# Redes Neurais Recorrentes no estudo de padrões climáticos sazonais na região Norte do Brasil

Juliana Aparecida Anochi
José Demisio Simões da Silva

Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12227-010 - São José dos Campos - SP, Brasil
{juliana.anochi, demisio}@lac.inpe.br

**Abstract.** In this work a recurrent neural network model for climate forecasting is presented. The model is built by training a neural network with available reanalysis data. The assessment of the model is preceded by the use of data reduction strategies to eliminate data redundancy to minimize the complexity of the models, considering the hypothesis that some variables are drivers for the process of weather forecasting. The results presented in this paper considered the use of Rough Sets Theory principles to derive relevant information from the available data that simplified the development process of the model. The paper presents results of climate prediction using the recurrent neural network based model for the North of Brazil. The input of the recurrent neural network are composed of two kinds of data: a subset of variables chosen by applying the rough sets theory based method and the complete set of variables. The two sets are used to train separate models to learn the seasonal behavior of the precipitation. The results obtained in the conducted experiments show the effectiveness of the methodology, presenting estimates similar to climatological situations in the data considered as the available observations for comparison purposes.

**Palavras-chave:** climate forecasting, rough sets theory, recurrent artificial neural, previsão climática, teoria dos conjuntos aproximativos, redes neurais recorrentes.

## 1. Introdução

Este trabalho propõe o desenvolvimento de modelos empíricos de previsão climática, a partir de dados de reanálise, utilizando redes neurais recorrentes com treinamento supervisionado, usando dados processados pela Teoria de Conjuntos Aproximativos (TCA), em uma abordagem de descoberta de conhecimento para a redução de dados.

O desenvolvimento de modelos de previsão climática a partir de dados considera a hipótese de que é possível extrair dos dados históricos informações sobre o comportamento das condições climáticas. Para isso, o desenvolvimento de tais modelos necessita de uma grande quantidade de dados que garanta a expressividade dos fenômenos em uma vasta faixa de situações. Entretanto, apesar de poder garantir, a princípio, maior robustez aos modelos derivados, a manipulação de grandes volumes de dados pode exigir muito poder computacional, inviabilizando o uso da metodologia de desenvolvimento para a calibração de modelos locais, ou a redundância existente nos dados pode polarizar os modelos durante o processo de especificação.

A disponibilidade de dados tem crescido com o surgimento de novas tecnologias de desenvolvimento de sensores e com a capacidade de processamento e armazenamento de informações em banco de dados. Desta maneira, a análise e extração de conhecimento a partir de dados requerem novas abordagens que gerem resultados em tempo hábil para uso deste conhecimento em processos de tomada de decisão, principalmente naqueles que envolvem situações críticas para o ser humano, como nos processos de análise de clima e de tempo em Meteorologia.

Desta forma, a Descoberta de Conhecimento em Banco de Dados (*Knowledge Discovery in Database* - KDD) surge como uma alternativa que pode ser usada para garantia da qualidade dos dados a serem utilizados durante a extração de informações úteis associadas a padrões escondidos nos dados, podendo ainda propor uma representação reduzida do volume de dados que viabiliza o desenvolvimento não polarizado dos modelos de previsão.

Este trabalho aborda o problema propondo a derivação de modelos de previsão baseados em redes neurais recorrentes, a partir de dados históricos de reanálise, que possam ser executados em qualquer tipo de máquina, mantendo o desempenho computacional. A questão do uso de um grande volume de dados é abordada através de uma técnica de redução de dados, considerando a hipótese de que o processo de eliminação de redundâncias nos dados mantém o conhecimento necessário para o entendimento dos fenômenos relacionados e para realização da climatologia, cujos objetivos consistem em prognosticar, descobrir e explorar o comportamento atmosférico, visando os benefícios para os seres humanos.

Para o desenvolvimento do modelo de previsão climática, este trabalho utiliza as redes neurais recorrentes Elman e Jordan, com treinamento supervisionado, seguindo a mesma metodologia apresentada em (Anochi; Silva, 2009), realizado com dados históricos, processados pela TCA para identificar os atributos mais relevantes para o processo de previsão, permitindo assim uma redução da complexidade do problema e a possibilidade de criar modelos de previsão climática a partir de um conjunto simplificado de dados, para realizar prognósticos do comportamento climático, utilizando modelos construídos diretamente dos dados.

O trabalho está organizado da seguinte forma: a Seção 2 faz uma breve caracterização do problema de previsão climática; a Seção 3 traz os princípios teóricos da TCA; a Seção 4 introduz as redes neurais artificiais, destacando os modelos perceptron de múltiplas camadas, Elman e Jordan; a Seção 5 apresenta a metodologia utilizada no trabalho; a Seção 6 mostra os resultados obtidos e; a Seção 7 traz as conclusões do trabalho.

## 2. Previsão Climática

A previsão climática consiste na estimativa do comportamento médio da atmosfera com alguns meses de antecedência. Por exemplo, em uma escala de tempo sazonal, pode-se prever se o próximo inverno será mais frio que a média, ou se haverá mais chuva. Cabe ainda à previsão climática analisar as ondas de calor e a friagem no inverno, visando prever as propriedades estatísticas do estado climático (Vianello, 2006).

Tempo e clima são conceitos distintos usados em meteorologia para se entender o comportamento da atmosfera em diferentes intervalos de tempo. A previsão do tempo consiste na descrição detalhada de ocorrências futuras, baseadas nas médias de diversas medições, sobre como o tempo estará em determinado local e instante, em um intervalo de tempo muito curto. A previsão do tempo depende de medições meteorológicas provenientes de estações meteorológicas, imagens de radar, balões atmosféricos, bóias marítimas, entre outras. Após a coleta, os dados são processados em modelos executados em supercomputadores, os quais geram possibilidades da evolução do tempo, gerando previsões baseadas em probabilidades. Por outro lado, o estudo do clima avalia o comportamento médio da atmosfera em um intervalo de tempo maior, consistindo de uma integração das condições do tempo em certo período representando uma caracterização mais abstrata.

## 3. Teoria dos Conjuntos Aproximativos

A Teoria dos Conjuntos Aproximativos (TCA) foi proposta no início da década de 80 pelo matemático polonês Zdzislaw Pawlak, como um formalismo matemático, para o tratamento de informações incertas e imprecisas em aplicações de Inteligência Artificial, por meio de aproximações de um conjunto de dados (Pawlak, 1982).

A TCA baseia-se na similaridade entre objetos calculada por uma relação de indiscernibilidade, a qual essencialmente considera que se tal relação existe entre dois ou mais objetos, significa que esses possuem os mesmos valores para todos os atributos que os caracterizam, portanto não podem ser discernidos entre si. Esta relação permite o uso da TCA na construção de subconjuntos de atributos a partir de uma base de dados, os quais são

capazes de representar o conhecimento da base de dados com seus atributos iniciais, através de um processo de eliminação de atributos irrelevantes proporcionando a redução dos dados, através dos chamados redutos (Øhrn; Komorowski, 1997).

Para o desenvolvimento deste trabalho o processamento dos dados pela TCA foi realizado no sistema ROSETTA (*Rough Set Toolkit for Analisys of Data*), que é um conjunto de algoritmos para análise tabular de dados sob a abordagem da TCA, construído como ferramenta de apoio a processos de mineração de dados, cobrindo as diferentes etapas do processo de KDD (Øhrn, 1999).

### 3.1 Sistema de Informação

A TCA considera o conjunto de dados como um Sistema de Informação (SI) organizado em formato de tabela, em que cada linha representa um objeto e as colunas são os respectivos atributos (Komorowski; Øhrn, 1999).

Formalmente um SI é um par ordenado SI = (*U*, *A*) em que *U* é um conjunto finito de elementos, não vazio chamado de universo, e *A* é um conjunto finito, não vazio, de elementos chamados atributos. A existência de um atributo de decisão (*d)* torna um SI um Sistema de Decisão (SD). É importante observar que *d* não pertence ao conjunto de atributos *A*. Formalmente, $SD = (U, A \cup \{d\})$, onde $d \notin A$. A Tabela 1 apresenta um exemplo de um SD.

Tabela 1. Sistema de decisão

| ***U*** | **Atributos Condicionais** | | | **Atributo de decisão** |
|---|---|---|---|---|
| | Estação do Ano | Temperatura | Vento | Chuva |
| $o_1$ | Outono | média | forte | Sim |
| $o_2$ | Inverno | baixa | moderado | Não |
| $o_3$ | Primavera | alta | moderado | Sim |
| $o_4$ | Verão | alta | forte | Sim |
| $o_5$ | Outono | média | forte | Não |

### 3.2 Indiscernibilidade

A relação de indiscernibilidade mede a similaridade entre dois ou mais objetos. A relação estabelece que os objetos que possuem os mesmos valores para seus atributos são idênticos e não podem ser discernidos entre si. A Equação 1 define a relação de indiscernibilidade, estabelecendo que dois objetos *x* e *x'* do conjunto *U*, são indiscerníveis para um subconjunto de atributos B $\subseteq$ A, se para cada atributo *a*, de *x* e *x'*, em *B*, os valores forem iguais (Pawlak, 1994):

$$IND_A(B) = \{(x, x') \in U^2 \mid \forall a \in B, a(x) = a(x')\} \quad (1)$$

Considerando o exemplo da Tabela 1 e tomando o subconjunto *B* = {Mês, Temperatura, Vento}, observa-se que os objetos $o_1$ e $o_5$ apresentam os mesmos valores para os seus atributos, portanto existe uma relação de indiscernibilidade, ou seja, não podem ser discernidos entre si, dessa forma, é possível reduzi-los, formando assim a classe C1 apresentada na Tabela 2.

Tabela 2. Classe para IND(B)

| *U* | Estação do Ano | Temperatura | Vento |
|---|---|---|---|
| $C_1$ | Outono | média | forte |
| $C_2$ | Inverno | baixa | moderado |
| $C_3$ | Primavera | alta | moderado |
| $C_4$ | Verão | alta | Forte |

## 4. Redes Neurais Artificiais

As Redes Neurais Artificiais (RNAs) são métodos computacionais cujo princípio de funcionamento é regido por um modelo matemático inspirado no funcionamento dos elementos básicos que formam a estrutura neural de organismos inteligentes, que adquirem conhecimento através de experiência. O comportamento inteligente resulta das interações entre as unidades de processamento, a partir de seu ambiente através de um processo de aprendizagem.

Computacionalmente, as RNAs são sistemas paralelos distribuídos, compostos por neurônios ou unidades de processamento, que implementam funções matemáticas, normalmente não-lineares. Esses neurônios podem ser distribuídos em uma ou mais camadas interligadas por conexões (pesos sinápticos), os quais armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da rede (Haykin, 2001).

Existem vários tipos de RNAs que diferem em arquitetura ou forma de treinamento. Para o problema de construção de modelos de previsão, como proposto neste trabalho, utilizam-se as redes MLP e os modelos recorrentes de Elman e Jordan, com treinamento supervisionado realizado pelo algoritmo de retropropagação do erro, como mostrado na Figura 1.

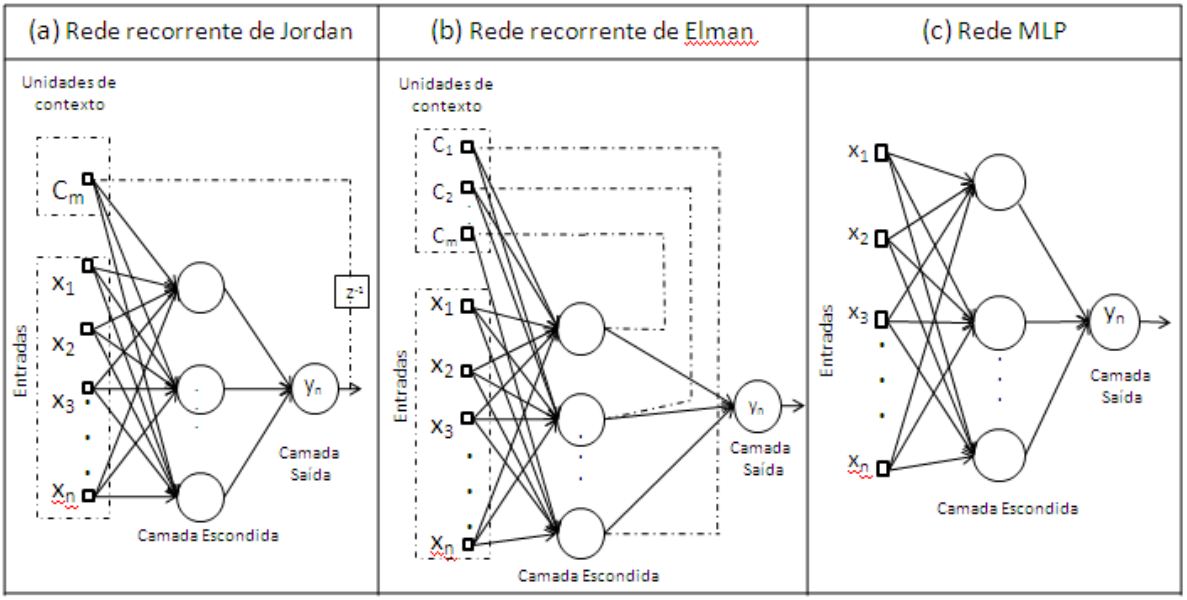


Figura 1: (a) Rede de Jordan; (b) Rede de Elman; (c) Rede MLP.

## 5. Metodologia

Nesse trabalho foram usadas duas abordagens para realização do processo de previsão climática na escala sazonal, formando dois modelos de previsão por RNA usada, no primeiro a RNA é treinada com todas as variáveis disponíveis; e na segunda abordagem, os dados disponíveis para treinamento são processados para gerar reduções que são usadas na aprendizagem de redes neurais para realizar a previsão climática. A redução usa a TCA, que identifica os atributos mais significativos para o processo de previsão climática segundo a relação de indiscernibilidade.

Os dados utilizados nos experimentos foram coletados da base de dados de reanálise do NCEP/NCAR (National Centers for Environmental Prediction / The National Center for Atmospheric Research) [http://ww.ncep.noaa.gov]. Os dados compreendem uma janela de tempo de 10 anos entre janeiro de 2000 e dezembro de 2009, em uma área contida entre as latitudes [10°N, 35°S] e longitudes [80°W, 30°W], referente à América do Sul. A resolução espacial, em ambas as dimensões da grade é de 2.5° e resolução temporal (t) de 1 mês. A área

de estudo para realização dos experimentos abrange a região Norte (N) do Brasil. As coordenadas geográficas estão compreendidas entre as longitudes [75°W, 45°W] e entre as latitudes [10°N , 15°S], compreendendo 140 pontos de grade (14 latitudes x 10 longitudes). As variáveis contidas na base de dados são: temperatura do ar (temp), componentes do vento zonal em: 300hPa (u300), 500hPa (u500), 850hPa (u850), componentes do vento meridional em: 300hPa (v300), 500hPa (v500), 850hPa (v850), pressão (spres), umidade específica (shum) e precipitação (prec).

Do conjunto total de dados foram selecionados 7 anos (janeiro de 2000 a dezembro de 2006) para realização dos treinamentos das RNAs e como entrada para o processamento pela TCA. Para a validação dos modelos foram utilizados os 3 anos restantes (janeiro de 2007 a dezembro de 2009), esses não foram usados no treinamento.

As topologias dos modelos de redes usadas nesse trabalho foram configuradas durante testes preliminares, variando-se de maneira *ad hoc*, o número de neurônios nas camadas escondidas e o número máximo de épocas de treinamento. Os testes conduzidos proporcionaram uma arquitetura constituída por apenas uma camada escondida com 16 neurônios, uma camada de saída, submetida ao máximo de 10000 épocas, onde cada neurônio foi configurado com a função de ativação do tipo logística sigmoidal.

No treinamento de cada rede, foi utilizada a estratégia de parada antecipada com erro calculado sobre o conjunto de teste e comparado com o erro de treinamento durante cada época de treinamento.

Após o treinamento das redes estas foram validadas (generalizadas) utilizando todos os dados e as reduções para gerar previsões sobre os dados de validação. O conjunto de dados de validação é composto por dados do período Janeiro de 2007 a Dezembro de 2009. A métrica para quantificar o desempenho da previsão foi o erro quadrático médio *E* dado por:

$$E = \frac{1}{N}\sum_{k=1}^{N}(y_k - \hat{y}_k)^2 \qquad (2)$$

Em que $N$ é o número de pontos da grade, $y_k$ é o valor real no ponto de grade e $\hat{y}_k$ é a estimação produzida pela rede neural.

No processo de redução dos atributos, os dados são discretizados e em seguida são submetidos ao algoritmo de redução. Os atributos que possuem no mínimo 70% de presença na função de discernimento, formam as reduções para o treinamento das redes neurais (Øhrn, 2001).

Para a visualização e análise dos resultados utilizou-se a ferramenta GrADS (*Grid Analysis and Display System*) que é um *software* utilizado para visualização e análise de dados em pontos de grades (Doty, 2009).

## 6. Resultados

Nesta seção são mostrados os resultados obtidos com os modelos de Jordan, Elman e MLP, utilizando os dados com as todas as variáveis disponíveis na base de dados, e os resultados obtidos para as mesmas redes utilizando os dados processados pela TCA, para as quatro estações dos anos de 2007 e 2008.

Seis modelos de redes neurais foram construídos: três dos modelos são gerados a partir do treinamento das redes MLP, Elman e Jordan, usando todas as variáveis disponíveis na base de dados, e os outros três modelos (MLP, Elman e Jordan), foram usados os dados pré-processados pela TCA, ou seja, os dados reduzidos. Esses modelos foram desenvolvidos para desempenhar a previsão da variável de precipitação na escala.

Na Tabela 3 são apresentados os atributos mais relevantes obtidos por meio da TCA. Observa-se que das dez variáveis mencionadas na Seção 5, sete são necessárias para realizar

previsão sazonal de precipitação. Estas reduções são utilizadas então para o treinamento das RNAs, na busca pelo modelo de previsão.

Tabela 3: Variáveis extraídas pela TCA

| **Variáveis** | **%** |
|---|---|
| temp | 74% |
| u300 | 81% |
| u500 | 81% |
| v300 | 76% |
| v500 | 78% |
| pres | 77% |
| shum | 88% |

A Tabela 4 mostra os erros quadráticos médios dos modelos MLP, Elman e Jordan, para as previsões que utilizam os dados com todos os atributos disponíveis na base de dados, para as quatro estações dos anos de 2007 e 2008.

Tabela 4: Erros quadráticos médios para as quatro estações usando todos os dados

| | **Outono** | | **Inverno** | | **Primavera** | | **Verão** | |
|---|---|---|---|---|---|---|---|---|
| **RNA** | **2007** | **2008** | **2007** | **2008** | **2007** | **2008** | **2007** | **2008** |
| Jordan | 0,0134 | 0,0808 | 0,0118 | 0,0124 | 0,0106 | 0,0233 | 0,0045 | 0,0166 |
| Elman | 0,0007 | 0,0281 | 0,0021 | 0,0069 | 0,0037 | 0,0055 | 0,0002 | 0,0135 |
| MLP | 0,0689 | 0,0668 | 0,0320 | 0,0006 | 0,1237 | 0,0557 | 0,0203 | 0,0188 |

A Tabela 5 apresenta os erros quadráticos médios dos modelos de previsão MLP, Elman e Jordan, que utilizam os dados pré-processados pela TCA, para as quatro estações dos anos de 2007 e 2008.

Tabela 5: Erros quadráticos médios nas quatro estações da região Norte usando a TCA.

| | **Outono** | | **Inverno** | | **Primavera** | | **Verão** | |
|---|---|---|---|---|---|---|---|---|
| **RNA** | **2007** | **2008** | **2007** | **2008** | **2007** | **2008** | **2007** | **2008** |
| Jordan | 0,0253 | 0,0011 | 0,0425 | 0,0080 | 0,0279 | 0,2689 | 0,0547 | 0,1359 |
| Elman | 0,0401 | 0,0094 | 0,0008 | 0,0002 | 0,0161 | 0,0274 | 0,0156 | 0,0148 |
| MLP | 0,0572 | 0,1499 | 0,0601 | 0,0585 | 0,0669 | 0,1299 | 0,0135 | 0,0602 |

A Figura 2 apresenta os resultados das estimativas de precipitação observada ou real, os resultados obtidos com os modelos de Jordan, Elman e MLP, utilizando os dados com todas as variáveis disponíveis na base de dados, e os resultados obtidos para as mesmas redes utilizando os dados processados pela TCA, na estação inverno de 2007. Observa-se que os resultados obtidos com os modelos recorrentes de Elman usando todos os atributos disponíveis na base de dados Figura 2(d), Jordan e Elman usando os dados reduzidos através da TCA, nas Figuras 2(c) e 2(e), apresentaram padrões semelhantes à situação real (observado), o que pode ser verificado pelos erros na Tabela 5.

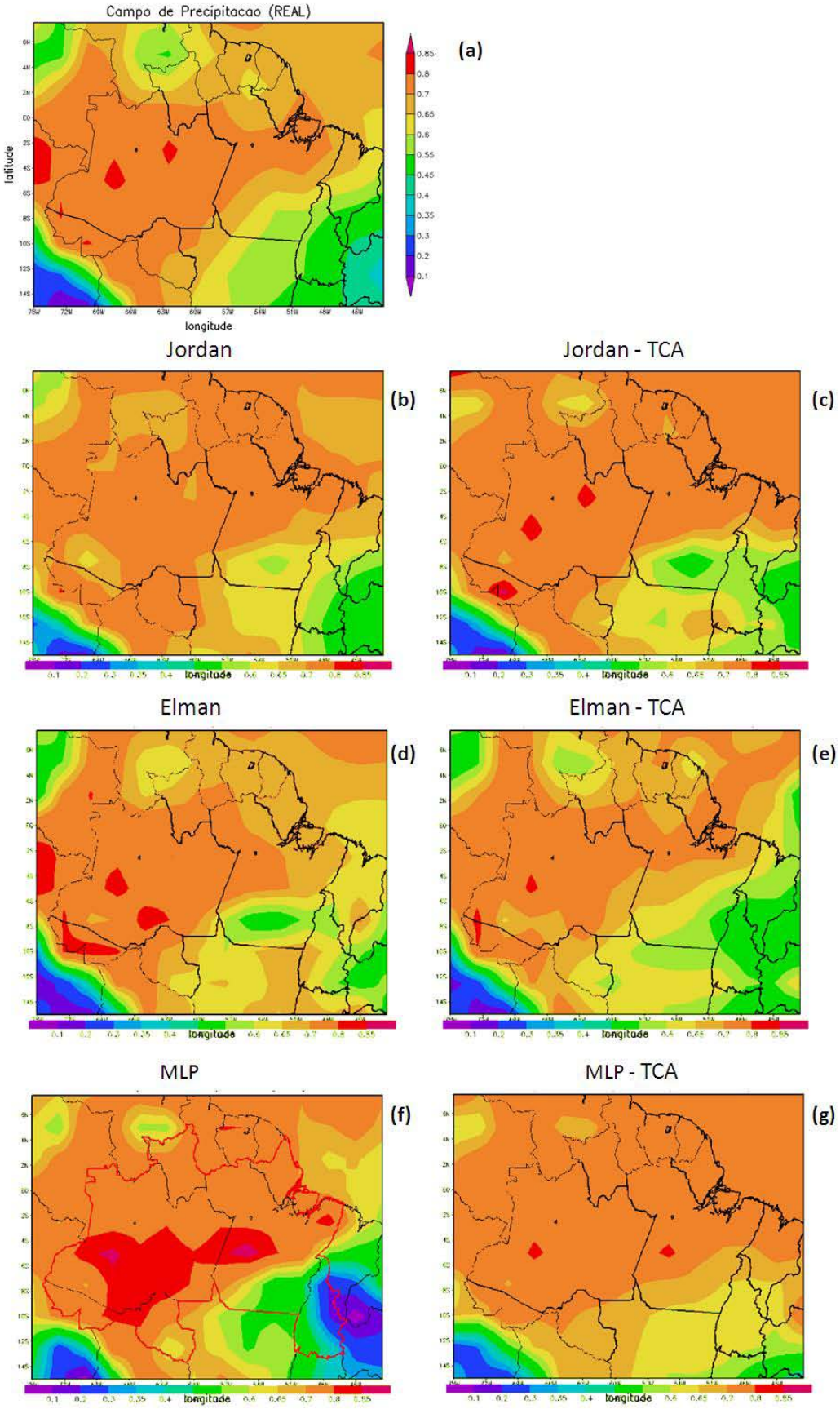


Figura 2: (a) Precipitação real; (b) Jordan usando todos os dados; (c) Jordan usando TCA; (d) Elman usanto todos os dados; (e) Elman usando a TCA; (f) MLP utilizando todos os dados; (g) MLP com TCA.

## 7. Considerações Finais

Neste trabalho utilizou-se uma metodologia para redução de dados, aplicada no estudo de padrões climáticos sazonais, utilizando as abordagens teoria dos conjuntos aproximativos e redes neurais na derivação de modelos de previsão com o propósito de estimar o estado futuro da variável de precipitação, na região Norte do Brasil.

As previsões geradas pelos modelos mostraram que as previsões realizadas usando as redes recorrentes Elman e Jordan, utilizando os dados reduzidos por meio da TCA, apresentaram padrões mais semelhantes àqueles presentes na situação real.

A principal vantagem no uso da metodologia proposta para derivação de modelos de previsão, usando redes neurais, é o uso da técnica de mineração de dados que reduz a complexidade do problema e no fato de poder usar métodos de previsão a partir de modelos derivados diretamente dos dados.

## Agradecimentos

Os autores agradecem ao INPE e a CAPES pelo apoio financeiro para a realização do presente trabalho.

## Referências Bibliográficas

Anochi, J. A.; Silva, J.D.S. Mineração de dados meteorológicos pela teoria dos conjuntos aproximativos para aplicação na previsão de precipitação sazonal. In: Congresso Nacional de Matemática Aplicada e Computacional, 32., 2009. **Anais**... Cuiabá, Mato Grosso: CNMAC, 2009., p. 204-207. On-line. ISSN 1984-820X. Disponível em: http://www.sbmac.org.br/eventos/cnmac/xxxii_cnmac/pdf/278.pdf. Acesso em: 06 nov.2010.

Doty, B. **Grid Analysis and Display System (GrADS)**. Maryland: Center for Ocean- Land-Atmosphere Studies. Disponível em: <http://grads.iges.org/grads/head.html>, Acesso em: 23.fev.2009.

Haykin S. **Redes Neurais Princípios e Práticas**. Porto Alegre: Bookman, 2001. 900p.

Komorowski, J.; Øhrn, A. Modelling prognostic power of cardiac tests using rough sets. Artificial Intelligence in Medicine, v.15, 1999, p. 167-191.

Øhrn A. Discernibility and Rough Sets in Medicine Tools and Applications. 1999. 239 p. Tese de Doutorado (Department of Computer and Information Science) - Norwegian University of Science and Technology, NTNU.

Øhrn, A.; Komorowski, J. Rosetta a rough set toolkit for analysis of data. In: International Joint Conference on Information Sciences, 3., 1997, **Proceedings**... Durham, USA: Institute of Computer Science, 1997. p. 403 - 407.

Øhrn, A.; **ROSETTA technical reference manual**. Department of Computer and Information Science. Norway, 2001. Disponível em: <http://www.lcb.uu.se>. Acesso em 4.ago.2010.

Pawlak Z. Rough sets present state and further prospects. In: International Workshop on Rough Set, 3., 1994. **Proceedings**... Poland: Institute of Computer Science, 1994. p. 72-76.

Pawlak, Z. Rough sets. In: International Journal of Computer and Information Sciences, 11., 1982. p. 341-356.

Vianello, R. L.; Alves, A. **Meteorologia básica e aplicações**. Viçosa: UFV, 2006. 449 p.