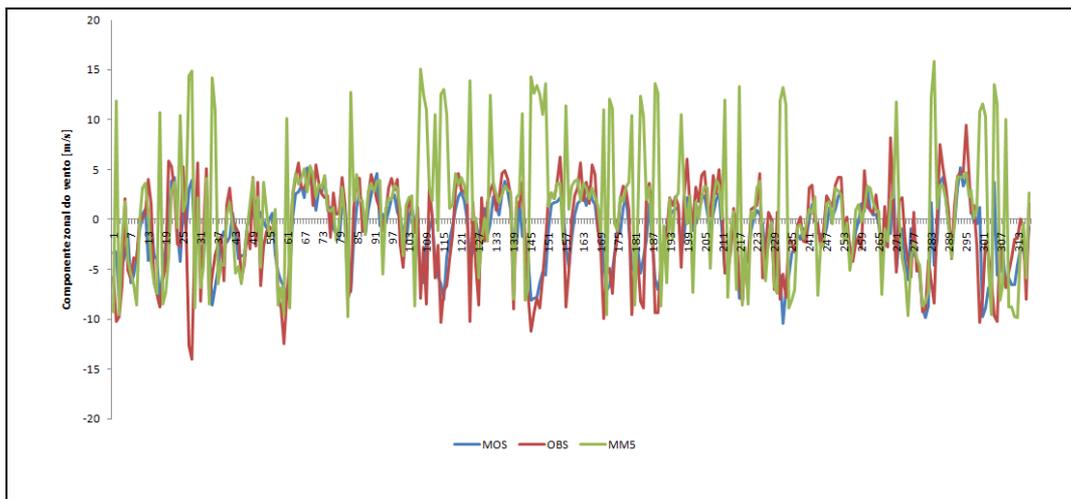


Figura 4.26 – Previsões MM5, MOS e observações da componente zonal do vento, da estação de Mogadouro.

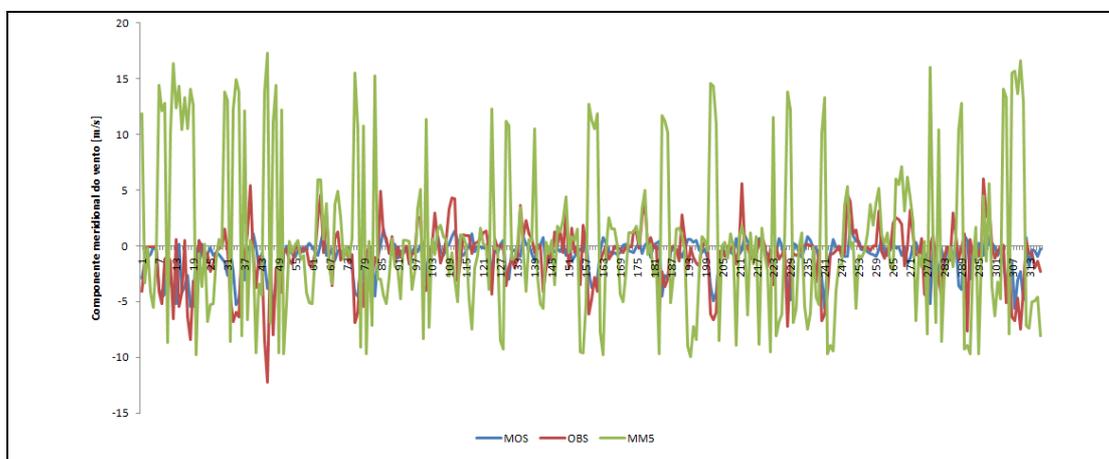


Figura 4.27 – Previsões MM5, MOS e observações da componente meridional do vento, da estação de Covilhã.

## 5. Breve apresentação e discussão dos resultados do MOS para diferentes prazos de previsão

Para os prazos de previsão de 48 e 72 horas, foram desenvolvidos modelos estatísticos MOS para os dois primeiros casos de estudo (caso A e B), com selecção da ordem dos predictores para o caso B. Na tabela 5.1, estão apresentadas as medidas estatísticas para as previsões de ambos os modelos de previsão.

Tabela 5.1 – Resumo dos parâmetros estatísticos de avaliação da qualidade da previsão dos modelos MOS e MM5, para a temperatura e cada componente horizontal do vento.

<i>Casos</i>	Prazo	Variáveis	<i>Parâmetros Estatísticos</i>				
			<b>BE</b>	<b>r</b>	<b>RMSE</b>	<b>MSE</b>	<b>SC [%]</b>
MM5	48	T	3.73	0.90	5.04	25.45	----
		u	-0.33	0.41	3.98	15.82	
		v	0.50	0.49	3.54	12.53	
MM5	72	T	3.65	0.87	5.00	25.00	----
		u	-0.08	0.40	3.77	14.18	
		v	0.37	0.46	3.62	13.13	
MOSA	24	T	0.04	0.92	2.30	5.31	41
		u	-0.02	0.50	1.77	3.14	79
		v	0.02	0.53	1.69	2.87	83
MOSA	48	T	0.14	0.90	2.70	7.31	71
		u	-0.06	0.41	1.88	3.54	78
		v	0.003	0.49	1.75	3.05	76
MOSA	72	T	0.20	0.89	2.83	7.99	68
		u	-0.04	0.40	1.90	3.60	75
		v	0.019	0.46	1.78	3.19	76
MOSB	24	T	0.04	0.92	2.30	5.31	41
		u	-0.02	0.50	1.77	3.14	79
		v	0.02	0.53	1.69	2.87	83
MOSB	48	T	0.64	0.90	2.80	7.82	69
		u	-0.02	0.37	1.98	3.91	75
		v	0.04	0.46	1.50	2.24	82
MOSB	72	T	0.47	0.89	2.83	8.01	68
		u	-0.03	0.39	1.91	3.65	74
		v	0.04	0.46	1.79	3.22	75
MOSB*	24	T	0.308	0.93	2.29	5.25	41
		u	0.018	0.54	1.75	3.05	79
		v	0.046	0.58	1.63	2.67	84
MOSB*	48	T	0.20	0.91	2.53	6.42	75
		u	-0.07	0.45	1.86	3.45	78
		v	0.007	0.51	1.72	2.96	76
MOSB*	72	T	0.30	0.90	2.72	7.38	70
		u	-0.03	0.41	1.89	3.58	75
		v	0.021	0.47	1.77	3.13	76

\*Com selecção da ordem dos predictores

Em geral, os resultados mostram que a previsão do MOS piora com o aumento do prazo de previsão, devido aos dados de input associados às previsões do modelo MM5. De acordo, com o verificado nos resultados anteriormente apresentados, a selecção estatística da ordem dos predictores melhora a previsão do modelo, assim como a adição de predictores. Dos casos estudados, os melhores resultados mostram uma melhoria na previsão a 48 e 72 horas de cerca de 70% para todas as variáveis, com erros médios quadráticos de 7°C e 3 m/s, para a temperatura e componentes horizontais do vento, respectivamente.

Como resumo, são apresentados, nas figuras 5.1 a 5.3, os mapas da evolução do prazo de previsão com o modelo MM5 e com a aplicação do MOS (caso B com selecção de predictores).

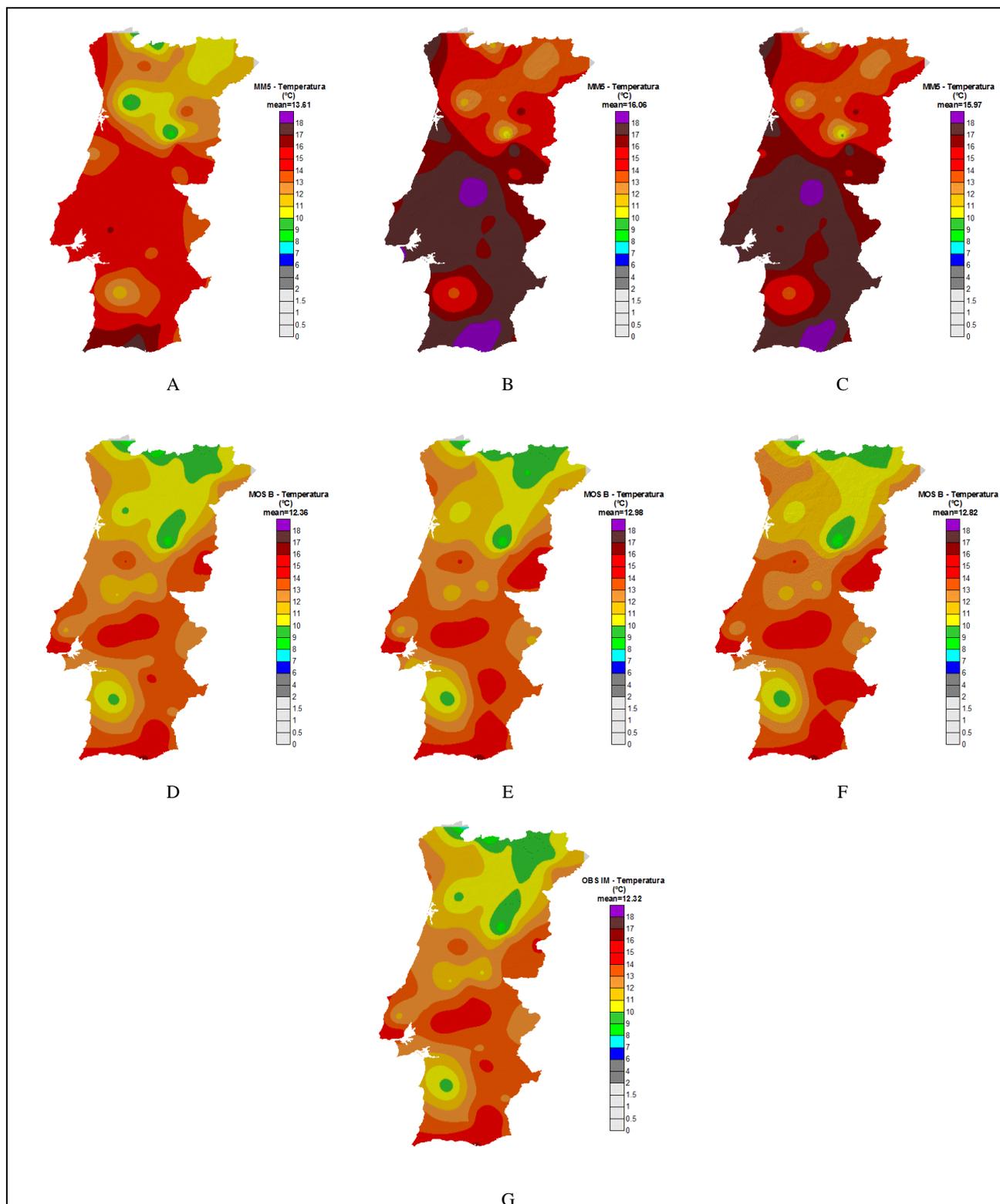


Figura 5.1 – Resultados da previsão da temperatura nos diferentes prazos, para o modelo MM5 e modelo MOS. A) MM5 - 24 horas; B) MM5 – 48 horas ; C) MM5 – 72 horas; D) MOS – 24 horas; E)MOS – 48 horas; F)MOS- 72 horas; G) Observação.

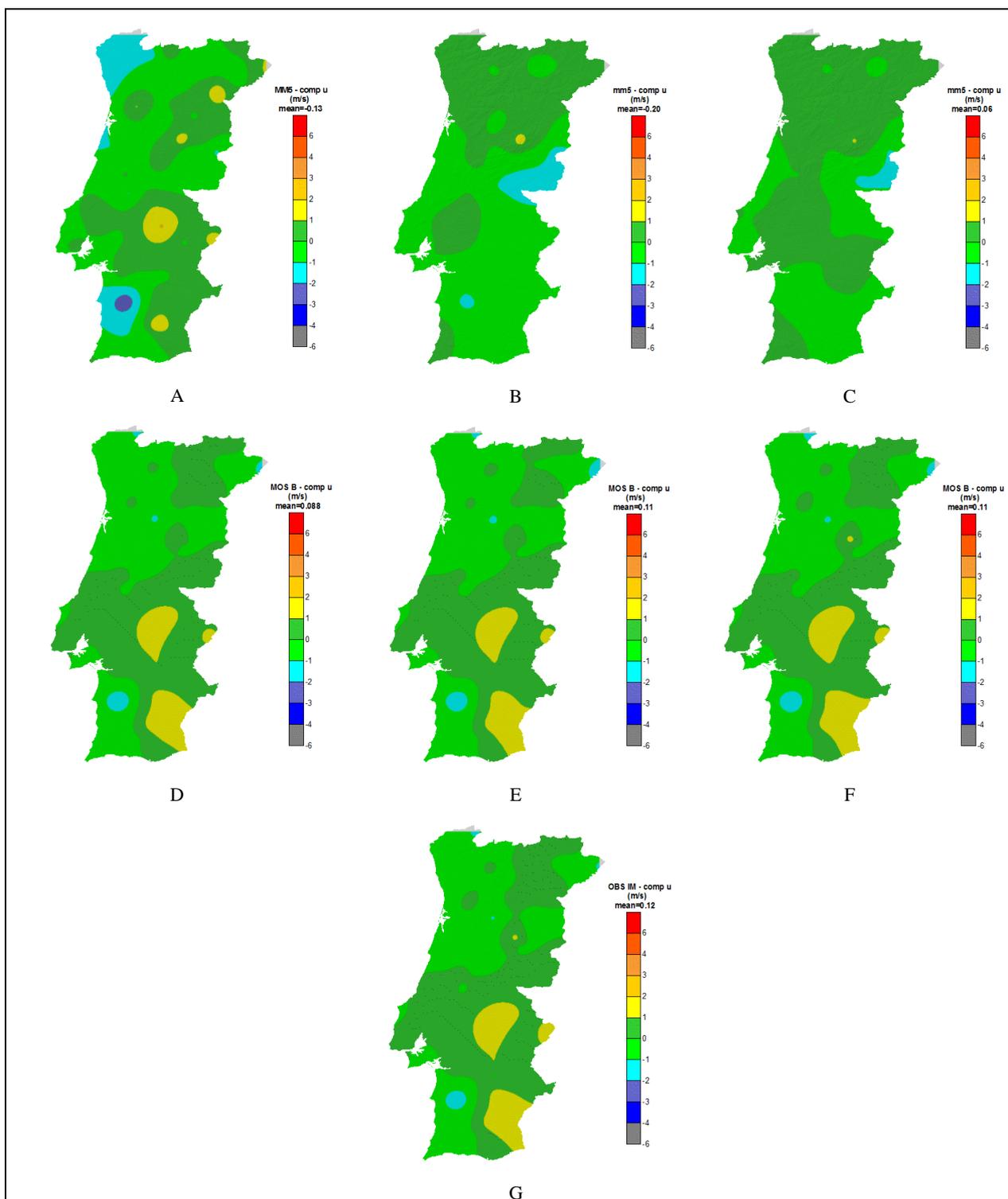


Figura 5.2 – Resultados da previsão da componente zonal do vento nos diferentes prazos, para o modelo MM5 e modelo MOS. A) MM5 - 24 horas; B) MM5 – 48 horas ; C) MM5 – 72 horas; D) MOS – 24 horas; E)MOS – 48 horas; F)MOS- 72 horas; G) Observação.

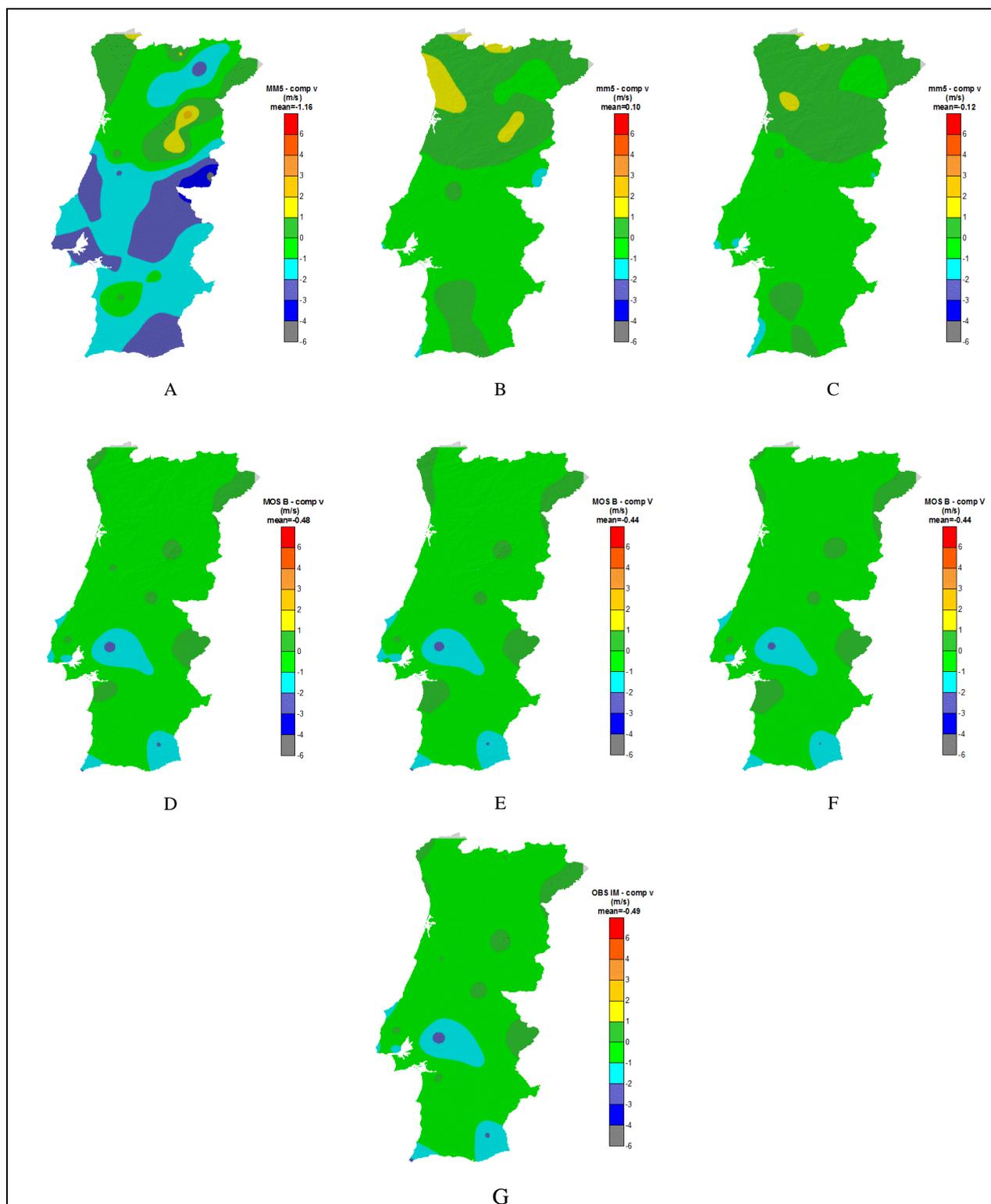


Figura 5.3 – Resultados da previsão da componente meridional do vento nos diferentes prazos, para o modelo MM5 e modelo MOS. A) MM5 - 24 horas; B) MM5 – 48 horas ; C) MM5 – 72 horas; D) MOS – 24 horas; E)MOS – 48 horas; F)MOS- 72 horas; G) Observação.

## 6. Conclusões

O MOS (model output statistics) é um modelo estatístico usado em todo o mundo para melhorar as previsões meteorológicas. A sua metodologia consiste na correcção da saída dos modelos de previsão numéricos baseada na informação de previsões passadas. Os modelos estatísticos, são em geral, equações de regressão linear múltipla, que utilizam as previsões como predictores e as observações como predictandos. O caso mais simplificado deste método é a equação da regressão linear simples, que estabelece a relação entre um único predictor e um predictando.

Este trabalho teve como objectivo elaborar o MOS para previsão estatística das 24 horas, com o intuito de melhorar as previsões do modelo numérico MM5. A metodologia foi aplicada às variáveis da temperatura e às componentes horizontais do vento ( $u$  e  $v$ ).

A implementação do modelo estatístico teve por base a ordem e selecção dos potenciais predictores. Foram elaborados cinco modelos com diferentes números de predictores pré-seleccionados. Para estes modelos foram adoptadas duas abordagens: a pré-selecção da ordem dos predictores e a selecção da ordem segundo o método stepwise regression.

Os resultados para a previsão ao prazo de 24h mostram que a selecção da ordem dos predictores melhora significativamente os resultados da previsão. Dos cinco casos analisados os melhores resultados foram no caso que apresentou mais predictores (caso E com 5 predictores), embora os scores praticamente estacionem ao escolher 3 predictores: previsão a prazo 24h, observação a 0h (instante de lançamento da previsão) e previsão a prazo de 12h. Em geral, a previsão da temperatura melhorou cerca de 60%, com erros médios quadráticos inferiores a 4°C, e as componentes horizontais do vento melhoraram na ordem dos 80%, com erros médios de 2 m/s. Contudo, o modelo testado com apenas um predictor apresentou os resultados mais significativos (caso A). A adição contínua de predictores apenas foi refinar a previsão melhorando cerca de 3 a 4% a previsão das variáveis.

Da análise dos valores extremos das medidas estatísticas (RMSE, MSE e correlacção) das previsões do modelo numérico, pode-se concluir que quanto maior é o erro de previsão do MM5 melhor é a previsão do modelo MOS. Os locais onde o modelo numérico prevê com melhor exactidão, o erro associado é menor, resultando numa percentagem, menos significativa da previsão do modelo MOS em relação ao MM5.

De uma forma geral, todas as medidas estatísticas melhoraram com a aplicação do modelo MOS. Apesar dos valores iniciais da correlacção, das componentes horizontais do vento, serem na ordem dos 50% , estes valores subiram em média 10% com a aplicação da equação de regressão com cinco predictores.

Uma breve análise, dos resultados do modelo MOS para as 48 e 72 horas de previsão, foi elaborada e permitiu concluir que a previsão do modelo numérico degrada-se com o aumento do prazo de lançamento de previsão. Contudo, os resultados da aplicação do modelo estatístico, ao output do modelo MM5, apresentam uma melhoria de cerca de 70% para as 48 e 72 horas de previsão.

Os bons resultados obtidos confirmam que a aplicação do MOS é um método necessário para melhorar as previsões dos modelos numéricos, e um incentivo a torná-lo operacional, assim como, desenvolver o método para outras variáveis.

## 7. Referências

- Azcarraga, R. & Ballester, A.J.G., 1991. Statistical system for forecasting in Spain. In Programme on Short- and Medium-Range Weather Prediction Research, H. R. Glahn, A.H. Murphy, L.J. Wilson and J.S. Jensenius, eds., WM/TD No. 421. World Meteorological Organization, Geneva, XX23-25.
- Brunet, N., Verret, R. & Yacowar, N., 1988. An objective comparison of model output statistics and "perfect prog" systems in producing numerical weather element forecasts. *Weather and Forecasting*, 3: 273-283.
- Carter, G.M., Dallavalle, J.P. & Glahn, H.R., 1989. Statistical forecasts based on the National Meteorological Center's numerical weather prediction system. *Weather and Forecasting*, 4: 401-412.
- Cheng, W. Y. Y., and W. J. Steenburgh, 2007. Strengths and weaknesses of MOS, running-mean bias removal, and Kalaman filter techniques for improving model forecasts over the western United States. *Wea. Forecasting*. 22, 1304-1318.
- Conte, M., de Simone, C. & Finizo, C., 1980. Post-processing of numerical models: Forecasting the maximum temperature at Milano Linate. *Rev. Meteorol. Aeronaut.*, 40: 247-265.
- Costa, P., 2004. Atlas do Potential Eólico para Portugal Continental. Dissertação para a obtenção do grau de Mestre em Ciências e Engenharia da Terra, Lisboa, pp 144.
- Cressman, G.P., 1981. Man and the machine in weather forecasting of the future. *Proc. Symp. On Current Problems of Weather Prediction*, Viena, Austrian Meteor. Soc., 3-6.
- Dallavalle, J. P., M. C. Erickson, and J. C. Maloney III, 2004. Model output statistics (MOS) guidance for short-range projections. Seattle, WA, Amer. Meteor. Soc.
- Esterle, G.R., 1992. Adaptive, Self-learning statistical interpretation system for the Central Asian region. *Annals Geophysical* 10. p. 924-929.
- Francis, P.E., Day, A.P. & Davis, G.P., 1982. Automated temperature forecasting, and application of Model Output Statistics to the Meteorological Office numerical weather prediction model. *Meteorol. Mag.*, 111: 73-87.
- Glahn, H. R., and D. A. Lowry, 1972. The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, 11:1203-1211.
- Klein, W. H., and H. R. Glahn, 1974. Forecasting local weather by means of model output statistics. *Bull. Amer. Meteor. Soc.*, 55: 1217-1227.
- Lemecke, C., and S. Kruizinga, 1988. Model output statistics forecasts: Three years of operational experiences in the Netherlands. *Mon. Wea. Rev.* 116: 1077-1090.
- Jacks, E., Bower, B., Dagostaro, V.J., Dallavalle, P.J., Erickson, M.C. & Su, J.C., 1989. New NGM-Based MOS Guidance for Maximum/Minimum Temperature, Probability of Precipitation, Cloud Amount, and Surface Wind. *Weather and Forecasting*, November 7-13, vol. 5, p.128-138.
- Laurence J. Wilson and Marcel Vallée, 2003. The Canadian Updateable Model Output Statistics (UMOS) System: Validation against Perfect Prog. *Weather and Forecasting*, 18, 288-302.
- Lemcke, C. & Kruizinga, S., 1988. Model output statistics forecast: Three years of operational experience in the Netherlands. *Mon. Weather Rev.*, 116: 1077-1090.
- Lu, R., 1991. The application on NWP products and progress of interpretation techniques in China. In Programme on Short- and Medium-Range Weather Prediction Research, H. R. Glahn, A.H. Murphy, L.J. Wilson and J.S. Jensenius, eds., WM/TD No. 421. World Meteorological Organization, Geneva, XX19-22.
- Sokol, Z., 2003. MOS-based precipitation forecasts for river basins. *Wea. Forecasting*. 18, 769-781.
- Storch, H.V., Zwiers, F.W., 1999. *Statistical Analysis in Climate Research*, University Press, Cambridge, UK.
- Tapp R.G., Woodcock, F. & Mills G.A., 1986. The Application of Model Output Statistics to Precipitation Prediction in Australia. *Monthly Weather Review*, vol. 114: N 1, p. 50-61.
- Taylor, A. A., and L. M. Leslie, 2005. A single-station approach to model output statistics temperature forecast error assessment. *Wea. Forecasting*. 20: 1006-1020.
- Termonia, P., 2003. Monitoring and improving the temporal interpolation of lateral-boundary coupling data for limited-area models. *Mon. Wea. Rev.* 131,3450-2463.
- Tomé, R., Previsão do tempo com modelos de mesoscala: Casos de estudo com o modelo MM5 nos Açores. Dissertação para a obtenção do grau de Licenciatura em Ciências Geofísicas – especialização meteorologia, Lisboa, pp 83.
- Vasiliev, P.P., 1991. Medium-range air temperature and precipitation Model Consensus. *Bull. Of pitation forecasting over Eurasia. Meteorologia i Hidrologia*, N 2, p. 13-23.
- Vislocky R.L. & Fritsch J.M., 1995. Improved Model Output Statistics Forecasts th the Am. Met. Soc. Vol. 76, N 9, P. 1157-1164.
- Vislocky R.L. & Fritsch J.M., 1997. An automated, Observation-Based System for Short-Term Prediction of Ceiling and Visibility. *Weather and Forecasting*, Vol. 12: N 3 P. 31-43.

Wilks, D.S., 1990. Probabilistic quantitative precipitation forecast derived from PoPs and conditional precipitation amount climatologies. *Mon. Wea. Rev.*, 118: p. 874-882.

Wilks, D.S., 1995. *Statistical Methods in Atmospheric Sciences An Introduction*, Academic Press, San Diego, California, U.S.A.

Yuval and William W. Hsieh, 2003. An Adaptive Nonlinear MOS Scheme for Precipitation Forecasts Using Neural Networks. *Weather and Forecasting*, 18, 303–310.