6/10/95

Universidade de São Paulo
Instituto de Física

Departamento de Física Geral

Embriogênese e Diferenciação Celular

João da Costa Chaves Júnior

Tese apresentada ao Instituto de Física da Universidade de São Paulo para obtenção do título de Doutor em Física

UNIVERSIDADE DE SÃO PAULO
SERVIÇO DE BIBLIOTECA E INFORMAÇÃO
INSTITUTO DE FÍSICA
2334
e.2

Banca Examinadora:

Prof. Dr. Paulo Murilo Castro de Oliveira (UFF)
Prof. Dr. Henrique Lins de Barros (Museu de Astronomia - RJ)
Prof. Dr. Paulo Mascarello Bisch (UFRJ)
Profa. Dra. Tânia Tomé Martins de Castro (IFUSP)

Orientador: Prof. Dr. Amando Siuiti Ito

Prof. Iberê Luiz Caldas
Presidente da Comissão de Pós-Graduação

SBI-IFUSP
305M810T2334

571.4
C512e
D
ex.2

FICHA CATALOGRÁFICA
Preparada pelo Serviço de Biblioteca e Informação do Instituto de Física da Universidade de São Paulo

Chaves Jr., João da Costa
Embriogênese e diferenciação celular. São Paulo, 1995.

Tese (Doutoramento) - Universidade de São Paulo. Instituto de Física. Departamento de Física geral
Área de Concentração: Física do Estado Sólido
Orientador: Prof. Dr. Amando Siuiti Ito

Unitermos: 1. Autômato celular; 2. Biofísica; 3. Criticalidade auto-organizada; 4. Embriogênese e diferenciação celular.

USP/IF/SBI - 44/95

Dedico este trabalho às
mulheres da minha vida
Marília
Tata
Lourdes

Agradecimentos

Este trabalho só está sendo concluído, porque pude contar com o apoio de muitas pessoas especiais. Pessoas às vezes bem próximas, às vezes um pouco distantes, mas sempre se colocando de forma a colaborar e apoiar.

Fica difícil citar todos, não porque não me lembre dos seus nomes, pelo contrário, os amigos e companheiros que estiveram juntos comigo durante a realização deste trabalho, estarão sempre presentes em minha memória.

Todos os que contribuiram sabem dos meus sentimentos.

Não posso deixar de fazer um agradecimento especial ao Amando, a quem muito admiro pela integridade profissional e pessoal, por ter me garantido a possibilidade de desenvolver este trabalho.

Sem a utilização dos recursos computacionais do Centro Nacional de Processamento de Alto Desempenho, e do Departamento de Física Geral, não teria sido possível a conclusão desta tese.

Finalmente agradeço ao CNPQ e à CAPS que ofereceram o indispensável suporte finaceiro.

RESUMO

Construimos um modelo de caráter probabilístico para a embriogênese e a diferenciação celular baseado nas redes genéticas aleatoriamente conectadas e no processo de criticalidade auto-organizada. Estudamos redes com 3, 4, 5, 10, 20, 30, 64 e 128 genes para as quais fizemos um levantamento do perfil da distribuição populacional da diversidade de atratores, e avaliamos a frequência com que foram gerados. Obtivemos resultados exatos para redes com 3 e 4 genes. Para redes maiores, foram obtidos resultados quantitativos através de varreduras aleatórias. Construimos autômatos de pilha de areia não direcionais, numa rede quadrada (L^2), e determinamos os máximos populacionais a eles associados para L=3, 4, 5, 6, 7, 8 e 9. Mostramos que os parâmetros do *Caenorhabidits Elegans*, cuja história desenvolvimental é bem conhecida, são compatíveis com o modelo proposto.

ABSTRACT

A probabilistic model for Embriogenesis and Cellular Diferentiation, has been constructed, based on the random genetic networks and the self organized criticality process. Networks with 3, 4, 5, 10, 20, 30, 64 e 128 genes were studied. A survey of the population distribution profile was made and their frequency of generation were evaluated. We obtained exact results for nets with 3 and 4 genes. Quantitative results were obtained for the larger ones by means of random scanning. We constructed a undirected sand pile model in a square net (L^2) and determined the population maxima for L = 3, 4, 5, 6, 7, 8 e 9. We showed that parameters for *Caenorhabidits Elegans*, that has a well known developmental story, are in agreement with the proposed model.

Índice

1. Introdução

(i) Embriogênese e Diferenciação Celular 1

(ii) Aspectos Biológicos 2

2. Parâmetros Físicos para a Embriogênese e Diferenciação Celular

(i) O Modelo de Kauffman 10

(ii) A Questão do Crescimento e sua Moderação 16

(iii) O Sistema Imune 16

(iv) Criticalidade Auto-Organizada 20

3. A Construção de um Modelo Global para a Embriogênese e Diferenciação Celular

(i) Pressupostos 27

(ii) O *Caenorhabditis Elegans* 28

(iii) Simulação Computacional 30

4. Resultados

(i) As Redes Pequenas 37

(ii) As Redes Maiores 40

(iv) Saturação Populacional 42

5. Conclusão 45

6. Anexos 48

(i) Gráficos 51

1. Introdução

(i) Embriogênese e Diferenciação Celular

Os seres vivos pluricelulares podem se desenvolver a partir de uma única célula que, através de um complexo processo de divisões sucessivas, vai aumentando a população celular e produzindo células com diferentes performances biológicas, até que finalmente se configura o organismo adulto. Este processo, designado genericamente como desenvolvimento foi considerado classicamente restrito à embriogênese - que vai desde o estágio de ovo fertilizado até o nascimento, fase durante a qual o organismo é denominado embrião - embora atualmente compreenda fenômenos que vão além do nascimento (Gilbert, 1994).

O ovo ou zigoto, célula que se obtém da fusão dos gametas (óvulo e espermatozóide), contém a informação básica que no decorrer da embriogênese vai produzindo um grande número de tipos celulares diversificados tais como neurônios, células musculares, linfócitos, osteoblastos, etc. Este processo é denominado diferenciação celular e ocorre concomitantemente com a morfogênese durante a qual as células diferenciadas vão sendo agrupadas em tecidos e órgãos. Segue-se o crescimento onde ocorre um aumento de tamanho, culminando com a última fase que fecha o ciclo - a reprodução, que desencadeia novamente o processo.

Apesar das inúmeras variações a que está sujeita a embriogênese no reino animal, obedece fundamentalmente a uma sequência de quatro eventos:

(1) Clivagem - constitui a fase em que ocorre uma série de divisões (mitoses) rápidas que se seguem à fecundação e produzem numerosas células, chamadas blastômeros, sem aumentar significativamente o volume citoplasmático do zigoto, formando no final do processo, uma esfera conhecida como blástula. (fig.1.1.e)

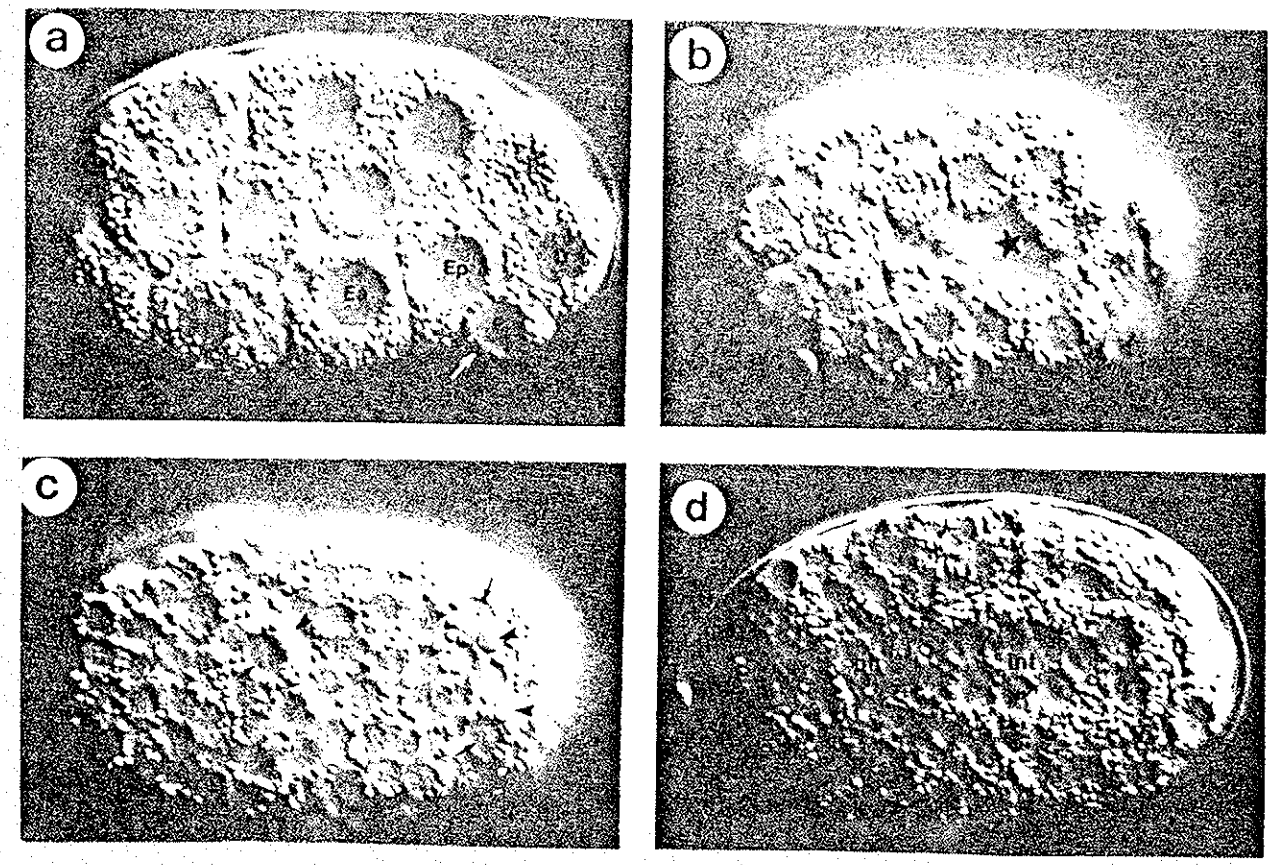
(2) Gastrulação - onde ocorrem intensas movimentações dos blastômeros culminando por produzir uma estrutura com três regiões distintas de células: a

ectoderme - camada externa da qual deriva a epiderme e o sistema nervoso; a endoderme - camada interna que gera o revestimento do sistema digestivo e órgãos como pâncreas, intestino, etc. ; e a mesoderme - que se localiza entre as anteriores e origina o coração, rins e as gônadas, além dos tecidos conjuntivos (ossos, músculos, tendões) e as células do sangue.

(3) Organogênese - ao final da gastrulação, começam a se configurar os órgãos como resultado de complexas interações entre as células. O processo é desencadeado a partir da formação do tubo neural que se origina de uma dobradura da ectoderme através de intensas movimentações celulares.

(4) Gametogênese - constitui o processo de produção de gametas a partir das células germinativas. Isto geralmente só é concluído quando o organismo estiver fisicamente maduro, tornando-o apto a realizar a fecundação e começar uma nova vida.

Fig. 1.1.



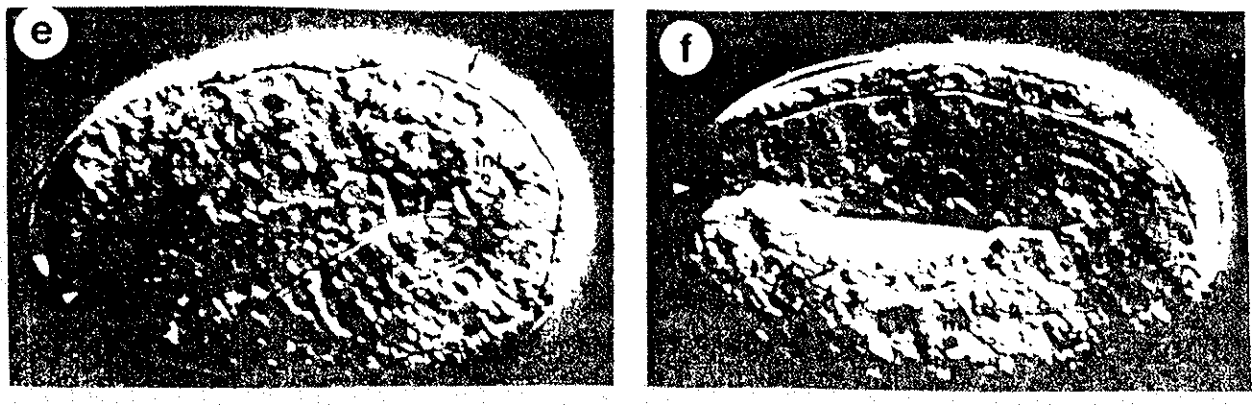


Fig.1.1.: Fotomicrografias de embrião de *Caenorhabidits Elegans* em diferentes estágios. (a) Início da gastrulação. (b) fase final da gastrulação. (c) Superfície ventral coberta por pequenos neuroblastos. (d) Vista dorsal mostrando cilindro de células intestinais e da faringe (ph). (e) Embrião nos seus primeiros movimentos - (int) intestino. (f) Embrião com o intestino (int) já melhor configurado.

Durante a embriogênese o sistema vivo passa por uma quantidade enorme e variada de transformações que dependem fortemente do conteúdo informacional de que suas células são portadoras. Parte da informação genética vai sendo expressa determinando a emergência de atividades bioquímicas específicas, as quais condicionam a expressão ulterior dessa informação, eventualmente desencadeando a expressão de genes até então inativos, e/ou bloqueando a expressão de outros que estejam ativos (Guimarães e col., 1994). A medida que o processo avança, vai se configurando paulatinamente o padrão de distribuição de tipos celulares, de tecidos e órgãos, que finalmente se arranjam numa estrutura organizacional característica da espécie.

Pela riqueza e complexidade de seu objeto de estudo, a biologia do desenvolvimento tem assumido um papel unificador dentro das ciências biológicas, integrando "biologia molecular, fisiologia, biologia celular, anatomia, pesquisa sobre o câncer, imunologia e até estudos sobre evolução e ecologia" (Gilbert, 1994) e, como veremos no decorrer deste trabalho, oferece possibilidades interessantes na sua interface com a física.

(ii) Aspectos Biológicos

As estruturas básicas ligadas ao armazenamento, transmissão e expressão da informação genética nos seres vivos são os ácidos nucleicos: o DNA (ácido desoxirribonucleico) e o RNA (ácido ribonucleico). Ambos são macromoléculas resultantes da polimerização de tipos distintos de monômeros, que são constituídos tipicamente de uma base nitrogenada heterocíclica (nucleotídios derivados da purina ou pirimidina), uma pentose (açúcar com cinco carbonos) e uma molécula de ácido fosfórico. O RNA possui as bases púricas, adenina e guanina, e as bases pirimídicas, citosina e uracila - derivados da pirimidina (Fig.1.2.) (Voet, 1990; Lehninger, 1977), sendo que sua molécula é constituída em geral por uma fita simples dos seus monômeros ligados sequencialmente. O DNA possui as bases púricas, adenina e guanina, e as bases pirimídicas citosina e timina, exibindo uma estrutura mais complexa que compreende um par de fitas geralmente arranjadas em forma de dupla hélice (Fig.1.3.), submetidas a um processo complexo de empacotamento.

Nos organismos eucariontes (que têm membrana nuclear em suas células), aproximadamente 99% da massa total de DNA se localiza no núcleo formando os seus cromossomos. Cada cromossomo contém uma molécula de DNA de tal modo empacotada que se esticarmos um cromossomo humano que tem entre 700 e 800 μm de comprimento, obteremos uma fibra duplex de aproximadamente 4 cm. O 1% restante de DNA que se encontra no citoplasma, constitue o DNA mitocondrial que tem peso molecular em torno de 10 milhões e tem uma estrutura consideravelmente mais simples que o DNA nuclear.

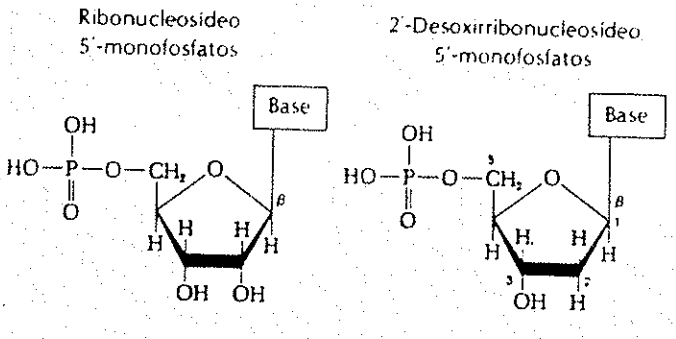


Fig. 1.2.: Representação esquemática de monômeros de RNA, ribonucleosídeo e de DNA, desoxirribonucleosídeo.

Fig.1.3.: Representação esquemática de uma dupla hélice de DNA

O conjunto de polipeptídios que uma célula efetivamente sintetiza é determinado pelo pool de genes que ela possue, de forma que podemos associar a cada peptídio um gene que é portador da informação cuja expressão nele se materializa. Além da eleição da sequência de bases que constitui o gene, e da produção de vários tipos de RNAs como o RNA mensageiro (mRNA) e o RNA transportador (tRNA), o processo de produção de um polipeptídio desde o seu início até a construção do produto final acabado, passa por uma série de etapas que podem assim ser esquematizadas:

- ativação: ocorre no citoplasma e corresponde ao momento em que os aminoácidos sofrem uma reação de esterificação com os seus tRNA correspondentes (Fig.1.4.), catalisada por enzimas denominadas aminoacil-tRNA sintetases, cuja especificidade é tal que cada uma atua exclusivamente num único par aminoácido - tRNA, formando o que se convencionou chamar de aminoácido ativado,

- iniciação: corresponde ao momento em que uma outra fita de RNA, o mRNA, que determina a ordenação sequencial dos aminoácidos que constituirão o peptídio, se liga à molécula de aminoacil-tRNA que contém o primeiro aminoácido através de pontes de hidrogênio entre sítios bastante específicos de ambos - o anti-codon no

tRNA e o codon no mRNA que correspondem a sequências de três bases nitrogenadas complementares,

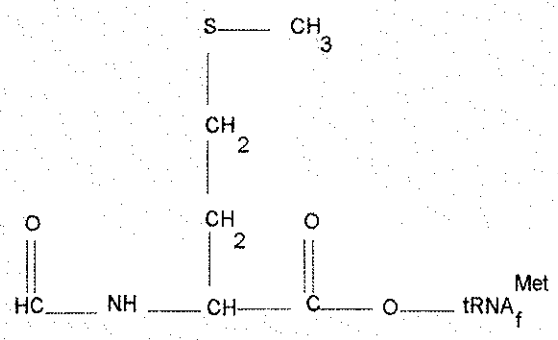


Fig.1.4.: Formilmetionina-tRNA

- alongamento: uma vez estabelecida a ligação entre o aminoacil-tRNA e o mRNA, havendo dois aminoácidos ativados consecutivos, intervém a enzima peptidil transferase, que catalisa a reação de formação de uma ligação peptídica entre o oxigênio terminal do primeiro aminoácido ativado e o grupo amino do aminoacil-tRNA vizinho formando um dipeptídio (Fig.1.5.); tão logo se consolida a ligação peptídica, o primeiro tRNA se desliga do codon do mRNA ficando livre no citoplasma, podendo ocorrer mais uma ligação anticodon-codon com eventualmente um outro aminoácido ativado, e outra ligação peptídica é estabelecida, de modo que o peptídio vai aumentando sua cadeia linear,
- terminação: quando o polipeptídio está completo, com todos os seus aminoácidos, o processo é interrompido por um complexo conjunto de sinalizações comandadas pelo mRNA.

Quando dizemos que existem células diferenciadas nos seres vivos pluricelulares, nos referimos ao fato delas sintetizarem durante a sua vida apenas uma certa fração dos peptídeos cuja informação certamente está armazenada no conjunto de seus genes. De fato, isto torna umas células diferentes das outras no que

se refere às interações que cada tipo celular poderá estabelecer com as outras células presentes no conjunto do organismo, e são as quantidades relativas de cada tipo que vão no final das contas determinar o padrão organizacional e funcional, bem como a "geometria corporal" característica de cada espécie.

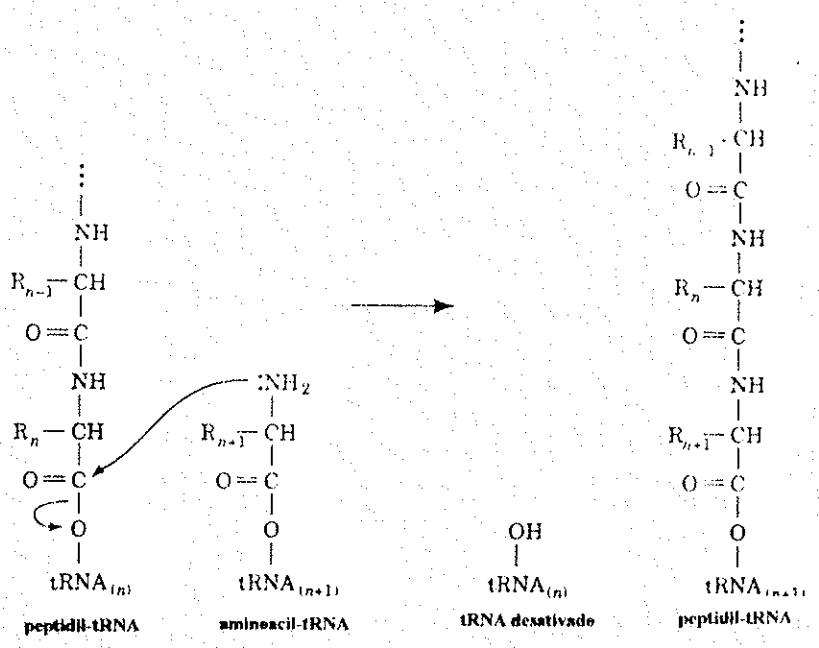


Fig.1.5.: Formação da ligação peptídica.

A informação codificada nos genes desempenha um papel de extrema importância no modo como as células se reproduzem, se diferenciam, se juntam determinando a emergência de padrões teciduais específicos, e finalmente em todo o processo de crescimento, que culmina com construção de um indivíduo pluricelular. Certamente a embriogênese ocorre sob controle genético, embora, o conhecimento dos mecanismos bioquímicos do controle genético, não seja suficiente para explicar o desenvolvimento (Edelman, 1987 e 1988; Lewin, 1994).

Uma das questões fundamentais da biologia do desenvolvimento é estabelecer uma ligação entre os fenômenos genéticos que ocorrem em nível intracelular e o processo de construção da forma e do padrão organizacional, no qual interferem interações entre coletividades celulares, que podem induzir modificações na atividade gênica. São as interações locais (locus-dependentes) que vão determinar a construção da configuração celular que originará um determinado tecido ou, em dado momento propiciará a migração de grupos de células de um local para outro do embrião. Sabe-se que essas interações ocorrem em sequências cuja regulação temporal ultrapassa a duração daqueles eventos submetidos a controle intracelular (Edelman, 1988), de modo que é forçoso admitir-se que durante a embriogênese deve haver algum tipo de concatenação entre os eventos que ocorrem em escala intracelular e os que ocorrem em escala intercelular, e mesmo entre coletividades de células. Isto ocorre através de uma rede de sinalizações que envolvem além das moléculas de adesão celular - ligadas a emergência de padões teciduais, a participação de um grande número de peptídios conhecidos como fatores de crescimento (Russo et al, 1992) - que ora estimulam, ora inibem divisões celulares. As interações entre essas moléculas, tanto interna quanto externamente à célula, com seus transdutores de sinais constitui uma rede que envolve um grande número de componentes na transmissão de sinais entre o meio interior e exterior das células, que em última instância vão afetar a "maquinaria" de transcrição genética. Muitos detalhes deste processo ainda não são bem conhecidos, mas sabe-se que o mesmo fator de crescimento pode ativar diferentes vias bioquímicas e diferentes fatores podem ativar uma determinada via.

Toda essa miríade de interações bioquímicas, de vias metabólicas, de ativações e inibições de crescimento celular constituem a dimensão microscópica do que em nível macroscópico constitue o processo de embriogênese e diferenciação celular.Os eventos descritos nesses níveis estão necessariamente relacionados, de

modo que para construirmos um modelo para este fenômeno temos que levar isto em conta.

O objetivo deste trabalho é desenvolver um modelo de crescimento que mimetize algumas propriedades do processo de desenvolvimento como o estabelecimento do contingente populacional dos diversos tipos celulares efetivamente presentes no indivíduo adulto e sua moderação, e também permita que sejam estabelecidas relações entre o número de genes de um determinado organismo e a composição da população celular presente no adulto.

2. Parâmetros Físicos para a Embriogênese e Diferenciação Celular.

(i) O Modelo de Kauffman

S. A. Kauffman (1969) propôs, com base no modelo de Monod (1963) que o funcionamento da maquinaria gênica dos seres vivos poderia ser mimetizado por certos autômatos celulares, que denotaremos aqui por autômatos de Kauffman (AK). Um AK é um objeto constituído de N sítios contendo zeros e uns, e um conjunto de funções f_i, que determinam a evolução temporal da sequência de sítios através da aplicação da equação:

$$S_i(t+1)=f_i(S_{j1(i)}(t), ..., S_{jk(i)}(t)) \qquad (1)$$

onde $S_i(t+1)$ é o valor do i-ésimo sítio no instante t+1, que por sua vez, é determinado pela função f_i, escolhida entre as 2^{2^K} funções Booleanas possíveis de K variáveis, aplicada aos valores dos sítios $S_{j1(i)}$, ..., $S_{jk(i)}$ no instante t. O conjunto dos índices $j_1(i)$, ..., $j_k(i)$ é escolhido aleatoriamente entre os N sítios, representando exatamente os K sítios que determinam a evolução temporal do i-ésimo sítio (Derrida, 1988). Assim, dada uma sequência inicial de zeros e uns (em t=0), o conjunto de funções f_i aplicadas aos valores atuais dos sítios, irá gerar uma nova sequência de zeros e uns eventualmente distinta da anterior, e assim por diante.

Cada sítio representaria um gene, que poderia estar em on (valor um) ou em off (valor zero) e, as funções f_i determinariam as sucessivas mudanças de estados da "rede genética aleatoriamente conectada" ao longo do tempo, incorporando as influências de outros K "genes" no funcionamento do i-ésimo "gene".

Examinemos um exemplo de AK, para que possamos, em seguida, explicitar melhor as equações utilizadas, bem como, mais adiante, o esquema do programa computacional.

Consideremos inicialmente que dispomos de 4 sítios (N=4), e que o estado inicial (t=0) da "rede genética" seja (1,1,0,1). Além disso, suponhamos que :

(i) K=2, e o estado funcional do gene 1 depende dos estados funcionais dos genes 2 e 3, o do gene 2 dos genes 1 e 4, o do gene 3 dos genes 2 e 4, o do gene 4 dos genes 1 e 3;

(ii) as funções Booleanas escolhidas sejam as seguintes:

TABELA I-1

valor		f_1	f_2	f_3	f_4
0	0	1	0	1	1
0	1	0	0	0	1
1	0	0	0	1	0
1	1	1	1	1	0

onde f_i é utilizada para determinar a evolução temporal do estado funcional do i-ésimo "gene".

O estado inicial da rede é:

GENE: 1 2 3 4

ESTADO DA REDE

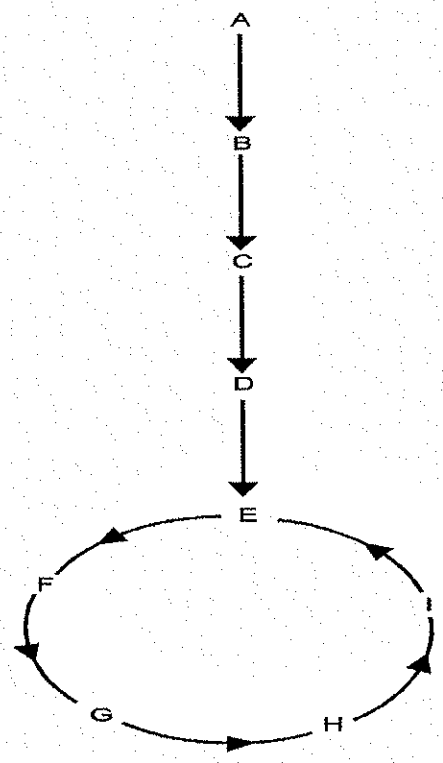
NO INSTANTE t=0: (1, 1, 0, 1)

isto, juntamente com as condições anteriormente expressas determinam a seguinte sequência:

INSTANTES (MOMENTOS) t	GENES 1 2 3 4	ESTADOS DA REDE
0	(1, 1, 0, 1)	≡ A
1	(0, 1, 1, 0)	≡ B
2	(1, 0, 1, 1)	≡ C
3	(0, 1, 0, 0)	≡ D
4	(0, 0, 1, 1)	≡ E
5	(0, 0, 0, 1)	≡ F
6	(1, 0, 0, 1)	≡ G
7	(1, 1, 0, 0)	≡ H
8	(0, 0, 1, 0)	≡ I
9	(0, 0, 1, 1)	≡ E

onde em t=9 é repetido o estado E e podemos facilmente concluir que daqui para frente nada de novo acontecerá, pois seguindo as mesmas regras, serão gerados novamente os estados F, G, H, I, E, F, G, H, I, ... e assim indefinidamente. Dizemos então que o autômato de Kauffman está descrevendo um ciclo constituído de 5 estados que pode ser representado da seguinte maneira:

Fig.2.1



Nos termos de (1), as equações aqui utilizadas foram as seguintes:

$$
\begin{aligned}
S_1(t+1)&=f_1(S_2(t),S_3(t))\\
S_2(t+1)&=f_2(S_1(t),S_4(t)) \qquad (2)\\
S_3(t+1)&=f_3(S_2(t),S_4(t))\\
S_4(t+1)&=f_4(S_1(t),S_3(t))
\end{aligned}
$$

Repetindo-se este procedimento para muitas outras escolhas, tanto dos estados iniciais quanto das funções Booleanas, obtêm-se uma coleção de sequências de estados da rede. Como o número de estados acessíveis a cada sequência é finito

(2^N para N genes), cada uma delas não pode ter mais que 2^N estados, de modo que todas elas apresentam ciclos, ou seja, a partir de um certo instante, começa haver repetição de uma sub-sequência de estados. Em casos limites, podemos ter ciclos de comprimento 1 ou 2^N. O ciclo de comprimento 1 ocorre quando é repetido indefinidamente um único estado; e o de comprimento 2^N, quando o ciclo percorre todos os estados possíveis. Kauffman (1969) mostrou que para K=2 ou K=3 (número de genes de que depende o estado de cada gene), uma rede de N genes pode gerar um certo número de ciclos estáveis (sob perturbação unitária) cujo comprimento é compatível com a duração de um ciclo de biossíntese celular real. Ciclos distintos (que passam por diferentes conjuntos de estados), estão associados a perfis diversos de síntese proteica, e portanto a células com características diferentes (diferenciadas).

Assim, a partir dos AKs é possível obter-se compartimentos distintos (células diferenciadas), cuja estabilidade e "tempo de vida", mimetizam aspectos importantes de processos fundamentais que ocorrem nos seres vivos, de modo que eles são capazes de gerar, desde que consideremos todos os estados iniciais possíveis, todos os tipos distintos de "células" com um dado número N de genes.

O número de estados internos dos ciclos (comprimento do ciclo) gerados, para N fixo, é função de K. Simulações numéricas indicam que, para $K \leq 2$, o comprimento dos ciclos é em geral pequeno, da ordem de $\beta(K)N^{1/2}$ enquanto que, para $K \geq 3$ cresce com $\exp[\alpha(K)N]$, tendo sido demonstrado que existe um valor crítico $K_c=2$, onde $\beta(K)$ e $\alpha(K)$ têm um comportamento singular que marca a transição entre dois tipos muito diferentes de comportamentos . Para $K \geq 3$, o comprimento dos ciclos mais populosos depende fortemente das condições iniciais, caracterizando um regime caótico enquanto que, para K=2, o tamanho desses ciclos depende apenas de N. Este último resultado, obtido analiticamente (Derrida e Pomeau, 1986; Derrida e col., 1992; Sole e Luque, 1995), bem como um resultado

numérico obtido anteriormente (Kauffman, 1969 e 1984; Derrida e Stauffer, 1986; Stauffer, 1987), constituem os motivos pelos quais passamos a utilizar K=2. A interpretação que pode ser dada a isto, é que estamos supondo que o estado funcional de cada gene deve depender em média dos estados de outros dois genes. O que temos até agora, nos permite afirmar o seguinte sobre os AKs:

(i) todos os 2^N estados iniciais possíveis de uma rede com N genes são equiprováveis, uma vez que não há qualquer razão para que exista um estado preferencial;

(ii) todas as 2^{2^K} funções Booleanas (f_i) possíveis são também equiprováveis por razões análogas,

(iii) dados N e K, podemos, em princípio, determinar a probabilidade PG com que cada ciclo é gerado.

Para representar uma célula diferenciada, que é um objeto dotado de alguma estabilidade, o ciclo deve ser estável, pelo menos em relação a perturbações unitárias. Isto quer dizer que, ao trocarmos o valor de um sítio de um estado do ciclo escolhido ao acaso, deve ser expressiva a probabilidade dele reconstituir algum estado original do ciclo depois de submetido à mesma dinâmica (f_is). Simulações indicam que em geral o ciclo perturbado retorna ao original cerca de 85% a 95% das vezes.

Para representar o estágio "final" da dinâmica ontogenética, vamos levar em conta todas as "células" que evolvem no sistema com estabilidade acima 85% - cada uma com uma probabilidade característica igual a P_G - e considerar que a produção, tanto de novas células iguais às que já estão no sistema, quanto de tipos celulares ainda não presentes, continua sem cessar, e consiste num processo de sorteio que ocorre a intervalos discretos de tempo de modo equivalente a retiradas de uma bola de uma urna.

Ciclos instáveis representarão células indiferenciadas que certamente ocorrerão durante o desenvolvimento do indivíduo, e ciclos estáveis representarão

células em seu estágio final de diferenciação. Esperamos que estejam presentes no adulto um número grande de células diferenciadas, e eventualmente um pequeno número de células indiferenciadas.

Para que se conclua a construção do modelo para a embriogênese e a diferenciação celular, é preciso pensar um mecanismo moderador do crescimento populacional e de tipos celulares.

(ii) A Questão do Crescimento e sua Moderação

O processo de desenvolvimento embrionário dos seres vivos sexuados, é muito complexo (Gilbert, 1994). Geralmente ocorre um grande crescimento logo após a fecundação, tendo a sua taxa diminuída a medida que o tempo passa, chegando finalmente a situação de crescimento zero. A partir daí, o número total de células presentes no indivíduo permanece aproximadamente constante, oscilando em torno de um certo valor médio característico para cada espécie, embora continue havendo produção de células novas no indivíduo adulto. Isto pode, em princípio, ocorrer de duas maneiras: morte celular e/ou retirada do sistema da célula recém gerada. Em ambos os casos ocorre a participação do Sistema Imune (SI) - por isto ele é o primeiro candidato a fornecer as pistas para que se possa pensar um dispositivo de moderação do crescimento.

(iii) O Sistema Imune

Sabemos que uma das funções mais importantes do SI consiste em discriminar, por meio de complexas sinalizações (Kindt & Capra, 1984), quais células fazem parte de um certo grupo (self) e devem permanecer no indivíduo, e quais (non-self) devem ser destruídas, ou retiradas como por exemplo: bactérias, vírus e outros agentes patogênicos, além das células tumorais que eventualmente emerjam no sistema vivo. Assim, partimos para uma avaliação das possibilidades de

compatibilização formal de alguns modelos de Sistema Imune (SI) com o modelo de embriogênese e diferenciação celular que tem como ponto de partida os AKs.

Com este objetivo, foi examinado o modelo KUT (Kaufman et al., 1985; Neumann, 1989) com a esperança de que fosse possível descrever adequadamente as condições de contorno capazes de moderar o crescimento populacional e de tipos celulares na embriogênese. Este modelo pareceu atraente uma vez que considera como unidade de interação do SI, um grupo de 5 objetos com especificidade bem determinada, que por sua vez, é descrito pela quíntupla [B1(t),B2(t),H(t),S(t), A(t)], onde cada elemento adquire valor zero ou um, conforme sua concentração assuma valores acima de um certo limiar fisiologicamente significativo. Estão representados respectivamente: linfócitos B virgens, linfócitos B maduros, linfócitos T auxiliares, linfócitos T supressores, e anticorpos. Ele estabelece também um conjunto de equações que regem a evolução temporal de um autômato celular que representa as transformações experimentadas pelos diversos objetos que constituem o SI quando "perturbado" pela introdução de um antígeno, cuja concentração atual é designada por $E_n(t)$, onde: $E_1(t)=0$ se ela estiver abaixo de um primeiro limiar, $E_1(t)=1$ e $E_2(t)=0$ se estiver acima do primeiro e abaixo de um segundo limiar, e finalmente, $E_1(t)=1$ e $E_2(t)=1$ se estiver acima do segundo limiar. As equações são:

$$
\begin{aligned}
B_1(t+1) &= \#E_1(t) \\
B_2(t+1) &= B_1(t)*E_2(t)*H(t) \\
H(t+1) &= E_1(t)*\#S(t) + H(t) \qquad (3) \\
S(t+1) &= H(t) + S(t) \\
A(t+1) &= B_2(t)*E_2(t)*H(t)
\end{aligned}
$$

onde o símbolo # troca o valor da expressão que o sucede, * é o produto usual de dois números e + corresponde ao operador lógico *ou*. Na verdade, estas equações

determinam as transformações pelas quais irá passar um compartimento muito pequeno do SI quando perturbado pela presença de antígenos. Por exemplo, o estado inicial (10000), que corresponde à existência de um linfócito B específico presente, é levado ao estado (10110) pelas equações (3), se considerando a adição de um antígeno em nível intermediário. Isto mimetiza uma resposta de sensibilização imunológica, pois se (10110) for perturbado novamente, produz o estado (11111) que corresponde à situação na qual emergem no sistema, linfócitos B, células T-helper, células T-supressoras e anticorpos.

Pensou-se inicialmente na utilização deste esquema de SI para promover a retirada de células, de modo a limitar não só o número delas, mas também, a sua diversidade efetivamente presente no ser vivo. Aqui deparou-se com algumas questões fundamentais:

(a) a partir de que momento da embriogênese poderíamos contar com este mecanismo?

(b) caso fosse conhecida a resposta de (a), como operacionalizar o modelo KUT num contexto de AKs ?

(c) como poderíamos pensar o desenvolvimento de seres vivos que não possuem um SI complexo o suficiente de tal modo que façam sentido os pressupostos do modelo KUT?

A questão mais simples é esta última, pois poderíamos dizer que estamos produzindo um modelo de desenvolvimento apenas para seres que já tivessem adquirido (filogeneticamente) um SI semelhante, em grau de complexidade, ao dos vertebrados. No entanto isto de nada adiantaria, pois a primeira questão ainda não foi respondida, se é que pode ser colocada naqueles termos. Quanto à segunda, é preciso que se imagine um dispositivo especial de crescimento - de geração de tipos celulares novos e de aumento populacional - cuja estrutura formal, de descrição matemática, seja compatível com as equações (3).

Behn e van Hemmen (1989) propuseram uma descrição do SI considerando o modelo de rede idiotípica cuja dinâmica é determinada por equações do tipo Lotka-Volterra, admitindo a existência de um "segundo" mecanismo de memória imune: as células B "dormentes". Toda vez que um linfócito B de determinada especificidade fosse estimulado a se reproduzir, algumas células da geração seguinte entrariam no estado dormente, cujo tempo de vida seria várias ordens de grandeza maior que o usual.

As equações propostas para descrever as interações de um dado conjunto de constituintes do SI são:

$$\dot{x}_i = x_i\left(\sum_{j=1}^{N} M_{ij}x_j - g\right) + d_i \sum_{j=1}^{N} m_{ij}x_j + (d_i + x_i) \sum_{j=N}^{N+R} m_{ij}y_i, \qquad i=1,\ldots,N \qquad (3.1)$$

e,

$$\dot{y}_i = y_i\left(a - \sum_{j=1}^{N} m_{ji}x_j\right), \qquad i=N+1,\ldots,N+R \qquad (3.2)$$

onde:

x_i é o número de anticorpos de uma dada especificidade (i),

y_j é o número de antígenos caracterizados pelo epitopo e_j, sendo $j=N+1,\ldots,N+R$,

m_{ij} a "constante" de reação entre o epitopo e_i e o paratopo p_j,

k um fator que introduz uma assimetria entre a estimulação e a inibição da produção de anticorpos,

g é o inverso da vida média dos anticorpos na ausência de estímulo,

d_i representa a fonte geradora de células B dormentes do tipo i, e

$M_{ij}=m_{ij}-km_{ji}$.

Levando em conta o fato de que a lista das variáveis e parâmetros destas equações está submetida a um processo dinâmico, (3.1) e (3.2) devem ser concebidas como imersas num esquema hierárquico que controla a geração das novas variáveis

e a dinâmica dos parâmetros. Isto coloca esta representação do SI dentro do figurino das redes idiotípicas, tornando-a, neste sentido, mais completa que o modelo KUT. Parece possível também cogitar-se um caminho para a construção de uma resposta à questão (b) levantada anteriormente, admitindo-se que as novas variáveis pudessem eventualmente ser geradas a cada passo com uma probabilidade igual à P_G. Infelizmente, apenas isto não é suficiente para que se conclua o modelo de desenvolvimento - as outras questões ainda estão postas e não têm solução à vista.

O fracasso em conciliar os modelos de SI com um modelo de desenvolvimento baseado nos AKs, fez com que fosse relaxada a exigência de um nexo imediato e visível com a fisiologia. Com este espírito, passou-se a investigar a utilização do modelo de criticalidade auto-organizada que oferece, algumas possibilidades de moderação, ou de limitação de contingente populacional, como veremos a seguir.

(iv) Criticalidade Auto-Organizada

Bak e col. (1988) propuseram um modelo, que posteriormente passou a ser designado como modelo BTW, para descrever a evolução espaço-temporal de sistemas dinâmicos complexos, com uma grande quantidade de graus de liberdade, que evoluem naturalmente para estados críticos através de um processo denominado criticalidade auto-organizada. O motivo pelo qual o sistema evolue para o estado crítico auto-organizado é porque ele é "um atrator da dinâmica".

Para compreender o modelo BTW, consideremos um autômato celular, que se presta a representar, entre outras coisas, um modelo de transporte em uma dimensão. Bak, Tang e Wiesenfeld sugeriram que a construção de um monte de areia, colocando-se um grão de cada vez, é um bom exemplo de processo que exibe criticalidade auto-organizada (SOC). Ele pode ser visualizado, em uma dimensão, como um conjunto de N pilhas a serem construídas lado a lado, ao longo de uma linha sobre a qual serão rotuladas N posições. A dinâmica do processo é

determinada pela adição, a cada instante, de um grão de areia no sítio (n) escolhido aleatoriamente fazendo com que o sistema se modifique assim:

$$h(n) \rightarrow h(n)+1 \qquad (4.1a)$$

feito isto, verifica-se a diferença entre o número de grãos de areia do n-ésimo sítio e o seu vizinho da direita :

$$z(n)=h(n)-h(n+1)$$

se z(n) for maior que um certo valor *lim*, ocorrerá o deslizamento de um grão de areia para a direita, de acordo com:

$$h(n) \rightarrow h(n)-1$$

$$h(n+1) \rightarrow h(n+1)+1$$

e

$$z(n) \rightarrow z(n)-2$$

$$z(n+1) \rightarrow z(n+1)+1 \qquad (4.2)$$

$$z(n-1) \rightarrow z(n-1)+1$$

caso contrário, será sorteado mais um sítio ao qual será adicionado um grão de areia, de modo que:

$$z(n) \rightarrow z(n)+1$$

$$z(n-1) \rightarrow z(n-1)-1$$

Pode-se impor as mais diversas condições de contorno, como por exemplo:

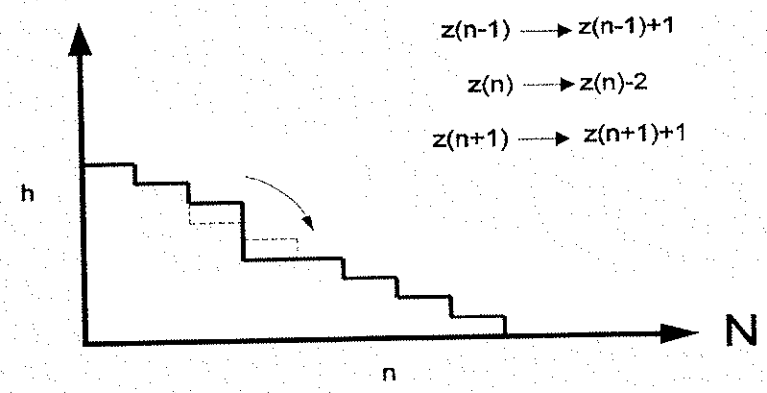
$$z(0)=0$$

e

$$h(N)=0 \qquad (4.3)$$

que representa a situação em que a fronteira esquerda está fechada e a direita aberta, conforme ilustra a fig. 2.2.

fig. 2.2. [reproduzida do art. cit. pg.15]



A medida que vão sendo colocados aleatoriamente mais grãos de areia, poderão formar-se eventualmente pequenos montes de areia. Se, em alguns deles, o limiar *lim* tiver sido atingido para muitas pilhas vizinhas, a adição de mais um grão de areia poderá desencadear uma avalanche cujo alcance poderá se estender até a outras partes do sistema. No caso limite, quando z(n)=*lim* para todo n, a adição de mais um grão de areia provocará um deslocamento de um sítio para o seu vizinho, e assim por diante, até que chegue um grão de areia ao N-ésimo sítio - este é certamente o estado menos estável de todos os estados estacionários, uma vez que qualquer perturbação produzirá necessariamente uma avalanche, isto é, uma sucessão de quedas de grãos de areia que culminará com a retirada de um grão de areia através da borda direita do sistema caso ela seja aberta. Para Per Bak, este estado final, que é denominado estado minimamente estável, é crítico no sentido estrito em que qualquer pequena perturbação poderá eventualmente se propagar "infinitamente" pelo sistema, enquanto que uma inclinação menor poderia evitar isto. Este estado crítico não tem uma estrutura espacial e as funções de correlação são triviais. No entanto, é possível obter-se autômatos de pilha de areia com

comportamento bem mais rico modificando-se a função z(n) e a regra que determina o modo como o sistema se estabiliza. De maneira geral, sempre que z(n) superar *lim* devemos ter:

$$h(n) \rightarrow h(n)-n_f(i) \qquad (4.4)$$

onde $n_f(i)$ representa o número de grãos de areia que cai do n-ésimo sítio. Definindo:

$$n_f(i)=N_F=cte \quad \text{(modelos limitados)} \qquad (4.5)$$

ou

$$n_f(i)=z(i)-N_S \quad \text{(modelos ilimitados)} \qquad (4.6)$$

Kadanoff et al. (1989) estudou o comportamento de quatro tipos de autômatos unidimensionais obtidos da combinação de (4.5) e (4.6) utilizando duas maneiras diferentes para a distribuição dos grãos de areia que caem. Na primeira eles são todos alocados para o vizinho a direita de acordo com:

$$h(i+1) \rightarrow h(i+1)+n_f(i) \quad \text{(modelos locais)} \qquad (4.7)$$

Na segunda, eles são distribuídos a direita da seguinte maneira:

$$h(i+j) \rightarrow h(i+j)+1 \quad \text{para } j=1,2,...,n_f(i) \qquad (4.8)$$

(modelos não locais)

Para cada avalanche, foram calculados o número de grãos D que abandona o autômato pelas suas bordas, e o número total de sítios envolvidos F. O comportamento numérico das funções de distribuição $\rho(D,N)$ e $\rho(F,N)$ para modelos limitados locais e não locais, e modelos ilimitados locais e não locais, indica a existência de um comportamento dinâmico não trivial, tendo sido possível determinar-se inclusive seus expoentes críticos .

O autômato de pilha de areia também pode ser definido para duas ou mais dimensões. Em duas dimensões ele pode ser representado como uma rede quadrada

de lado N, onde a pilha é construída sobre um "tabuleiro" quadrado de modo que $1 < i,j < N$, onde existem N^2 sítios aos quais são associadas as variáveis h(i) que correspondem ao número de grãos de areia que cada um contém num dado instante. São colocados aleatoriamente grãos de areia, sendo que aqui, a função z(i,j) é, em geral, definida de uma maneira um pouco mais complicada do que no caso unidimensional.

Tanto em uma quanto em duas dimensões, o comportamento do autômato depende de como é definida a função z(*), e dos critérios que irão determinar a forma como deverá ocorrer o transporte (a queda) de um grão de areia quando o valor da função z(*) superar o limiar *lim*.

Pode-se, por exemplo definir:

$$z(i,j)=4h(i,j)-h(i+1,j)-h(i,j+1)-h(i-1,j)-h(i,j-1) \qquad (5.1)$$

e adotar um critério não direcional para a queda de grãos de areia, impondo que caia um grão em cada um dos primeiros vizinhos de (i,j) de acordo com:

$$\begin{aligned} h(i,j) &\rightarrow h(i,j) - 4 \\ h(i+1,j) &\rightarrow h(i+1,j) + 1 \\ h(i,j+1) &\rightarrow h(i,j+1) + 1 \\ h(i-1,j) &\rightarrow h(i-1,j) + 1 \\ h(i,j-1) &\rightarrow h(i,j-1) + 1 \end{aligned} \qquad (5.2)$$

Há outras definições de z(*) possíveis que também levam a SOCs, como por exemplo a utilizada por Dhar (1990) que considera as alturas "locais" invés do gradiente.

Embora haja controvérsia se de fato esses autômatos se comportam da mesma maneira que pilhas de areia reais (Jaeger e col., 1989; Bretz e col. 1992; Rosendahl e col., 1993), eles têm sido úteis na compreensão de um grande número de fenômenos até então não relacionados como o transporte dissipativo em sistemas abertos (Hwa e Kardar, 1989; Hentschel e Family, 1991), a propagação de incêndios em florestas (Bak e Tang, 1990; Grassberger e Katz, 1991; Packuski e Bak, 1993), a propagação de abalos sísmicos (Bak e Tang, 1989; Sornette e Sornette, 1989; Prado e Tibúrcio, 1992), bioconvecção (Noever e col., 1994), processos Robin Hood (Zaitsev, 1992) entre outros.

Para este trabalho, um dos aspectos mais importantes dos processos que exibem criticalidade auto-organizada é que, uma vez atingido o SOC, qualquer "ruído", como o deslizamento de um grão de areia diagonal abaixo, se propagará eventualmente pelo sistema como um todo num efeito dominó. Se forem utilizadas condições de contorno abertas, temos que um grão de areia será necessariamente retirado do sistema, como se tivesse caído de uma mesa sobre a qual se colocam aleatoriamente grãos de areia. Mesmo que se continue a adicionar grãos de areia ao monte que está sobre a mesa, ele não cresce mais, pois já atingiu a sua inclinação limite, de modo que a população total de grãos de areia não sofrerá mais modificação substancial, devendo permanecer aproximadamente constante. A partir daqui, novos grãos de areia introduzidos no sistema serão transportados para fora dele.

Isto será utilizado neste trabalho para representar uma situação de crescimento com moderação, na qual, a partir de um certo instante a probabilidade de que as partículas sejam retiradas do sistema se iguala a probabilidade de que sejam geradas dentro dele, e ele pára de crescer sem que cesse a entrada de "partículas" novas.

UNIVERSIDADE DE SÃO PAULO
SERVIÇO DE BIBLIOTECA E INFORMAÇÃO
2334
e.2
INSTITUTO DE FÍSICA

Como será mostrado a seguir, este é um cenário compatível com a construção de um modelo para a embriogênese e a diferenciação celular baseado nos AKs.

3. A Construção do Modelo Global para a Embriogênese e Diferenciação Celular

(i) Pressupostos

Conforme foi sugerido no final do capítulo anterior, as redes genéticas aleatoriamente conectadas e a dinâmica da criticalidade auto-organizada, constituem as bases fundamentais nas quais se assenta o modelo que desenvolvemos para a embriogênese e diferenciação celular, mais explicitamente os pressupostos fundamentais do modelo são:

(i) o contingente populacional dos diversos tipos celulares distintos (diferenciados), presentes no estágio final do processo ontogenético, será determinado pela distribuição de probabilidades com que cada um deles é gerado por uma rede genética aleatoriamente conectada.

(ii) o crescimento do número total de células, independente dos tipos (diferenciados), é moderado pelo processo de criticalidade auto-organizada.

O modelo de Kauffman será utilizado para estimar a probabilidade com que um determinado tipo celular emerge no final do processo ontogenético. O número total de células presentes, independente da especificidade, será limitado pela dinâmica da criticalidade auto-organizada. O conjunto das células que compõem um indivíduo adulto será obtido dos NC tipos celulares que uma rede genética formal é capaz de produzir. A esses tipos, serão associados NC sítios de um plano sobre o qual se construirá um autômato de "pilha de areia".

Esta pilha de areia deverá atingir um estado crítico, a partir do qual a adição de novos grãos não produzirá mais um aumento da sua inclinação média, nestas circunstâncias, a escolha de condições de contorno "abertas", estabelecerá um limite para o número total NT de grãos que permanecerá no sistema, retirando os excedentes - estes serão transportados para fora do tabuleiro. Este estado

estacionário é um estado crítico auto-organizado e representará a fase final da ontogênese - o indivíduo adulto.

Uma limitação importante deste modelo é que ele não contém os meios para se calcular o limiar (*lim*) que determina a inclinação da pilha, e consequentemente a população total de células no final do processo. Isto terá que ser feito "ad hoc" através do estudo do desempenho do modelo aplicado a uma situação real conhecida. Será construído um autômato de "pilha de areia" para um animal (*Caenorhabdits Elegans*) cujo número total de tipos celulares distintos, bem como a população total de células na fase adulta é conhecida, para que se obtenha o valor de *lim* que o modelo prevê para ele.

(ii) O *Caenorhabdits Elegans*

O *C. elegans* (Sulston, J. E. & Horvitz, H. R. , 1977; Sulston et. al., 1983; Wood, W. B. ,1988) é um nematodo de vida livre de aproximadamente 1 mm de comprimento, cujo ciclo de vida dura 3,5 dias (a 20°C). Ele tem cerca de 1000 células formando tecidos diferenciados (nervos, músculos, hipoderme, intestino e gônadas), e é anatomicamente simples (Fig.3.1.) - seu corpo tubular consiste numa parede hipodérmica com a musculatura subjacente, que contém seus sistemas digestivo e reprodutivo. Uma das suas características mais interessantes é que ele é translúcido à luz visível, e o contorno dos núcleos de todas as suas células pode ser visto, ao microscópio óptico (Fig.3.2.) com o animal vivo, tornando-o um objeto privilegiado para a pesquisa da biologia do desenvolvimento, já que é possível, por exemplo, gravar em vídeo e acompanhar diretamente o destino de quase todas as suas células nos diversos estágios de desenvolvimento, desde o ovo recém fecundado até a fase adulta, passando pelos quatro estágios larvares que a precedem.

Graças a essas propriedades, o *C. elegans* é hoje o animal mais conhecido do ponto de vista desenvolvimental. Estão mapeadas as linhagens germinativas de

todas as suas células, quer no hermafrodita (811 células no adulto), quer no macho (971 células no adulto).

Estima-se que ele tenha cerca de 3000 genes presentes no seu genoma, dos quais aproximadamente 2000 sejam funcionais e os restantes aparentemente não funcionais. Uma vez conhecidos esses parâmetros, pode-se construir o cenário que o modelo prescreve para o *C. elegans*, e a partir dele, obter-se o valor de *lim*.

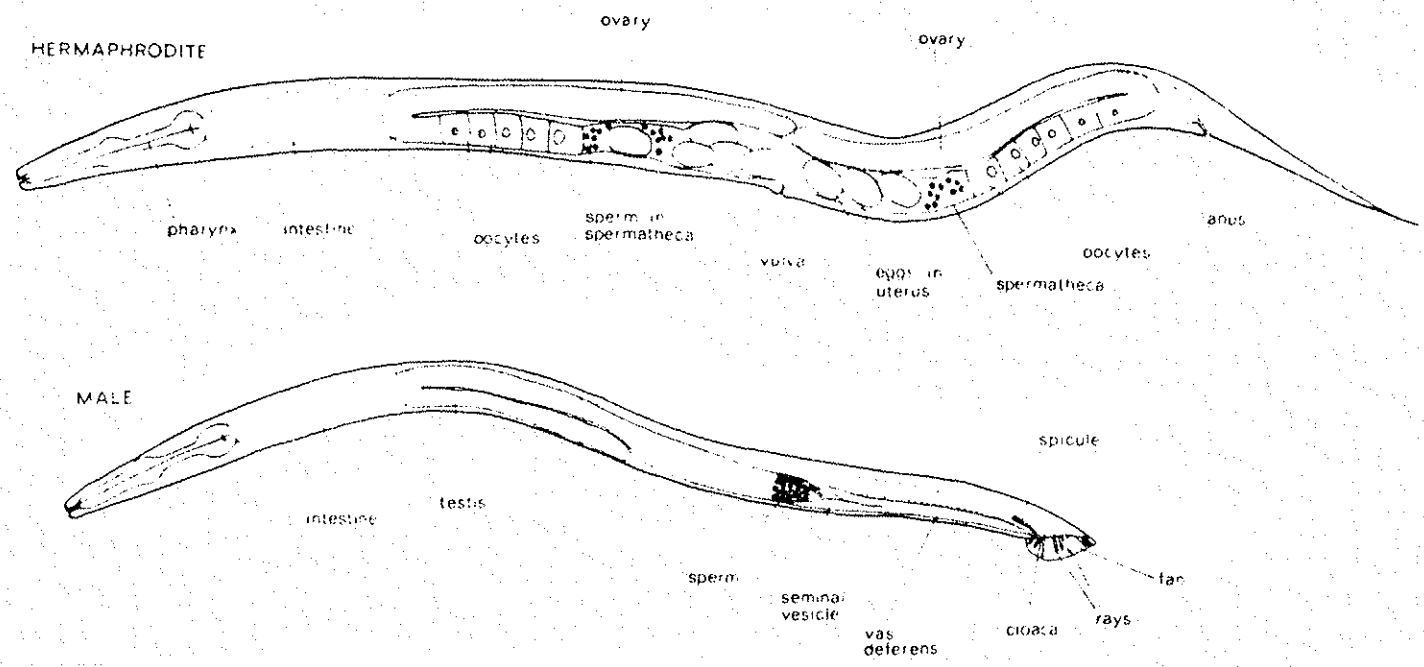


Fig.3.1.: Diagrama esquemático do *C. elegans* hermafrodita e macho

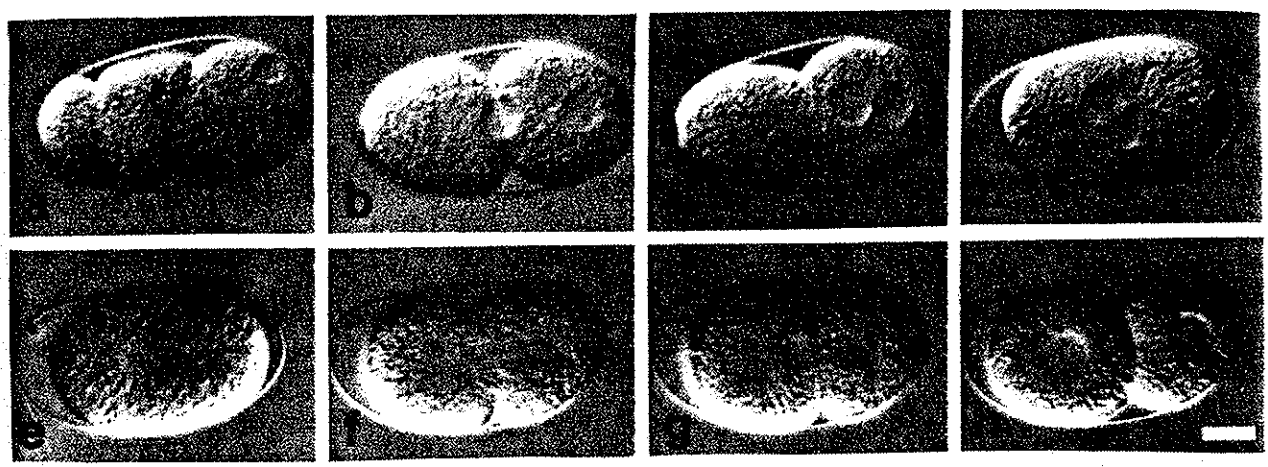
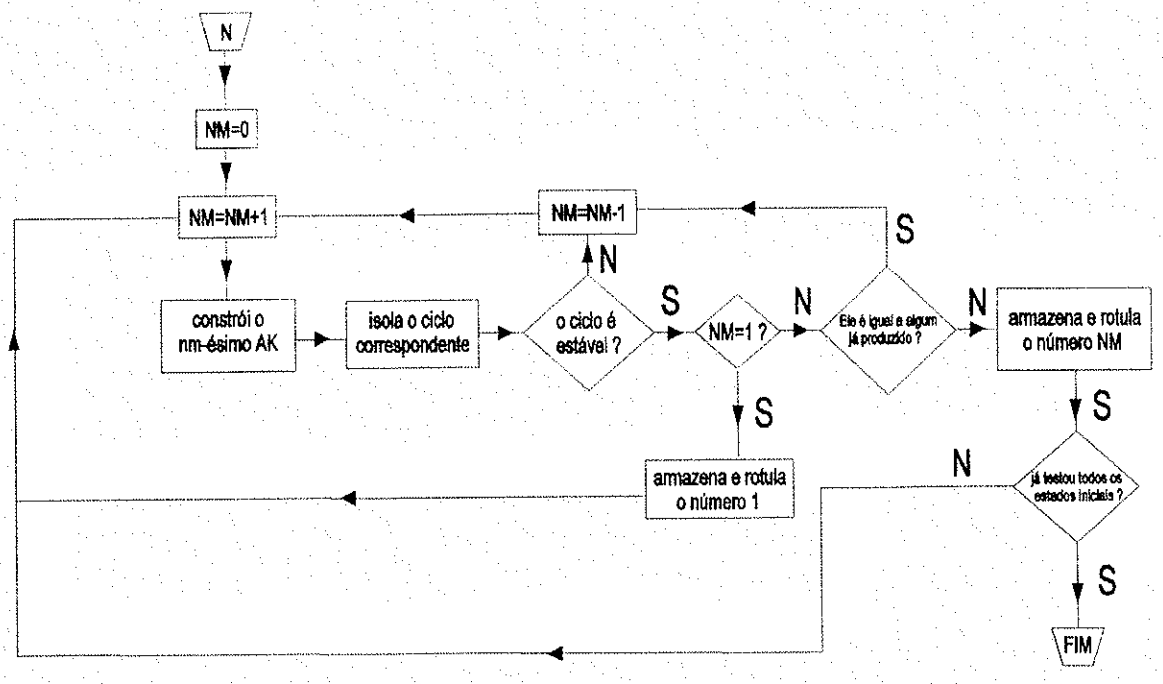


Fig.3.2.: Contorno das células do *C-elegans* nos primeiros estágios de desenvolvimento

(iii) Simulação Computacional

Do ponto de vista computacional, o trabalho foi dividido em duas etapas; uma para o estudo dos autômatos de Kauffman e outra para a obtenção dos parâmetros necessários à conclusão do modelo para o *C. elegans* que envolveu uma dinâmica de criticalidade auto-organizada. Os programas foram desenvolvidos em linguagem fortran 77 estruturado e foram processados num computador HP 9000 do DFGE e nos computadores do Centro Nacional de Processamento de Alto Desempenho (CENAPAD) na UNICAMP.

Na primeira etapa, o nosso objetivo foi identificar todos os ciclos estáveis distintos que os autômatos de Kauffman são capazes de gerar para um dado N. Para isto é necessário levarmos em conta todos os estados iniciais da rede genética, e determinarmos, para cada um, os ciclos estáveis que são gerados utilizando-se todas as combinações possíveis de funções Booleanas de duas variáveis. O esquema geral do programa é o seguinte:



Retomando o exemplo dado no item (iii)-a, onde N=4, o programa começa com a construção das matrizes:

$$A(IA,JA)=\begin{pmatrix} 0\,0\,0\,0 \\ 0\,0\,0\,1 \\ 0\,0\,1\,0 \\ \cdot\;\cdot\;\cdot\;\cdot \\ 1\,1\,0\,1 \\ 1\,1\,1\,0 \\ 1\,1\,1\,1 \end{pmatrix} \quad , \quad XIS(IX,JX)=\begin{pmatrix} 1\;1\;1\;1 \\ 1\;1\;1\;2 \\ 1\;1\;1\;3 \\ \cdot\;\cdot\;\cdot\;\cdot \\ 1\;2\;3\;4 \\ \cdot\;\cdot\;\cdot\;\cdot \\ 12\;12\;12\;11 \\ 12\;12\;12\;12 \end{pmatrix}$$

e,

$$R(IR,JR)=\begin{pmatrix} (2,3) & (1,3) & (1,2) & (1,2) \\ (2,3) & (1,3) & (1,2) & (1,3) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ (2,3) & (1,4) & (2,4) & (1,3) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ (3,4) & (3,4) & (2,4) & (2,3) \end{pmatrix}$$

que correspondem respectivamente ao conjunto de todos os estados iniciais possíveis - matriz A(IA,JA); ao conjunto de todas as combinações de 12 das 16 funções Booleanas possíveis - matriz XIS(IX,JX); e ao conjunto de pares ordenados que representam todas as possibilidades de pares de genes que constituem as variáveis das funções Booleanas - matriz R(IR,JR). Foram retiradas 4 funções cujos valores dependem apenas de uma variável, ou de nenhuma delas, como a tautologia e a

negação. Isto foi feito com o objetivo de diminuir o tempo de processamento do programa, sem perder de vista que a expressão de cada gene deve depender em média de outros dois.

Concluída esta etapa, entramos na construção do AK propriamente dito. A cada uma das IA linhas de A(IA,JA), matriz dos estados iniciais, são associadas cada uma das IX sequências de JX rótulos (JX=JA) das funções Booleanas, de modo que, a cada função é associada uma das IR linhas de R(IR,JR), cujos elementos são os pares de rótulos que indicam os sítios que terão seus valores utilizados como variáveis das funções Booleanas.

Em nosso exemplo, foi associada à linha (1,1,0,1) de A(IA,JA), a linha (1,2,3,4) de XIS(IX,JX), a qual, por sua vez foi associada à linha ((2,3),(1,4),(2,4),(1,3)) de R(IR,JR). Exatamente aqui, começa a construção da matriz AK(IK,JK), cuja primeira linha é a primeira linha de A(IA,JA), e a segunda é obtida, da primeira, do seguinte modo:

$$AK(2,JK)=f_{XIS(IX,JK)}[R(IR,JK)] \quad , \qquad JK=1,2,3 \text{ e } 4$$

onde: XIS(IX,1)=1, XIS(IX,2)=2, XIS(IX,3)=3 e XIS(IX,4)=4; R(IR,1)=(2,3), R(IR,2)=(1,4), R(IR,3)=(2,4) e R(IR,4)=(2,4). A matriz AK(IK,JK) contém os estados do autômato celular que está sendo gerado, e não pode ter linhas repetidas, por isso, é necessário que se compare a linha recém criada, com a(s) anterior(es), pois se houver duplicidade de linhas, teremos construído um ciclo, e deveremos interromper o processo. A matriz AK(IK,JK) obtida é a seguinte:

$$AK(IK,JK) = \begin{pmatrix} 1\,1\,0\,1 \\ 0\,1\,1\,0 \\ 1\,0\,1\,1 \\ 0\,1\,0\,0 \\ 0\,0\,1\,1 \\ 0\,0\,0\,1 \\ 1\,0\,0\,1 \\ 1\,1\,0\,0 \\ 0\,0\,1\,0 \\ 0\,0\,1\,1 \end{pmatrix}$$

que "pára" na décima linha, uma vez que encontrou-se uma linha idêntica à quinta, o que quer dizer que se completou a geração de um ciclo. Para que ele seja devidamente rotulado e armazenado, ele precisa ser estável sob perturbação unitária, de modo que mudando-se, um de cada vez, o valor de cada um dos sítios que o constituem, e aplicando-lhe as mesmas funções Booleanas, devemos ter boas chances de recuperar algum estado do ciclo original. Passando neste teste, é necessário ainda saber se não estamos gerando uma duplicata de um ciclo que já foi armazenado, para isto comparamos o ciclo recém gerado com todos os outros que já existem. Caso seja distinto daqueles anteriormente gerados, este ciclo será armazenado na matriz:

$$C(I,J,NM) = \begin{pmatrix} 5 & NM & NR & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

representada apenas para um dado valor de NM, que é o número total de ciclos estáveis distintos, produzidos até o momento. A matriz C(I,J,L) constitui uma espécie de arquivo, onde serão armazenados todos os atratores (ciclos estáveis) gerados, indicando o rótulo NM, o número de estados internos, 5 neste caso, e o número NR de vezes que cada um é gerado.

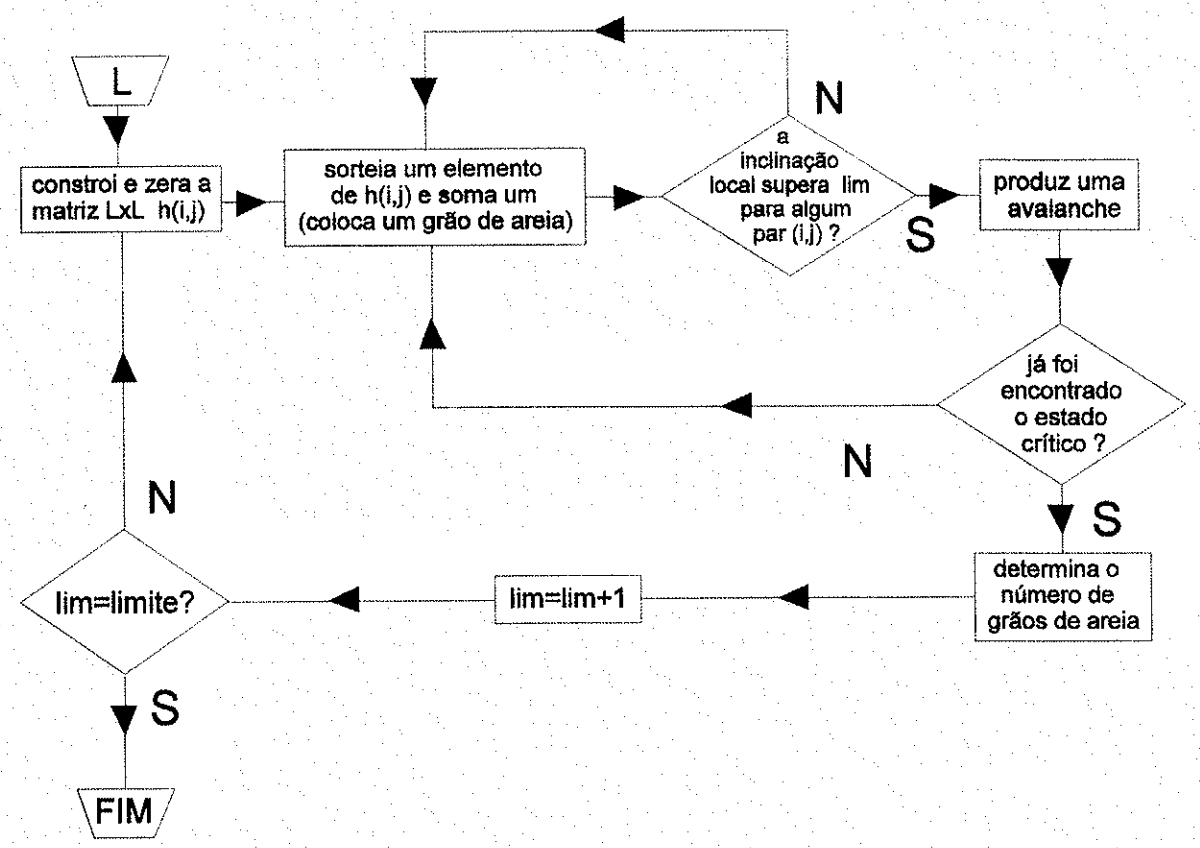
Para N genes há exatamente 2^N estados iniciais, N^{12}-uplas de funções Booleanas sendo utilizadas, e NP ($NP=N^2-3N+2$) conjuntos distintos de N pares de variáveis. O processamento para N=3 utiliza uma matriz XIS de 4096 linhas, sendo que para N=4 esse número sobe para 65536. A partir de N=5 (5 genes formais), é impraticável a simulação determinística da rede, pois o tempo de processamento cresce exponencialmente com N.

Assim, o estudo das redes maiores teve que ser feito através de uma adaptação do programa, incorporando a ele um processo de amostragem e varredura aleatória que consiste na realização de várias sequências de um certo número de sorteios de estados iniciais, onde cada estado inicial consiste em: (a) uma N-upla ordenada contendo apenas zeros e uns, (b) uma sequência ordenada de N funções Booleanas escolhidas entre as que estão sendo utilizadas e (c) um conjunto de N pares ordenados correspondentes às variáveis das funções Booleanas. No exemplo utilizado no capítulo 2, o estado inicial é constituído da quádrupla (1,1,0,1),

pela sequência de funções f_1, f_2, f_3 e f_4 e pelos pares ordenados (2,3), (1,4), (2,4) e (1,3)

Para N=10, 20, 30, 64 e 128, foram feitos testes para ajustar o número de varreduras e de sorteios de modo que não se dispendesse muito tempo.

A segunda etapa do trabalho computacional consistiu na elaboração de um programa para construir um autômato de pilha de areia com L^2 sítios e fronteiras abertas, que determinasse para cada valor do limiar (lim) *lim*, a população de grãos de areia no estado crítico auto-organizado. O esquema do programa é o seguinte:



Para construirmos o autômato de pilha de areia, consideramos uma região plana contendo os L^2 sítios correspondentes aos tipos celulares distintos do organismo. Associamos a ela a matriz h(i,j), cujos elementos representam o número

de grãos de areia presentes no lugar com rótulo (i,j), e uma função z(i,j) que mede a inclinação local em (i,j), de acordo com a expressão (5.1), e procedemos da seguinte maneira:

(i) sorteamos um sítio sobre o qual é colocado mais um grão de areia,

(ii) verificamos se a inclinação local z(i,j) supera *lim* para cada sítio (i,j),

(iii) se z(i,j)>*lim*, deverá ocorrer uma avalanche, que consiste na transferência de grãos de areia para as pilhas primeiras vizinhas mais baixas, até que não reste qualquer sítio com z(i,j)>*lim*,

(iv) se z(i,j)≤*lim*, voltamos ao item (i) até que seja atingido o estado crítico,

(v) os procedimentos (i)-(iv) serão repetidos para valores crescentes de *lim* até que este atinja um certo máximo previamente estipulado.

4. Resultados

(i) As Redes Pequenas

Foram feitas simulações determinísticas em redes com 3 e 4 genes formais, utilizando-se todos os estados iniciais possíveis. Os resultados obtidos estão representados nas tabelas 4.1.1 e 4.1.2 e nos gráficos 4.1.1 e 4.1.2. Para N=3, há uma expressiva maioria dos ciclos com 5 estados. A mediana da distribuição se situa em 4 estados e coincide com a média. Para N=4, o máximo se localiza em 6 estados, que corresponde à média do número de estados, ficando a mediana em 5 estados.

Tabela 4.1.1
(Simulação determinística para 3 genes)

Comprimento do Ciclo	Número de Ciclos
1	8
2	10
3	24
4	51
5	96
6	6
7	12

Tabela 4.1.2
(Simulação determinística para 4 genes)

Número de Estados	Número de Ciclos
1	16
2	88
3	544
4	3316
5	8736
6	9936
7	9384
8	6624
9	1248
10	384
11	0
12	0
13	0
14	336
15	336

Com a intenção de estabelecermos parâmetros para o estudo das redes maiores, para as quais não foi possível uma simulação determinística, fizemos várias sequências de varreduras aleatórias para redes com 3 e 4 genes. Começamos por considerar que não há razão para supormos a existência de estados iniciais

privilegiados, de modo que admitimos que tanto as N-uplas iniciais de zeros e ums, quanto as funções Booleanas e as sequências de variáveis têm igual probabilidade *a priori*. Pensamos inicialmente que o problema de obter os ciclos estáveis, fosse semelhante ao de conhecer o número de rótulos distintos de um dado com uma quantidade enorme de lados rotulados, com um número de rótulos bem menor que o de lados, lançando-o um certo número de vezes. Parece plausível esperar que a confiabilidade do resultado seja tanto maior, quanto maior for o número de lançamentos, e que exista um número crítico de vezes que o dado é lançado, para o qual é máxima a probabilidade dos resultados terem percorrido todos os rótulos diferentes. Tendo isto em mente, processamos o programa para um número limitado de estados iniciais equiprováveis, escolhidos aleatoriamente.

Fizemos então a simulação efetuando varreduras com $5x10^2$, 10^3, $2x10^3$, $4x10^3$, $6x10^3$, $8x10^3$, 10^4 e $1,2x10^4$ sorteios de estados iniciais para N=3; e 10^3, $2x10^3$, $4x10^3$, $8x10^3$, $16x10^3$, $32x10^3$, $64x10^3$, $128x10^3$, $256x10^3$, $512x10^3$, $1024x10^3$, $2048x10^3$, $4x10^6$, $5x10^6$, $6x10^6$, $7x10^6$ e $8x10^6$ sorteios para N=4, com o objetivo de estabelecer o número mínimo de sorteios para chegarmos a distribuição populacional da situação determinística.

Os resultados, apresentados nos gráficos 4.1.3 e 4.1.4, mostram que foram necessários cerca de $4x10^3$ sorteios para N=3 - 29% do total dos estados iniciais - e $5x10^6$ para N=4, que corresponde a um número quase duas vezes e meia maior que o número de estados iniciais (1 990 656). Assim, fica evidenciado que o método das verreduras aleatórias é muito pouco "econômico" para determinar o número total de ciclos estáveis produzidos por essas redes. No entanto, esses resultados indicam a existência de uma distribuição populacional aparentemente "bem comportada", na medida que todas as simulações para N=3 mantiveram os respectivos picos no mesmo lugar, e as para N=4 que tiveram um deslocamento do pico de 4 estados, obtidos com um número de sorteios de estados iniciais entre 1000 e 8000 , para 5

estados entre $16x10^3$ e $1024x10^3$ sorteios, chegando finalmente em 6 estados acima de $1024x10^3$ sorteios. Para N=5, as varreduras efetuadas com $2x10^4$, $5x10^4$, 10^5, $5x10^5$, 10^6 e $2x10^6$ sorteios de estados iniciais mantiveram o pico em 5 estados a partir de $5x10^5$ sorteios. Desses resultados, ilustrados nos gráficos 4.1.3, 4.1.4 e 4.1.5, podemos inferir duas consequências importantes: (i) que o número de ciclos estáveis obtidos antes de atingida a saturação, depende do número de sorteios realizados, e (ii) que os extremos da distribuição conservam sua localização dentro de uma certa faixa (5 estados para N=3, de 4 a 6 estados para N=4, e em 5 estados para N=5).

(ii) As Redes Maiores

A intuição obtida do estudo das redes menores (3, 4 e 5 genes formais), fez com que começássemos a investigação das redes maiores (10, 20, 30, 64 e 128 genes formais), empregando o método das varreduras aleatórias, na esperança de detectar alguns padrões de comportamentos dessas redes, uma vez que há indicações na literatura (cit. pg. 11, 16; Kauffman, 1974) de que seus atratores ocupam uma região extremamente pequena do seu espaço de estados.

Os resultados para N=10, apresentados nos gráficos 4.2.1 e 4.2.2, foram obtidos fazendo-se a superposição de 47 sequências de varreduras de 50000 sorteios, e como se pode ver, sugerem que a forma de cada um parece ser independente dos conjuntos de estados iniciais utilizados para obtê-los. O mesmo estudo preliminar foi feito para N=20 (50 varreduras de 5000 sorteios), N=30 (50 varreduras de 1000 sorteios), N=64 (30 varreduras de 50 sorteios), e para N=128 (13 varreduras de 50 sorteios), tendo sido obtidos resultados análogos que podem ser verificados nos gráficos 4.2.3, 4.2.4, 4.2.5 e 4.2.6.

Foi feito um estudo mais detalhado deste fenômeno com 98, 91 e 71 varreduras respectivamente para N=10, 20 e 30, para N=64 e para N=128, mantidas as quantidades de sorteios realizadas no início. Para todos esses valores de N foram

determinados os ciclos estáveis produzidos, e medidos os respectivos tamanhos. Com esses resultados construíram-se os gráficos 4.2.7, 4.2.8, 4.2.9, 4.2.10, 4.2.11 e 4.2.12, onde estão assinaladas as barras de erro para cada contingente populacional, indicando uma confirmação dos resultados preliminares. Observando esses gráficos, é possível verificar que os ciclos mais populosos estão numa faixa que compreende ciclos de 1 a 10 estados, centrada aproximadamente em torno de 4 estados para N=10; ciclos de 1 a 20 estados, centrada aproximadamente em 4 estados para N=20; em 5 para N=30; em 8 para N=64. Em N=128, a situação é bem menos nítida e por isso, não foram feitas estimativas a partir desse gráfico. Um outro resultado interessante destas simulações é que, a partir de N=10, não se verificou nenhuma repetição de ciclos.

O nosso estudo das redes genéticas aleatoriamente conectadas, foi realizado com uma sistemática diferente da utilizada por Kauffman. Em seu artigo de 1969 que serviu de referência para todos os seus estudos posteriores, ele relatou que para estudar a distribuição dos comprimentos dos ciclos em função do tamanho da rede (N) foram feitas 100 simulações para N=15, 50, 64, 100, 191, 400, 1024, 4086 e 8191 elementos; e para estudar o número de ciclos distintos produzidos por rede, foram feitas simulações com 50 sequências iniciais de zeros e uns para as redes com até 400 genes. Isto posto, afirmamos que o método das varreduras aleatórias (MVA) por nós utilizado, mesmo impondo uma restição maior ao tamanho das redes investigadas, já que o tempo de processamento parece crescer exponencialmente com N, proporciona resultados mais confiáveis, por que é capaz de percorrer uma região bem maior do espaço de estados das redes

Embora o MVA não seja adequado para a determinação das populações totais de ciclos estáveis produzidos, ele se mostrou útil na localização aproximada dos ciclos mais populosos, confirmando para as redes maiores (pelo menos até 64

genes), a tendência já verificada para redes pequenas de consevar a posição dos picos populacionais.

A conservação da posição dos picos populacionais ou, falando de outro modo menos preciso, porém mais ilustrativo, a conservação da forma dessas distribuições tem pelo menos dois significados muito importantes para este trabalho. O primeiro é que o número de estados iniciais utilizados foi estatisticamente significativo, e o segundo é que essas distribuições são bastante uniformes. Esta conclusão se deve ao fato que certamente não são os mesmos ciclos que estão sendo sorteados todas as vezes, de modo que a repetição do número de ciclos distintos reflete necessariamente a existência de uma distribuição uniforme desses ciclos em todo o espaço amostral. Para melhor compreender o significado disto, voltemos à analogia com o lançamento do dado onde ciclos distintos, com o mesmo número de estados, corresponderiam, por exemplo, a faces com o mesmo rótulo, mas com cores diferentes. Assim, a realização de *nx* sequências com *ny* estados iniciais significa fazer *nx* lançamentos de *ny* dados idênticos, e tabular para cada uma delas o número de cores distintas sorteadas para cada rótulo. Se lançarmos de cada vez um número suficientemente grande de dados, podemos obter sempre o mesmo número de cores distintas para o mesmo rótulo.

(iii) Saturação Populacional

Conforme já mencionamos, a saturação populacional será obtida através da construção de autômatos de pilha de areia com extremidades abertas. Nosso objetivo aqui é estudar a relação entre a variável *lim* e o número de grãos de areia presentes no momento em que ocorre a saturação do autômato de pilha de areia, e verificar se existem valores compatíveis com os os dois parâmetros mais importantes do *C. elegans* para o modelo que estamos desenvolvendo: o número total de células

presentes no indivíduo adulto - 959 no macho e 811 no hemafordita - e o número de tipos celulares distintos que ele apresenta - cerca de 70.

Foram feitas simulações para redes quadradas correspondendo a L^2 tipos celulares distintos presentes no indivíduo adulto - com L=3 para 15 valores de *lim*, L=4 para 19 valores, L=5 para 11 valores, L=6, 7, 8 e 9 para 7 valores. As pilhas foram construídas sorteando-se um sítio a intervalos discretos de tempo e colocando-se um grão de areia sobre ele, representando o aparecimento de uma nova célula no animal . Os L^2 sítios foram considerados equiprováveis conforme sugerem os dados obtidos com os autômatos de Kauffman.

Os resultados apresentados nos gráficos 4.3.1, 4.3.2, 4.3.3, 4.3.4, 4.3.5, 4.3.6 e 4.3.7, mostram claramente que há um patamar de saturação populacional bem definido, como era de se esperar. Este patamar corresponde ao que chamaremos de população crítica (Pc), que consiste num valor médio, em torno do qual vai flutuar a população de grãos de areia, mesmo que continuemos a colocá-los no sistema.

Esta situação é análoga a que ocorre na maioria dos seres vivos onde a população de células em indivíduos adultos oscila em torno de um valor médio, que é característico de sua espécie, embora ele continue a produzir células novas, de modo que o fluxo de saída de células tem média temporal igual ao fluxo de entrada. Estamos propondo que isto reflete uma propriedade coletiva do "cluster" de células que constitui um ser pluricelular, decorrente do fato de existirem interações entre elas que estão submetidas a uma dinâmica de criticalidade auto-organizada.

Na tabela 4.3.1 estão mostrados os valores de Pc para 7 valores de L e 6 de *lim*. O mais próximo de 959, compatível com o *C. elegans* é 1176 para *lim*=3 e L=8, que corresponde a 64 tipos diferentes de células. O fato do modelo prever um número de células maior do que ocorre, pode ser contornado se considerarmos que não pretendemos que a criticalidade auto-organizada seja o único moderador do crescimento da população celular. Propomos apenas que ela seja um dos

mecanismos importantes para o estabelecimento do patamar populacional. Certamente vários tipos diferentes de interações entre os muitos sub-sistemas que compõem os sistemas vivos influenciam o processo, estabelecendo limitações a ele.

Uma vez identificado os parâmetros do *C. elegans* na classe de autômatos com lim=3, julgamos conveniente estudar o comportamento de Pc em função de L com o objetivo de investigar as possibilidades de previsão do modelo. Para isto fizemos os gráficos 4.3.8, 4.3.9, 4.3.10, 4.3.11, 4.3.12, 4.3.13 e 4.3.14 que ajustaram retas com coeficientes de correlação acima de 0,99914. Assim, encontramos:

$$Pc(L)=C.L^{3,4}$$

para *lim*=3, C=0,9575.

Tabela 4.3.1
(Valores de Pc em várias situações)

lim / L	2	3	4	5	6	7	8
3	26±4	43±8	67±8	94±10	126±8	154±8	186±8
4	53±6	106±10	175±10	250±8	332±12	410±12	485±12
5	94±12	220±16	381±16	548±16	715±18	885±16	1049±16
6	156±14	411±20	728±18	1042±20	1357±16	1672±18	1993±10
7	247±18	718±26	1267±24	1813±30	2366±15	2912±12	3460±13
8	367±24	1176±12	2067±30	2952±19	3835±20	4718±17	5611±16
9	521±24	1799±36	3166±20	4526±20	5887±18	7248±20	8605±20

5. Conclusão

Um objeto que pretenda representar um sistema vivo deve exibir uma certa estabilidade em relação a "pequenas" perturbações oriundas do meio externo, de modo que possa constituir algo bem definido no espaço e no tempo. Todos os seres vivos têm uma identidade estrutural e funcional que se mantém durante algum tempo em que podem ser nomeados e distinguidos de seus semelhantes. Por exemplo, bactérias podem, em princípio, ser rotuladas, mantendo a identidade durante um intervalo de tempo, eventualmente muito curto. Isto ocorre com qualquer animal que possamos imaginar. No entanto, esta estabilidade, de curto alcance (ontogenética), se descaracteriza quando pensamos em intervalos de tempo muito maiores que o tempo médio de vida de um indivíduo, ou seja, em escala filogenética. Nesta escala, os seres vivos podem ser vistos como sistemas capazes de acumular variações adaptativas (mutações) capazes de mudar efetivamente o conteúdo informacional que determina a ontogênese, produzindo indivíduos com diferenças (genéticas) eventualmente fundamentais no processo de adaptação. A evolução, selecionando os indivíduos mais ajustados, vai através do tempo alterando o repertório de possibilidades dos processos ontogenéticos.

A embriogênese e a diferenciação celular são fenômenos que ocorrem em escala ontogenética, e culminam com a produção de um indivíduo adulto, que contém uma população de células que flutua em torno de um valor característico da espécie, distribuídas em um certo número de tipos diferentes, também característico da espécie. No entanto, ao atingir a fase adulta, quando cessa o crescimento, continua havendo a produção de células novas dentro do sistema, e se estabelece um processo compensatório com a "saída" de células, ou por que morrem, ou por que são retiradas através dos mecanismos fisiológicos do sistema imune.

Mostramos neste trabalho que é plausível que este fenômeno seja concebido como um processo de transporte num sistema aberto, que se torna estável do ponto

de vista populacional quando atinge o estado crítico auto-organizado, conforme indicam os resultados das simulações que estão de acordo com os parâmetros do *C. elegans*.

Os dois aspectos fundamentais da embriogênese e diferenciação celular, qual sejam, a produção da diversidade dos tipos celulares presentes no adulto; e o estabelecimento de um patamar populacional no final do processo, foram abordados inicialmente de maneiras distintas, posteriomente combinadas no modelo proposto.

Para a geração da diversidade foram utilizadas as redes genéticas aleatoriamente conectadas propostas no trabalho histórico de Kauffman de 1969, que constitui um marco no estudo das redes de autômatos Booleanos aleatórios, embora tenhamos optado por uma sistemática diferente para a abordagem dessas redes.

Foram feitas simulações determinísticas para redes genéticas aleatórias de 3 e 4 genes considerando-se todas as possibilidades de estados iniciais, e utilizando o método das varreduras aleatórias, foi feito um número maior de simulações do que as relatadas na literatura para redes com 5, 10, 20, 64 e 128 genes, o que permitiu que obtivéssemos uma melhor qualidade estatística dos resultados. Mostramos que: (i) para todas elas a distribuição populacional de ciclos estáveis com diferentes comprimentos é bastante regular; (ii) que todas exibem os máximos populacionais nítidos para ciclos de determinado comprimento, de modo que, o número de ciclos distintos produzidos por uma rede com N genes, se encontra numa faixa relativamente estreita que se distribui em torno de ciclos com $N^{1/2}$ estados internos, confirmando os resultados de Kauffman; (iii) há muito mais que $N^{1/2}$ ciclos distintos com qualquer número de estados em torno da média; ao contrário do que afirma Kauffman; e (iv) o número de ciclos distintos repetidos de mesmo tamanho cai drasticamente na medida em que N aumenta, de modo que a partir de N=10 já não ocorre um número significativo de repetições. Essas afirmações podem ser feitas

com boa segurança para redes com até 64 genes, embora os gráficos feitos para N=128 ofereçam algumas indicações de que o mesmo comportamento deva ocorrer.

Gostaríamos de frisar que em seu artigo de 1969, Kauffman fez um número muito menor de simulações do que fizemos no presente trabalho, e não há qualquer indicação em seus trabalhos posteriores (Kauffman, 1990 e 1992, além dos já citados) que tenha melhorado os resultados nos quais apoiou a afirmação de que existem $N^{1/2}$ ciclos distintos numa rede de N genes.

Considerando este último resultado (iv), foram feitas simulações com modelo de dinâmica de criticalidade auto-organizada (Per Bak et al, 1988) numa rede plana quadrada com L^2 sítios. Associando-se a cada sítio igual probabilidade, e variando-se os parâmetros, tamanho da rede (L), e o valor do limiar *lim*, foi possível mostrar que essa dinâmica é compatível com a embriogênese do *C. elegans* . Sugerimos que a expressão $Pc(L)=C.L^{3,4}$ seja válida para outros sistemas vivos que tenham parâmentros próximos aos desse animal, uma vez que possivelmente a criticalidade auto-organizada terá um peso específico, no conjunto de processo que determinam a população de células, semelhante ao que tem no *C. elegans* . No entanto, no que se refere a outros animais, os dados da literatura são escassos e muito pouco precisos. Há indicações de que sistemas vivos com parâmetros muito diferentes divergem do modelo, por exemplo, nos seres humanos onde ($N \sim 10^5$), o modelo prevê $Pc \sim 10^5$, quando o valor real é da ordem de 10^{12}.

Embora sua aplicabilidade esteja limitada a indivíduos que tenham poucos genes, este trabalho oferece uma contribuição para o entendimento da embriogênese e da diferenciação celular, no sentido em que estabelece a possibilidade de pensarmos que esse processo possa caracterizar-se por um comportamento crítico, de maneira semelhante a um grande número de processos que ocorrem em sistemas complexos.

6. Bibliografia:

Bak, P.; Tang, C. e Wiesenfeld, K. ; Phys. Rev. A 38(1), 364, (1988).

Bak, P. e Tang, C.; J. Geophys. Res. 94(B11),15635,(1989).

Bak, P. e Tang, C.; Phys Lett. 147(5/6), 297, (1990).

Bak, P. e Sneppen, K.; Phys. Rev. Lett. 71(24), 4083, (1993).

Behn, U. e van Hemmen, J. L.; J. Stat. Phys. 56(3/4), 533, (1989).

Brenner, S.; Genetics 77, 71, (1974) .

Brenner, S.; Dove, W.; Herskowits, I. e Thomas, R.; Genetics 126, 479, (1990).

Bretz, M.; Cunninbham, J., B.; Kurckynski, P. L. e Nori, F.; Phys. Rev. Lett. 69(16), 2431, (1992).

Chan, C. K. e Liang, N. Y.; Phys. Rev. Lett. 67, 1122, (1991).

Chau, H. F.; Mak, F. e Kwok, P. K.; Physyca A 215, 431, (1995)

Coulson, A.; Sulston, J.; Brenner, S. e Karn, J.; Proc. Natl. Acad. Sci. USA 83,7821, (1986).

Christensen, K.; Fogedbly, H. C. e Jensen, H. J.; J. Stat. Phys. 63(3/4), 653, (1991).

Derrida, B. in "Non Linear Evolution and Chaotic Phenomena" p. 213-244, ed. por Giovanni Gallavotti e Paul F. Zweifel, NATO ASI Series, (1988).

Derrida, B. e Pomeau, Y.; Europhys. Lett. 1(2), 45, (1986).

Derrida, B.; Domany, E. e Mukamel, D.; J. Stat. Phys. 69,(3/4), 667, (1992).

Derrida, B. e Stauffer,D.; Europhys. Lett. 2(10), 739, (1986)

Dhar, D. e Ramaswamy, R.; Phys. Rev Lett. 63, 1659, (1989).

Dhar, D.; Phys. Rev. Lett. 64, 1613, (1990).

Edelman, G.; "Neural Darwinism". Basic Books, New York. (1987).

Edelman, G.; "Topobilogy: an introduction to molecular embriology". Basic Books, New York. (1988).

Edelman, G.; Sci. Am. 260(5), 44, (1989).

Eder, J. e Rembold, H.; Naturwissenshaften 79,60, (1992).

Gilbert, S. F. "Biologia do Desenvolvimento" 3a ed. Sinauer Associates, (1994).

Grassberger, P. e Kantz, H.; J. Stat. Phys. 63(3/4), 685, (1991).

Greven, E. e Hollander, J.; J. Stat. Phys. 65(5/6), 1123, (1991).

Greven, E. e Hollander, J.; J. Stat. Phys. 65(5/6), 1147, (1991).

Guimarães, R. C.; Pereira Jr, A. e Chaves, J. C.; Cadernos do CLE, Unicamp (1994).

Hentschel, H. G. E. e Family, F.; Phys. Rev. Lett. 66(15), 1982,(1991).

Hwa, T. e Kadar, M.; Phys. Rev. Lett. 62(16), 1813, (1989).

Jacob, F. e Monod, J.; 21st Symp. Soc. Study of Development and Growth. London. Academic Press. (1963).

Jaeger, H. M.; Liu, C. e Nagel, S. R.; Phys. Rev. Lett. 62(1), 40,(1989).

Kadanoff, L. P.; Nagel, S. R.; Wu, L. e Zhou, S.; Phys. Rev. A 39(12), 6524, (1989).

Kauffman, S. A.; J. Theor. Biol. 22, 437, (1969).

Kauffman, S. A.; J. Theor. Biol. 44, 167, (1974).

Kauffman, S. A.; Physica D 10, 145, (1984).

Kauffman, S. A. in "Complexity, Entropy and the Physics of Information". Ed. by Zurek, W. H. Addison-Wesley Publ. Comp. p.152-192, (1990).

Kauffman, S. A. in "Spin Glasses and Biology". Ed. by Stein, D. L. Word Scientific. p. 61-100, (1992).

Kauffman, S. A. "The Origins of Order: Self-Organization and Selection in Evolution". Oxford University Press, (1993).

Kaufman, M.; Urbain, J.; Thomas, R.; J. Theor. Biol. 114,527, (1985).

Kindt, T. J. e Capra, J. D. "The Antibody Enigma". Plenum Press, N. Y. (1984).

Lewin, R.; Science 224, 1327, (1984).

Maruyama, I. N.; Miller, D. M. e Brenner, S.; Mol. Gen., Genet. 219, 113, (1989).

Neumann, A. U.; Physica A 162, 19, (1989).

Noever, D. A.: Matsos, H.C. e Cronise, R. J.; Phys. Lett. A 194,295, (1994).

Paczuski, M. e Bak, P.; Phys. Rev. E 48(5), 3214, (1993).

Prado, C.P.C e Tibúrcio, G. V.; Programas e Resumos do XV Encontro Nacional de Física da Matéria Condensada, Caxambu, (1992).

Rosendahl, J.; Vekic, M. e Kelley, J.; Phys. Rev. E 47(2), 1401, (1993).

Russo, V. E. A.; Brody, S.; Cove, D. e Ottolenghi, S. - ed. "The Molecular Genetic Approach". Springer-Verlag, (1992).

Solé, R. V. e Luque, B. Phys. Lett. A 196, 331, (1995).

Sornette, A. e Sornette, D.; Europhys. Lett. 9(3), 197, (1989).

Stauffer, D.; Phyl. Mag. B 56(6),901, (1987).

Stein, D. L. ed."Spin Glasses and Biology". Word Scientific Publishing Co. Pte. Ltd. London., (1992).

Sulston, J. E. e Horvitz, H. R.; Develop Biol. 56, 110, (1977).

Sulston, J. E.; Schierenberg, E.; White, J. G. and Thomson, J. N.; Develop Biol. 100, 64-119, (1983).

Watson J., D."Molecular Biology of the Cell". 2ª ed. Garland Publishing, (1989).

Wood, W. B. "The Nematode Caenhorhabditis Elegans". Cold Spring Harbor Laboratory, (1988).

Zaitsev, S. I.; Physica A 189, 411, (1992).

Zhang, Y.; Phys. Rev. Lett. 63, 470, (1989).

7. Anexos

(i) Gráficos

Grafico 4.1.1

SIMULACAO DETERMINISTICA PARA N=3

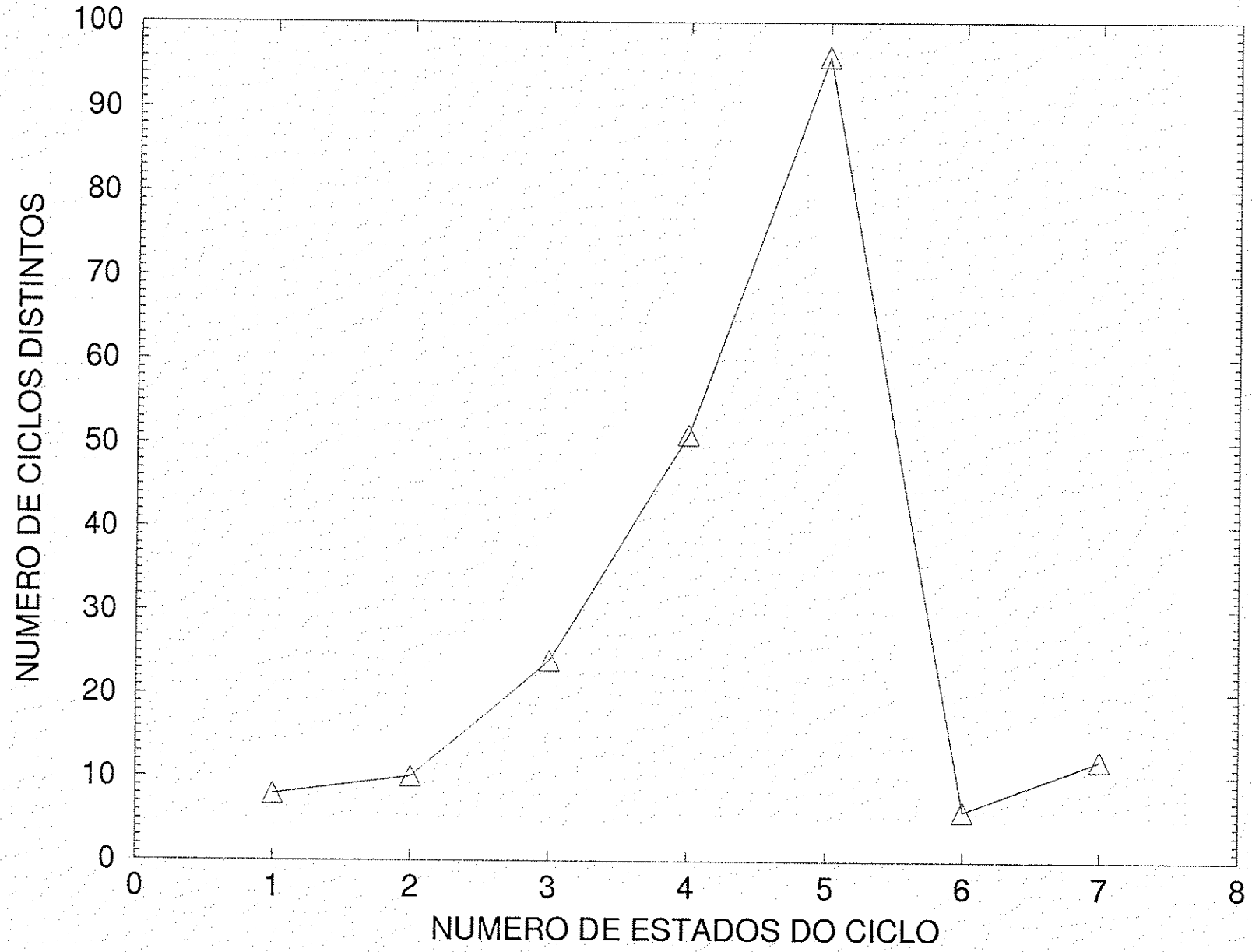


Grafico 4.1.2

SIMULACAO DETERMINISTICA PARA N=4

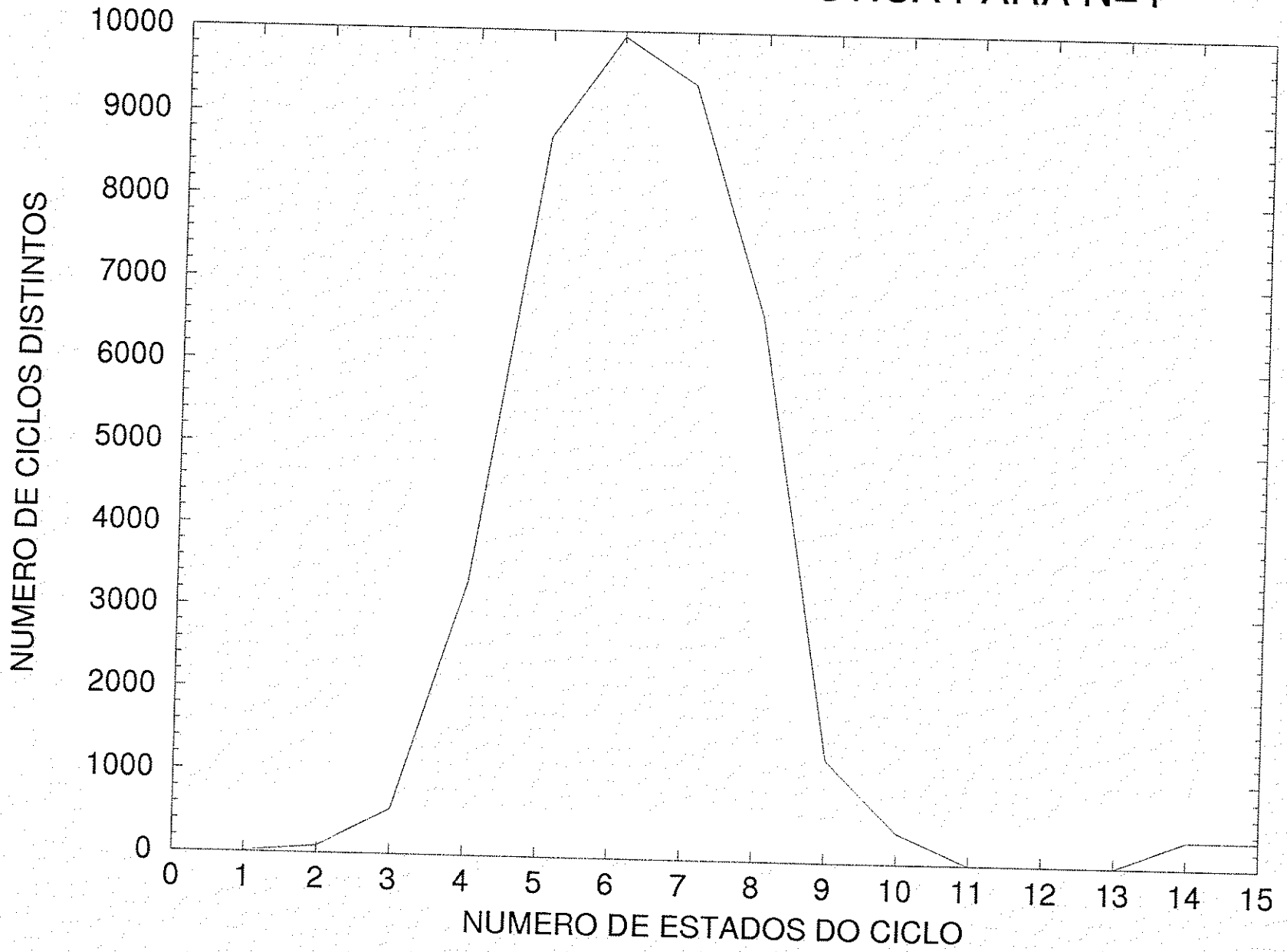


Grafico 4.1.3

VARREDURA ALEATORIA PARA N=3

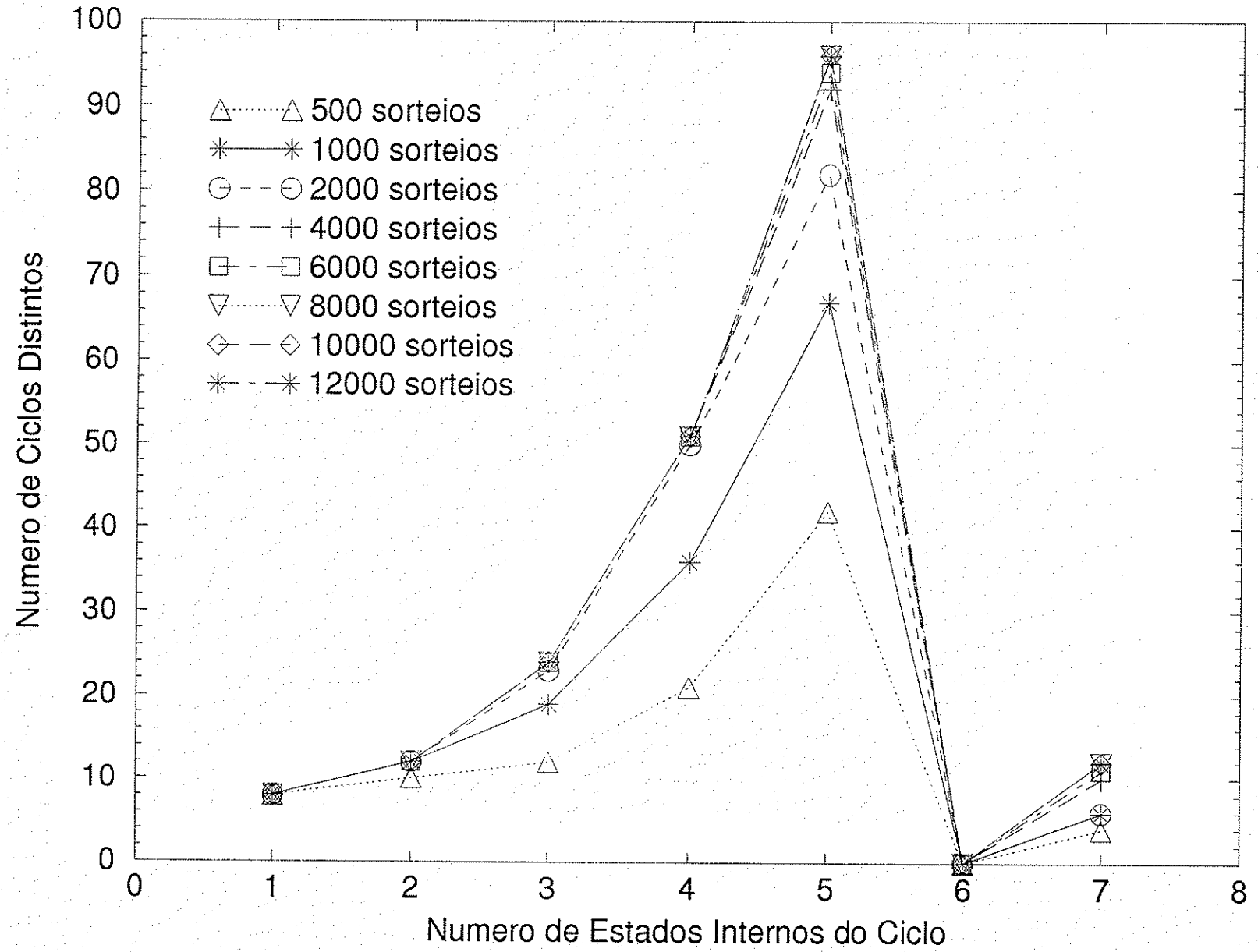


Grafico 4.1.4

VARREDURA ALEATORIA PARA N=4

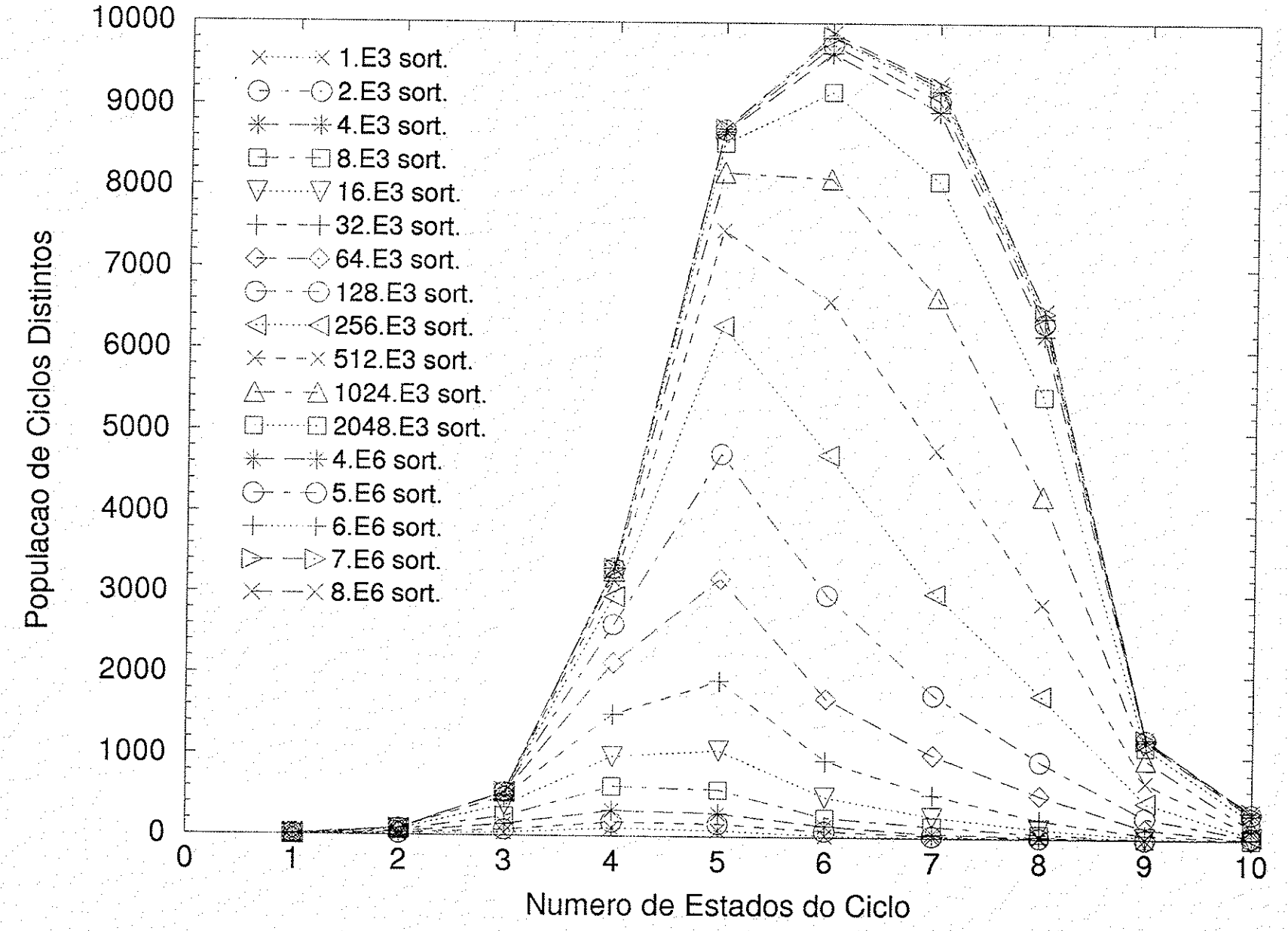


grafico 4.1.5

VARREDURA ALEATORIA PARA N=5

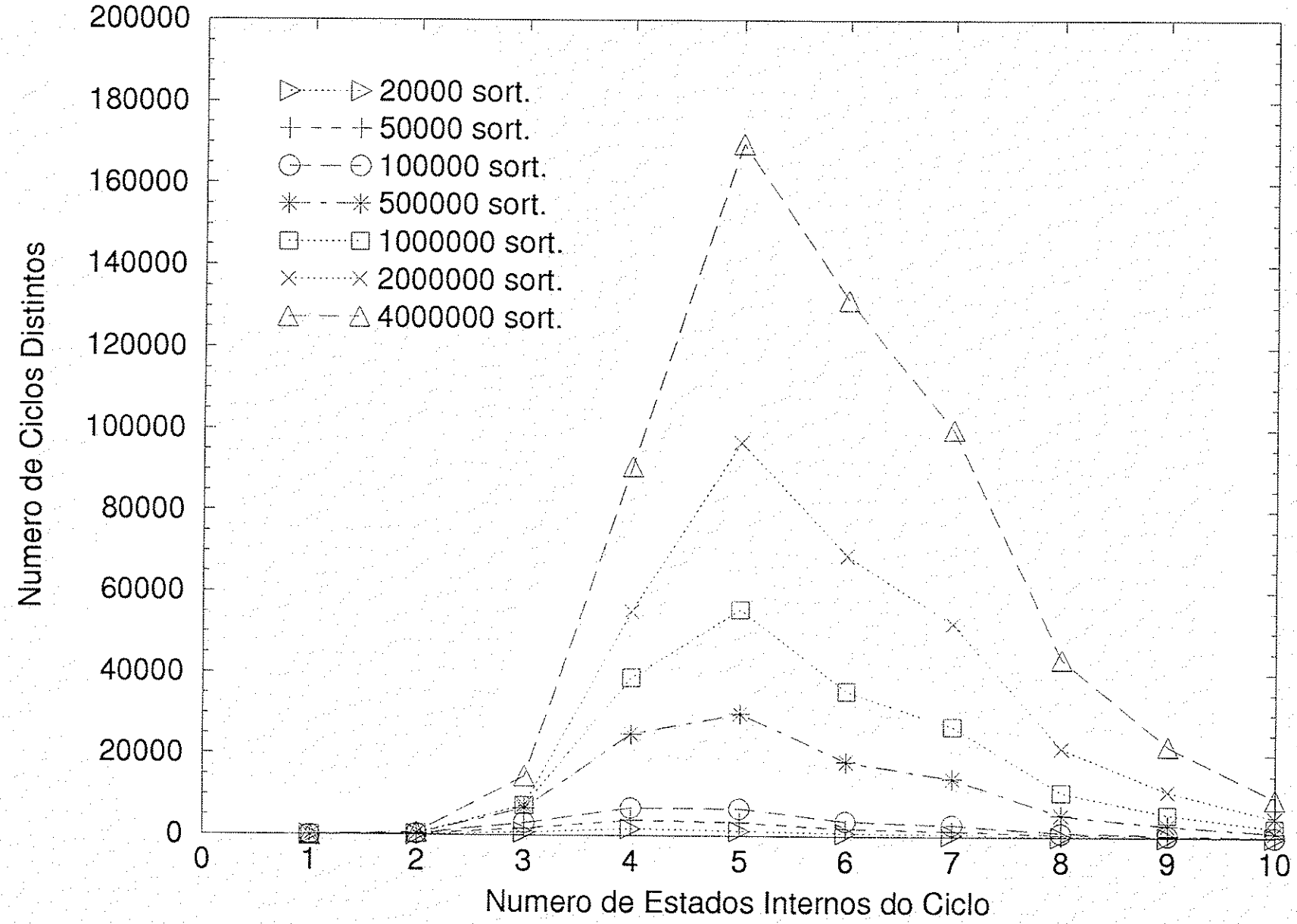


Grafico 4.2.1

VARREDURAS ALEATORIAS PARA N=10

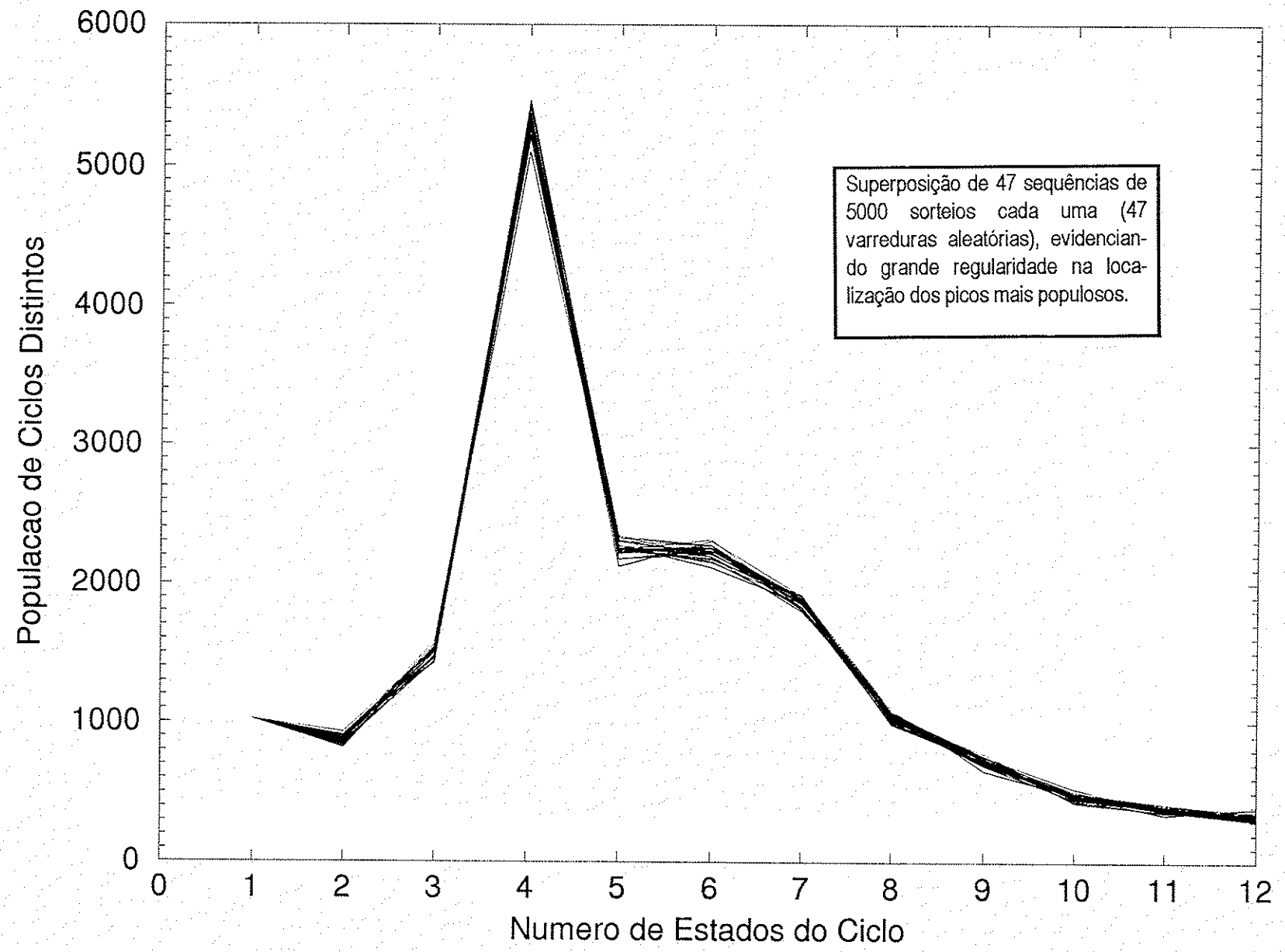


Grafico 4.2.2

VARREDURAS ALEATORIAS PARA N=10 (detalhe)

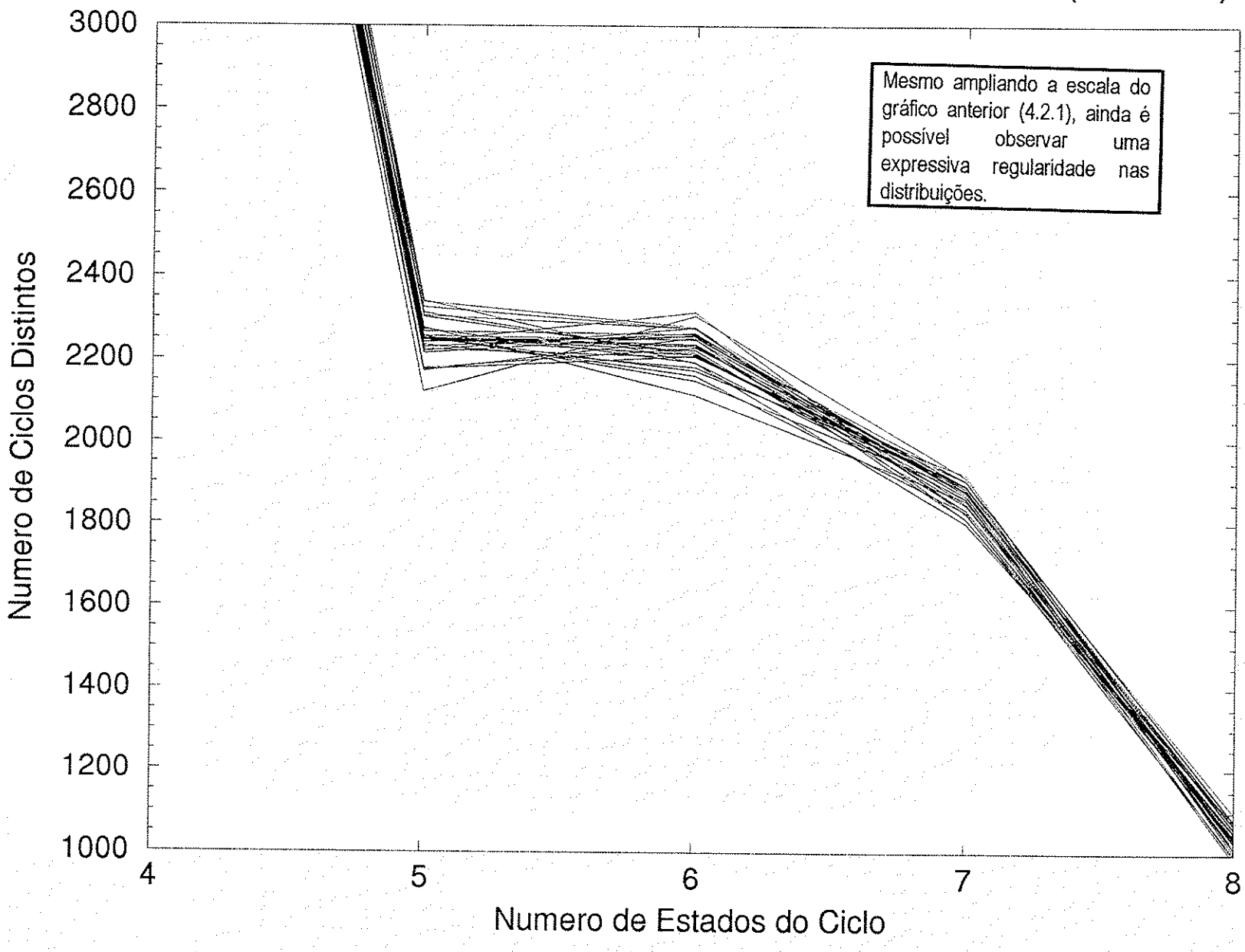


Grafico 4.2.3

VARREDURAS ALEATORIAS PARA N=20

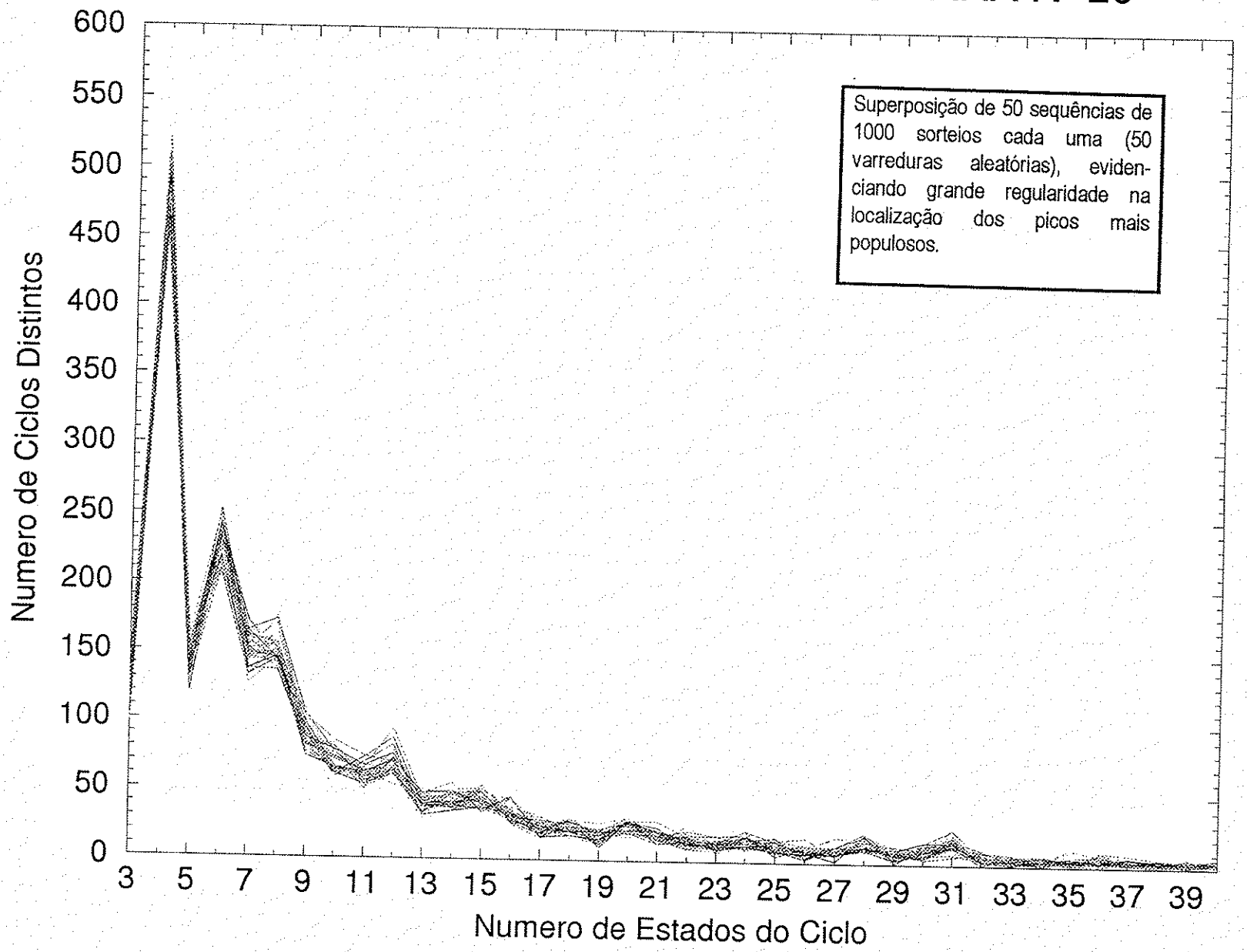


Grafico 4.2.4

VARREDURAS ALEATORIAS PARA N=30

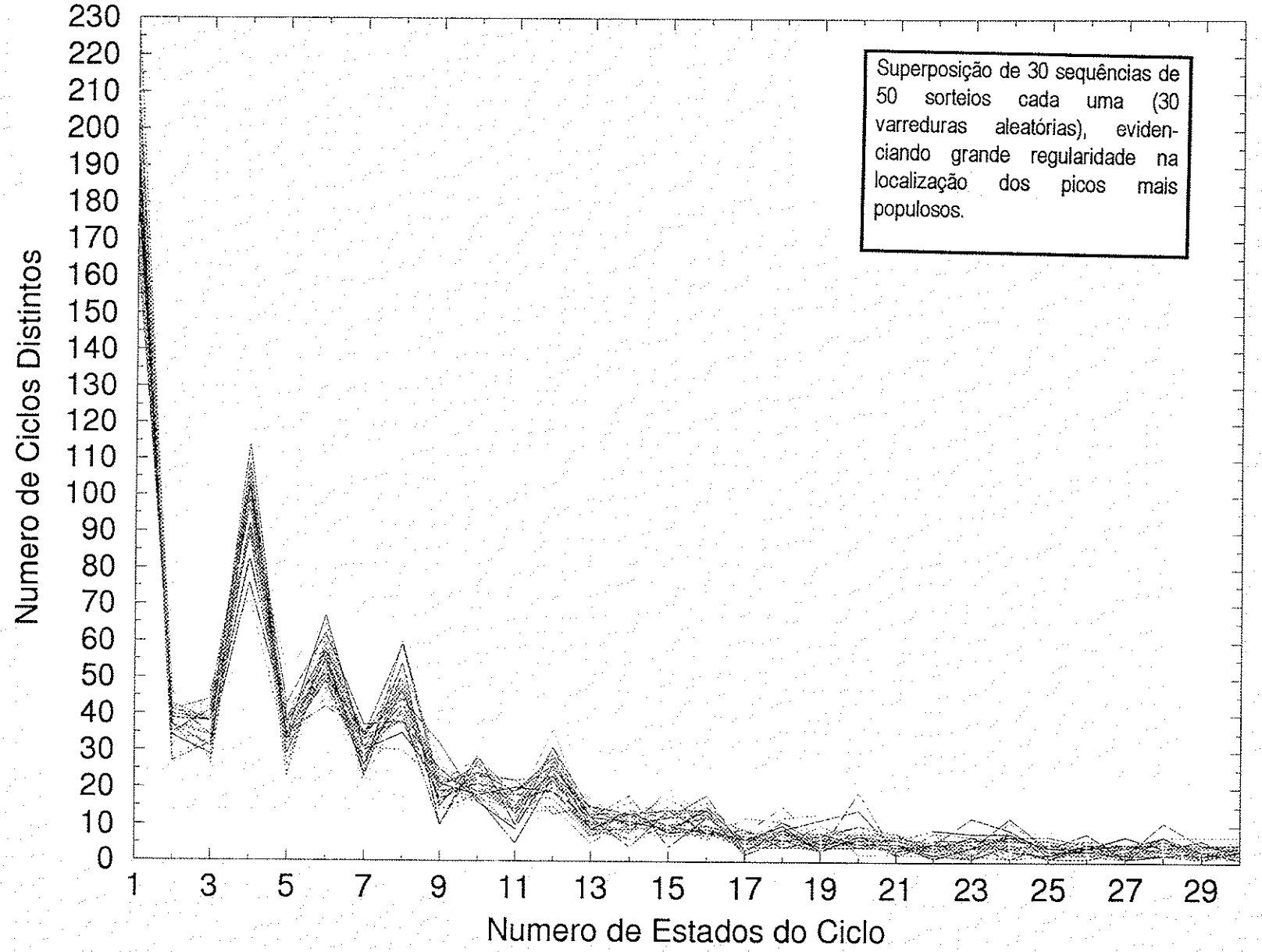


Grafico 4.2.5

VARREDURA ALEATORIA PARA N=64

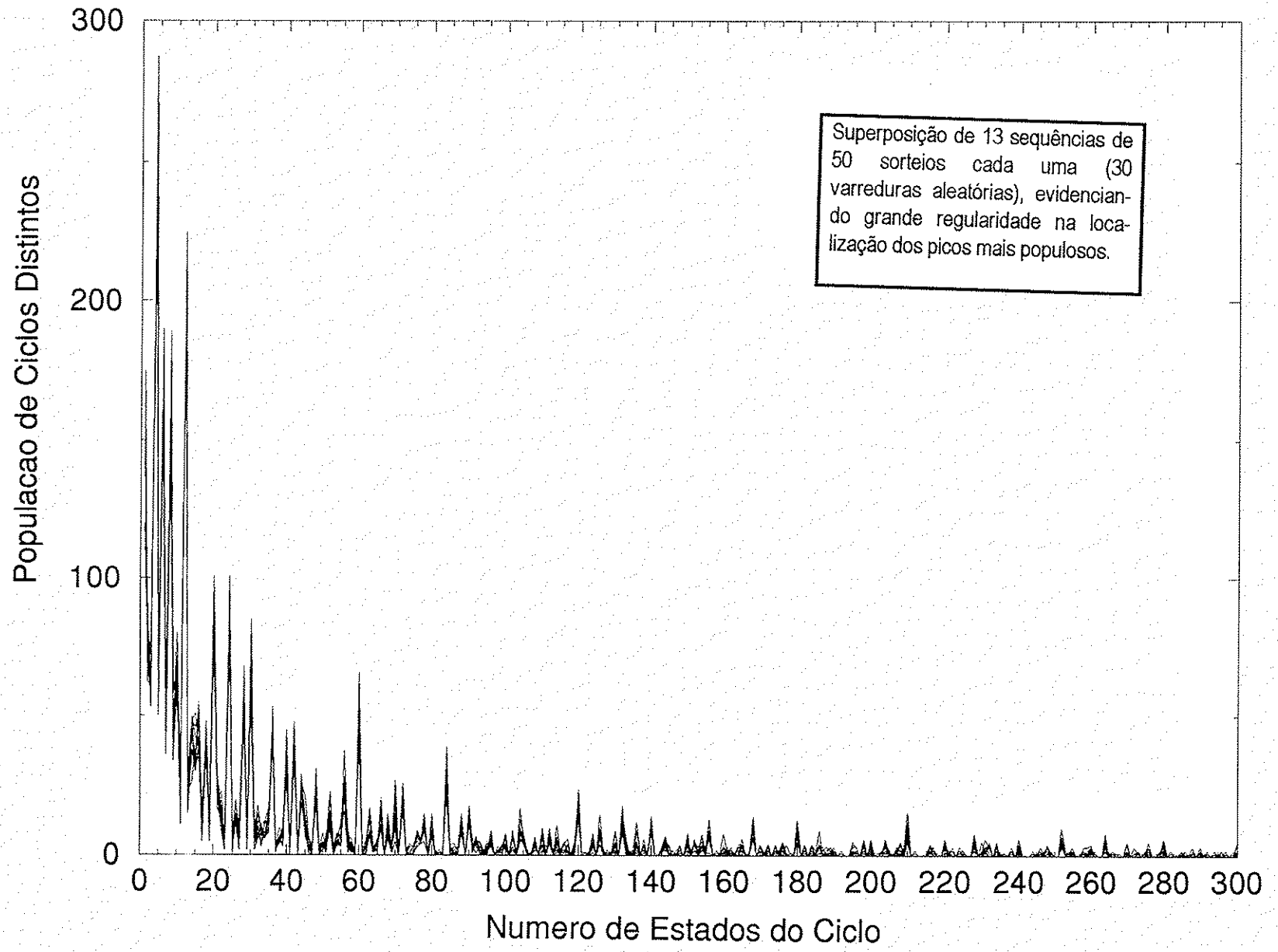


GRAFICO 4.2.6

VARREDURA ALEATORIA PARA N=128

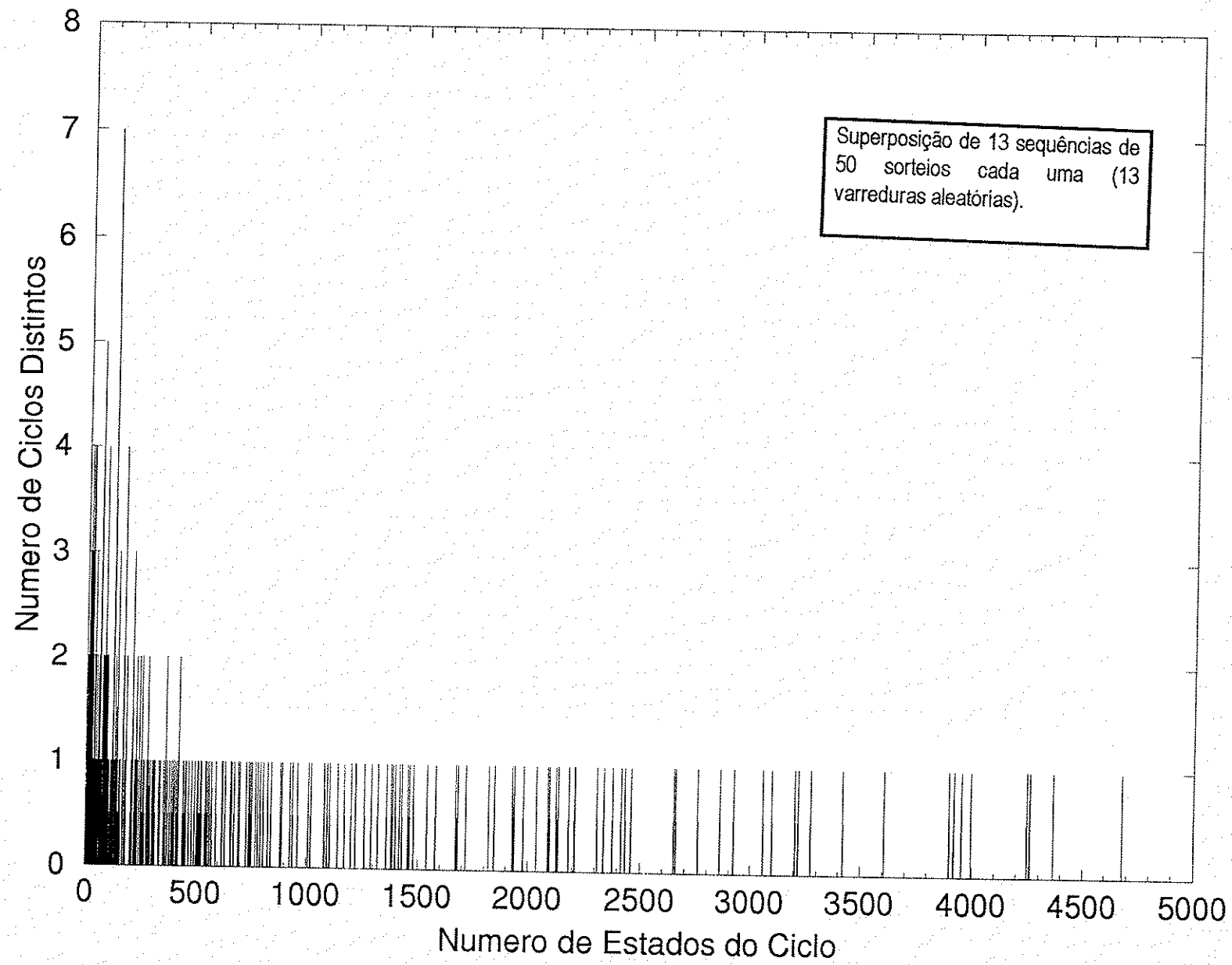


Grafico 4.2.7

VARREDURA ALEATORIA PARA N=10

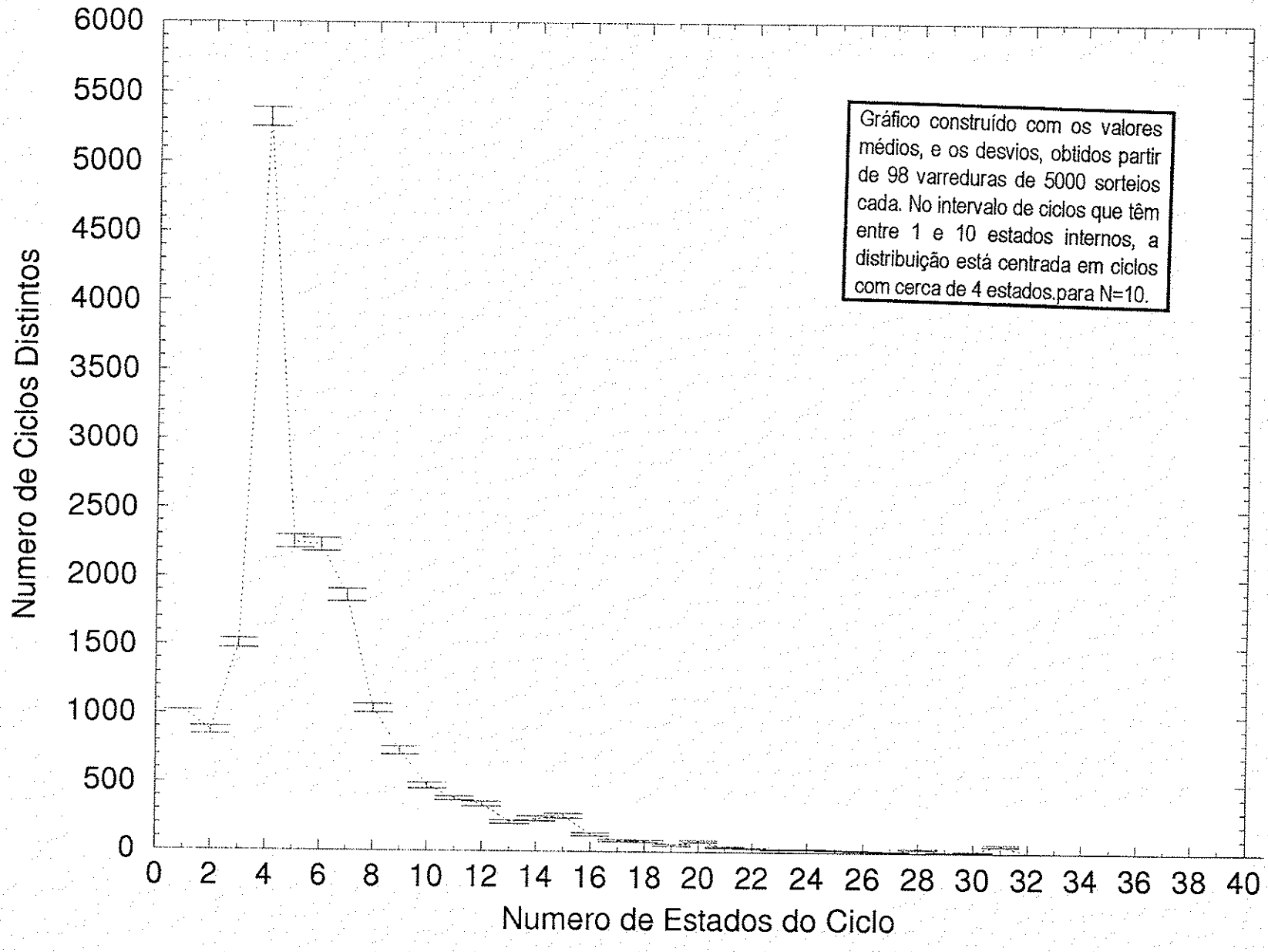


Gráfico construído com os valores médios, e os desvios, obtidos partir de 98 varreduras de 5000 sorteios cada. No intervalo de ciclos que têm entre 1 e 10 estados internos, a distribuição está centrada em ciclos com cerca de 4 estados.para N=10.

Grafico 4.2.8

VARREDURA ALEATORIA PARA N=20

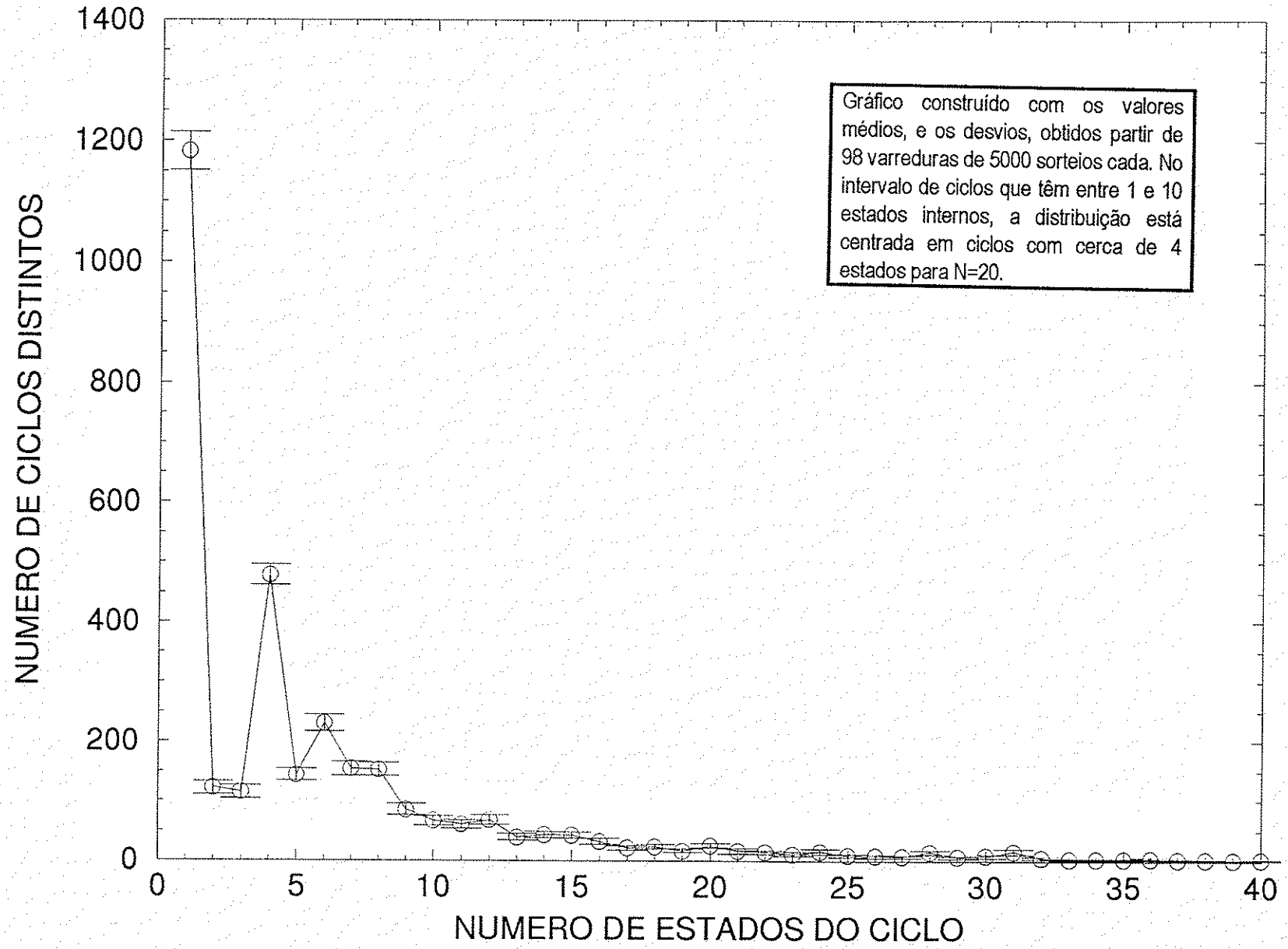


Gráfico construído com os valores médios, e os desvios, obtidos partir de 98 varreduras de 5000 sorteios cada. No intervalo de ciclos que têm entre 1 e 10 estados internos, a distribuição está centrada em ciclos com cerca de 4 estados para N=20.

Grafico 4.2.9

VARREDURA ALEATORIA PARA N=30

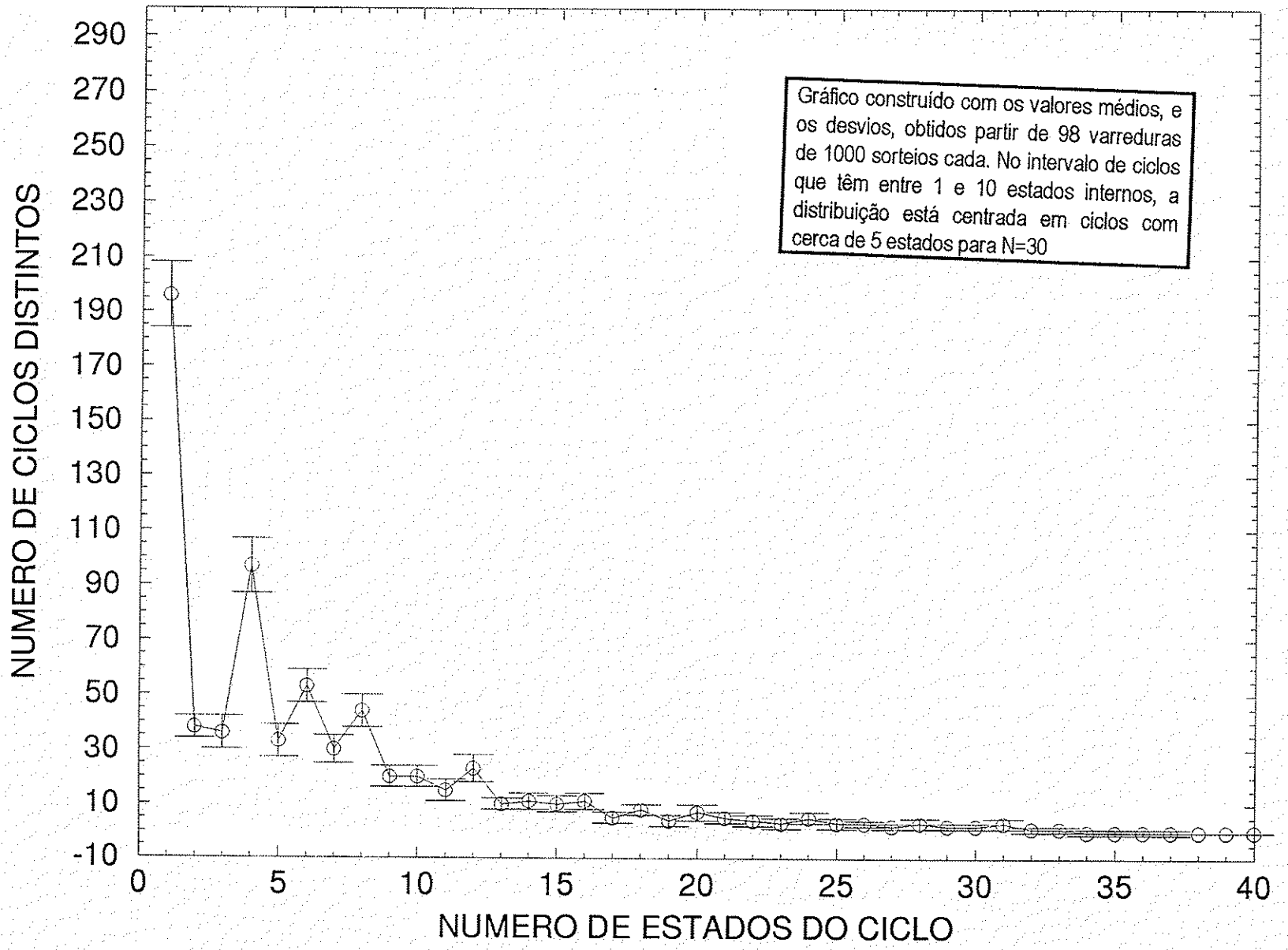


Grafico 4.2.10

VARREDURA ALEATORIA PARA N=64

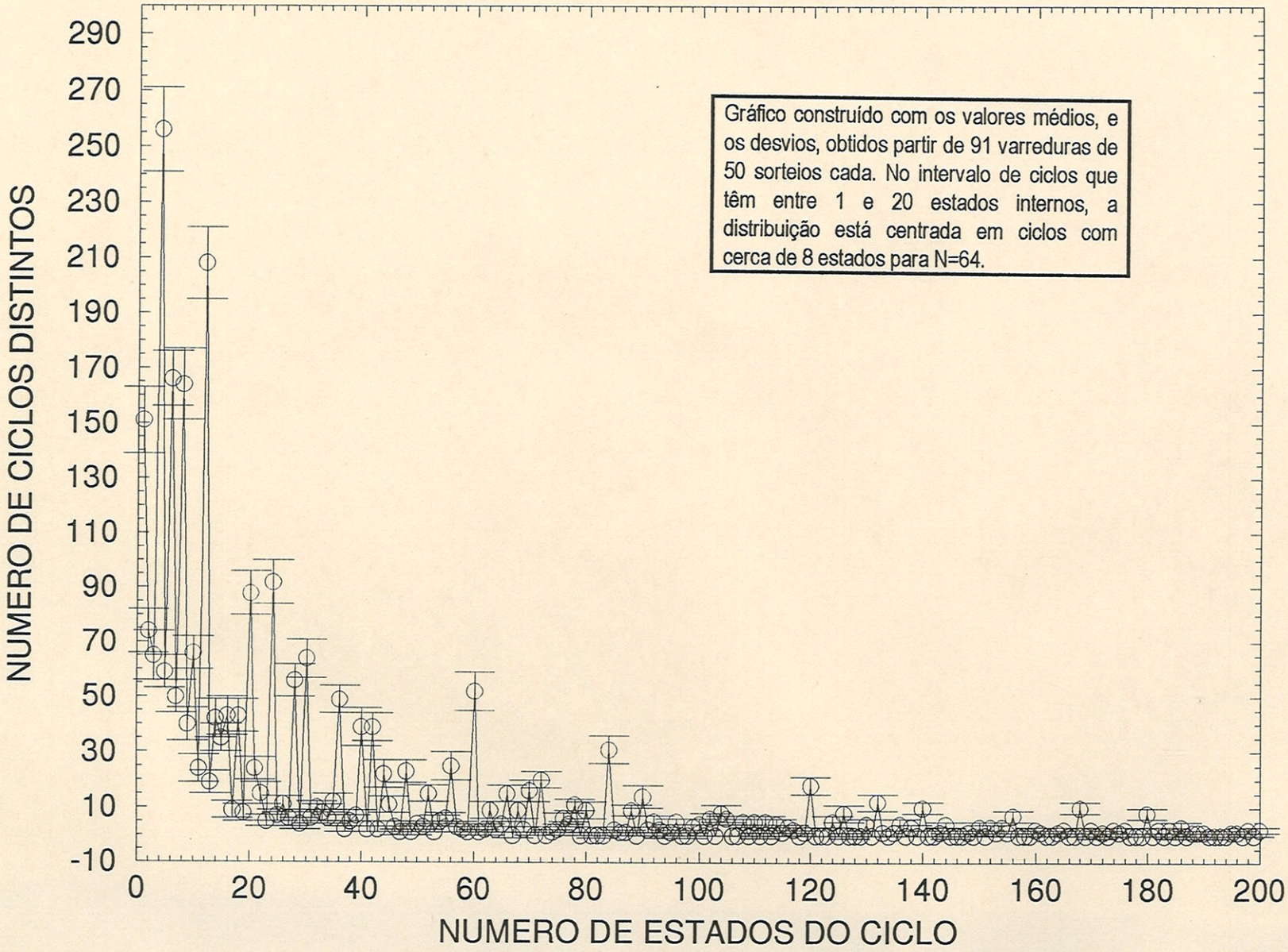


Grafico 4.2.11

VARREDURA ALEATORIA PARA N=64

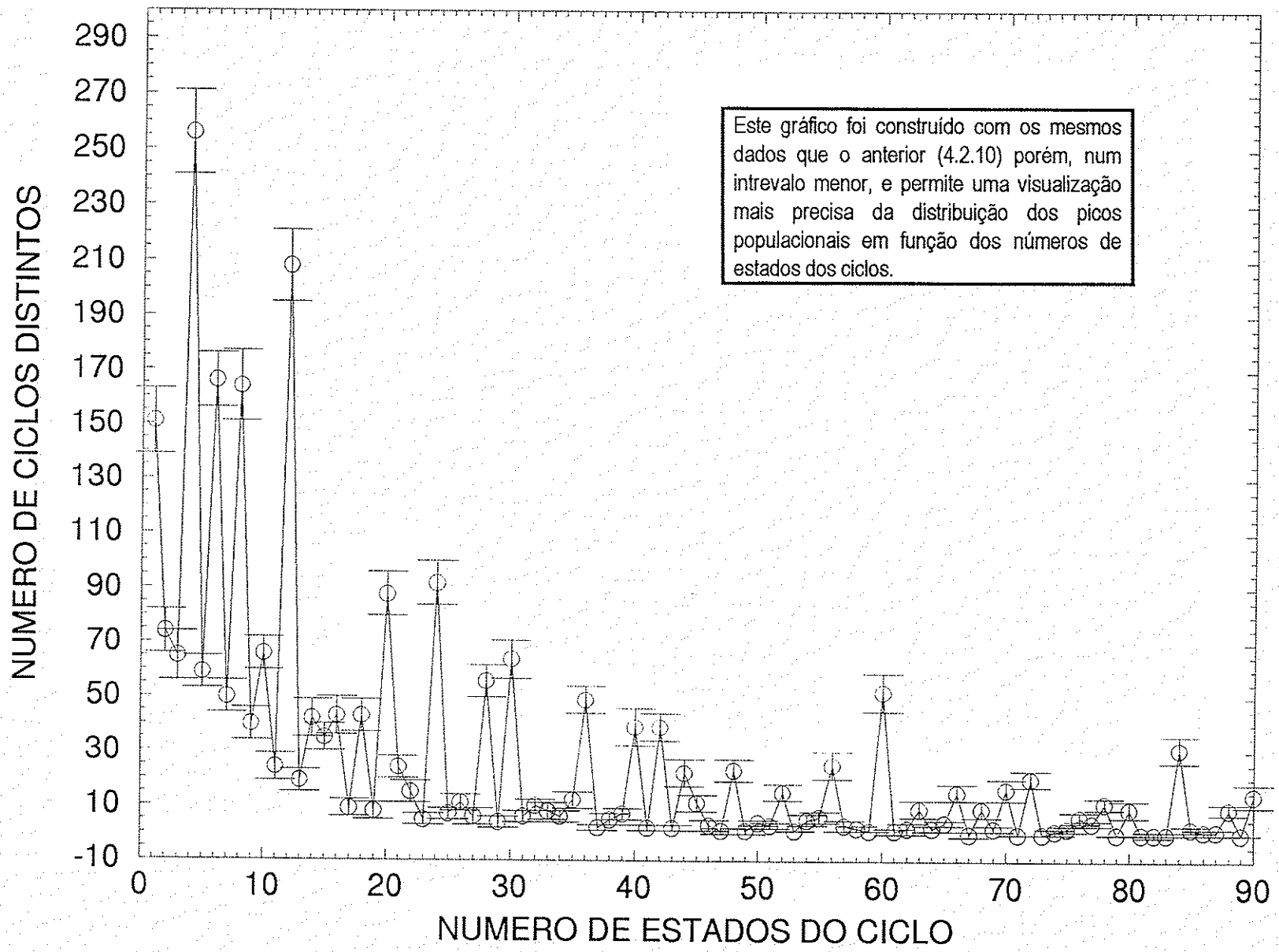


Grafico 4.2.12

VARREDURA ALEATORIA PARA N=128

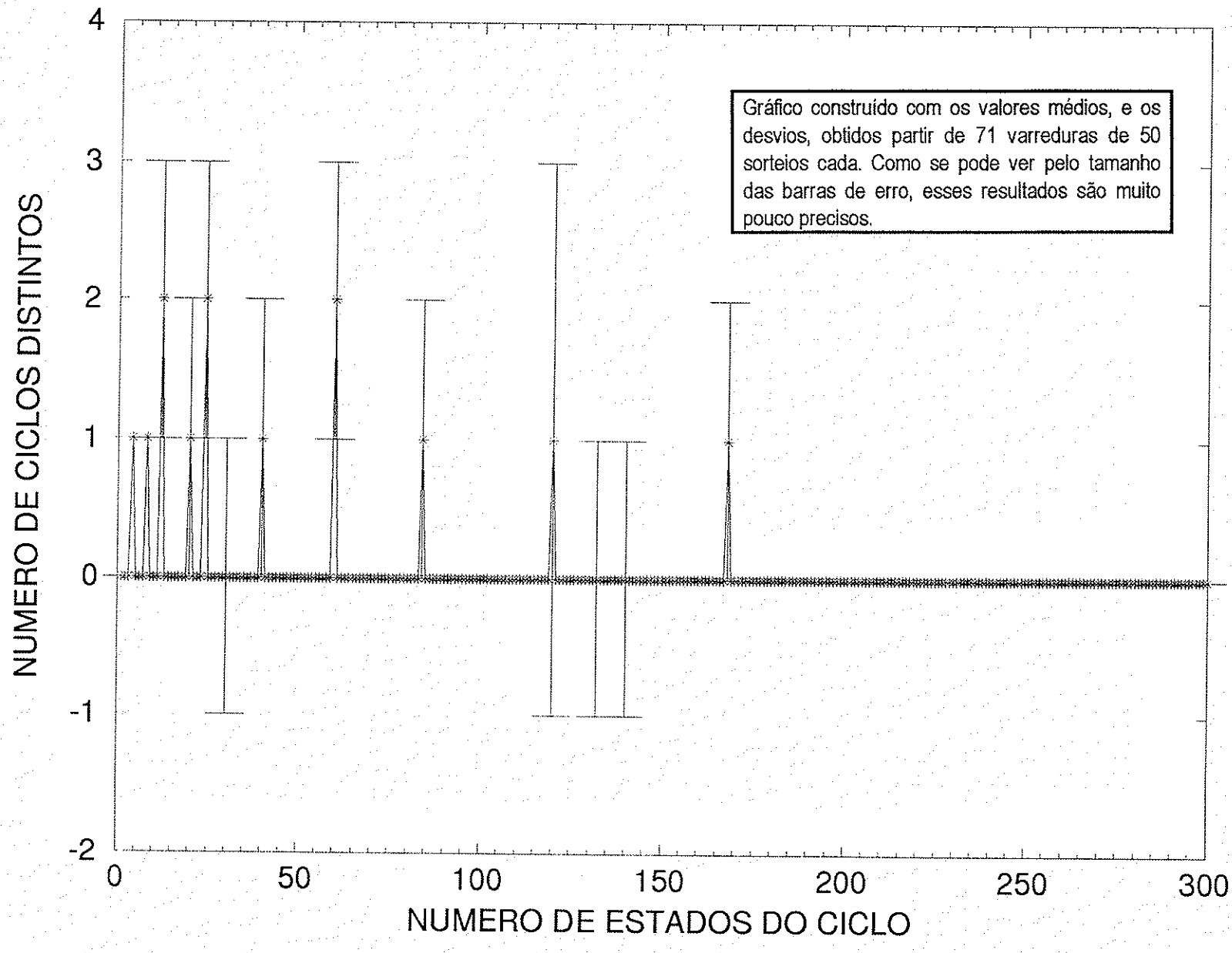


Grafico 4.3.1

Curva de Saturacao Populacional para L=3

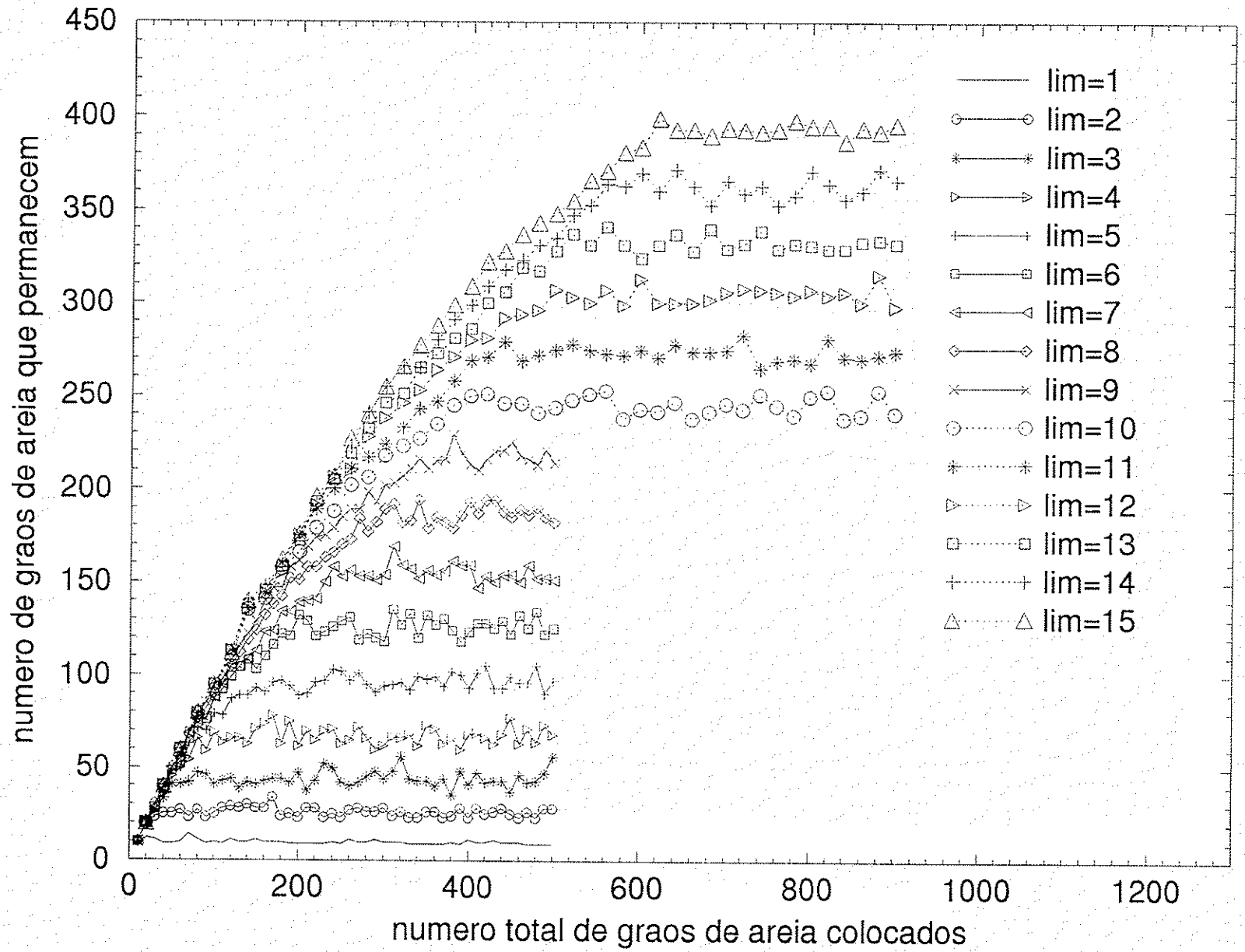


Grafico 4.3.2

Curvas de Saturacao Populacional para L=4

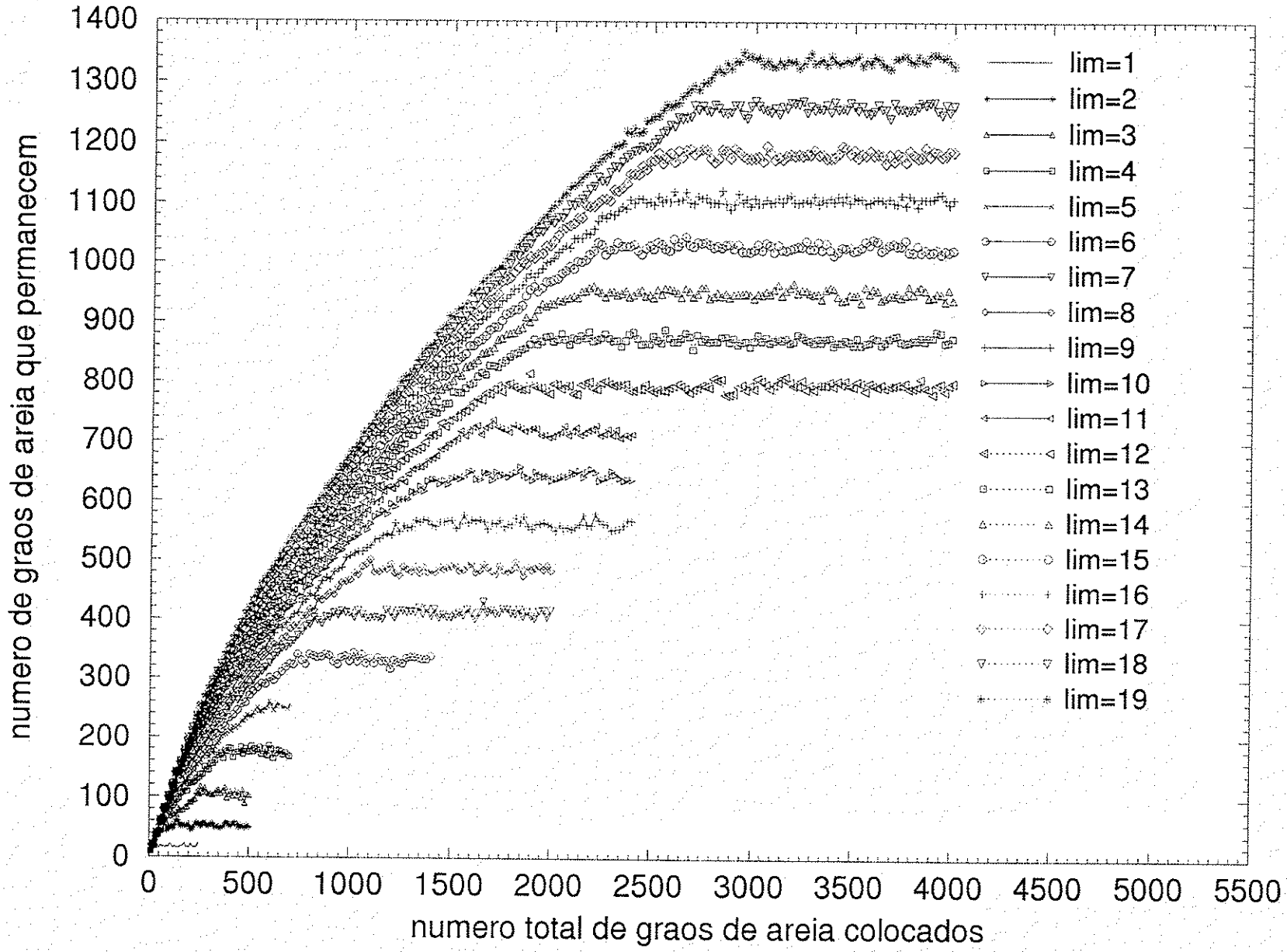


Grafico 4.3.3

Curvas de Saturacao Populacional para L=5

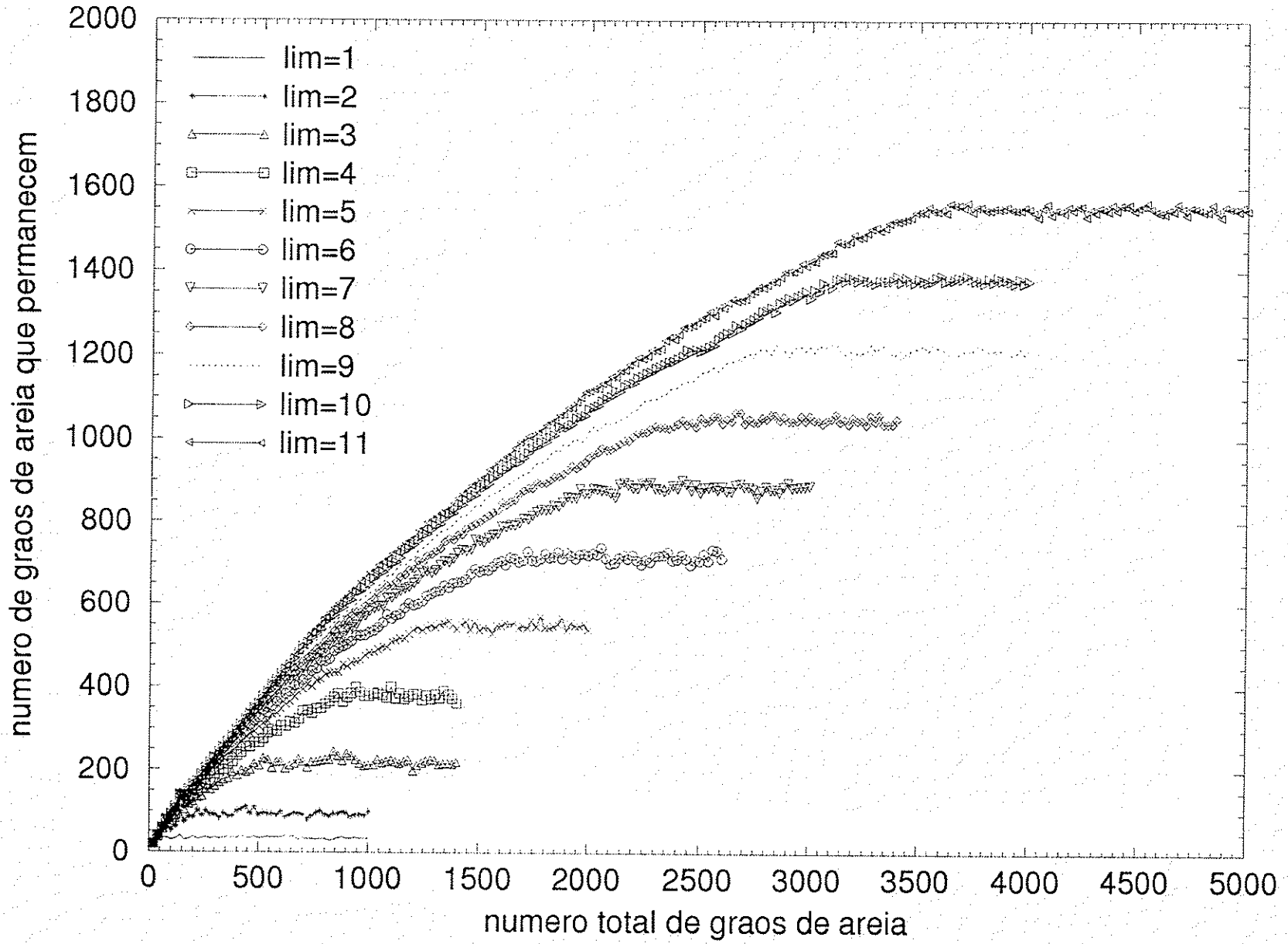


Grafico 4.3.4

Curvas de Saturacao Populacional para L=6

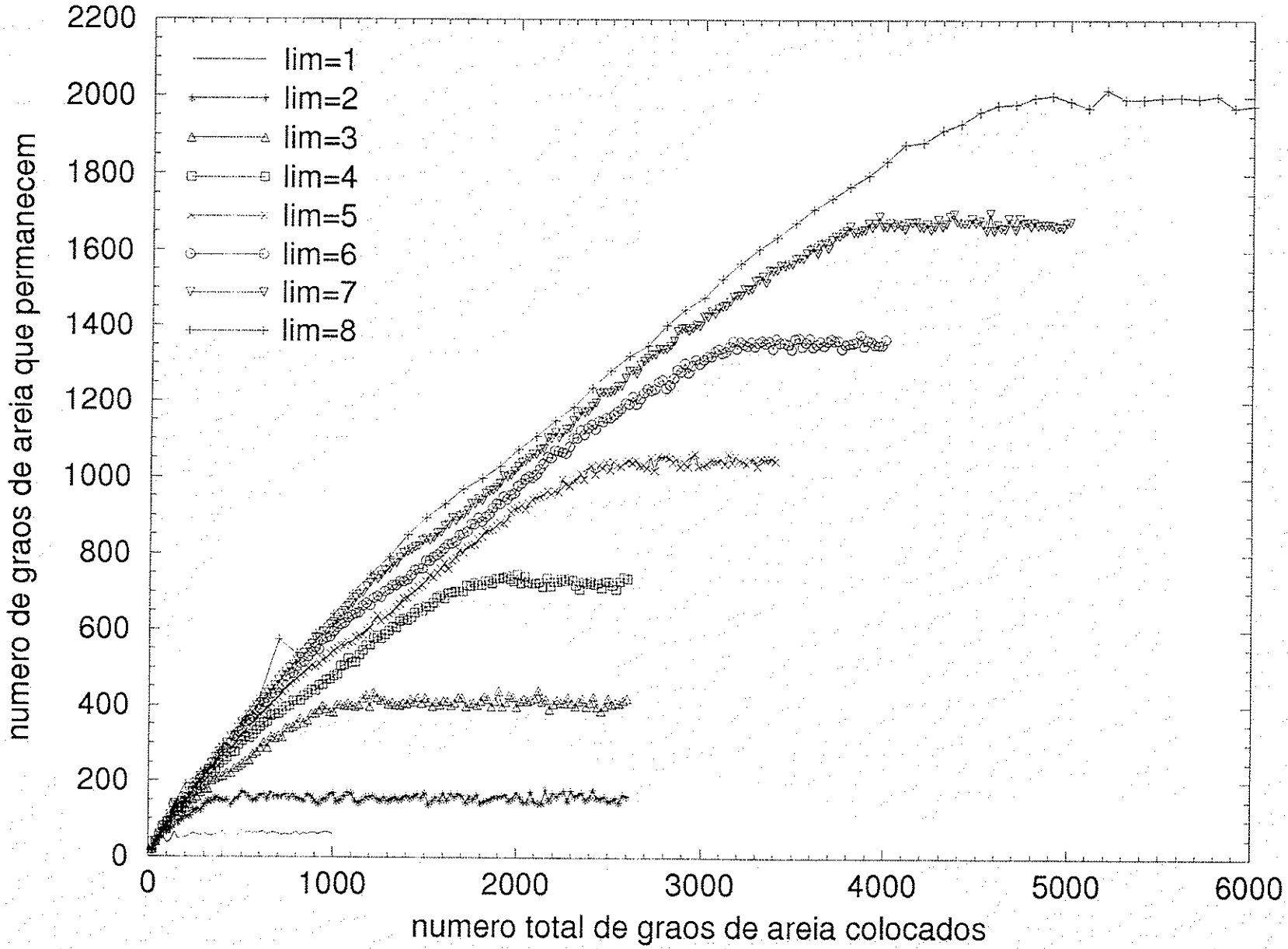


Grafico 4.3.5

Curvas de Saturacao Populacional para L=7

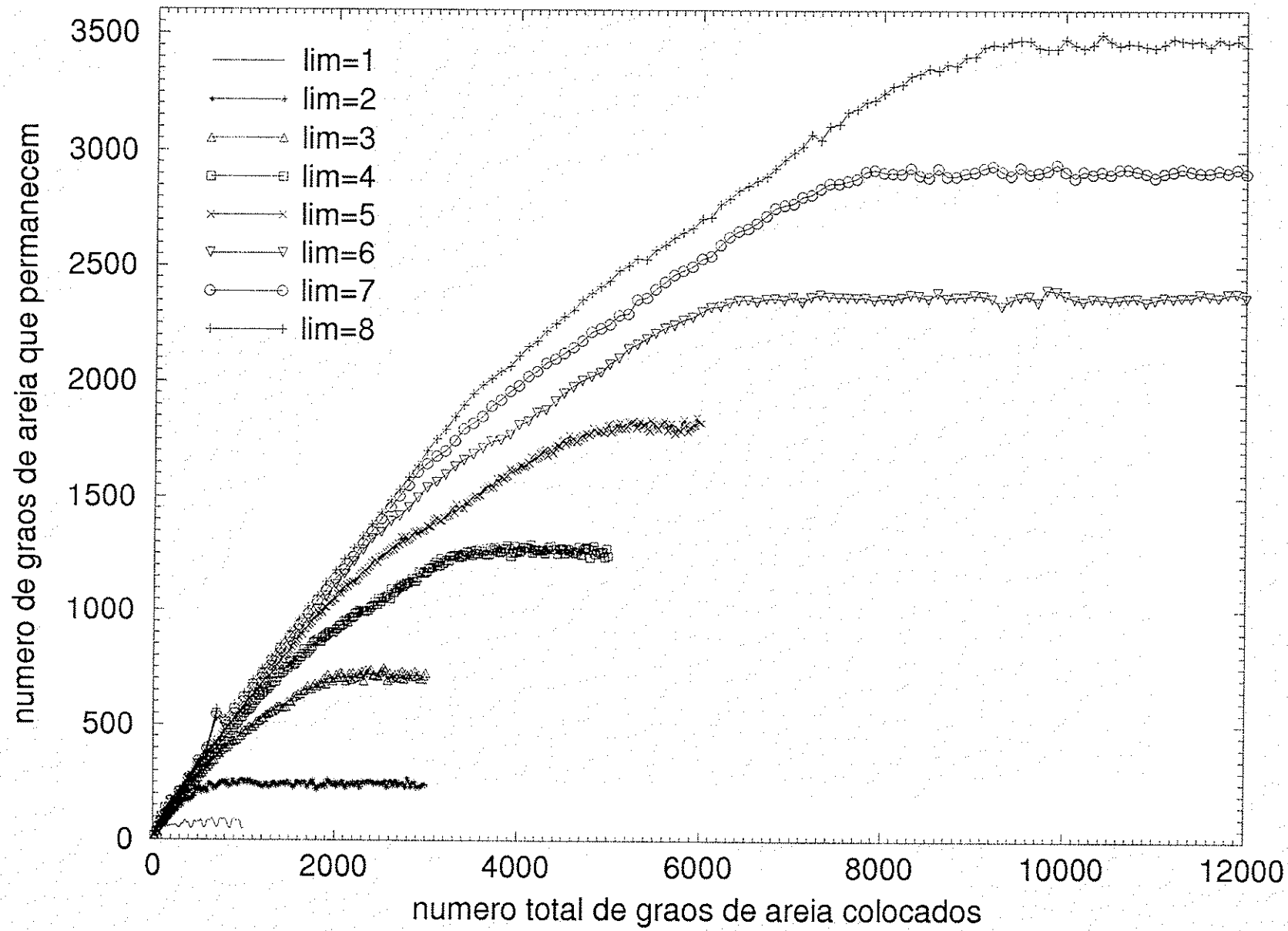


Grafico 4.3.6

Curvas de Saturacao Populacional para L=8

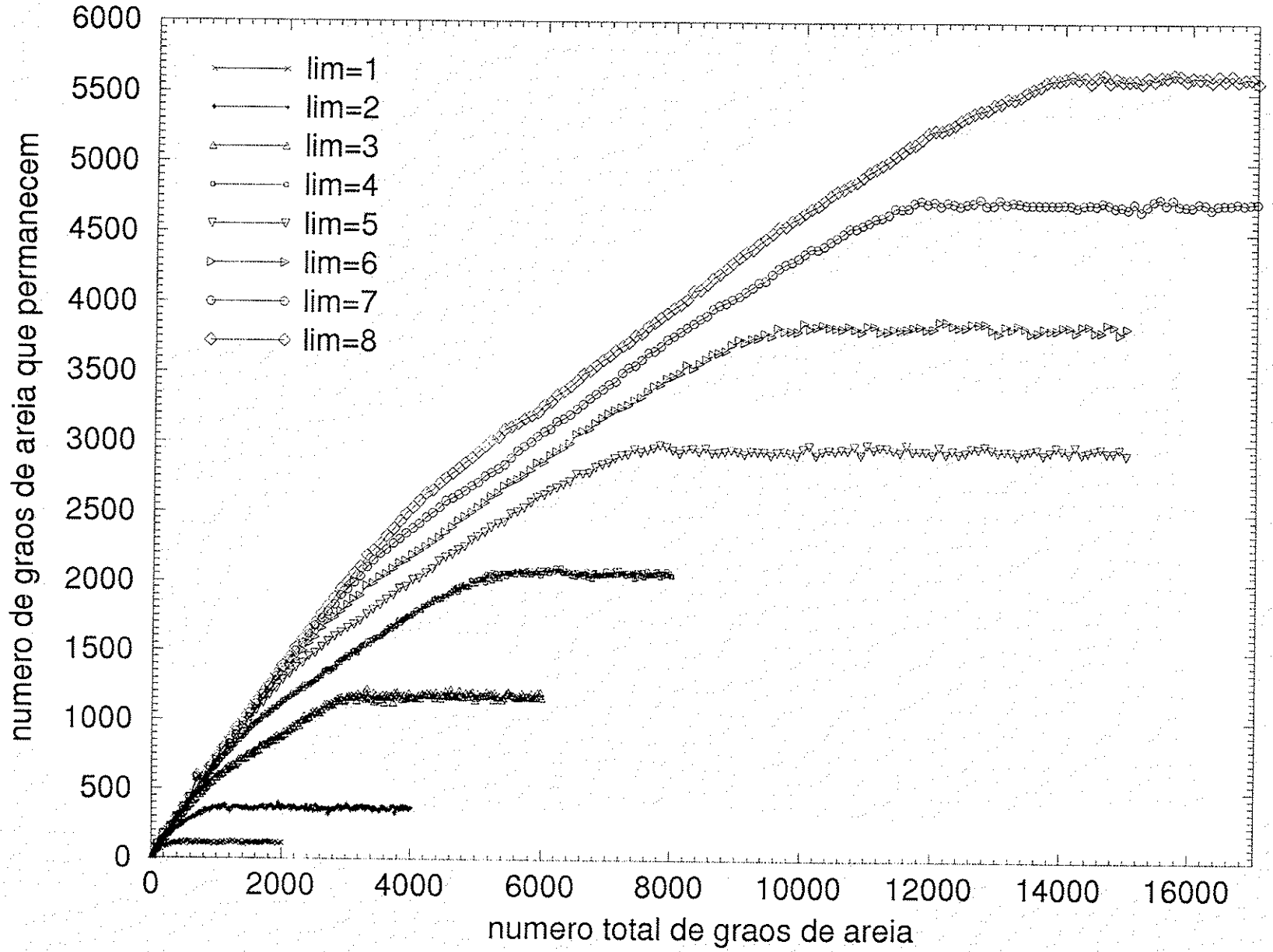


Grafico 4.3.7

Curvas de Saturacao Populacional para L=9

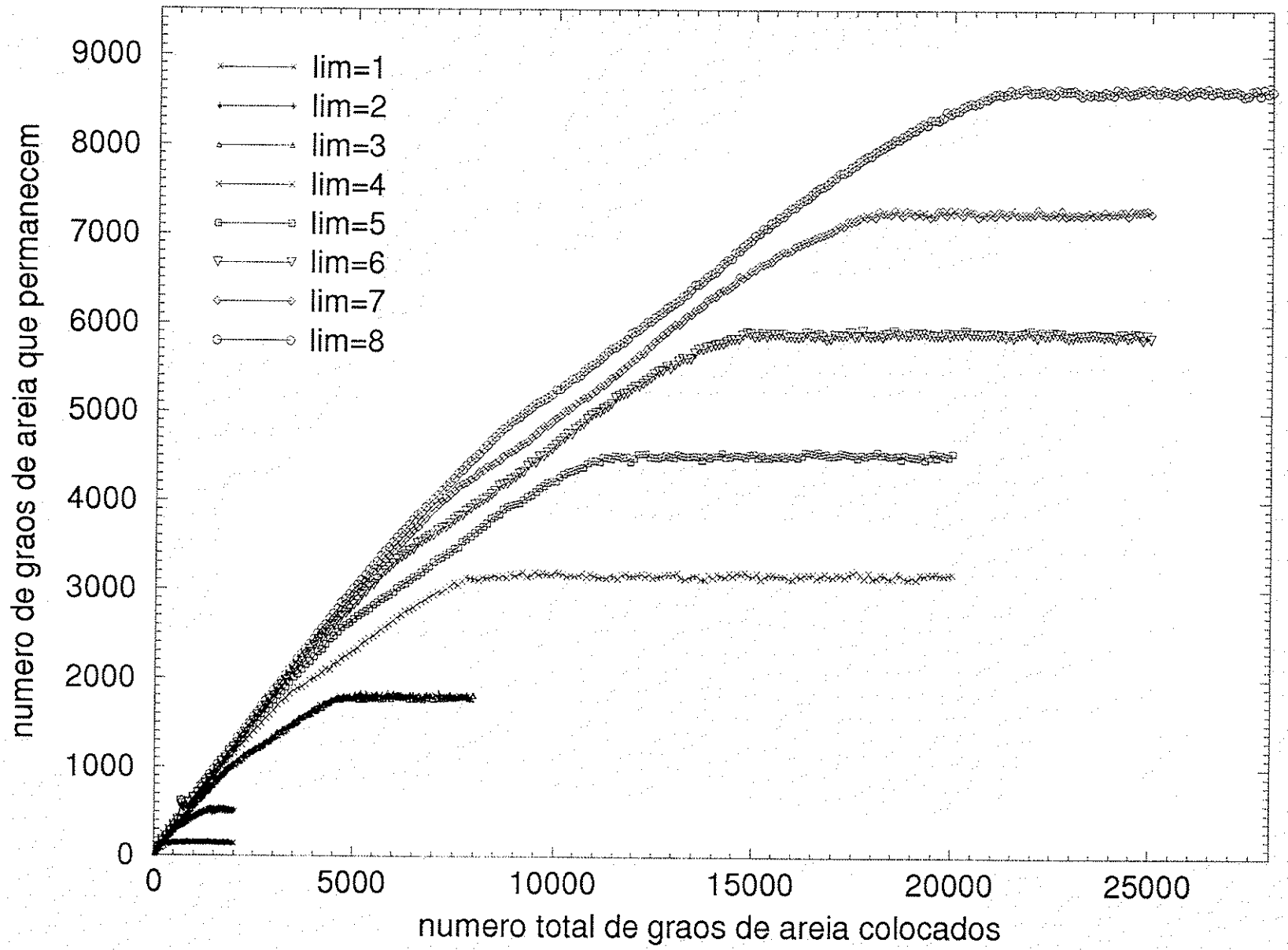


Grafico 4.3.8
Log Pc X log L
para lim=2

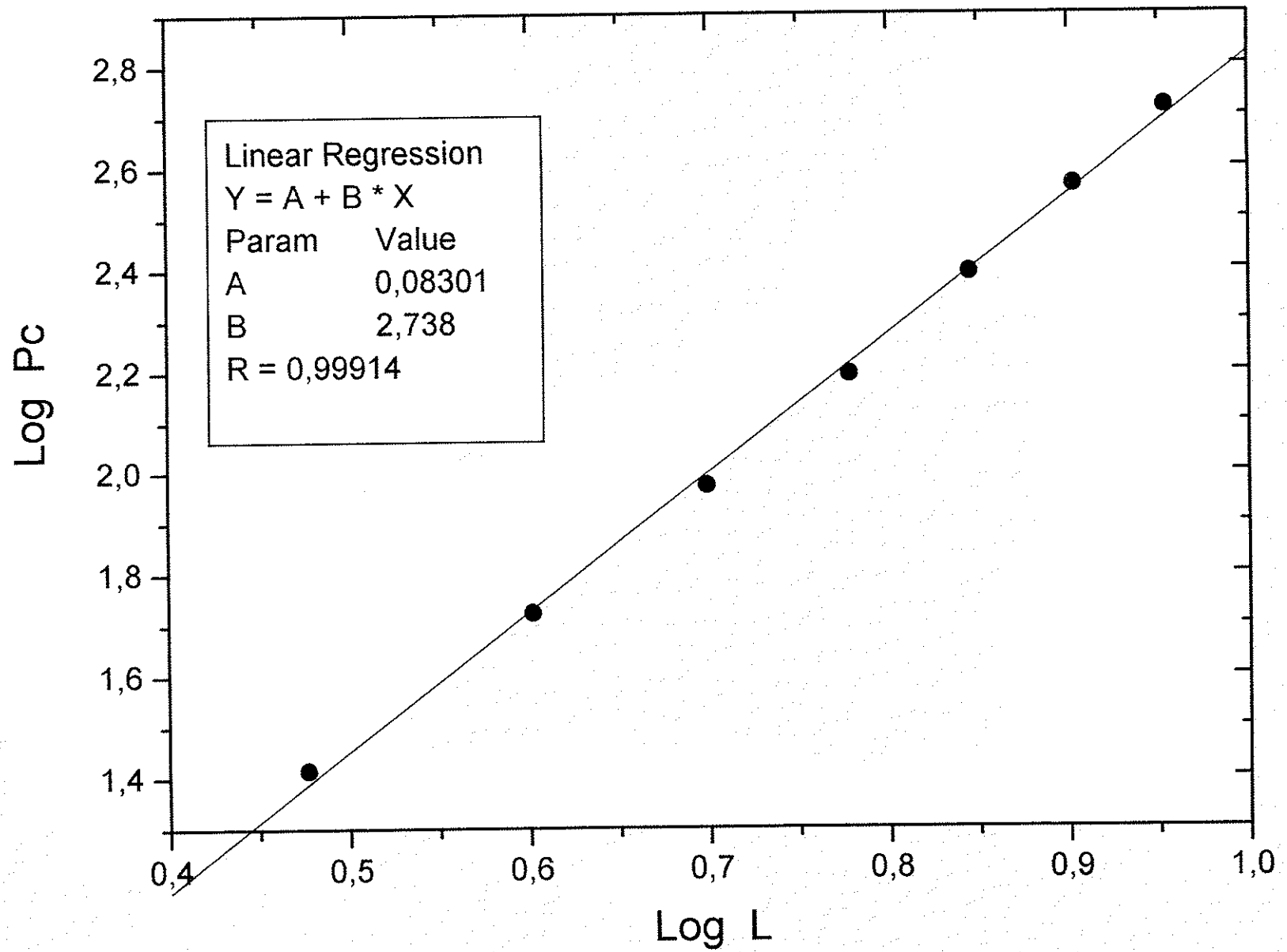


Gráfico 4.3.9
Log Pc X Log L
para lim=3

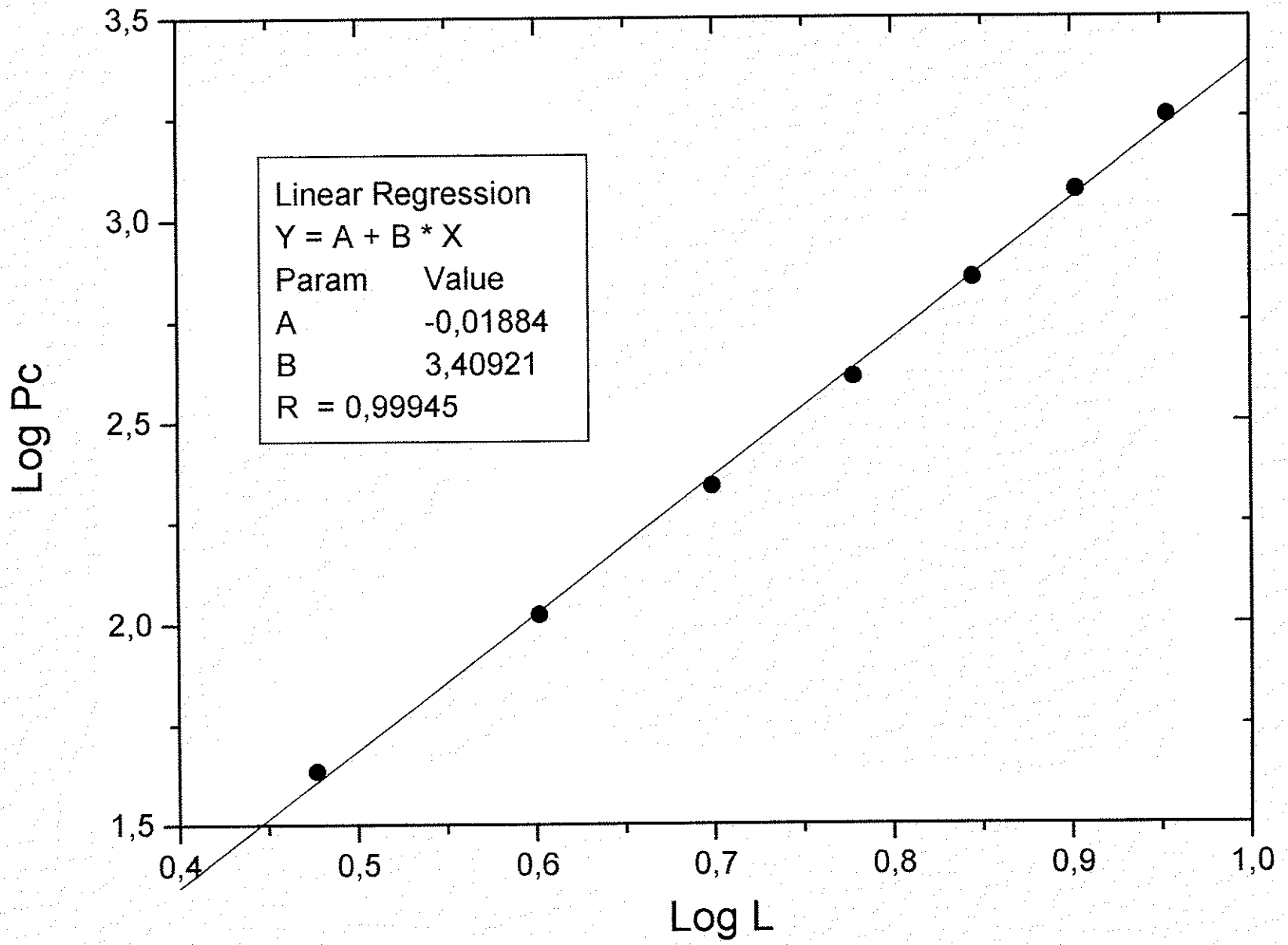


Gráfico 4.3.10
Log Pc X Log L
para lim=4

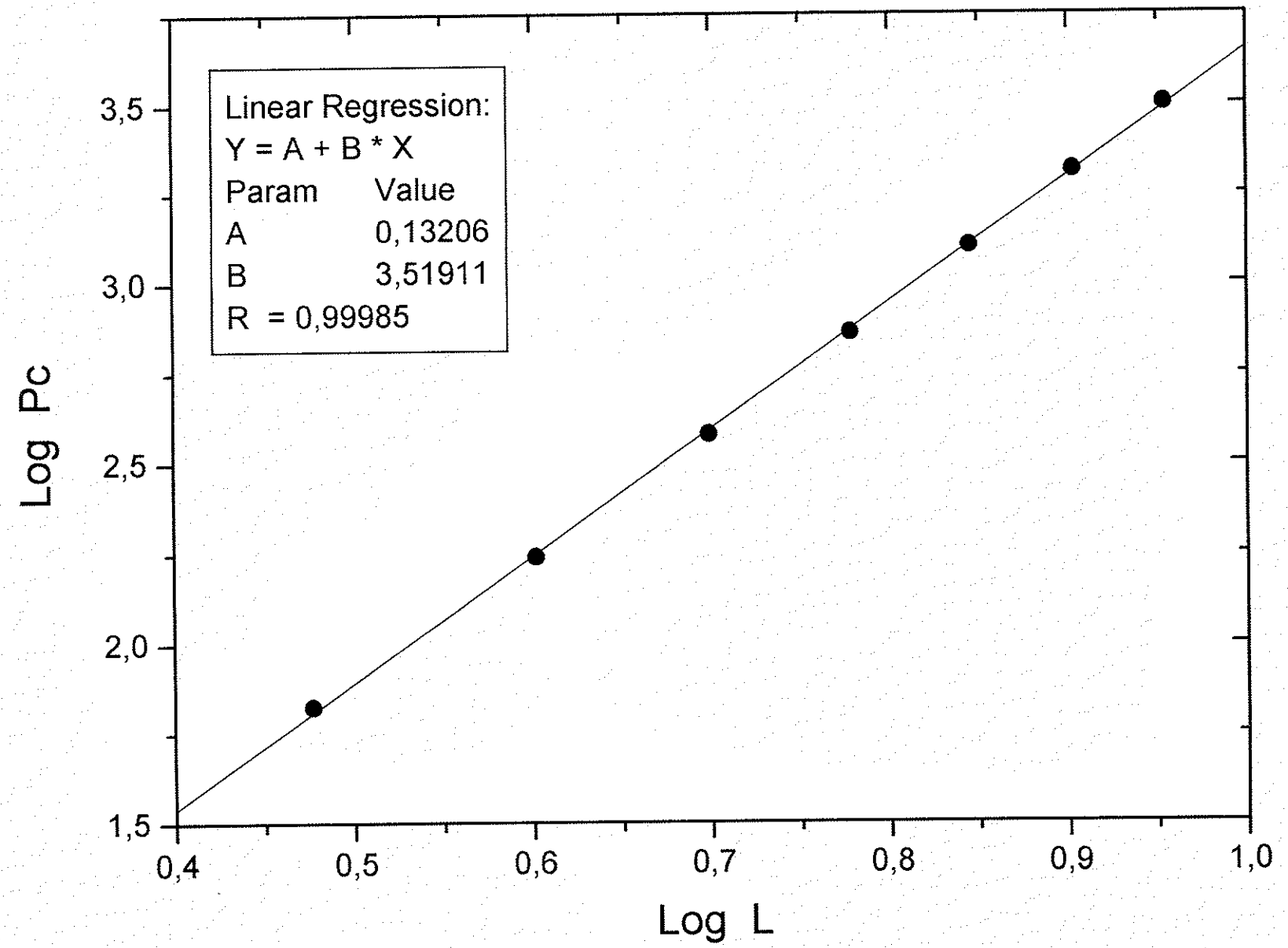


Gráfico 4.3.11
Log Pc X Log L
para lim=5

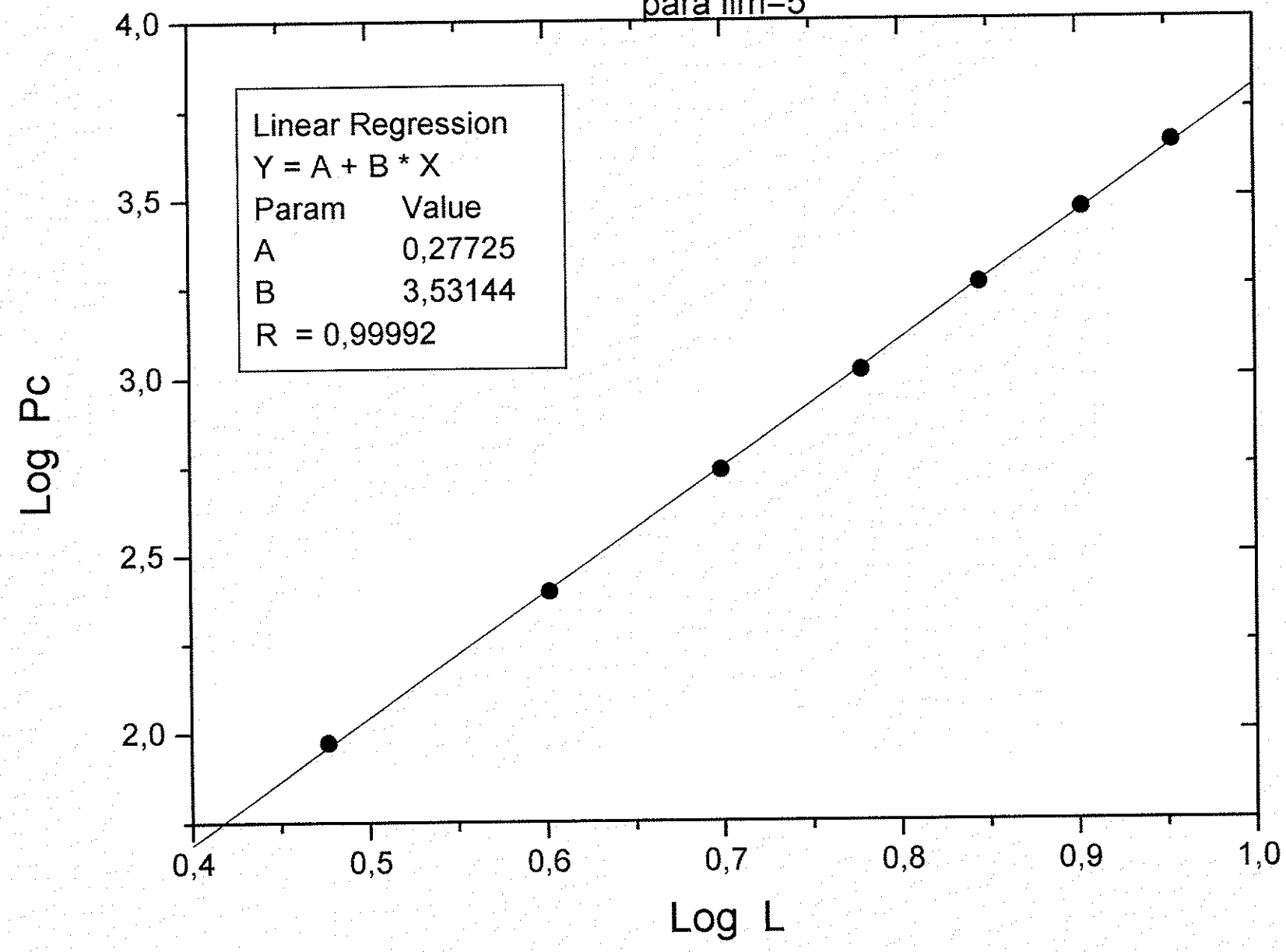


Gráfico 4.3.12
Log Pc X Log L
para lim=6

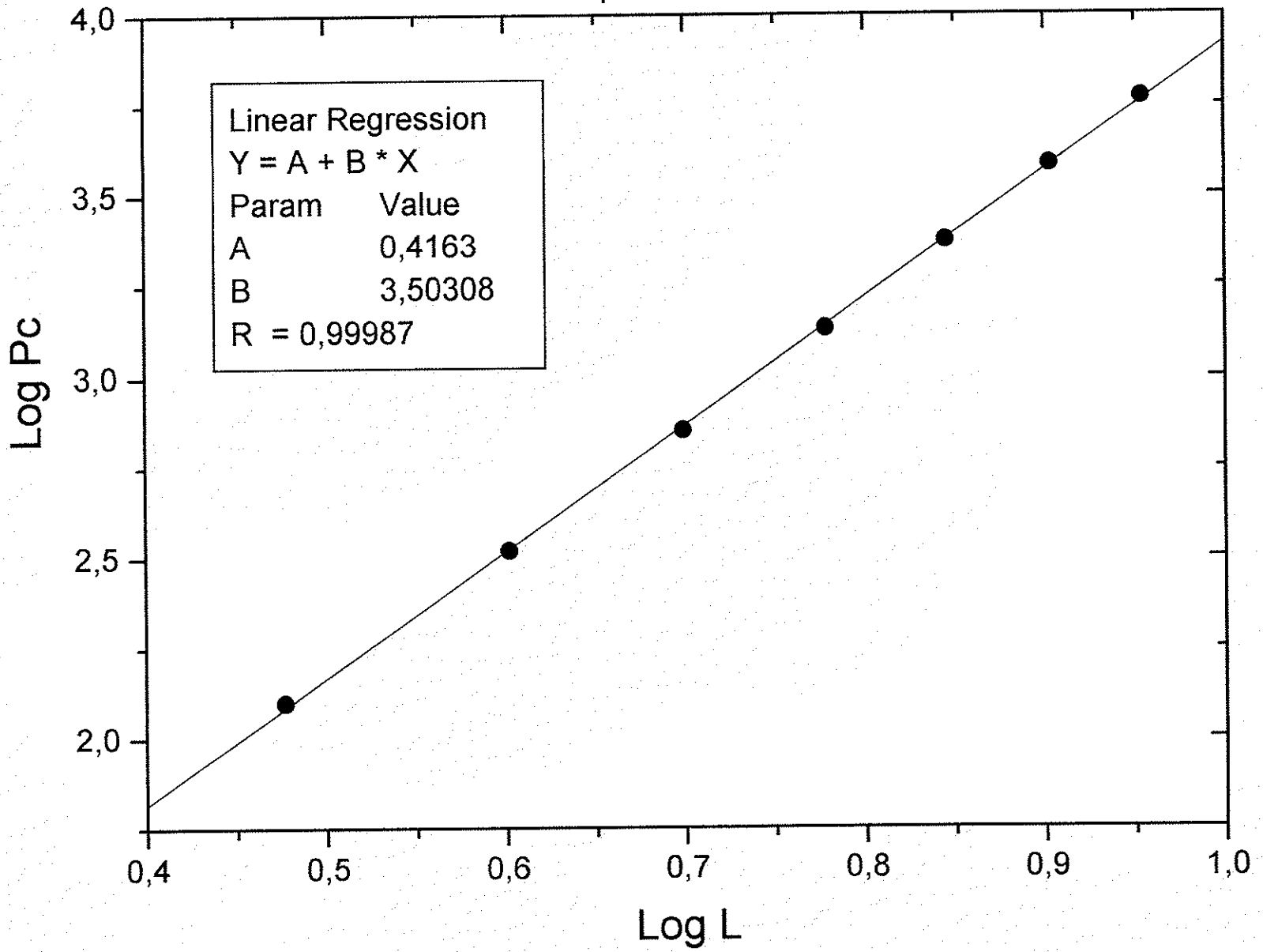


Gráfico 4.3.13
Log Pc X Log L
para lim=7

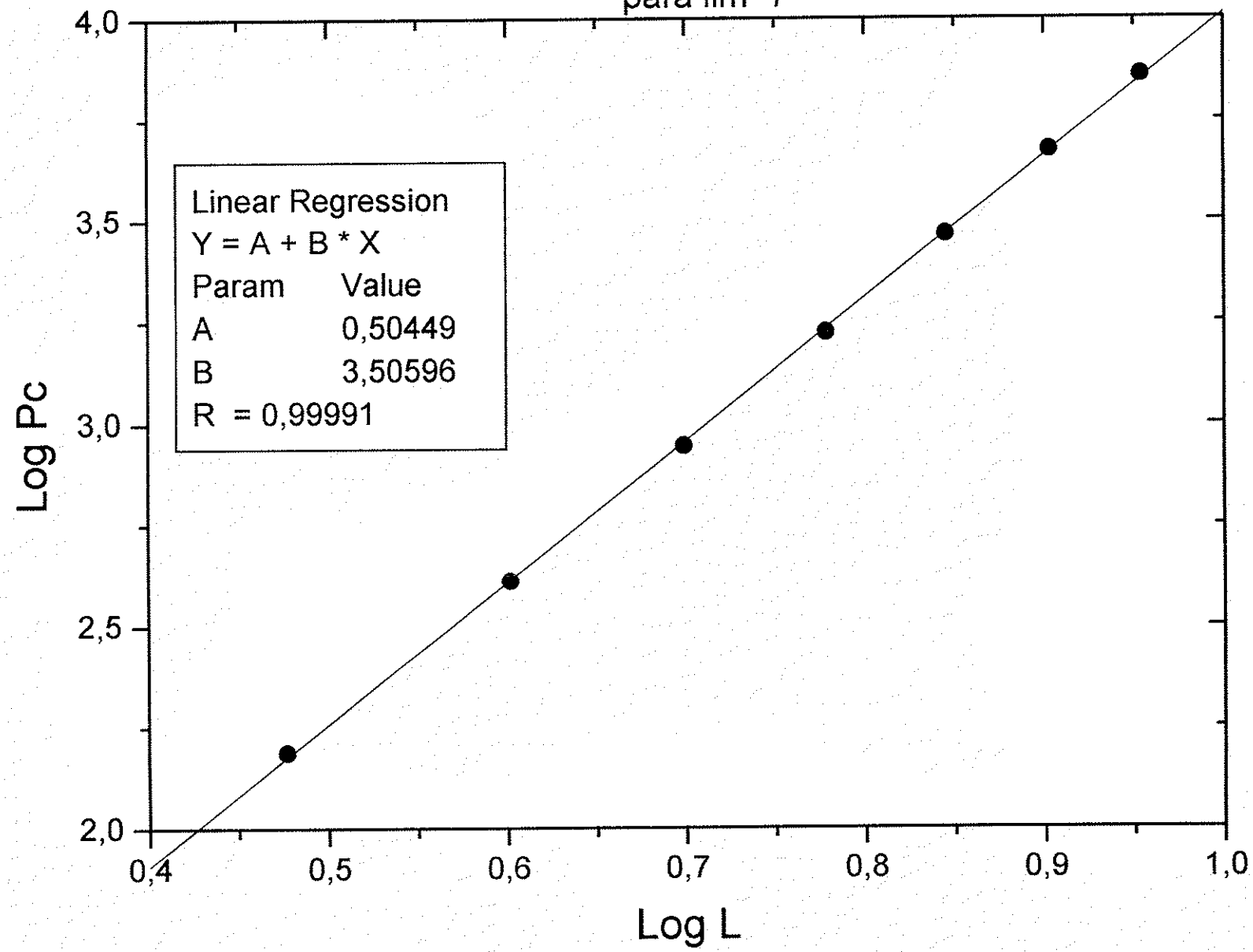
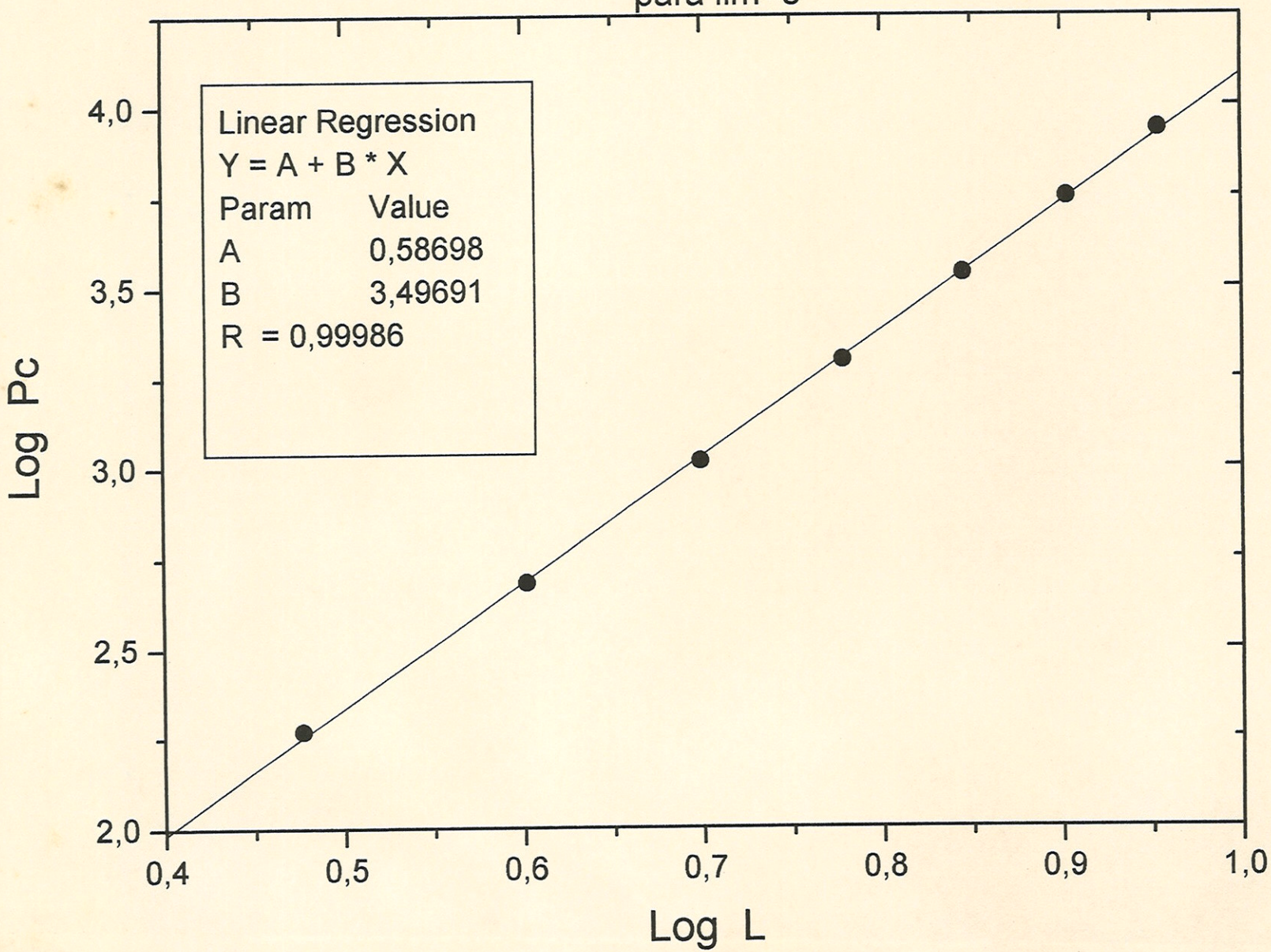



Gráfico 4.3.14
Log Pc X Log L
para lim=8