

SISTEMA AUTOMÁTICO DE VALIDAÇÃO DA INFORMAÇÃO DA REDE SAGRA

Maia, J.⁽¹⁾, Santos, M.⁽²⁾ & Neto, M.⁽³⁾

⁽¹⁾ Centro Operativo e de Tecnologia de Regadio

⁽¹⁾ jorge.maia@cotr.pt, Quinta da Saúde, Apt 354, 7801-904 BEJA, Portugal

⁽²⁾ Centro Operativo e de Tecnologia de Regadio,

⁽²⁾ marta.santos@cotr.pt, Quinta da Saúde, Apt 354, 7801-904 BEJA, Portugal

⁽³⁾ Agri-Ciência, Consultores de Engenharia, Lda

⁽³⁾ mneto@agriciencia.com, Rua dos Lusíadas, N.º 52 – 1º, 1300-372 Lisboa

Resumo

O Sistema Agrometeorológico para a Gestão de Rega no Alentejo – SAGRA é um serviço que está a disponibilizar informação meteorológica para o apoio na gestão da rega. O COTR consciente do facto do serviço estar a chegar a um número cada vez maior de utilizadores assumiu a necessidade de desenvolver metodologias, recomendadas internacionalmente, no controlo de qualidade da informação disponibilizada aos utilizadores.

No presente documento descreve-se o sistema automático que o serviço SAGRA está a implementar no controlo de qualidade. O sistema assenta sobre uma rotina automática que diariamente permite monitorizar todos os dados recolhidos e assinalar como erróneos ou suspeitos os valores que se possam desviar da parametrização normal de modo a alertar o técnico responsável pelo serviço de possíveis erros de registo.

Este procedimento foi concebido para os dados horários, sobre os quais foi feita uma análise estatística dos valores registados durante um período de cinco anos. Com base nessa análise, foram determinados intervalos de aceitação de variação entre registos. Sempre que há um desvio superior ao limite estabelecido sobre a variação entre observações, o administrador do sistema é notificado.

Compete ao administrador do sistema a validação final da informação, último passo antes da disponibilização da informação, cuja função principal é fazer um julgamento sobre os avisos emitidos. Assim se tenta garantir uma melhoria na qualidade da informação, e uma optimização dos serviços de apoio à tomada de decisão como MOGRA, unidades de frio, previsão de ocorrência de geadas e outros que poderão surgir entretanto.

Palavras-chave: Controlo de qualidade, informação agrometeorológica, validação de dados

Abstract

The Automatic Weather Station Network for Irrigation Scheduling in Alentejo - SAGRA is a service that is supplying weather information for support irrigation management strategies, being able, however, to be used for other agricultural purposes. This information has been reaching at a wide net of users evermore, due to that COTR assumed the necessity to implement methodologies, which are recommended internationally, in information quality assurance.

In the present document it is described the automatic system that SAGRA service is implementing in the quality control. The system is based on an automatic routine that daily allows, from the collected data, to flag all the suspicious or erroneous values in order to alert the service responsible technician for the data possible errors.

This procedure was conceived based on hourly data statistic analysis, on which the average and standard deviation for five years information of each weather station allowed to establish the lower and upper limit of variation between observations. Whenever there is an outlier of the limit established on the variation between stations, the administrator of the system is notified.

The system administrator has the responsibility for the information final validation, as the last step before the information get available for final user. The system administrator main function is to make a final judgment on the emitted flags. With that is tried to guarantee an improvement in the information quality that feed the decision support systems.

1) Introdução

O clima do sul de Portugal, embora com influência atlântica, é predominantemente caracterizado por Mediterrânico, com uma concentração da precipitação durante a estação fria e uma estação quente com altos valores de temperatura e de radiação solar, mas seca, durante a qual o desenvolvimento de espécies vegetais é favorecido com recurso a aplicação de água de forma artificial.

A baixa ocorrência de precipitação e sua concentração durante a estação fria torna a utilização da água em regadio numa vertente muito importante para as regiões da orla Mediterrânica. Principalmente, num período como o actual, em que a ocorrência de precipitação, inclusive na época fria tem sido extremamente reduzida em vários anos.

Por outro lado, a pressão exercida por outros sectores, na utilização pouco eficiente de água pela agricultura, tem crescido de forma linear. Na tentativa de minimizar esta pressão, várias metodologias têm sido desenvolvidas para melhorar a eficiência no uso da água em regadio. Uma delas está relacionada com implementação de redes agrometeorológicas com o objectivo de apoiar os regantes na correcta gestão da rega.

Para isso o Centro Operativo e de Tecnologia de Regadio (COTR) instalou uma rede de estações agrometeorológicas automatizadas, com estações localizadas nas principais

zonas de regadio do Alentejo, cujo objectivo principal é apoiar o desenvolvimento do regadio de forma mais eficiente.

Assim, o desenvolvimento do Sistema Agrometeorológico para a Gestão da Rega no Alentejo (SAGRA) tem disponibilizado, desde 2001, informação para uma vasta rede de utilizadores que vem crescendo anualmente.

Havendo várias fontes de erros, umas completamente aleatórias e outras possíveis de identificar (como por exemplo, a má aferição dos aparelhos e a falta de sincronização na recolha da informação), a questão da fiabilidade dos dados recolhidos nas Estações Meteorológicas Automáticas reveste-se de particular acuidade (Pinheiro, 2004).

O COTR ciente da necessidade de garantir uma qualidade mínima dos dados e do facto do serviço estar a chegar a um número cada vez maior de utilizadores, delineou uma linha de actividade que consiste no desenvolvimento de metodologias recomendadas internacionalmente no controlo de qualidade da informação disponibilizada, de modo a garantir essa necessidade.

2) A rede agrometeorológica automática

Em 2001 no âmbito do programa INTERREG II C – Ordenamento do Território e Luta Contra a Seca em Portugal, MEDIDA 2 – Reforço e Optimização do Uso de Água na Agricultura o COTR, em parceria com o Instituto de Hidráulica Engenharia Rural e Ambiente (IHERA, actual IDRHa), desenvolveu o projecto denominado SAGRA – Sistema Agrometeorológico para a Gestão da Rega no Alentejo cujo principal objectivo consistiu em instalar uma rede de estações agrometeorológicas automáticas tendo em vista iniciar a criação dum sistema de avisos de rega no Alentejo.

O SAGRA é composto, actualmente, por um conjunto de doze estações agrometeorológicas automáticas (EMA's) (Fig. 1) ligadas em rede, por GSM, através de uma unidade central de recolha, armazenamento e processamento de dados (concentrador regional) com arquivo de dados originais e corrigidos em bases de dados distintas no concentrador.



Figura 1 – Mapa Localização EMA's da Rede SAGRA

As leituras dos diversos sensores, como temperatura e humidade relativa do ar, velocidade e direcção do vento, radiação solar global e precipitação, são feitas em intervalos regulares de 10 s e integradas em relatórios horários e posteriormente em relatórios diários. Os dados são armazenados em loggers *Data Taker 500*. A alimentação energética das estações é feita por painel solar e armazenada em baterias de 12 V.

Destes dados, os diários são imediatamente disponibilizados em bruto no sítio do COTR. Internamente os dados são tratados com a finalidade de validação da informação meteorológica disponibilizada, de estudos climatológicos e do cálculo da evapotranspiração das culturas.

A crescente rede de utilizadores registados no serviço SAGRA-Net (Fig. 2) (este serviço consiste na possibilidade do utilizador aceder, através da Internet, à base de dados climática da rede SAGRA) e registados nos serviços de apoio à gestão da rega, que têm como base de dados climáticos a rede SAGRA, como o MOGRA ou Calendário de Rega (Fig. 3) têm aumentado a responsabilidade do COTR perante o utilizador.

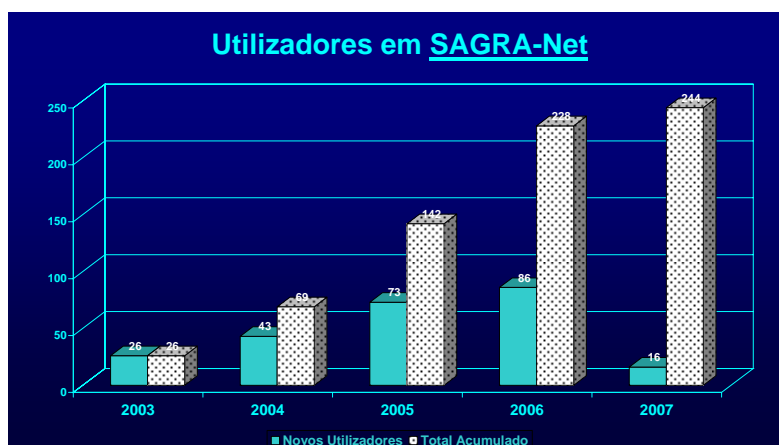


Figura 2 – Evolução dos Utilizadores SAGRA-Net

Assim, e conforme se pode verificar pela análise da Fig. 2, do ano de 2004 para 2005, registou-se um aumento de 100% no número de utilizadores, prevendo-se, a este ritmo, para o final de 2007 existirem cerca de 300 utilizadores registados, o que permitirá duplicar o número de registos entre 2005 e 2007.

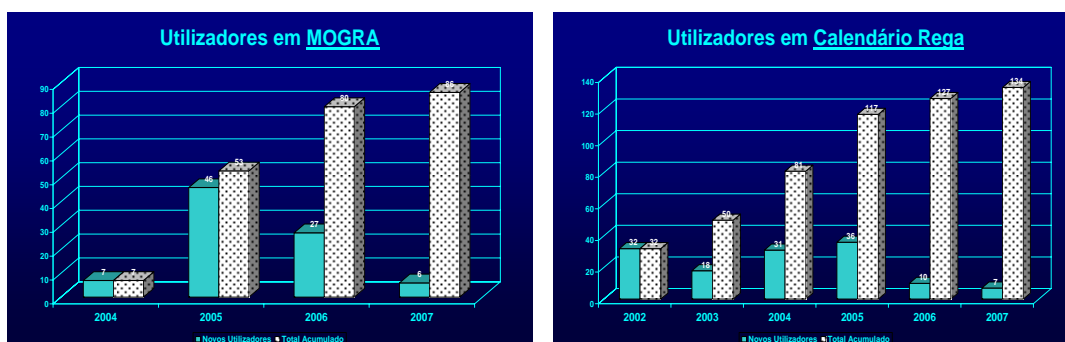


Figura 3 – Evolução dos Utilizadores Serviços Gestão Rega

Perante a realidade que vem sendo constatada, a área SAGRA do COTR, assumiu como uma das suas prioridades o desenvolvimento do serviço de controlo de qualidade da

informação, que permita garantir uma qualidade mínima da informação que fica acessível para o utilizador.

3) Controlo de qualidade da informação

De acordo com alguns autores (Shafer *et al.*, 2000) o controlo de qualidade de dados meteorológicos é um complexo processo composto por diversas actividades entre as quais se enumeram as mais importantes:

- Rotinas automáticas de verificação de dados;
- Calibração de sensores;
- Serviço de manutenção do equipamento de campo;
- Inspeção manual de validação de informação.

Relativamente à rede SAGRA, todo este processo foi descrito por Maia & Santos (2005) no anterior Congresso de Rega e Drenagem realizado em Beja. Neste sentido, pretende-se nos próximos pontos descrever de forma mais pormenorizada os mais recentes desenvolvimentos efectuados no âmbito das rotinas automáticas de validação de dados.

3.1) Situação actual

No âmbito do serviço prestado pelo Sistema Agrometeorológico para a Gestão da Rega no Alentejo (SAGRA), diariamente, no concentrador regional o técnico responsável valida a informação. Entende-se, nesta fase, por validação da informação, o processo de inspeção manual dos relatórios produzidos e verificação da necessidade de efectuar correcções dos dados.

Posteriormente, essa informação é tratada com a finalidade de validação da informação, de forma mais detalhada, com recurso ao desenvolvimento de estudos climatológicos, que vêm sendo efectuados desde 2001.

Continuam a ser desenvolvidas:

- Análise da radiação extraterrestre referente a cada estação, que permite identificar problemas de medição do sensor de radiação;
- Estudo climatológico do vento (directão e intensidade) referente a cada estação;
- Comparação de parâmetros meteorológicos com estações próximas pertencentes a redes exploradas por outras entidades, que poderão ter condições meteorológicas semelhantes, de modo a verificar se existem diferenças significativas entre estações;
- Análise de frequência da velocidade vento cujo objectivo está relacionado com a verificação do aumento do atrito, devido ao aumento de fricção que se verifica nos componentes de movimento do sensor. A análise das frequências absolutas da velocidade do vento é um indicador de decisão da necessidade de substituição dos rolamentos;

3.2) Processo automático de validação de dados

Toda a situação descrita até aqui, é feita de uma forma manual e com relativa base científica, por isso, no sentido de melhorar o trabalho e otimizar os recursos, desenvolveu-se o sistema automático de validação de dados, que permite fazer uma filtragem inicial a toda a informação recebida, com base estatística em vários anos de observações.

3.2.1) Procedimento de validação de dados

No controlo da qualidade dos dados podem-se considerar dois tipos de aproximações com os seguintes procedimentos:

1) Testes que usam uma única estação para detectar possíveis “outliers”, como sejam:

a. Testes para verificação de valores plausíveis (valores compreendidos entre máximo e mínimo, por exemplo no caso das temperaturas, os registos fora do intervalo -20°C e $+50^{\circ}\text{C}$, devem ser assinalados), ou

b. Testes para verificação da consistência dos dados através da taxa de variação dos dados de um momento de observação para o seguinte (por exemplo, assinalar variações da temperatura de uma hora para a anterior superiores a 2°C ou inferiores a $0,1^{\circ}\text{C}$. Nestes casos podem estabelecer-se limites para dados suspeitos e limites para os considerar erróneos. Estes métodos têm sido indicados e usados por vários autores (veja-se, por exemplo Meek and Hatfiel, 1994; ou Eischeld *et al.*, 1995).

2) Testes que usam dados de outras estações próximas ou de dados afins e comparam com os dados da estação em análise para controlo da qualidade dos dados da (Eischeld *et al.*, 1995; Hubbard, 2001). A estimativa usada para comparação pode ser obtida ponderando os dados inversamente à distância que separa as estações ou usando métodos estatísticos como a regressão linear simples (regressão espacial) (Hubbard, *et al.*, 2004) ou dados emparelhados, ou outro critério estatístico que possibilite aferir a fiabilidade dos dados da estação em análise. Em qualquer dos casos, o objectivo é testar o valor da observação com a estimativa e decidir se a observação é ou não um potencial “outlier”.

3.2.2) Fases do processo de validação de dados

A validação processa-se em duas fases ou momentos diferentes. Na primeira fase, usam-se apenas os dados da própria estação e eliminam-se dos erros grosseiros. Assim, encontraram-se, para cada parâmetro, valores que permitiram formar intervalos para classificar os dados que chegam à estação em:

- Plausíveis;
- Suspeitos;
- Erróneos.

Esta classificação refere-se tanto aos valores lidos (medidos) bem como à sua variação. Por exemplo, se chegar à estação um valor de 70°C para a temperatura média do ar, esse

valor deve ser imediatamente rejeitado por ser considerado erróneo. Do mesmo modo, se o valor da temperatura média do ar variar, no intervalo de uma hora, em 20°C ou mais, esse valor também deve ser rejeitado por erróneo.

Como conclusão da primeira fase estabeleceram-se os limites para os valores dos parâmetros e para a sua variação horária de modo a permitir a sua avaliação, e também, a plausibilidade da taxa de variação, ou seja, para poder avaliar a consistência interna dos dados.

Para cada parâmetro de cada estação e tendo em consideração a análise estatística efectuada sobre os dados recolhidos para a estação meteorológica automática instalada na Quinta da Saúde em Beja desde o início da actividade da rede, propôs-se a seguinte metodologia para encontrar os limites superiores e inferiores para os valores considerados plausíveis, suspeitos e erróneos.

Quadro 1 – Limites definidos para cada classificação de valores

| Valores Observados | Limites |
|--|--|
| Valores Plausíveis | Mínimo histórico \leq Valores plausíveis \leq máximo histórico |
| Valores Erróneos | Mínimo histórico $>$ Valores erróneos $>$ máximo histórico |
| Variação dos Valores Observados | Limites |
| Valores Plausíveis | (Média período - 2 x Desvio Padrão período) \leq Valores plausíveis \leq (Média período + 2 x Desvio Padrão período) |
| Valores Suspeitos | (Média período + 2 x Desvio Padrão período) $>$ Valores suspeitos \leq máximo histórico |
| Valores Erróneos | (Média período - 2 x Desvio Padrão período) $<$ Valores suspeitos \geq mínimo histórico |

Em função da metodologia seguida, apresenta-se no Quadro 2, o exemplo dos valores determinados como máximo e mínimo absoluto histórico para a estação meteorológica de Beja.

Quadro 2 – Valores máximos e mínimos absolutos (2001-2005) que definem o intervalo de aceitação dos parâmetros climáticos

| Parâmetro Climático | Valor Mínimo | Valor Máximo |
|--|---------------------|---------------------|
| Temperatura do Ar (°C) | -4,33 | 45,69 |
| Humidade Relativa do Ar (%) | 0,0 | 100,0 |
| Radiação Solar Global (kJ/m ²) | 0,0 | 4272,50 |
| Velocidade Média do Vento (m/s) | 0,0 | 14,78 |
| Direcção do Vento (graus) | 0 | 360 |
| Precipitação (mm) | 0,0 | 18,60 |
| Temperatura Sup. Solo (°C) | -2,99 | 67,90 |

Seguindo a metodologia apresentada no Quadro 1, apresenta-se, ainda, no Quadro 3 o exemplo dos valores determinados como variação máxima e mínima verificada entre observações horárias consecutivas para a estação meteorológica de Beja.

Quadro 3 – Valores máximos e mínimos que definem o intervalo de aceitação de variação dos parâmetros climáticos

| Parâmetro Climático | Valor Mínimo | Valor Máximo |
|--|---------------------|---------------------|
| Temperatura do Ar (°C) | - 3 | 3 |
| Humidade Relativa do Ar (%) | -10 | 10 |
| Radiação Solar Global (kJ/m ²) | - 820 | 820 |
| Velocidade Média do Vento (m/s) | -1 | 1 |
| Velocidade Máxima do Vento (m/s) | -20 | 20 |
| Precipitação (mm) | -1 | 1 |
| Temperatura Sup. Solo (°C) | -4 | -4 |

A segunda fase é composta pela decisão sobre metodologia a usar na detecção de valores erróneos ou na ausência de informação devido a, por exemplo, problemas na estação meteorológica.

Para isso, dividiram-se as estações em grupos de três ou quatro, em função do grau de proximidade ou de semelhanças entre características meteorológicas e analisou-se o comportamento das variáveis climáticas entre estações do mesmo grupo, com o objectivo de determinar o grau de correlação, de modo a corrigir valores com recurso a outra (s) estação (ões) do grupo.

O teste efectuado para um ano de dados, mostrou existir uma alta correlação entre os dados dos grupos homogéneos de estações, definidos pelo COTR, onde se obtiveram valores de correlação entre os dados altamente significativos. Desta forma, será de adoptar a seguinte metodologia simples e eficaz em vez da metodologia seguida por outros autores (como, por exemplo, Hubbard *et al.*, 2004):

- Para os dados erróneos utilizar a média aritmética das outras estações homólogas para os substituir;
- Para os dados suspeitos usar o intervalo: média das estações homólogas menos um desvio padrão e média estações homólogas mais um desvio padrão. Se a observação suspeita estiver neste intervalo deve considera-se plausível e ser validada. Se estiver fora do intervalo deve ser declarada errónea e substituída pela média aritmética das observações das outras estações homólogas.

É de salientar, no entanto, que esta abordagem não dá resposta às variáveis climáticas precipitação e velocidade do vento, pois são fenómenos com uma variabilidade espacial muito influenciada por condições locais e para as quais, por enquanto, continua a ser intervenção humana a melhor opção na resolução da ocorrência de dados erróneos, usando valores de estações próximas de redes independentes, por observação de fenómenos ou registos locais ou por interpolação entre observações de horas adjacentes.

Para além dos parâmetros descritos no parágrafo anterior, a situação aplica-se a todos os parâmetros meteorológicos no caso da estação de Odemira, que está significativamente afastada das restantes estações SAGRA e que pela proximidade ao mar, este lhe confere características climáticas significativamente diferentes das restantes estações, sendo, por isso, difícil estabelecer uma correlação alta que a permita integrar num grupo.

4) Conclusões e recomendações

Perante o que foi descrito anteriormente, o alargamento a todas as estações da rede SAGRA será a próxima fase do processo, de modo a permitir de forma generalista para todas as estações da rede SAGRA a avaliação da metodologia implementada.

A utilização destes procedimentos, não esgota a questão da validade de dados porque existem alguns outros procedimentos a que se pode recorrer, como a utilização de metadados, que permite caracterizar de forma mais eficaz determinado parâmetro meteorológico em determinado momento.

Assim, a questão em implementação pelo COTR, no âmbito do Projecto INTERREG III-A - *Optimización Agronómica y Medioambiental del Uso del Agua de Riego*, é a inclusão de metadados associado à informação meteorológica. Entendendo-se metadados como a descrição de tudo o que possa caracterizar determinado registo, como descrição de procedimentos de obtenção, processos utilizados, condições envolventes do local de obtenção dos dados. Este registo funciona como o bilhete de identidade dos dados e acompanhará sempre que existe disponibilização do dado.

O objectivo principal deste procedimento consiste em auxiliar o utilizador do dado na sua correcta interpretação, de modo a retirar deste a melhor utilização possível.

Com este conjunto de procedimentos espera o COTR contribuir, em certa medida, por um lado, para clarificar este assunto e dar-lhe alguma relevância, e por outro, mostrar o caminho que deverá ser percorrido por todas entidades responsáveis por redes de estações agrometeorológicas, de modo a unificar procedimentos. Para isso, continua a ser extremamente importante a criação de um grupo nacional de trabalho em agrometeorologia de modo a desenvolver normas de procedimentos nesta área.

5) Bibliografia

- Allen, R. G.; Raes, D.; Smith, M. & Pereira, L. S. (1998). Crop evapotranspiration: Guidelines for computing crop requirements. FAO Irrigation and Drainage Paper 56. Roma.
- Eischeld, J. K., C. B. Backer, T. Karl and H. F. Diaz (1995). The quality control of long-term climatological data using objective data analysis. J. Appl. Meteor. 34: 2787-2795.
- INTERREG III-A (2003). Optimización Agronómica y Medioambiental del Uso del Agua de Riego. SUBPROGRAMA: 5 Alentejo-Algarve-Andalucía. Eixo: 1 Dotação de infra- estruturas, ordenamento e desenvolvimento rural do espaço transfronteiriço MEDIDA : 1.3 Desenvolvimento Rural e Transfronteiriço.
- Hubbard, K.G. (2001). Multiple station quality control procedures. In Automated weather stations for applications in agriculture and water resource management. AGM-3 WMO/TD N° 1074. 248P.

- Hubbard, K.G., S. Goddard, W. D. Sorensen, N. Wells, and T. T. Ousugi, (2004). Performance of Quality assurance Procedures for an Applied Climate Information System. *J. Atmos. And Oceanic Tech.*
- Maia, J. & Santos, M. (2005). Garantia de Qualidade: o caso da informação da rede SAGRA. I Congresso de Rega e Drenagem. Beja.
- Meek, D. W. and J. L. Hatfiel (1994). Data quality checking for single station meteorological database. *Agrc. And Forest Meteor.* 69: 85-109.
- Pinheiro, A. (2004). Validação de Dados de Estações Meteorológicas Automáticas. Agriciência. Lisboa.
- Shafer, M. A., C.A. Fiebrich, D. S. Arndt, S. E. Fredrickson, T. W. Hughes (2000). Quality assurance procedures in the Oklahoma Mesonet, *J. Atmos. Oceanic Tech.*, 17, 474-494.