



***TERRATRÈMOLS
I TEMPORALS
DE LLEVANT:
DOS EXEMPLES
DE SISTEMES
COMPLEXOS***

SECCIÓ DE CIÈNCIES I TECNOLOGIA

15
sèrie jornades científiques



15

sèrie jornades científiques

Jornades
Científiques de
l'Institut d'Estudis
Catalans

SECCIÓ DE CIÈNCIES I TECNOLOGIA

*Terratrèmols i temporals de llevant:
dos exemples de sistemes complexos*

Editor

Antoni M. Correig

Institut d'Estudis Catalans. Jornades Científiques (2002)

Terratrèmols i temporals de llevant : dos exemples de sistemes complexos : Jornades Científiques de l'Institut d'Estudis Catalans, Secció de Ciències i Tecnologia. — (Sèrie jornades científiques ; 15)

Textos en català i anglès. — Bibliografia

ISBN 84-7283-679-7

I. Correig, Antoni M. (Correig Blanchar), ed. II. Títol

1. Catàstrofes naturals — Congressos 2. Terratrèmols — Predicció — Congressos

3. Temporals de llevant — Congressos 4. Previsió del temps — Congressos

55(061.3)

Aquest treball ha comptat amb el suport de la Comissió Interdepartamental de Recerca i Innovació Tecnològica (CIRIT) de la Generalitat de Catalunya

© dels autors de les ponències

© 2003, Institut d'Estudis Catalans, per a aquesta edició

Carrer del Carme, 47. 08001 Barcelona

Primera edició: octubre de 2003

Tiratge: 450 exemplars

Text revisat lingüísticament per l'Oficina de Correcció i Assessorament Lingüístics de l'IEC

Compost per Anglofort, SA

Carrer del Rosselló, 33. 08029 Barcelona

Imprès a Limpergraf, SL

Polígon industrial Can Salvatella. Carrer de Mogoda, 29-31. 08210 Barberà del Vallès

ISBN: 84-7283-679-7

Dipòsit Legal: B. 40977-2003

Són rigorosament prohibides, sense l'autorització escrita dels titulars del *copyright*, la reproducció total o parcial d'aquesta obra per qualsevol procediment i suport, incloent-hi la reprografia i el tractament informàtic, la distribució d'exemplars mitjançant lloguer o préstec comercial, la inclusió total o parcial en bases de dades i la consulta a través de xarxa telemàtica o d'Internet. Les infraccions d'aquests drets estan sotmeses a les sancions establertes per les lleis.

Sumari

<i>Presentació</i>	
Antoni M. Correig	7
<i>La física dels processos atmosfèrics</i>	
Josep Calbó	9
<i>Generació d'un temporal de llevant</i>	
Eliseu Vilaclara	25
<i>Turbulència en sistemes aquàtics naturals</i>	
Elena Roget	41
<i>Predicció numèrica del temps</i>	
Àngel Redaño	57
<i>The physics of earthquakes, seismic occurrence and its prediction</i>	
Vladimir G. Kossobokov	71
<i>Earthquake preparedness: the contribution of earthquake prediction and deterministic hazard research</i>	
Giuliano F. Panza, A. Peresan, F. Vaccari, L. Romashkova, V. Kossobokov, A. Gorshkov i I. Kuznetsov	91
<i>The mathematics and simulation of complex systems (I and II)</i>	
William I. Newman	117
<i>Anàlisi de les sèries temporals associades amb fenòmens naturals</i>	
Antoni M. Correig	147
<i>Prediction of critical transitions in complex systems: colliding cascades model and real disasters</i>	
Vladimir I. Keilis-Borok	175

Presentació

Antoni M. Correig

Laboratori d'Estudis Geofísics Eduard Fontserè, Institut d'Estudis Catalans
Departament d'Astronomia i Meteorologia, Universitat de Barcelona

Des de la més remota antigor, l'home ha estat preocupat per l'ocurrència de desastres naturals i la seva possible predicció. Per a poder predir un fenomen, però, l'hem de conèixer, hem de saber com s'inicia, com evoluciona i com es resol. És aquest el cas dels fenòmens naturals? No si els comparem, per exemple, amb el coneixement que tenim del sistema solar. Però, de fet, no són sistemes comparables: mentre que el moviment dels planetes obeeix a una llei simple, la llei de la gravitació universal, els fenòmens naturals es caracteritzen pel fet de ser sistemes altament no lineals, en què tots els seus components interaccionen entre si, alhora que són sistemes altament estructurats. És el que hom coneix com a *sistemes complexos*, aparentment tots diferents (en què s'assemblen els terratrèmols i els temporals de llevant?), però dels quals podem tractar amb la mateixa metodologia.

En efecte, els fenòmens naturals, com ara els terratrèmols i els temporals de llevant, són sistemes autoorganitzats allunyats de l'equilibri. Ambdós són caracteritzats per un aportament continu d'energia que s'allibera de manera sobtada en finalitzar el procés, el qual es reinicia un altre cop i dona lloc a un nou cicle. Les durades dels cicles són irregulars i el seu final, el que coneixem com a *terratrèmol* o *temporal de llevant*, és previsible/predictible només des d'un punt de vista probabilista. El procés físic que segueixen ambdós fenòmens és aparentment diferent. Podem assimilar els terratrèmols a canvis de fase ordre-desordre perquè abasten un punt crític determinat, mentre que els temporals de llevant estan associats amb processos de turbulència atmosfèrica. Tanmateix, aquests processos s'encreuen i s'entrellacen, atès que un temporal de llevant pot considerar-se com un fenomen crític, i la sismicitat, com la turbulència dels sòlids. Estem, doncs, en presència de fenòmens complexos que podem abordar amb una terminologia comuna, la pròpia dels sistemes complexos.

Per tal de presentar l'estat actual del coneixement que tenim sobre els fenòmens naturals, el gener de 2002, dins el seu programa de Jornades Científiques, l'Institut d'Estudis Catalans organitzà un seminari titulat Terratrèmols i temporals de llevant: dos exemples de sistemes complexos, si bé els temes de què s'hi tractà no es reduïren només a aquests dos aspectes. Hom intentava, amb aquest títol, de cridar l'atenció sobre els processos atmosfèrics i la dinàmica

mica de la terra sòlida partint dels exemples de l'ocurrència de terratrèmols i temporals de lleuant, fenòmens ben coneguts a Catalunya. Donada la impossibilitat de tractar dels temes amb tota la seva amplitud, hom caracteritzà cada un dels dos grans blocs en uns quants temes bàsics: la física dels processos atmosfèrics, la generació de temporals de lleuant, la turbulència (en sistemes aquàtics) i la predicció numèrica del temps pel que fa als processos atmosfèrics; i la física dels terratrèmols, l'ocurrència sísmica i la predicció, i el procés de preparació d'un terratrèmol. A aquests dos grans blocs se n'hi afegí un altre de dedicat a les matemàtiques i la simulació dels sistemes complexos, i l'anàlisi de les sèries temporals associades amb fenòmens naturals. La conferència inaugural anà a càrrec del doctor Jorge Wagensberg, del Museu de la Ciència, que tracta del tema «Cap a una nova teoria de la complexitat?», basat en el seu article «Complexity versus uncertainty: the question of staying alive», publicat a *Biology and Philosophy*, núm. 15 (2000), p. 493-508.

Cada conferenciant hi ha aportat la seva visió des d'un punt de vista molt ampli, moltes vegades interdisciplinari, i ha organitzat el seu escrit de la manera que ha cregut més adequada. Des d'aquest punt de vista, per tant, la publicació no és homogènia. Alguns conferencians s'han centrat en problemes concrets, mentre que uns altres han tractat d'un problema general; alguns n'han tractat de manera qualitativa, i uns altres, quantitativa. Cal no oblidar que en aquestes Jornades hom ha tractat de problemes «vius», que estan en ple desenvolupament, i que algunes de les conclusions aportades poden variar en el futur. Cal, per tant, que el lector s'hi apropi amb mentalitat oberta. La publicació finalitza amb un article del professor Vladimir Keilis-Borok, el qual no va poder assistir a les Jornades tal com era el seu desig. L'article, «Prediction of critical transitions in complex systems: colliding cascades model and real disasters», fou publicat en els Proceedings of the Jubilar Plenary Session of the Pontifical Academy of Sciences, *Scripta Varia*, núm. 99, p. 408-425, «Pontificia Academia Scientiarum» (2001). Agraïm sincerament tant al professor Keilis-Borok com a l'Academia Pontifícia que ens n'hagin autoritzat la publicació, ja que sintetitza l'esperit que vam voler donar a les Jornades.

Per acabar, vull expressar el meu agraïment al doctor Josep Enric Llebot, impulsor de les Jornades, pel seu encoratjament constant; a la senyora Neus Portet i al personal de l'Institut d'Estudis Catalans que tingueren cura de la logística de les Jornades, i a la senyora Mercè Castells, que s'ha encarregat de l'edició del present volum.

Resum

L'atmosfera, aquesta barreja de gasos que envolta la Terra i que és absolutament determinant en l'existència i el desenvolupament de la vida i de les nostres vides, pot ser estudiada des de diversos punts de vista. Un d'aquests punts de vista és l'aportat per la física: es tracta d'aplicar les lleis fonamentals de la física als diversos fenòmens que s'esdevenen en l'atmosfera, considerada el sistema objecte d'estudi. En aquest text ens centrarem en la dinàmica atmosfèrica. Per una banda, s'hi presentaran les equacions que típicament s'utilitzen per a descriure els moviments atmosfèrics, destacant-ne els punts crítics i les no-linealitats, i s'hi dedicarà una atenció especial al tractament dels fenòmens de mesoscala. Per l'altra, hi analitzarem una sèrie experimental de mesures del vent sobre un lloc determinat i hi veurem les dificultats d'ajustar els resultats teòrics a aquestes mesures experimentals.

Abstract

The atmosphere, that mixture of gases surrounding the earth, and which absolutely determines the existence and development of life and our lives, can be studied from various points of view. One of these points of view is that contributed by Physics: the application of the fundamental laws of Physics to the various phenomenons that take place in the atmosphere, considering the system as an object of study. In this text, we will focus on atmospheric dynamics. on one hand, the equations typically used for describing atmospheric movements will be presented, with emphasis on its critical points and nonlinearity. Special attention will be paid to the treatment of mesoscale phenomenons. On the other hand, we will also analyze an experimental series of measurements of wind on a specific place, and we will see the difficulties in adjusting theoretical results to these experimental measurements.

Introducció

L'atmosfera és la capa gasosa que forma la part més externa del planeta Terra. Es tracta d'una barreja de diversos gasos (bàsicament, nitrogen i oxigen en proporcions del 80 % i del 20 % aproximadament) anomenada *aire*, que s'estén des de la superfície de la Terra líquida o sòlida fins a diversos centenars de quilòmetres d'altura. L'atmosfera ha canviat en composició al llarg del temps (en escales geològiques), però roman gairebé inalterable en escales temporals curtes, almenys pel que fa als gasos que hi són presents en concentracions més elevades. L'atmosfera actual és, d'una banda, el resultat de la presència de la vida sobre la Terra, ja que la relativament elevada concentració d'oxigen només es pot explicar per l'efecte dels vegetals; de l'altra, la composició actual de l'atmosfera és indispensable per al manteniment de la vida o, almenys, de les formes de vida suposadament més avançades, com ara l'espècie humana. És ben sabut que petites variacions en la presència d'uns gasos determinats que apareixen com a traces (com ara el diòxid de carboni o l'ozó) poden provocar efectes ben perjudicials per a la vida.

Abans d'aprofundir el contingut d'aquest document, cal introduir una mica de terminologia: d'una banda, la que s'utilitza per a distingir diverses parts de l'atmosfera i, de l'altra, la que s'emptra per a classificar els fenòmens segons la seva extensió en l'espai i el temps. Pel que fa a la divisió vertical de l'atmosfera, s'utilitza com a base la variació de la temperatura amb l'altura (vegeu, per exemple, Andrews, 2000). En general (és a dir, fent una mitjana per a grans extensions i períodes de temps), la temperatura de l'atmosfera disminueix durant els primers deu quilòmetres per sobre de la superfície a la capa anomenada *troposfera*. Per sobre, i fins a uns cinquanta quilòmetres d'altura, la temperatura augmenta amb l'altura, fet que dificulta enormement els moviments verticals de l'aire (això es pot demostrar a partir de consideracions termodinàmiques) i dona peu al nom d'aquesta capa: *estratosfera*. Encara per damunt trobem una altra capa on la temperatura torna a disminuir amb l'altura, anomenada *mesosfera*. Cal esmentar que la majoria de fenòmens d'interès meteorològic (vents, núvols, precipitacions, dispersió de contaminants, etc.) s'acompleixen bàsicament a la troposfera.

Una altra possible classificació té un origen més dinàmic i es basa en la influència que té la superfície del planeta en el moviment de les masses d'aire que passen per sobre seu. Des d'aquest punt de vista, hom distingeix la *capa fronterera* o *capa límit* (on els efectes de la superfície són tan importants com els gradients de pressió o les forces de Coriolis, o més importants) de l'*atmosfera lliure* (on, en general, gradient de pressió i força de Coriolis estan aproximadament equilibrats i donen lloc al *vent geostrofic*).

Pel que fa a la distinció dels fenòmens segons la seva escala espacial i temporal, les propostes fetes per diferents autors són diverses tot i que presenten força punts en comú (un bon resum de les diferents propostes es troba a Thunis i Bornstein, 1996). De totes maneres, tots estan d'acord a parlar de fenòmens de *macroescala*, amb extensions superiors als 10^6 m i durades

superiors a un dia (10^5 s); de *mesoscala*, per a fenòmens que presenten extensions i durades entre 10^3 i 10^6 m i entre 30 min (10^3 s) i un dia respectivament; i de *microescala*, per a fenòmens d'extensió o durada inferior.

En aquesta conferència centrarem la nostra atenció en fenòmens que s'esdevenen a la troposfera i, encara més en particular, a la capa fronterera. A més, aquests fenòmens pertanyen típicament al domini de la mesoscala i, més en particular, al límit inferior de la mesoscala β (10^4 m i 10^4 s). Finalment, s'hi explicaran qüestions relacionades bàsicament amb la dinàmica, tot i que es faran petites incursions en la termodinàmica. En canvi, no s'entrarà en multitud d'uns altres fenòmens que s'acompleixen a l'atmosfera i que són igualment interessants (la transferència radiativa, els canvis de fase de l'aigua i la nucleació, els intercanvis superfície-atmosfera, l'electricitat atmosfèrica, etc.). D'aquesta manera, posem límits al títol de la conferència, que és massa general per a l'extensió aquí admissible.

Les equacions que descriuen el comportament de l'atmosfera

Les equacions bàsiques que regeixen el comportament de l'atmosfera són les que descriuen el comportament d'un fluid newtonià en un sistema de referència en rotació; i, per tant, les que descriuen els fenòmens principals que s'hi esdevenen, com són el seu moviment, i els intercanvis d'energia entre diferents punts del fluid o entre el fluid estudiat i la resta de l'univers. Són, per tant, les equacions de conservació de la massa, de la conservació del moment (o equació del moviment), de l'equació de conservació de l'energia i de l'equació d'estat. El que diferencia aquestes equacions de les que podrien aplicar-se a algun altre fluid són, d'una banda, les constants que caracteritzen l'aire i, de l'altra, les aproximacions que sobre aquelles es puguin fer posteriorment d'acord amb el fenomen que es pretengui representar. Per començar, veurem les equacions escrites respecte d'un sistema de coordenades cartesianes, cosa que vol dir que ja fem una primera aproximació (negligim la curvatura de la Terra a la zona d'aplicació). La presentació i el desenvolupament de les equacions que es faran aquí es poden trobar amb més extensió i detall en diversos textos (com ara Pielke, 1984, o Stull, 1988).

Tenim en primer lloc l'equació de conservació de la massa d'aire, o equació de continuïtat, que es pot expressar de la manera següent:

$$\frac{\partial u_i}{\partial x_i} = -\frac{1}{\rho} \frac{d\rho}{dt}, \quad [1]$$

on x_i ($i = 1, 2, 3$) són les tres coordenades espacials (nord-sud, est-oest i vertical respectivament), que també escriurem quan convingui com a (x, y, z) ; t és la variable temps, u_i són les tres components del vector vent en cada punt i en cada moment, i ρ és la densitat de l'aire.

L'equació de la conservació del moment, o equació del moviment, és la que correspon a un fluid newtonià en un sistema de referència en rotació:

$$\frac{\partial}{\partial t} \rho u_i + \frac{\partial}{\partial x_j} \rho u_j u_i = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j} \left[\mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) - \frac{2}{3} \mu \frac{\partial u_k}{\partial x_k} \delta_{ij} \right] + \rho g_i - 2\rho \varepsilon_{ijk} \Omega_j u_k, \quad [2]$$

on p és la pressió atmosfèrica, μ és la viscositat molecular de l'aire ($\mu = 1,789 \times 10^{-5} \text{ kg m}^{-1} \text{ s}^{-1}$), g_i és el vector acceleració de la gravetat ($0,0,-g$) ($g = 9,8 \text{ m s}^{-2}$), Ω_j és el vector velocitat angular de la Terra ($\Omega = 7,27 \times 10^{-5} \text{ rad s}^{-1}$), δ_{ij} és la delta de Kronecker i ε_{ijk} és el tensor de Levi-Civita. En el primer membre d'aquesta equació hi ha les acceleracions i , en el segon membre, les forces, que són, en el mateix ordre en què apareixen els termes: la força de pressió, la de la viscositat molecular, la de la gravetat i la de Coriolis. La força centrífuga, que presenta una direcció gairebé paral·lela a la de la gravetat però és molt més petita en mòdul, es negligeix en aquesta equació. Igualment, no s'hi considerarà el canvi de l'acceleració de la gravetat amb l'altura (habitualment, això es tracta mitjançant la introducció de l'anomenada *altura geopotencial*).

A aquestes equacions purament dinàmiques cal afegir-hi l'equació de conservació de l'energia, la forma més elemental de la qual és la que es desprèn del primer principi de la termodinàmica, és a dir:

$$\delta c = \delta w_T + du_e = \frac{S_c}{\rho} dt,$$

on δc és la calor intercanviada en un cert procés per unitat de massa d'un cert volum d'aire, δw_T indica el treball per unitat de massa efectuat per aquest volum en aquest procés, du_e és la variació de l'energia interna per unitat de massa del volum d'aire durant aquest procés i S_c és el terme de fonts i embornals d'energia per unitat de volum i de temps. De totes maneres, no és habitual utilitzar l'equació en aquesta forma, sinó que, dividint per la temperatura absoluta, utilitzant el segon principi de la termodinàmica, l'equació d'estat, la relació entre les calors específiques a pressió constant i volum constant i, finalment, definint la temperatura potencial com a:

$$\theta = T \left(\frac{p_{00}}{p} \right)^{r_d / c_p},$$

es pot arribar a escriure aquesta equació en una forma més usual:

$$\frac{d\theta}{dt} = \frac{1}{\rho c_p} \left(\frac{p_{00}}{p} \right)^{r_d / c_p} S_c, \quad [3]$$

on la derivada total de la temperatura potencial es podria descompondre encara en els termes local i advection (com s'ha fet, per exemple, en l'equació del moviment).

L'equació d'estat és la dels gasos ideals, on s'introdueix la temperatura virtual, a causa de la presència de vapor d'aigua en l'aire, amb la qual cosa s'aconsegueix estalviar una equació (la d'estat del vapor d'aigua) i un paràmetre (la constant termodinàmica del vapor):

$$p = \rho r_d T_v, \quad [4]$$

on r_d és la constant termodinàmica de l'aire sec ($r_d = 287 \text{ J kg}^{-1} \text{ K}^{-1}$) i T_v és la temperatura virtual de l'aire, que inclou el contingut d'humitat.

La dificultat —de fet, la impossibilitat— per a resoldre analíticament aquestes equacions és força òbvia, i no solament per la no-linealitat del terme advection. En aquest punt hi ha dues opcions, no necessàriament excloents: simplificar mitjançant aproximacions les equacions, per tal de permetre'n la solució analítica, o resoldre-les mitjançant tècniques numèriques (Calbó *et al.*, 1994; Calbó i Baldasano, 1995). Aquí presentarem un tractament de les equacions encaminat inicialment a la seva solució numèrica, tot i que al final acabarem fent prou aproximacions per a determinar una solució de manera analítica.

Quan les equacions anteriors es volen aplicar a la mesoscala, la primera cosa que fan la majoria dels autors és dividir el valor d'algunes de les variables en dues parts: la primera, associada amb l'estat sinòptic o de macroescala de l'atmosfera; la segona, la fluctuació sobre l'anterior motivada pel fenomen de mesoscala. En tot el que ve a continuació, la part associada a l'estat sinòptic (que sovint s'anomena *estat bàsic* de l'atmosfera) es considera constant en el decurs del fenomen de mesoscala encara que, lògicament, això no sigui estrictament cert. També és habitual suposar vàlida la hipòtesi d'equilibri hidrostàtic per a l'estat bàsic i, finalment, assumir la hipòtesi de Boussinesq, que fa referència a la importància relativa de les fluctuacions de la densitat.

El conjunt d'equacions que s'hi obté continua essent (igual que el conjunt d'equacions 1-4 presentat més amunt) un sistema tancat en el sentit que el nombre d'equacions i d'incògnites és el mateix (hi ha quatre equacions, amb quatre incògnites: la densitat, la temperatura, la pressió i el vent) sempre que suposem conegut l'estat bàsic. Però encara no és l'adequat per a estudiar l'atmosfera, ja que s'ha de tenir en compte que el flux atmosfèric sovint és turbulent. És a dir, aquestes equacions són característiques d'un flux laminar i, per tant, sempre donen com a resultat un flux d'aquest tipus, que és inestable per a petites perturbacions si les condicions són adequades per a la turbulència. En conseqüència, cap de les solucions que obtinguéssim no seria observable en l'atmosfera (especialment quan estudiem la capa més propera a la superfície de la Terra —la capa fronterera—, on la turbulència és dominant).

Una altra manera d'entendre aquesta mateixa qüestió és pensant que, de fet, és conseqüència de la impossibilitat tant de mesurar com de calcular les diferents variables amb prou resolució espacial, imposada per les condicions de validesa de les equacions. Per deduir-les, s'han fet (sense explicitar-ho) bàsicament dues hipòtesis inicials: que els increments de longitud infinitesimals considerats són molt més grans que la distància mitjana entre les molècules (per a poder

tractar de l'aire d'acord amb el seu comportament estadístic) i que són prou petits per a poder negligir els termes de segon ordre en els desenvolupaments en sèries de Taylor que s'efectuen de les diferents variables. Això porta a escales espacials de prop d'1 cm, i temporals de prop d'1 s. En conseqüència, si es vol treballar en una zona, per exemple, de 10 km × 10 km × 1.000 m, s'hauria de definir una xarxa de 10¹⁷ punts, que no pot ser tractada per cap ordinador actual (Zannetti, 1990).

La qüestió és, doncs, introduir la turbulència (o, cosa que és igual, el transport que s'esdevé en escales inferiors a les de la xarxa de treball) en les equacions anteriors. El que és més habitual en aquest sentit és seguir la convenció, deguda a Reynolds, de descompondre el flux (això és, les variables que el descriuen) en uns valors mitjans, que varien tant espacialment com temporalment amb una certa lentitud, i uns valors turbulents, que varien ràpidament de manera quasi aleatòria. En el nostre cas, per a les variables que anteriorment hem descompost en part sinòptica i en part de mesoscala, farem la nova descomposició sobre aquesta darrera.

Aleshores, totes les equacions es poden descompondre en dues: una per als valors mitjans (els del flux *observable* mitjà, per tant) que es poden resoldre definint una certa xarxa de discretització no gaire fina, i una altra per als valors turbulents, és a dir, per als fenòmens que no es poden resoldre sobre aquesta xarxa. A continuació, es donen les equacions corresponents al flux mitjà, que s'obtenen de les equacions bàsiques, on se substitueixen les variables per la seva descomposició corresponent, i es fa la mitjana a continuació. Per a la seva deducció s'ha tingut en compte que l'operació de derivar i la de fer la mitjana poden intercanviar l'ordre amb què s'efectuen.

$$\frac{\partial}{\partial x_i} (\rho_0 \bar{u}_i) = 0$$

$$\frac{\bar{T}_v}{T_{v0}} = - \frac{\bar{p}}{\rho_0}$$

$$\frac{\partial \bar{u}_i}{\partial t} = - \bar{u}_j \frac{\partial \bar{u}_i}{\partial x_j} - \frac{1}{\rho_0} \frac{\partial \rho_0 \overline{u'_i u'_j}}{\partial x_j} - \theta_0 \frac{\partial (\pi_0 + \bar{\pi})}{\partial x_i} - 2\varepsilon_{ijk} \Omega_j \bar{u}_k \quad (\text{per } i = 1, 2)$$

$$\frac{\partial \bar{u}_3}{\partial t} = - \bar{u}_j \frac{\partial \bar{u}_3}{\partial x_j} - \frac{1}{\rho_0} \frac{\partial \rho_0 \overline{u'_j u'_3}}{\partial x_j} - \theta_0 \frac{\partial \bar{\pi}}{\partial x_3} + \frac{\bar{\theta}}{\theta_0} g - 2\varepsilon_{3jk} \Omega_j \bar{u}_k$$

$$\frac{\partial (\theta_0 + \bar{\theta})}{\partial t} = - \bar{u}_i \frac{\partial (\theta_0 + \bar{\theta})}{\partial x_i} - \frac{1}{\rho_0} \frac{\partial \rho_0 \overline{u'_i \theta'}}{\partial x_i} + \frac{1}{\rho_0 \pi_0} \bar{S}_c.$$

En totes les equacions, el subíndex 0 indica l'estat bàsic, la barra per damunt indica l'estat mesoscalar mitjà i la prima indica la fluctuació turbulenta. Hi apareix també la pressió d'Exner, π ,

que està relacionada directament amb la pressió (d'una manera semblant a com la temperatura potencial ho està amb la temperatura).

Si es repassen les equacions anteriors per als valors mitjans, s'observa que, en introduir-hi la turbulència (o els fenòmens associats amb escales inferiors a la xarxa de discretització), han aparegut noves incògnites. És a dir, a les ja conegudes cal afegir-hi ara els fluxos turbulents $\overline{u'_i u'_j}$, $\overline{u'_i \theta'}$, que, fins i tot tenint en compte criteris de simetria, són un total de nou incògnites noves. S'hi pot veure que les equacions per als valors turbulents tampoc no són útils (ni suficients) per a la seva determinació. Si s'intenten establir noves equacions per als fluxos turbulents, resulta que apareixen uns nous termes amb productes de tres valors turbulents, la qual cosa significa que apareixen moltes més incògnites noves. Aquest fet és el que es coneix com a *problema de clausura*, que es pot enunciar dient que les equacions termodinàmiques d'un flux turbulent no formen un sistema tancat. Aquest problema és conseqüència de la no-linealitat del terme advection de les equacions i és un dels problemes de la física clàssica que continuen sense solució (Stull, 1988). La manera de tractar-ne és escollir un nombre finit d'equacions i aproximar la resta d'incògnites segons les quantitats conegudes (les que es determinen per a alguna de les equacions) i alguns paràmetres. En aquest sentit, es parla de la parametrització de les incògnites com a solució del problema de clausura.

No és l'objectiu d'aquesta conferència aprofundir més en què és la turbulència i com se'n tracta. Només hi introduïrem la parametrització més senzilla, anomenada *de primer ordre*, o, també, *teoria K*. Es tracta d'establir que hi ha una relació directa entre els fluxos turbulents d'una variable i els gradients dels valors mitjans de la mateixa variable i en la mateixa direcció. Aquesta aproximació té la seva justificació teòrica i experimental, però presenta també moltes limitacions en la seva aplicabilitat a una gran diversitat de situacions.

El perfil vertical del vent horitzontal. L'espiral d'Ekman

Per tal d'il·lustrar el que s'ha explicat fins a aquest moment, pot ser convenient determinar una solució analítica de les equacions anteriors o, més ben dit, d'una simplificació de les equacions anteriors. Aquest desenvolupament es pot trobar en diversos textos, com, per exemple, Arya (1988). Així, suposarem un conjunt d'hipòtesis: 1) estat estacionari, 2) homogeneïtat horitzontal, 3) absència de component vertical del vent. Les equacions per a les dues components horitzontals del vent queden:

$$0 = -\frac{1}{\rho_0} \frac{\partial \rho_0 \overline{u'_1 u'_3}}{\partial x_3} - \theta_0 \frac{\partial (\overline{\pi_0} + \overline{\pi})}{\partial x_1} - 2\varepsilon_{132} \Omega_3 \overline{u_2}$$

$$0 = -\frac{1}{\rho_0} \frac{\partial \rho_0 \overline{u'_2 u'_3}}{\partial x_3} - \theta_0 \frac{\partial (\overline{\pi_0} + \overline{\pi})}{\partial x_2} - 2\varepsilon_{231} \Omega_3 \overline{u_1} .$$

Si, a més, suposem: 4) atmosfera incompressible, i 5) teoria K, amb la K, anomenada *viscositat turbulenta*, constant; i escrivim les equacions que resulten amb una notació més senzilla (per exemple, eliminant les barres de sobre les variables mitjanes o tornant a escriure la pressió de la manera més habitual), tenim:

$$0 = K \frac{\partial^2 u}{\partial z^2} - \frac{1}{\rho_0} \frac{\partial p}{\partial x} + fv$$

$$0 = K \frac{\partial^2 v}{\partial z^2} - \frac{1}{\rho_0} \frac{\partial p}{\partial y} - fu,$$

on f és el paràmetre de Coriolis ($f = 2 \Omega \sin\varphi$, essent φ la latitud del lloc).

Encara més, si suposem que el camp de pressions no canvia amb l'altura, podem escriure els termes del gradient de pressió d'acord amb el vent geostròfic (constant i uniforme) a què donaria lloc i resulta:

$$0 = K \frac{\partial^2 u}{\partial z^2} - fv_G + fv$$

$$0 = K \frac{\partial^2 v}{\partial z^2} - fu_G - fu.$$

Aquestes equacions es poden resoldre, utilitzant com a condicions de contorn:

$$u = v = 0 \quad , \quad a z = 0$$

$$u = u_G \quad ; \quad v = v_G \quad , \quad a z = \infty \quad (\text{és a dir, a l'atmosfera lliure, lluny de la superfície}).$$

El resultat que s'obté és:

$$u - u_G = -e^{-az} [u_G \cos(az) + v_G \sin(az)]$$

$$v - v_G = -e^{-az} [u_G \sin(az) - v_G \cos(az)] ,$$

amb $a = \sqrt{f/2K}$. Per tal de poder analitzar el resultat més fàcilment (però sense perdre generalitat), considerem un cas en què el vent geostròfic imposat a l'atmosfera lliure només té component en la direcció x (per exemple, $v_G = 0$). Aleshores:

$$u = u_G [1 - e^{-az} \cos(az)]$$

$$v = u_G e^{-az} \sin(az) .$$

Si representem com varia el vector vent amb l'altura (fig. 1), obtenim la forma d'una espiral que porta el nom de *espiral d'Ekman* en honor de l'oceanògraf suec que va proposar aquesta solució (de fet, el seu desenvolupament original va ser fet per als corrents oceànics). Normalment, es pren com a altura de la capa d'Ekman el valor $z_E = \pi / a$, ja que en aquest nivell el vent ja pren un valor prou semblant al geostròfic. Considerant latituds mitjanes com la nostra ($f = 10^{-4} \text{ s}^{-1}$) i valors típics de la viscositat turbulenta ($K = 10 \text{ m}^2 \text{ s}^{-1}$), l'altura de la capa d'Ekman seria d'uns 10^3 m .

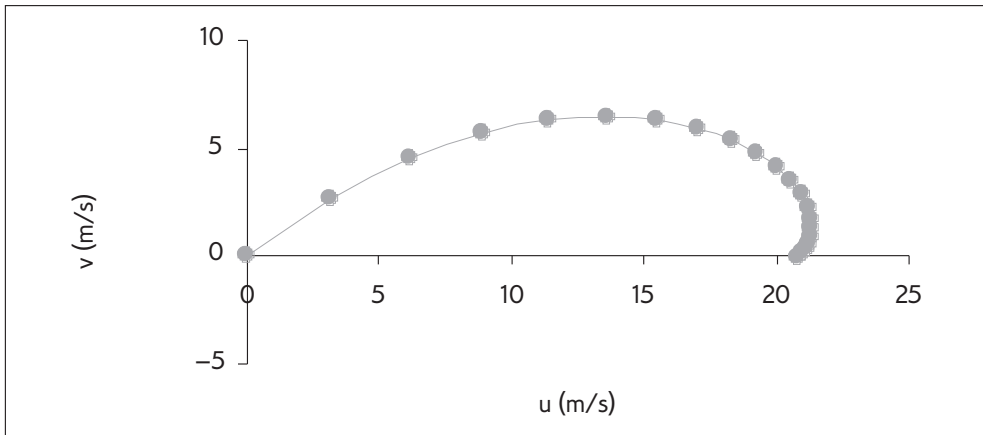


FIGURA 1. Hodògrafa corresponent a una espiral d'Ekman, obtinguda amb $f = 10^{-4} \text{ s}^{-1}$, $K = 5 \text{ m}^2 \text{ s}^{-1}$, i $u_G = 20 \text{ m/s}$. Els punts hi són representats a intervals de 50 m d'altura, entre 0 i 1.000 m.

Pel que fa al mòdul d'aquest vector, és a dir, pel que fa a la velocitat, tenim:

$$U^2 = u^2 + v^2 = M_G^2 (1 + e^{-2z} - 2e^{-az} \cos(az)) .$$

No obstant això, molt poques vegades s'ha observat a l'atmosfera real variacions del vent amb l'altura que s'ajustin a un comportament com el que prediu l'espiral d'Ekman. Sens dubte, l'explicació d'aquest fet cal buscar-la en la gran quantitat d'hipòtesis que s'han suposat per a deduir-la. En particular, l'homogeneïtat del terreny subjacent i la uniformitat i la constància de la viscositat turbulenta són dues de les hipòtesis que més rarament s'acompleixen. Ateses les hipòtesis, l'espiral d'Ekman s'hauria d'ajustar millor a la part de la capa fronterera que hi ha per sobre de la capa superficial, ja que en aquesta darrera, en general, el terme de turbulència és molt gran comparat amb el de Coriolis i, a més, varia ràpidament amb l'altura. També s'hauria d'ajustar a situacions de capa superficial molt estable (des del punt de vista de l'estratificació tèrmica), atès que aleshores la viscositat turbulenta té un valor més petit i més uniforme en aquesta capa. Sigui com sigui, la deducció i els resultats anteriors permeten destacar alguns fets que són patents: la important variació de la magnitud del vent amb l'altura especialment

en els primers metres damunt de la superfície i la variació també de la direcció d'aquest vent, que, en principi, a l'hemisferi nord, hauria de ser en sentit horari a mesura que pugem en la capa fronterera atmosfèrica.

El vent real: algunes dades experimentals

Per tal de mostrar la complexitat d'estudiar el vent a la capa fronterera, presentarem a continuació algunes dades experimentals de què, com veurem, ni que sigui qualitativament, no és fàcil de tractar amb la teoria desenvolupada més amunt. Les dades van ser mesurades en una torre instrumentada amb anemòmetres a 10, 20 i 45 m d'altura. El lloc de mesura és un altiplà petit a Navarra. La sèrie de dades és llarga, tot i que aquí en presentarem només una part, corresponent al mes de març de 2001. Les mesures que s'hi enregistren són els valors mitjans de la velocitat en intervals de 10 min. A la figura 2 es presenta la sèrie de dades per a l'anemòmetre situat a 45 m (que és el que *a priori* hauria de presentar menys influència dels efectes del terreny). Tot i això, i malgrat que són valors mitjans en 10 min, hi observem la gran variabilitat en la velocitat del vent. A la figura 3 es presenta la mateixa sèrie, on s'ha efectuat una mitjana mòbil en finestres de 3 h. Fins i tot després d'efectuar aquesta mitjana, la variabilitat del vent continua essent molt gran. Un tractament possible per a analitzar una sèrie d'aquest tipus és efectuar una anàlisi espectral. En efecte, una anàlisi d'aquesta mena, que no s'ha fet en aquest cas, acostuma a presentar molta energia en el període diari (lògicament, atès que la radiació solar és el mecanisme principal de «forçament» del sistema). També acostuma a aparèixer-hi, sig-

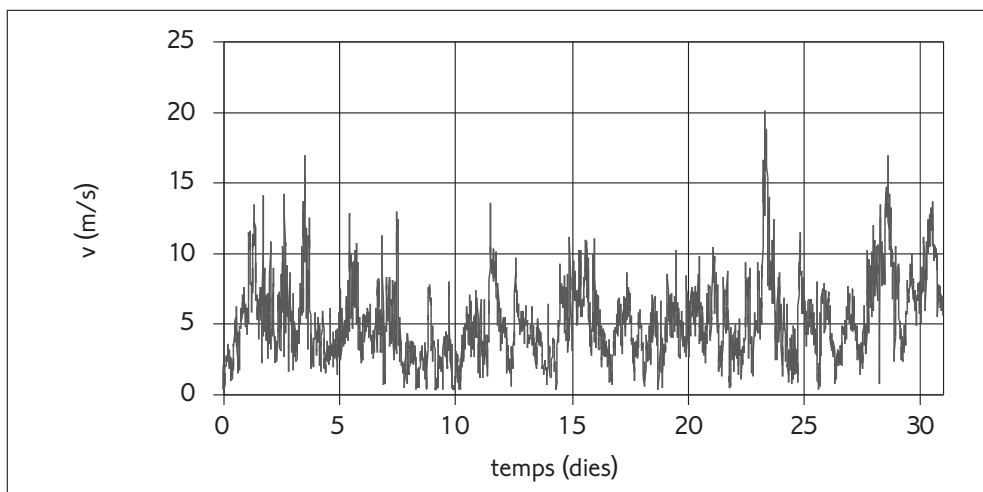


FIGURA 2. Velocitat del vent mesurada a 45 m sobre el terra durant el mes de març de 2001 en una torre situada en un altiplà a Navarra. Els valors representats són mitjanes per intervals de 10 min.

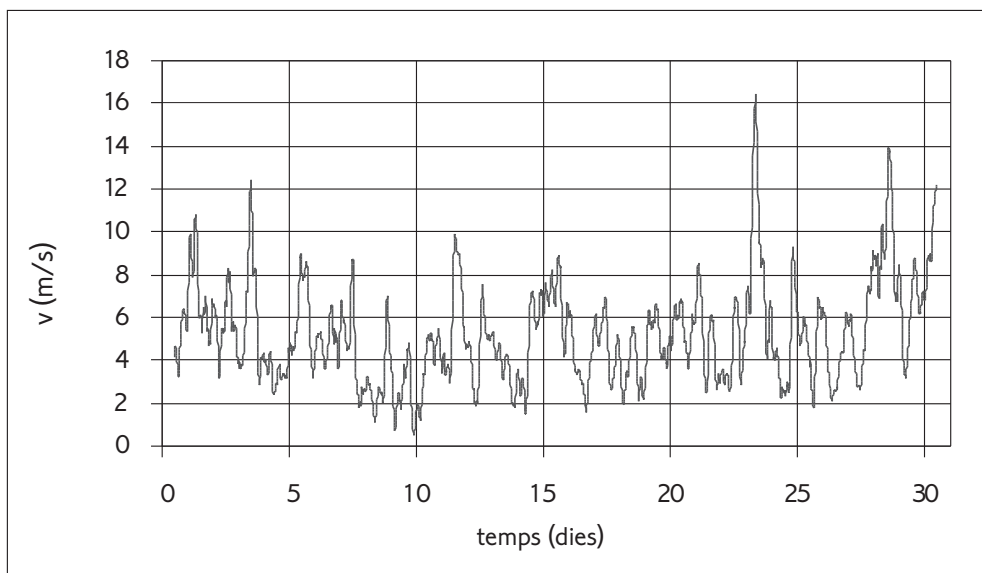


FIGURA 3. Velocitat del vent mesurada a 45 m sobre el terra durant el mes de març de 2001 en una torre situada en un altià a Navarra. S'hi representa un valor cada 10 min, que correspon a la mitjana mòbil efectuada per intervals de 3 h.

nificativament, una freqüència més baixa, és a dir, un període d'uns quants dies, que es pot associar amb el pas dels grans sistemes sinòptics. A l'espectre sovint apareixen unes altres freqüències (més altes) significatives, associades amb la turbulència, i normalment hi ha un conjunt de freqüències intermèdies amb menys energia associada. A la figura 4 es representa la diferència entre els valors mesurats (per exemple, els de la figura 2) i els amitjanats en 3 h (per exemple, els de la figura 3). D'alguna manera, i sense ser excessivament rigorosos, aquests valors poden estar relacionats amb les fluctuacions turbulentes que més amunt s'han esmentat. En tot cas, torna a posar-se en evidència la dificultat d'analitzar l'origen físic d'aquestes fluctuacions, i també s'hi observa que aquestes presenten sovint unes amplituds semblants als valors mitjans.

Analitzem ara la variabilitat espacial i, més concretament, com el vent varia amb l'altura en aquesta torre. Per exemple, a la figura 5 s'han representat els valors enregistrats (cada 10 min) durant dos dies concrets i en els tres nivells de mesura. S'hi observa que, en alguns moments, el vent en tots tres nivells és molt semblant. Podem aventurar que aquests moments corresponen a situacions d'estratificació neutra o inestable, és a dir, quan hi ha més barreja vertical. Per exemple, això podria succeir abans del migdia del dia 23, quan el vent molt fort permet preveure una situació neutra; o després del migdia del dia 24, quan hipotèticament les condicions podrien ser favorables per a la convecció. Per contra, en uns altres moments hi ha diferències ben apreciables entre el vent a diferents altures (per exemple, durant la nit del 24 al 25 o durant la tarda i el ves-

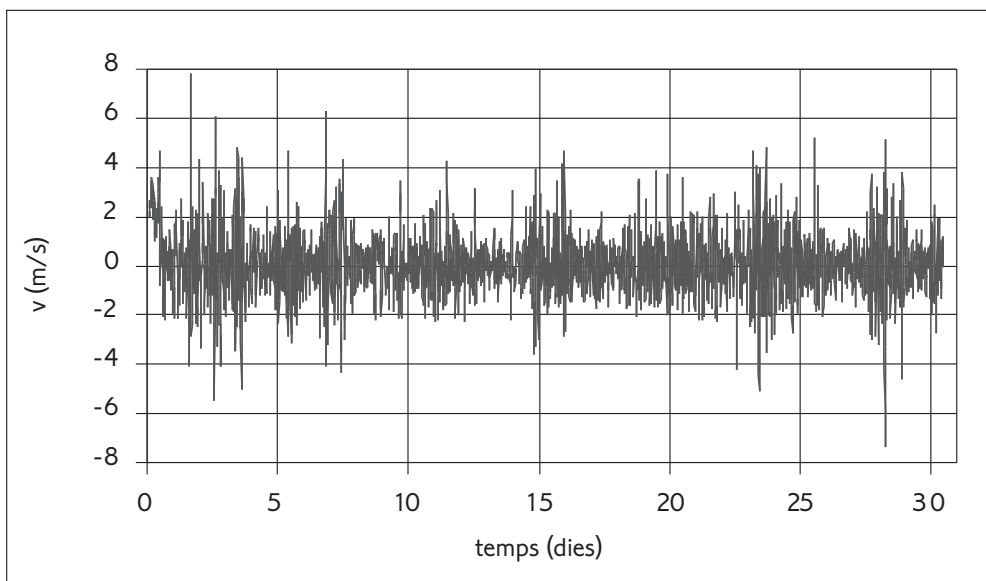


FIGURA 4. Fluctuacions de la velocitat del vent mesurada a 45 m sobre el terra durant el mes de març de 2001 en una torre situada en un altià a Navarra. Els valors corresponen a la diferència entre el valor enregistrat (mitjanes per intervals de 10 min) menys la mitjana mòbil en 3 h.

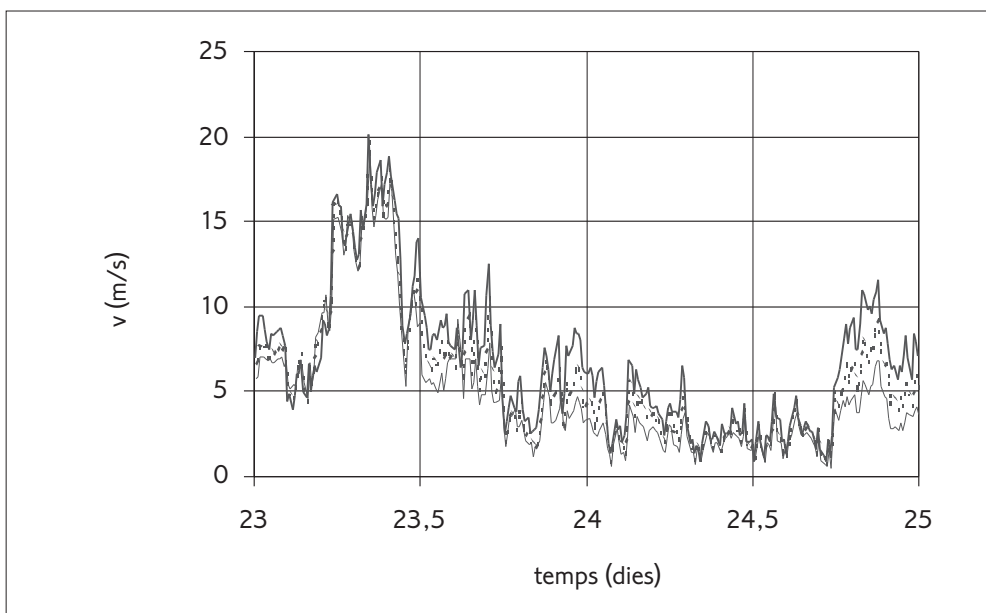


FIGURA 5. El registre del vent durant dos dies (23 i 24 de març) en els tres nivells de mesura de la torre: 45 m (línia gruixuda), 20 m (línia discontinua), 10 m (línia prima).

pre del dia 25). En aquests moments, i com caldria esperar, el vent creix amb l'altura. Previsiblement, en aquells moments l'atmosfera presentava una estratificació més estable.

Podem ara analitzar amb una mica més de detall aquestes variacions amb l'altura. Per això hem representat a les figures 6 i 7 unes situacions en particular. A la figura 6 s'han representat tots els perfils que complien la condició que, a 45 m, la velocitat del vent fos propera

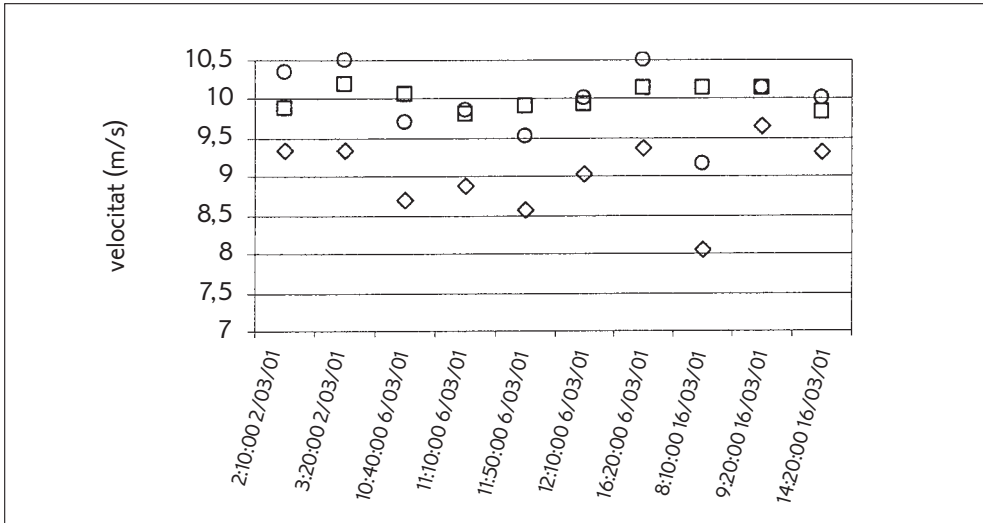


FIGURA 6. El vent mesurat en els tres nivells en les situacions en les quals el vent a 45 m (quadrats) era en l'interval $10 \pm 0,2$ m/s i la seva direcció en l'interval $150 \pm 5^\circ$. El vent a 20 m és indicat pels cercles, i a 10 m, pels rombes.

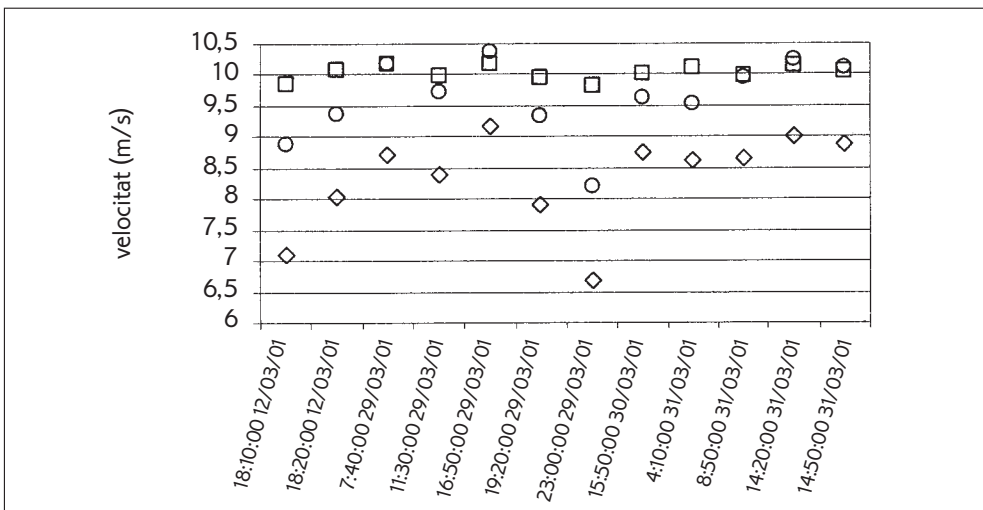


FIGURA 7. Com a la figura anterior, però per a la velocitat en l'interval $10 \pm 0,2$ m/s i la direcció en l'interval $335 \pm 5^\circ$.

a 10 m/s i la direcció fos propera als 150° (SSE). De les deu situacions que hem trobat, hi podem observar que en tres casos el vent a 20 m és més fort que a 45 m. En quatre, el vent a 20 m és gairebé igual que a 45 m. Només en els altres tres casos el vent segueix un perfil creixent amb l'altura. En realitat, les deu situacions es donen tan sols en tres dies del mes de març, dies, però, que evolucionen de maneres diferents. A la figura 7 s'ha seguit la mateixa idea, però ara per a un vent a 45 m que presenti una direcció propera als 335° (NNW). Aquí els perfils que es troben són o bé creixents amb l'altura (set casos), o bé presenten vents iguals a 20 m i 45 m. En definitiva, el que diem és de nou que la variabilitat que presenta el vent, fins i tot en situacions que haurien de ser més o menys semblants *a priori*, és molt elevada.

Hom podria argumentar que les situacions definides només pel vent en un punt (a 45 m) no tenen cap raó per a ser semblants, ja que el vent depèn, per exemple, de l'estratificació tèrmica de la capa superficial, que al seu torn depèn, per exemple, de la radiació solar present. Si restringim l'anàlisi a la nit, tampoc no resollem el problema, atès que, òbviament, el vent que fa en un moment determinat depèn també del que ha passat (dinàmicament i tèrmica) en temps anteriors. Tot això no fa més que avalar, però, la idea de la gran dificultat (o complexitat) que comporta estudiar amb detall el vent a la capa fronterera.

Per acabar, hem mirat d'ajustar un dels perfils (el primer de la figura 7) als tres perfils analítics comentats anteriorment (logarítmic, potencial, d'Ekman). Com s'observa a la figura 8, tots

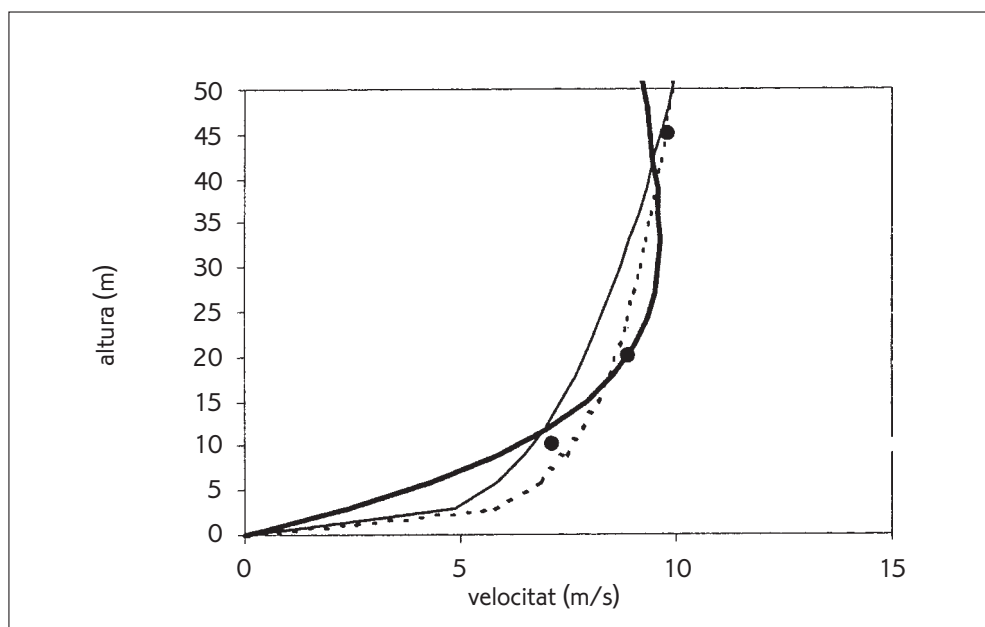


FIGURA 8. Els tres perfils teòrics comentats més amunt, amb uns valors determinats dels seus paràmetres (vegeu el text), i els punts experimentals d'un moment concret. Línia prima: perfil potencial; línia discontinua, perfil logarítmic; línia gruixuda, perfil d'Ekman.

tres es poden ajustar prou bé, però sempre mitjançant la modificació dels paràmetres. Els que s'hi mostren són els perfils potencial (amb l'exponent igual a 0,25 i el vent a 50 m igual a 10 m/s), logarítmic (amb velocitat d'escala igual a 0,5 m/s i paràmetre de rugositat igual a 0,05 m) i d'Ekman (amb vent geostrofic igual a 9 m/s i viscositat turbulenta igual a 0,01 m²/s). Aquests ajustos vénen a dir que per al coneixement del que passa a la capa fronterera la mesura del vent tan sols en tres punts és insuficient o, si més no, ha d'anar acompanyada de més informació (del terreny, del perfil tèrmic, de l'evolució temporal, etc.).

Agraïments

Vull expressar el meu agraïment a l'Institut d'Estudis Catalans per haver-me convidat a participar en aquestes Jornades Científiques i al coordinador d'aquestes Jornades, doctor Antoni M. Correig, pel seu esforç. Agraïxo a Ecotècnia, SCCL, la possibilitat de presentar dades mesurades en una de les seves torres de mesura del vent.

Bibliografia

- ANDREWS, D. G. (2000). *An introduction to atmospheric physics*. Cambridge: Cambridge University Press.
- ARYA, S. P. (1988). *Introduction to micrometeorology*. San Diego: Academic Press Inc.
- CALBÓ, J.; BALDASANO, J. M. (1995). «PROMETEO: an hydrostatic mesoscale model applied to the simulation of land-sea breeze in the Barcelona area». *Environmental Software*, núm. 10, 3, p. 139-155.
- CALBÓ, J.; COSTA, M.; CREMADES, L.; BALDASANO, J. M. (1994). «Dispersión de contaminantes atmosféricos: modelos meteorológicos». *Mundo Científico*, núm. 149, p. 724-733.
- PIELKE, R. A. (1984). *Mesoscale meteorological modeling*. Orlando: Academic Press.
- STULL, R. B. (1988). *An introduction to boundary layer meteorology*. Dordrecht: Kluwer Academic Publishers.
- THUNIS, P.; BORNSTEIN, R. (1996). «Hierarchy of mesoscale flow assumptions and equations». *Journal of Atmospheric Sciences*, núm. 53, p. 380-397.
- ZANNETTI, P. (1990). *Air pollution modeling: Theories, computational methods and available software*. Southampton: Computational Mechanics Publications.

Generació d'un temporal de llevant

Eliseu Vilaclara i Ribas

Responsable de predicció del Servei de Meteorologia de Catalunya

Resum

Catalunya es troba a la conca mediterrània. Això li confereix un clima mediterrani en general, tot i que, estant situada en una costa oriental, té alguns trets diferencials respecte de la major part d'indrets de clima mediterrani del món.

La Mediterrània és un mar gairebé tancat i es troba en una latitud prou baixa per a emmagatzemar una gran quantitat de calor durant l'època estival. A més, està protegit de la circulació general sinòptica per cadenes muntanyoses, especialment a la meitat occidental. Això afavoreix la creació de la massa d'aire mediterrània, en absència de vents sinòptics forts, que arriba a un gruix d'uns 1.500 m.

Tot això fa que en les estacions de transició (primavera i tardor), quan les ones associades amb la circulació general als nivells mitjans de la troposfera arriben sovint fins a la Mediterrània occidental, les condicions siguin molt favorables per a la creació de depressions. Aleshores, la forma de la costa afavoreix la ciclogènesi en llocs determinats, com són el golf de Lleó, el de Gènova o el de València i el mar d'Alborán.

Si les depressions se situen entre el golf de València i les illes Balears, impulsen vents forts de l'ENE sobre Catalunya que alteren molt l'estat de la mar i poden ocasionar precipitacions importants no pas pel que fa a la seva intensitat, sinó pel que fa a l'acumulació. Són els temporals de llevant.

Abstract

Catalonia is located in the Mediterranean basin. This gives it a Mediterranean climate in general, although as it is located on an eastern coast, it has some differential features with respect to the majority of the places in the world with Mediterranean climates.

The Mediterranean is a sea that is almost closed, and at a low enough latitude to store a

great amount of heat during the summer season. Also, it is protected from general synoptic circulation by mountain chains, especially on the eastern half. This favors the creation of the Mediterranean mass of air in the absence of strong synoptic winds, which reach a thickness of about 1.500 meters.

All this means that in the transitional seasons (spring and autumn), when the waves associated with general circulation at the average levels of the troposphere often reach the eastern Mediterranean, conditions are very favorable for the creation of depressions. Then, the form of the coast favors cyclogenesis in certain places, such as the Gulf of León, that of Genoa, that of Valencia, or that of the Sea of Alborán.

If the depressions are located between the Gulf of València and the Balearic Islands, they drive strong winds of the ENE towards Catalonia, which affect the state of the sea a great deal, and they can cause significant precipitation, not so much with regard to intensity as in accumulation. These are the eastern storms.

Meteorologia mediterrània

El clima mediterrani

Catalunya gaudeix de clima mediterrani. Aquesta mena de clima, caracteritzat per hiverns suaus i plujosos i estius eixuts i no gaire calorosos, no és patrimoni de la Mediterrània, mar que li dóna el nom, sinó que és comú a llocs tan distants com una part del litoral xilè, el litoral de Califòrnia, l'extrem sud del continent africà o una part del litoral australià.

En tots aquests casos, el fet de tractar-se de costes occidentals i de predomini dels ponents, comuns a les àrees temperades dels dos hemisferis terrestres, afavoreix un flux terra-mar que dóna lloc a aquest clima temperat. El descens del front polar durant l'hivern fa que les depressions i els fronts que duen associats les visitin en aquesta època de l'any i hi originin l'estació humida.

La vegetació té uns trets lògicament semblants i la producció vinícola d'aquestes àrees n'és una bona mostra.

Ara bé, tothom sap que a Catalunya l'hivern no és especialment humit, i és que, tot i ésser banyada pel mar que dóna nom al clima, n'és una excepció. I no tan sols Catalunya, sinó tota la façana mediterrània de la península Ibèrica.

Per què? Doncs senzillament perquè la costa mediterrània de la península Ibèrica és una costa oriental i no pas occidental, de manera que el ponent, humit a la resta de zones de clima mediterrani, és un règim de vents eixuts al nostre país, a la resta dels Països Catalans i a la costa est d'Andalusia. Per tant, el descens de latitud de les depressions a l'hivern no va acompanyat, en general, d'un increment de les precipitacions.

Els vents de ponent creuen més de 700 km de terra abans d'arribar a les nostres costes i puguen i baixen muntanyes, que en alguns casos superen els 2.000 m d'altura. En aquest camí

s'assequen i s'escalfen. S'assequen perquè s'esdevé precipitació en els ascensos i s'escalfen per l'efecte Föhn en els descensos, amb la calor latent aportada en el canvi de fase gas - líquid de l'aigua que contenia al principi.

A la primavera, l'escalfament progressiu del sòl contrasta amb una atmosfera encara fresca i crea tempestes a les muntanyes i a l'interior, que de vegades són empeses fins a la costa pels vents dominants a capes superiors. Aquest fet dona lloc a un màxim secundari de precipitació, que als Pirineus arriba a ser el màxim absolut entre el final de la primavera i el principi de l'estiu.

L'ascens de latitud de les altes pressions a l'estiu, amb l'assentament de l'anticicló de les Açores estenent-se en forma de falca cap a la península Ibèrica, produeix l'estabilització de l'atmosfera i l'absència de pluja a les planes. La calor no és excessiva a causa que la Mediterrània, relativament fresca al costat d'un terra càlid, crea la marinada, una circulació local de l'aire de mar a terra que refresca no únicament el litoral i el prelitoral, sinó fins i tot molts punts de la Depressió Central. La marinada mediterrània té un gruix enorme i mobilitza, de vegades, una capa d'aire de prop de 2 km per sobre de la qual s'estableix el corrent de retorn, de terra cap a mar. El terral nocturn, de terra cap a mar, completa el cicle. El conjunt dels dos vents tèrmics s'anomena *brisa*.

Durant la tardor s'acompleix el màxim de les precipitacions al litoral i al prelitoral de Catalunya. Cal buscar-ne el motiu en l'existència de la massa mediterrània, ben diferenciada de la resta.

La massa mediterrània

La mar Mediterrània en general, però especialment la conca mediterrània occidental, és envoltada de muntanyes i està protegida dels vents dominants a les capes més baixes (J. M. Jansà, 1961).

Si ens fixem en la Mediterrània occidental, hi trobem l'Atlas al sud; els sistemes Bètic, Penibètic i Ibèric a l'oest; els Pirineus i els Alps al nord, i els Apenins i la muntanyosa illa de Sicília a l'est (fig. 1). Aquestes serralades, amb altures sovint superiors als 3.000 m, protegeixen l'aire que se situa immediatament sobre el mar de la influència dels corrents dominants superiors i fan que no es renovi si no hi arriben vents sinòptics prou forts (J. M. Jansà, 1959). La Mediterrània crea una massa d'aire d'uns 1.500 m de gruix amb unes característiques pròpies d'humitat i temperatura conferides pel contacte amb el mar.

Podríem dir que la Mediterrània occidental és una gran olla que no és plena, amb una paret d'uns 1.500 m d'altura, com a mitjana, al seu voltant.

Aquesta massa d'aire té propietats diferents segons l'època de l'any i la interacció amb les masses d'aire veïnes també és, lògicament, diferent. Durant l'estiu, l'aigua es va escalfant, però de manera molt lenta, ja que la calor específica de l'aigua, molt gran, fa que necessiti molta escalfor per a augmentar lleugerament la seva temperatura. A partir de mitjan agost, la temperatura de l'aire sobre el sòl comença a baixar, gairebé amb dos mesos de retard respecte del sols-

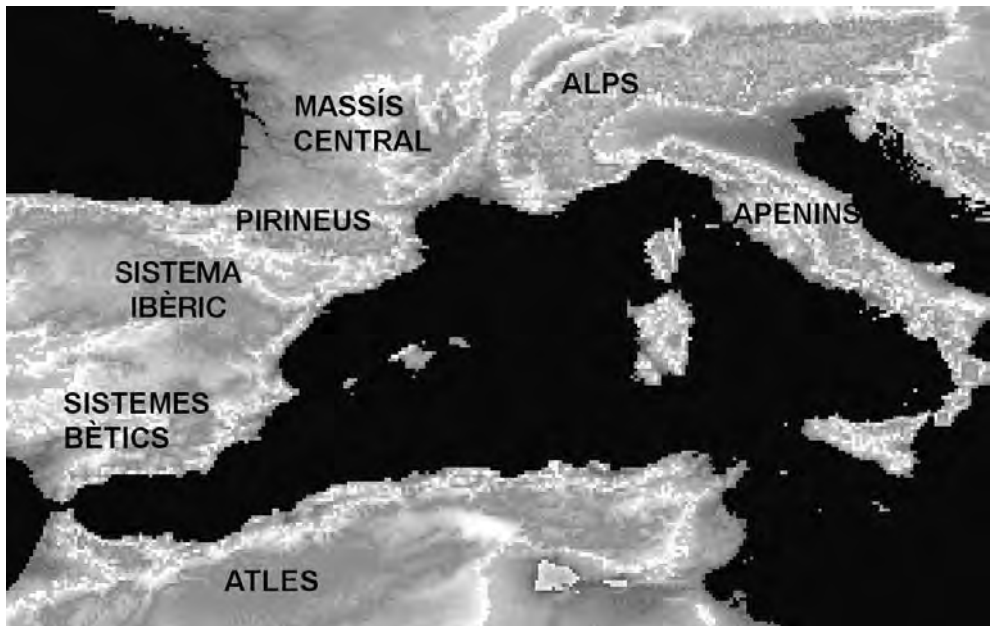


FIGURA 1

tici d'estiu, però el mar encara s'escalfa lleugerament fins al setembre. Per aquestes dates s'esdevé el primer encreuament entre els gràfics de la temperatura de l'aire i la del mar: el mar passa a ésser més calent que l'aire.

Aquesta situació es manté durant tota la tardor, amb un descens ràpid de la temperatura de l'aire sobre el sòl i un lent descens de la temperatura del mar, la qual cosa provoca una diferència cada cop més gran entre totes dues temperatures. La gran massa marítima, amb molta inèrcia tèrmica, la seva baixa conductivitat i la baixa calor específica fan que el refredament sigui molt lent.

En aquesta època és quan «l'olla mediterrània bull». S'hi crea una massa d'aire càlida i humida capaç de generar ascensos fàcilment dins les masses d'aire més fresques que s'hi comencen a superposar. La massa càlida i humida és potencialment molt més inestable que no pas una massa d'aire freda i seca, ja que la calor latent després pel vapor d'aigua quan els ascensos produeixen condensació en manté elevada la temperatura i afavoreix que l'ascens continuï fins a generar corrents convectius molt importants i, consegüentment, grans núvols de tempesta. No calen influències sinòptiques importants perquè, de tant en tant, el mar es desprengui de la calor amb temporals que consumeixen unes grans quantitats d'energia. En aquests moments la Mediterrània governa el temps tota sola. L'inici d'aquest període es pot detectar pel canvi de situació de les nuvolades: fins a mitjan agost, els cels amenaçadors que es veuen sovint a la tarda són cap a l'interior; a partir d'aleshores (alguns anys una mica abans, uns altres una mica

després) les negrors apareixen sobre el mar, especialment al matí o a la nit (quan el mar es manifesta més clarament com a font càlida en comparació de la temperatura de l'aire sobre el sòl).

A mesura que el mar es refreda, perd la capacitat de generar grans tempestes, especialment a partir del desembre, però fins a mitjan abril la relació entre la temperatura del mar i la de l'aire sobre el sòl no canvia. El mar, que s'ha refredat fins a uns tretze graus en ple hivern, s'escalfa molt lentament a la primavera, però el sòl s'escalfa ràpidament. A partir d'aquest moment la massa mediterrània passa a convertir-se en una font freda que inhibeix la convecció i es mostra molt estable. Aleshores tendeix a crear capes d'inversió tèrmica, sovint acompanyades de núvols baixos.

Altes i baixes mesoscalars

Quan el règim de vents sinòptics és feble, hem dit que la Mediterrània genera el seu propi ritme meteorològic, que sovint té poc a veure amb el que indiquen els mapes del temps tradicionals. Ara, quan apareixen fluxos generals importants, l'entorn de la conca mediterrània és també capaç d'alterar-los i de fer néixer mesoaltes i mesobaixes a sobrevent i a sotavent de les serralades principals (per exemple, Vázquez, 1995).

El dipol orogràfic

És ben conegut en la literatura meteorològica el fenomen del *dipol orogràfic*, generat a partir d'un flux perpendicular a una serralada prou alta. A sobrevent de la serralada l'aire es comprimeix i la pressió augmenta (contribució hidrodinàmica). A sotavent l'aire s'enrareix i la pressió disminueix. Si l'aire incident és prou humit, a sotavent de la serralada l'aire s'escalfa per l'efecte Föhn, disminuint la seva densitat i augmentant l'efecte de descens de pressió (contribució tèrmica). A més, l'acceleració del vent sobre la serralada per efecte d'estrangulació fa que la força de Coriolis esdevingui localment més important que la bàrica (Campins *et al.*, 1995) i hi hagi un desviament del flux cap a la seva dreta al nostre hemisferi que confereix al vent una curvatura ciclònica (contribució dinàmica).

D'aquesta manera, es creen mesoaltes i mesobaixes que poden originar salts de pressió de prop dels 5 hPa entre un vessant i l'altre d'una serralada prou alta.

Mesobaixes principals

A la Mediterrània occidental hi ha nius de mesobaixes on se'n formen més d'un centenar a l'any (Pettersen, 1956). Els principals són la costa nord-africana a sotavent de l'Atlas oriental, amb fluxos marcats del sud o del sud-oest; el golf de València, a sotavent de l'extrem sud del

Sistema Ibèric, amb fluxos de l'oest; la costa catalana, a sotavent dels Pirineus (Bessemoulin *et al.*, 1993), amb fluxos del nord; i el golf de Gènova, a sotavent dels Alps (Palmen *et al.*, 1969), també amb fluxos del nord (fig. 2).

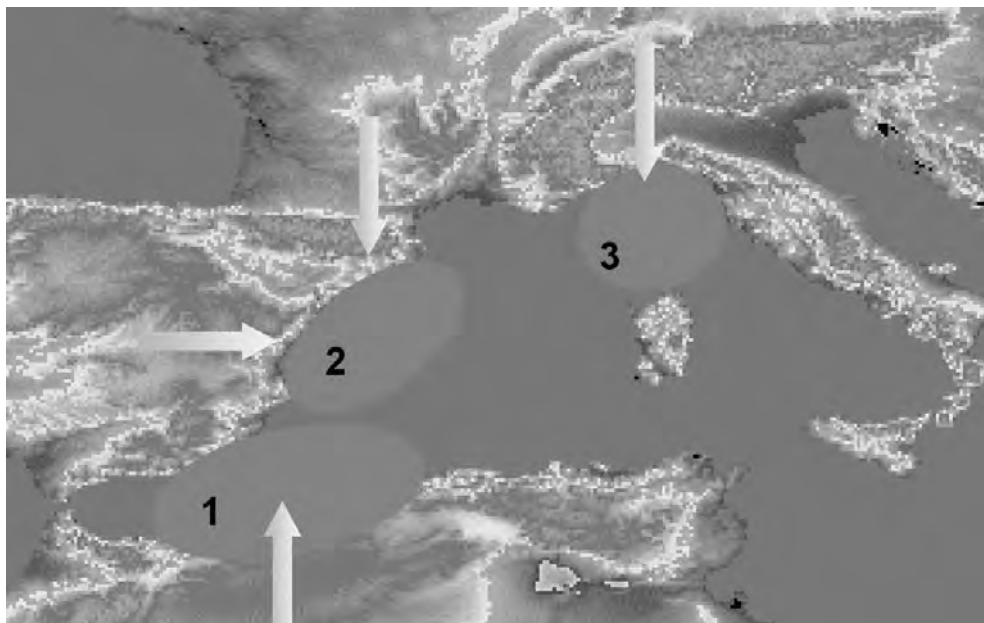


FIGURA 2

En alguns d'aquests casos (golf de València i golf de Gènova), la forma de la costa afavoreix la ciclogènesi a causa de la seva curvatura.

Convergència de fluxos a la Mediterrània occidental

La facilitat de la formació de mesoaltes i mesobaixes en uns punts determinats i els passos entre muntanyes que canalitzen el vent a la Mediterrània occidental produeixen l'efecte d'una convergència de fluxos.

Si ens fixem en les proximitats de Catalunya, existeixen dos passos principals entre muntanyes que canalitzen en direccions fixes els vents sinòptics: el pas entre els Alps i els Pirineus, i la vall de l'Ebre. Més lluny trobem l'estret de Gibraltar, capaç també de fixar en una sola direcció (i dos sentits) vents d'origen sinòptic molt diferent, i el canal de Sicília.

L'efecte que fa l'orografia és similar al que fa un embut amb raigs d'aigua que incideixen en diferents punts de la seva boca. L'embut els acanala tots i els fa sortir sempre pel mateix lloc, de manera que diferents fluxos entrants acaben en un únic flux sortint (fig. 3). Això és el que

d'alguna manera fan les muntanyes al nostre entorn: diverses situacions sinòptiques lleugerament diferents acaben donant un mateix patró mesoscalar.

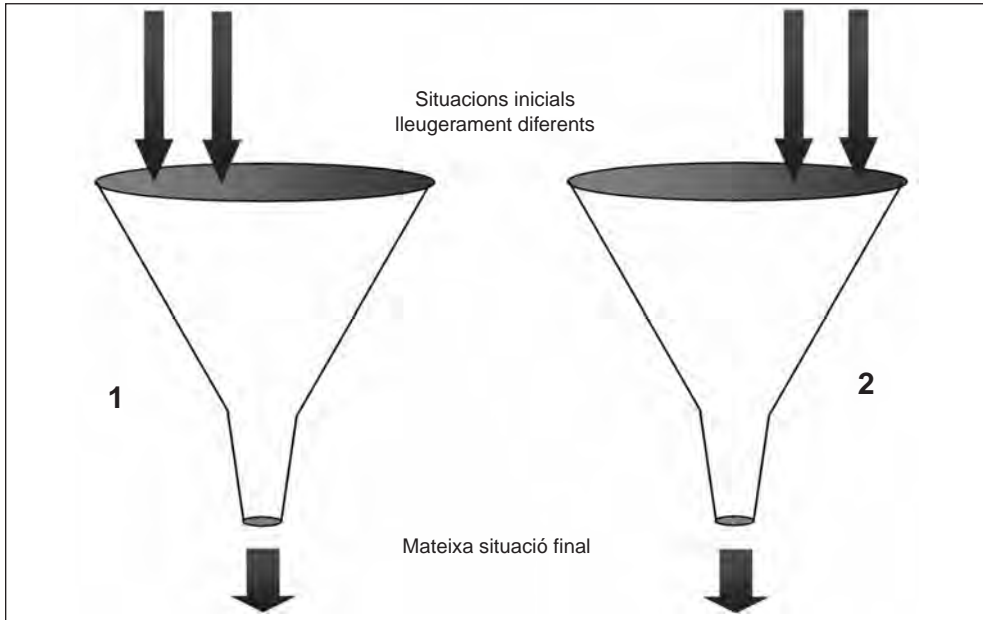


FIGURA 3

Un exemple clar d'aquest «efecte embut» són les situacions del nord. Vents de component nord amb angles d'incidència diferents acaben donant sempre tramuntana a l'Empordà, mestral a la vall de l'Ebre (A. Jansà, 1987), llebeig a part de la Depressió Central i xaloc o migjorn a la costa central (Fontserè, 1950). La tramuntana esdevé molt forta a causa de l'estrenyiment entre els Alps i els Pirineus i l'augment de la força bàrica que produeix el dipol orogràfic. El dipol crea corrents de retorn amb vents contraris al general a la costa central, mentre que les muntanyes prelitorals meridionals desvien una part del corrent de mestral de la vall de l'Ebre com a vent de l'oest o del sud-oest.

Aquesta convergència de situacions és un avantatge per al predictor, que coneix els principals patrons mesoscalars, cadascun dels quals serà originat per una família de situacions sinòptiques.

Els fronts superficials

Un altre dels fenòmens que es manifesten sovint a la Mediterrània és l'aparició de fronts superficials. El fet que la massa mediterrània no és desplaçada fàcilment de les zones protegides

per muntanyes, però que sí que pot ser empesa pel vent sinòptic en els passos naturals entre aquelles, crea fronts superficials en posicions diferents dels fronts sinòptics que ens marquen els mapes del temps (que són certs per sobre de la superfície de 850 hPa però no pas per sota). En una situació del nord com la que comentàvem abans, es crea un front superficial entre la massa freda que entra per la Costa Blava francesa i la massa marítima que es manté enfront de la costa catalana. Aquest front, sovint acompanyat de núvols baixos, es detecta com una traça que, partint de l'extrem oriental dels Pirineus, es pot perllongar fins a l'illa de Sardenya (Vilaclara, 1999) (fig. 4).

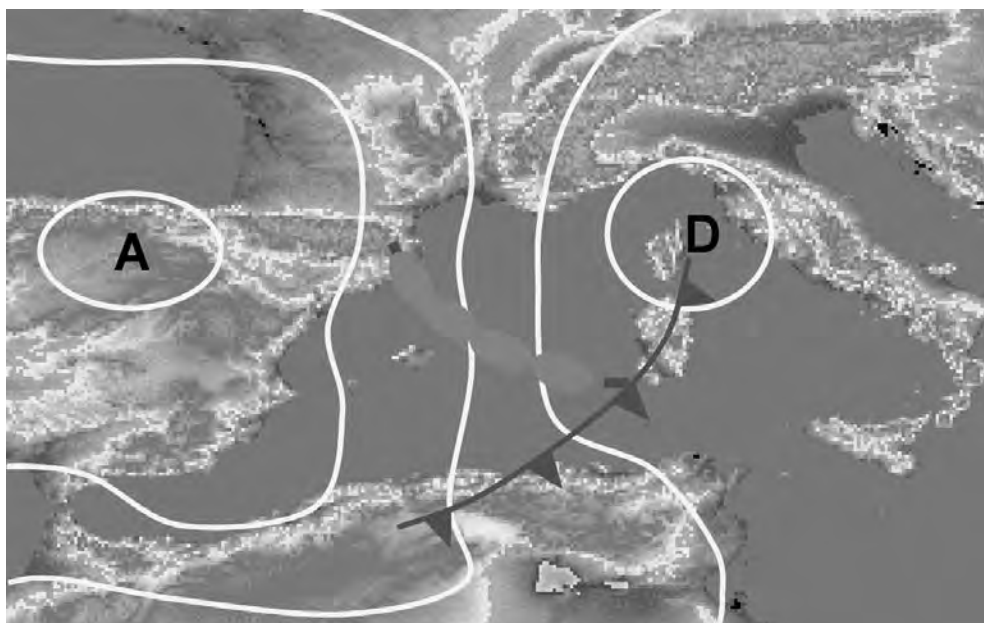


FIGURA 4

Quan una depressió se situa davant del delta de l'Ebre, un front superficial menor apareix a la Catalunya central. La posició de la depressió produeix vent del nord a l'Empordà, xuclat del continent per l'únic pas obert. Genera llevant a la costa central, empès sinòpticament per la depressió, i mestral a la vall de l'Ebre, que es dirigeix cap al centre de la baixa perpendicularment a les isòbares. Aquest vent de mestral es troba amb la Serralada Prelitoral meridional, d'una manera anàloga a la situació del nord descrita al paràgraf anterior, i crea un ramal secundari de vent aponenat (contrari a les isòbares) que es troba amb el vent de l'est que entra per la zona central del litoral a l'àrea de la Catalunya central i intensifica allà les precipitacions (Vilaclara *et al.*, 2000).

Generació d'un temporal de llevant

La gènesi d'un temporal de llevant pot tenir unes quantes causes. Hi ha, però, un marc general sinòptic que gairebé sempre es compleix, i és que parteix d'una situació de bloqueig amb un anticicló fort al nord d'Europa. Aquest bloqueig pot donar lloc a tres situacions que desembocuin en la formació d'una depressió mediterrània:

a) L'entrada d'una depressió atlàntica per l'estret de Gibraltar a causa del bloqueig.

b) La progressió en altura d'est a oest d'un solc pel marge sud de l'anticicló que provoqui l'aparició d'una depressió a la superfície.

c) La retirada de l'anticicló cap a l'Atlàntic que deixi pas a una ona que progressi de nord a sud pel seu marge oriental i que formi una depressió en arribar a la Mediterrània, moment en el qual l'anticicló pot tornar a ocupar la seva posició inicial al nord d'Europa.

En tots tres casos, la situació final mostra una depressió ben formada a la Mediterrània, amb pressions al voltant dels 1.010 hPa, i un anticicló potent, sovint amb pressions de més de 1.040 hPa, al nord del continent.

Si la depressió «s'encaixa» al golf de València i situa el seu centre al voltant de les illes Balears, a Catalunya apareix un patró mesoscalar típic que és conegut com a *temporal de llevant*. Malgrat aquest nom, el vent a la superfície no acostuma a ser de l'est, sinó més aviat de l'ENE (Fontserè, 1929) o, fins i tot, del NE. Aquesta situació va acompanyada de vents forts a la cos-

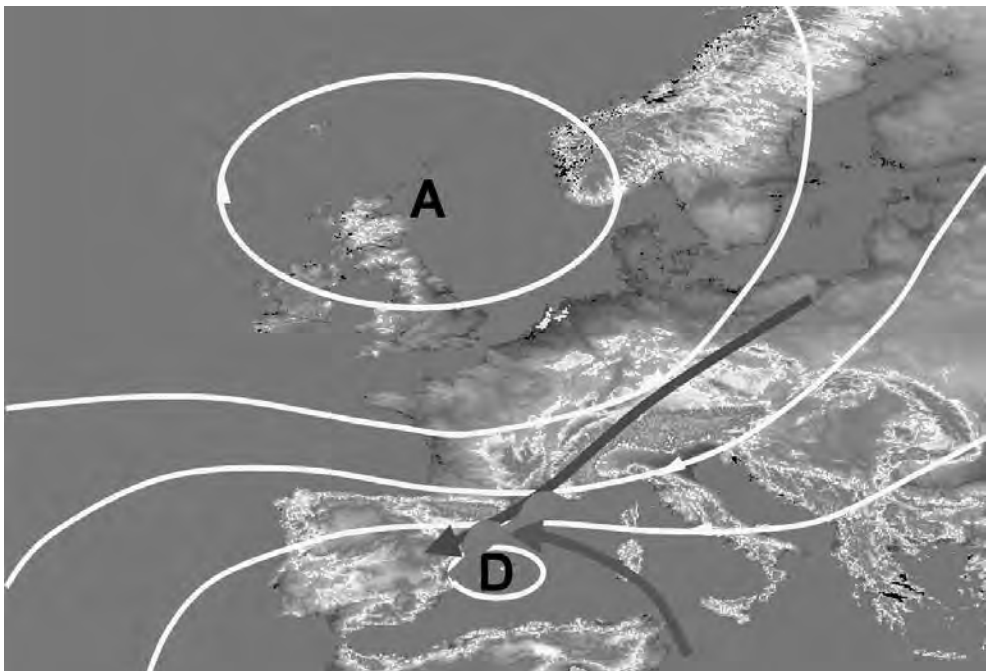


FIGURA 5

ta, de notable alteració marítima i de precipitacions abundants, si bé en general d'intensitat feble o moderada i no pas forta. La seva durada pot allargar-se fins a uns tres dies, en els quals es poden acumular fàcilment quantitats de més de 200 mm d'aigua i sovint a punts de la Garrotxa, del Ripollès, d'Osona, de la Selva o del Vallès Oriental (al vessant est del Montseny, de la Serralada Transversal i dels Pirineus orientals).

El model conceptual corresponent a aquesta situació és marcat per una entrada freda a la superfície, acompanyada de vents més aviat eixuts del NE, que s'escola pel pas entre els Alps i els Pirineus, si bé també creua per sobre de l'extrem oriental dels Pirineus. Com més al sud, el vent és més alleu, ja que la posició de la depressió l'empeny en aquesta direcció. Al marge oriental de la baixa pressió, la massa mediterrània, en forma de vent del sud-est, avança cap a Catalunya, gira a llevant prop del golf de Lleó i es troba amb el ramal fred de vent de gregal. Aleshores li passa per sobre, pujant com per una rampa en un típic efecte de front càlid, i s'esdevenen la condensació i la formació de núvols estratiformes que generen precipitacions contínues (fig. 5). Si la massa mediterrània és inestable, s'hi poden intercalar núvols de tempesta. La pluja o la neu pot afectar gairebé tot Catalunya, si bé la més abundant acostuma a enregistrar-se a la meitat nord del prelitoral i als Pirineus orientals. Pot ser que la precipitació no arribi a l'extrem nord-oest, especialment a la Vall d'Aran, l'Alta Ribagorça i els dos Pallars.

Els llevants secs

A vegades, la configuració sinòptica afavoreix llevant sense precipitació apreciable. De fet, llevant no és sinònim de precipitació, ja que és necessari l'ascens d'una massa d'aire prou humida, de vegades sobre les muntanyes, però normalment per xoc amb una massa d'aire diferent.

Quan en altura hi ha un solc pronunciat sobre l'Atlàntic que arriba fins a les proximitats de les illes Canàries, s'estableix un flux de vent del sud sobre la Mediterrània que, en creuar l'Atlas, crea una baixa a la superfície sobre la costa nord-africana (Ramis i Jansà, 1983). Aquesta, reforçada pel solc a la seva esquerra, genera llevant a la Mediterrània occidental, però va acompanyada d'una massa molt càlida al nivell de 850 hPa que inhibeix la formació de núvols des de sota. Aleshores el cel és tapat per núvols mitjans associats amb el solc atlàntic i per núvols molt baixos associats amb el llevant superficial, però no es formen núvols de prou gruix per a precipitar-se d'una manera important. La situació, típica de la primavera o del principi de l'estiu, amb la massa mediterrània actuant com a font freda, acaba en pluja feble de fang i notable temporal marítim.

El temporal del 15 de novembre de 2001

Els vuit dies transcorreguts entre el 10 i el 17 de novembre de 2001 seran recordats molt temps, especialment pels meteoròlegs. S'hi van succeir dos temporals de llevant molt violents:

el primer, amb el seu màxim el dia 11, i el segon, el dia 15. El dia 17, les restes de la depressió en altura i la relativa calma en superfície van afavorir la creació de tempestes a la Costa Daurada que van ocasionar alguns estralls, amb precipitacions molt intenses i abundants a poblacions com el Vendrell, Coma-ruga o Calafell.

Els temporals de l'11 i del 15 van ser deguts a depressions profundes a la superfície creades a partir de dues ones a la troposfera mitjana que es van desplaçar de nord a sud i es van escolar pel marge oriental d'un anticicló atlàntic.

El temporal del dia 11, amb vent molt fort del NE i del NNE, va ocasionar poca pluja, però un temporal de mar excepcional, amb un registre d'onada màxima de 9,92 m al cap de Tortosa que no s'havia mesurat mai en els dotze anys d'observacions en aquell punt, dutes a terme pel Laboratori d'Enginyeria Marítima de la Universitat Politècnica de Catalunya.

El dia 15, després de la formació de la segona baixa a la Mediterrània, l'anticicló es va col·locar al nord d'Europa i es va configurar una situació més típica de temporal de llevant.

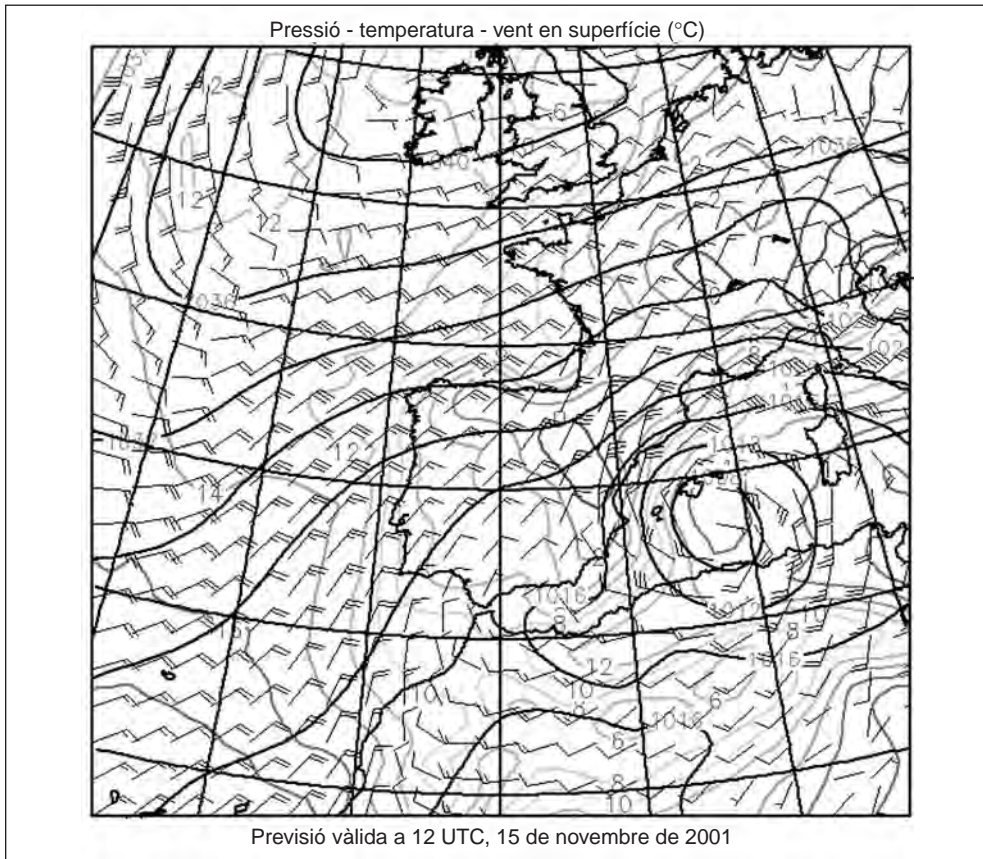


FIGURA 6

A la figura 6 s'aprecia aquesta configuració, amb una entrada de vent del NE fort a la superfície sobre el sòl però més alleuavat sobre el mar. La depressió mediterrània està centrada una mica al SE de les illes Balears, amb una pressió lleugerament inferior a 1.008 hPa.

A la figura 7, corresponent al radiosondatge de Barcelona del dia 15 a les 00 TU (el de les 12 TU no es va poder fer a causa del fort temporal de vent), s'aprecia la saturació gairebé a tota la columna atmosfèrica entre 400 i 850 hPa, però no a la superfície, on l'entrada freda continuava. Entre 850 i 750 hPa s'hi observa una entrada càlida, corresponent a l'advecció mediterrània que va passant per damunt de la massa més seca i fresca de superfície. L'altitud de la base dels núvols és responsable de la relativament escassa precipitació que es va recollir a la costa central (entre 10 i 20 mm).

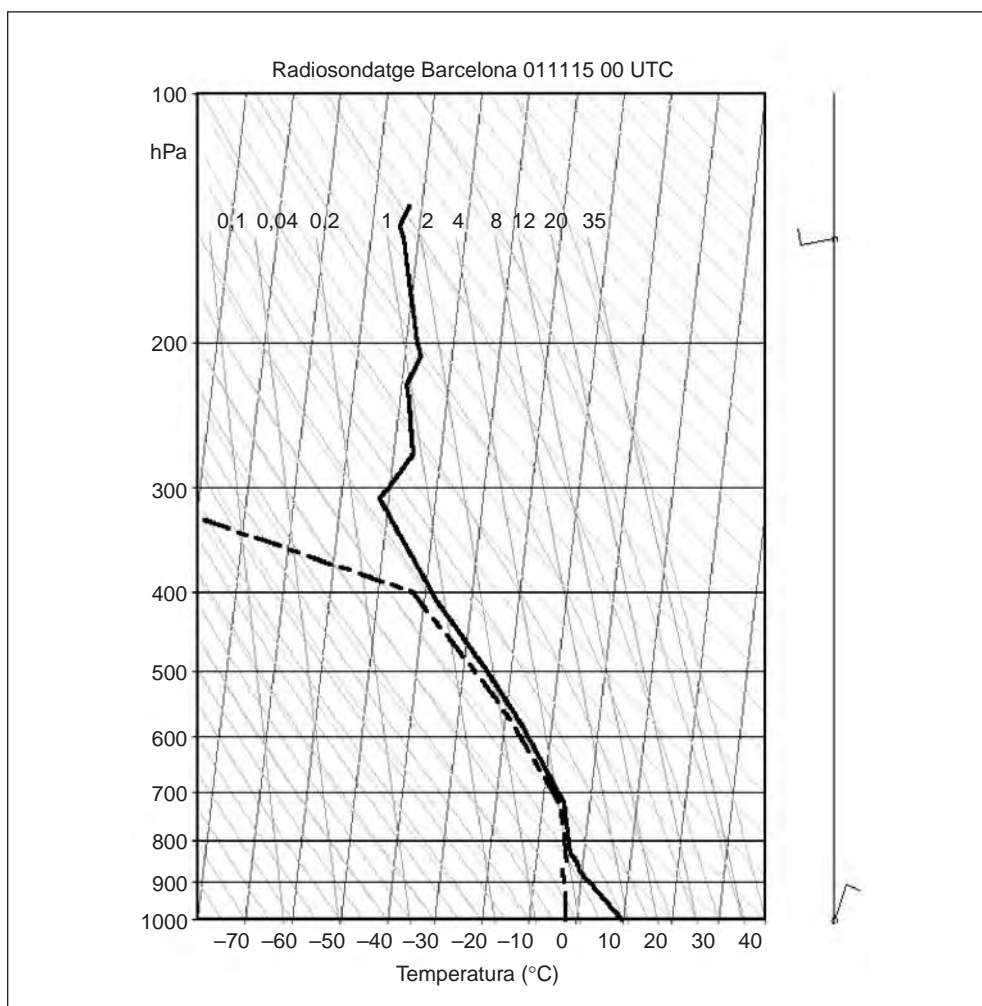


FIGURA 7

La figura 8 correspon a la sortida del model meteorològic MASS (Mesoscale Atmospheric Simulation System) per a les 12 TU del 15 de novembre de la simulació de 15 km de resolució. Mostra el vent a la superfície i l'advecció de temperatura, també a la superfície.

Aquest mapa és enormement il·lustratiu del front que es crea davant la costa catalana amb una advecció càlida a la superfície. Sobre el mar, el vent és fort de l'ENE, però sobre el sòl és del NNE i s'allevanta cap al sud. A la sortida de 850 hPa, en canvi, el vent és de lleuant fort a tot arreu, amb velocitats de més de 50 Kt sobre el sòl, la qual cosa demostra que la massa càlida que hi ha damunt el mar a la superfície és la mateixa que trobem sobre el sòl a 850 hPa i que s'ha elevat per sobre de la massa freda de superfície en un efecte de front càlid.

Al mapa de gran resolució de la figura 8 es detecta el centre de la depressió prop del golf de València i no gaire lluny del delta de l'Ebre, fet que crea un corrent de mestral per la vall de l'Ebre i un front superficial cap al sud de Catalunya, responsable de màxims secundaris de precipitació entre 20 i 30 mm en aquella zona.

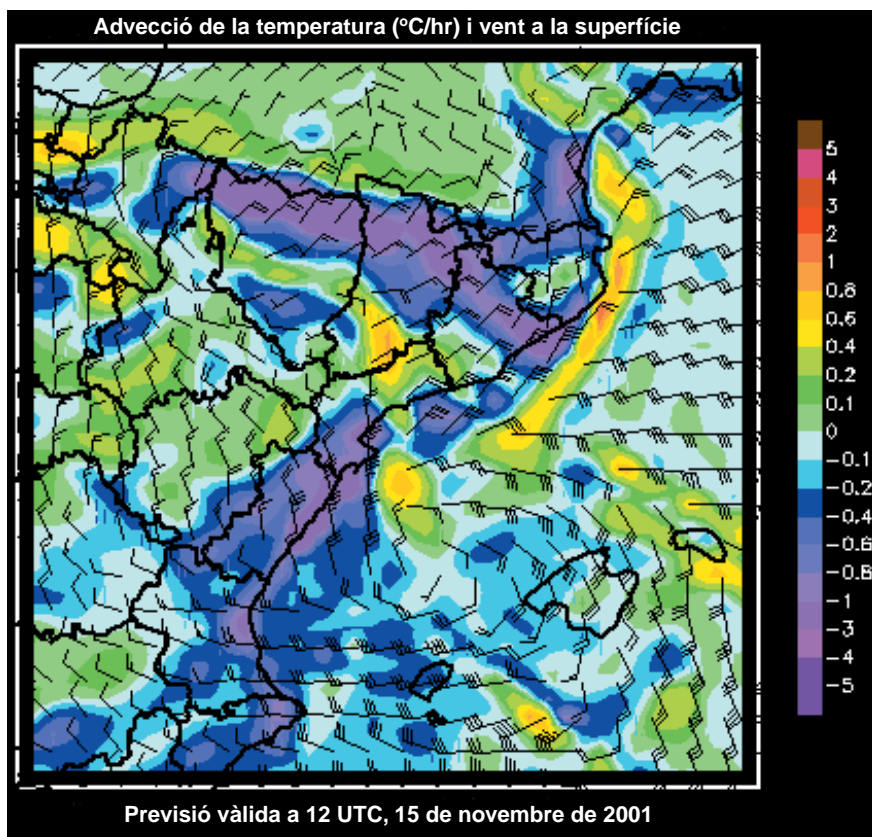


FIGURA 8

Els màxims de precipitació recollits en aquest episodi van ser 94 mm a Roses i 85 mm a Portbou, a la zona més activa de contacte entre les dues masses d'aire. El vent, però, va tornar a ser el fenomen més destacat, amb ratxes de fins a 140 km/h a l'Observatori Fabra de Barcelona o de 119 km/h a Vilassar de Mar. L'onatge, en canvi, no va superar les marques de quatre dies abans.

Com a conclusió, cal dir que la influència de l'orografia, fins i tot de les serralades menors, és determinant en la localització de les àrees de major precipitació en un temporal de llevant no pas perquè l'aire ha de remuntar les muntanyes, sinó pel fet que crea passadissos de vent de direccions diferents que acaben enfrontant-se i creant diversos fronts superficials. La massa mediterrània exerceix, evidentment, el paper de font de calor i humitat necessària per a la producció de la precipitació.

La meteorologia mediterrània i els terratrèmols

Si volem buscar una analogia entre temporals mediterranis i terratrèmols, els fenòmens naturals més violents als quals pot ésser sotmès el nostre entorn geogràfic, la cosa més raonable és comparar les plaques tectòniques i les masses d'aire.

Un terratrèmol es genera a partir de les tensions entre plaques tectòniques veïnes que tenen moviments diferents. Un temporal a la Mediterrània (no pas obligatòriament un temporal de llevant) es genera a partir de l'energia acumulada a la massa mediterrània, que ha anat acumulant calor durant mesos, quan aquesta entra en contacte amb masses veïnes d'humitat i temperatura diferents. Aquest xoc pot provocar núvols convectius, capaços d'alliberar, en poca estona, grans quantitats d'energia, de la mateixa manera que l'energia potencial elàstica de la Terra es pot alliberar de cop en la fractura que dona origen a un terratrèmol.

Bibliografia

- BESSEMOULIN, P.; BOUGEALT, P.; GENOVÈS, A.; JANSÀ, A.; PUECH, D. (1993). «Mountain pressure drag during PYREX». *Beitr. Phys. Atmospheric*, núm. 66, p. 305-325.
- CAMPINS, J.; JANSÀ, A.; BENECH, B.; KOFFI, E.; BESSEMOULIN, P. (1995). «PYREX observation and model diagnosis of the tramuntane wind». *Meteorol. Atmospheric Physics*, núm. 56, p. 209-228.
- FONTSERÈ, E. (1929). «Die "Llevants" der Katalanischen Küste». *Beitr. Phys. Atmospheric*, núm. 15, p. 55-60.
- (1950). *La tramuntana empordanesa i el mestral del golf de Sant Jordi*. Ginebra: Institució Patxot.
- JANSÀ, A. (1987). «Distribution of the mistral: a satellite observation». *Meteorol. Atmospheric Physics*, núm. 36, p. 201-214.

- JANSÀ, J. M. (1959). «La masa de aire mediterráneo», *Rev. Geofísica*, núm. 18, p. 35-50.
- (1961). *Nociones de climatología general y de Menorca*. Maó: Sintés.
- PALMEN, E.; NEWTON, C. W. (1969). *Atmospheric circulation systems: Their structure and physical interpretation*. Nova York; Londres: Academic Press.
- PETTERSEN, S. (1956). *Weather analysis and forecasting*. Nova York: McGraw-Hill.
- RAMIS, C.; JANSÀ, A. (1983). «Condiciones meteorológicas simultáneas a la aparición de oscilaciones del nivel del mar de amplitud extraordinaria en el Mediterráneo occidental». *Rev. Geofísica*, núm. 39, p. 35-42.
- VÁZQUEZ, L. A. (1995). «Tramuntana i mestral a Catalunya». *Butlletí de la Societat Catalana de Ciències*, vol. XIV/1, p. 69-75.
- VILA CLARA, E. (1999). «Situacions del nord sobre Catalunya: "l'estela pirinenca"». A: *V Jornades de Meteorologia Eduard Fontserè*. Cornellà de Llobregat: Associació Catalana de Meteorologia.
- VILA CLARA, E.; SAIROUNI, A.; BECH, J. (2000). «Contribució del model meteorològic MASS a l'estudi de les causes de l'aiguat del 10 de juny de 2000». A: *VI Jornades de Meteorologia Eduard Fontserè*. Cornellà de Llobregat: Associació Catalana de Meteorologia.

Resum

Els sistemes aquàtics naturals són sistemes oberts que intercanvien energia i matèria amb l'exterior i que presenten dinàmiques no lineals i estructures complexes, des de l'escala dels mil·límetres fins a centenars de quilòmetres. Aquí, però, només tractarem de processos complexos a petita escala, i més concretament de la turbulència i la microestructura. Tanmateix, la turbulència, ubiqua en els sistemes naturals, té un paper crucial en els processos de transferència local i determina, en últim terme, els balanços —i, per tant, la dinàmica— a gran escala. Dels fluxos turbulents es diu que són no lineals, caòtics, tridimensionals, altament difusius, però essencialment són vorticals; si un flux presenta estructures complexes i sembla turbulent però és irrotacional, no és turbulent.

Dividirem aquesta comunicació en dues parts. En la primera parlarem de la vorticitat i hi veurem les escales turbulentes més importants, les propietats espectrals de la turbulència —especialment pel que fa a la transferència d'energia— i el paper que exerceixen l'estratificació i la rotació en la seva dinàmica. També hi deduirem els paràmetres de barreja que es poden inferir a partir de mesures directes de microestructura.

Plantegem la segona part des d'un punt de vista més aplicat, començant per presentar els instruments de mesura i les diverses metodologies per a abordar el càlcul de les variables turbulentes a partir de les dades obtingudes. Fet això, es presentaran els resultats de diferents campanyes a l'estany de Banyoles, al transecte 53 N de l'Atlàntic Nord i a l'embassament de Boadella, centrant-nos principalment en l'estudi de la dinàmica convectiva d'una surgència subterrània, en el primer cas, i en la transferència d'energia i la barreja a la capa superficial, en els altres dos casos.

Abstract

Natural aquatic systems are open systems that exchange energy and matter with the outside and that present nonlinear dynamics and complex structures, on a scale ranging from millime-

ters to hundreds of kilometers. Here, however, we will only handle small-scale complex processes, more specifically turbulence and microstructure. Turbulence, ubiquitous in natural systems, plays a crucial role in the processes of local transfer and determines, in the end, balance —and therefore dynamics— on a large scale. It is said of turbulent flows that they are nonlinear, chaotic, three-dimensional, highly diffusive, but essentially vortical; if a flow presents complex structures and seems turbulent, but is irrotational, it is not turbulent.

This talk will be divided in two parts. In the first part we will speak of vorticity, and we will see the most important turbulent scales, the spectral properties of turbulence, especially regarding the transfer of energy, and the role played in its dynamic by stratification and rotation. We will also deduce the parameters of mixture that can be inferred from direct measurements of microstructure.

The second part is oriented from a more applied point of view, beginning by presenting the measurement instruments and various methodologies for approaching the calculation of turbulent variables from the information obtained. This done, the results of different campaigns in the Banyoles lake, on the crossing of 53 N of the North Atlantic and the reservoir of Boadella will be presented, focusing mainly on the study of the convective dynamic of a subterranean uprising in the first case, and in the transfer of energy and mixture in the surface layer in the other two cases.

Introducció

La turbulència és impredecible perquè és caòtica, no pas perquè sigui aleatòria. De fet, en el principi de la teoria del caos o de la complexitat es troba el problema de l'atracció gravitatòria de tres cossos, i aquest és un problema que sabem plantejar perfectament de manera determinista. Es tracta d'un sistema discret governat per equacions diferencials no lineals molt sensibles a les equacions inicials. Això mateix passa en el cas d'un flux turbulent. Les equacions que descriuen el moviment d'un fluid són la de continuïtat, les de conservació de la quantitat de moviment i de les propietats escalars que es considerin (temperatura del fluid, salinitat...) i la d'estat corresponent. Si tinguéssim un coneixement exacte de les condicions inicials del sistema i poguéssim resoldre aquestes equacions a escales temporals i espacials prou petites, llavors la turbulència seria determinista. De fet, aquests càlculs són possibles per a fluxos poc turbulents (nombre de Reynolds < 2000) i formen la disciplina de la dinàmica de fluids computacional. Els fluids geofísics, però, presenten valors del nombre de Reynolds molt elevats: de prop de 10^6 a la capa superficial de barreja a l'oceà, i s'hi poden donar encara valors molt més alts. Per tant, calen unes altres aproximacions per a estudiar-los, com ara les aproximacions estadístiques que presentarem, les quals no solament han permès avançar en la comprensió de la turbulència, sinó que han proporcionat la base teòrica que permet inferir les magnituds reals de la barreja, en sistemes naturals, a partir de mesures *in situ*. Abans, però, presentarem conceptual-

ment les equacions que governen el comportament d'un fluid en general i, a partir d'aquestes, comentarem algunes característiques de la turbulència.

1. Conceptes bàsics de física de fluids i sistemes aquàtics naturals

1.1. Equació constitutiva d'un fluid

Sobre una partícula imaginària i contínua de fluid hi ha forces de contacte, que actuen sobre la seva superfície, i forces a distància, com ara la gravetat, que actua sobre la seva massa. És fàcil d'entendre que les forces de contacte estiguin relacionades amb la deformació de la partícula. Si, per exemple, pressionem perpendicularment la superfície de la partícula, la comprimirem i, per tant, la deformarem linealment; si hi actuem tangencialment (hi apliquem una tensió), la deformarem angularment.

Un cas simple d'aplicació d'una tensió sobre un fluid és el flux de Poiseuille, que es genera, per exemple, desplaçant una superfície plana per sobre d'una capa fina de fluid. Aquesta superfície arrossega les partícules de fluid que hi estan en contacte, alhora que aquestes n'arrosseguen d'altres de més allunyades, però de manera que l'arrossegament decreix proporcionalment a la distància de la partícula fluida a la placa. Per tant, dins de la capa de fluid es crea un gradient de velocitats, dv/dz , que també s'anomena *cisallament*. Aquest gradient ens dona, de fet, una mesura de la deformació que experimenten les partícules fluides, ja que la deformació és donada directament per la diferència de velocitats entre la part superior de la partícula (més gran, en el flux de Poiseuille) i la inferior (més petita). Més concretament, la velocitat de deformació angular de la partícula és igual al gradient de velocitats que s'hi estableix. La constant de proporcionalitat entre la tensió, τ , i la velocitat de deformació, dv/dz , és la viscositat dinàmica, μ , que depèn de cada fluid. Per tant, en el cas del flux de Poiseuille podem escriure que $\tau = \mu dv/dz$. La idea intuïtiva que podem associar amb el concepte de viscositat és si el fluid és més o menys espès. Si el fluid és més espès, llavors el transport de la quantitat de moviment en la direcció perpendicular a aquella en què actua la tensió és més efectiu.

En un cas general, cada component, j , de les tensions aplicades sobre les cares imaginàries d'una partícula fluida perpendiculars a les tres direccions de l'espai i, τ_{ij} , es relaciona linealment amb les nou possibles components de la deformació de la partícula fluida sobre la qual actuen, e_{ij} . Aquesta relació és el que s'anomena *equació constitutiva del fluid*. Per als sistemes aquàtics, una bona aproximació de l'equació constitutiva és la que defineix els fluids newtonians i que s'escriu

$$\tau_{ij} = -p\delta_{ij} + 2\mu e_{ij} - (2/3)\mu \nabla \cdot \mathbf{v} \delta_{ij},$$

on p és la pressió; $\nabla \cdot \mathbf{v}$, la divergència de la velocitat, que, en cas que negligim la compressibilitat del fluid, és zero (el que hem destacat en negreta ens n'indica el caràcter vectorial), i e_{ij} té la forma $e_{ij} = (\partial u_i / \partial x_j + \partial u_j / \partial x_i) / 2$.

1.2. L'equació de Navier-Stokes per a sistemes aquàtics naturals

L'equació de Navier-Stokes és l'equació de conservació de la quantitat de moviment per a fluids newtonians que es dedueix a partir de la generalització de la llei de Newton per a una partícula fluida (equació de Cauchy). La suma de les variacions de les tensions que actuen en la direcció j , quan ens movem seguint els tres eixos de referència, $d\tau_{ij}/dx_i$, és precisament la component de la força en la direcció j . Per tant, l'equació de Newton generalitzada es pot escriure com a $\rho a_j = \rho f_j + d\tau_{ij}/dx_i$, on ρ és la densitat de la partícula; a_j , la seva acceleració en la direcció j , i f_j , les forces a distància que actuen sobre la partícula. Si considerem conjuntament les equacions de Cauchy i la constitutiva per al cas de fluids newtonians, tenim l'equació de Navier-Stokes. Considerant que $e_{ij} = (\partial u_i/\partial x_j + \partial u_j/\partial x_i)/2$, llavors $a_j = \rho f_j - \partial p/\partial x_j + \mu \nabla^2 u_j + \partial (\mu \nabla \cdot \mathbf{v}/3)/\partial x_j$. Si, a més, considerem que l'única força a distància que actua sobre la partícula és la gravitatòria ($f_j = \mathbf{g}$) i que el flux és incompressible ($\nabla \cdot \mathbf{v} = 0$), llavors l'equació de Navier-Stokes en forma vectorial s'escriu $D\mathbf{v}/Dt = \mathbf{g} - \nabla p/\rho + (\mu/\rho) \nabla^2 \mathbf{v}$.

Fins aquí hem parlat sempre d'una partícula de fluid imaginària sobre la qual actuen les forces, i, per tant, l'equació anterior descriu la dinàmica d'aquesta partícula. Tanmateix, en sistemes naturals, les mesures habitualment no es fan pas marcant les partícules del flux i seguint-les (representació lagrangiana), sinó que les fem en un punt fix de l'espai per on flueix el fluid que estudiem (representació euleriana). L'una representació i l'altra es relacionen, ja que la variació d'una propietat (b) d'una partícula determinada depèn de la velocitat en què es mou dintre d'un flux, que pot ser no homogeni ($\mathbf{v} \cdot \nabla b$), i de la contribució no estacionària local ($\partial b/\partial t$). Segons això, tenim

$$\partial \mathbf{v}/\partial t + \mathbf{v} \cdot \nabla \mathbf{v} = \mathbf{g} - \nabla p/\rho + (\mu/\rho) \nabla^2 \mathbf{v},$$

que és l'equació de Navier-Stokes en la representació euleriana.

En el cas dels geofluids, però, cal considerar el fet que els eixos de referència en què fem les mesures no són inercials, ja que estan fixats a la Terra, que gira a una velocitat angular Ω . Per tant, en l'equació de la conservació de la quantitat de moviment haurem d'introduir l'anomenada *força de Coriolis*, que, com se sap, no és res més que una força fictícia que ens permet considerar el nostre sistema com a inercial. Així, doncs, tenim

$$\partial \mathbf{v}/\partial t + \mathbf{v} \cdot \nabla \mathbf{v} = -2\Omega \times \mathbf{v} + \mathbf{g} - \nabla p/\rho + (\mu/\rho) \nabla^2 \mathbf{v}.$$

Cal tenir present, però, que les forces fictícies no fan cap treball sobre el sistema i, per tant, no apareixeran a les equacions de l'energia.

Les dificultats implícites en la resolució de les equacions fa que sovint no es treballi amb l'equació anterior, sinó que es procura introduir-hi tantes simplificacions com sigui possible mentre es descriu el sistema raonablement bé. Una de les simplificacions més habituals amb què

es treballa és l'*aproximació de Boussinesq*, que es basa en el fet que les variacions de densitat a l'oceà normalment són molt petites, menys del 3 %. Aquestes diferències de densitat, tanmateix, són importants pel que fa a les forces de flotabilitat (gravitatòries) que actuen sobre les partícules fluides quan es mouen en un fluid estratificat. En canvi, pel que fa a la massa inercial, són prou petites per a poder-les negligir. Així, doncs, l'aproximació de Boussinesq permet substituir la densitat, ρ , en tots els termes de l'equació de Navier-Stokes per un valor de referència ρ_0 excepte en el terme de la gravetat. És a dir, considerant aquesta aproximació, tenim que

$$d\mathbf{v} / dt + \mathbf{v} \cdot \nabla \mathbf{v} = -2 \boldsymbol{\Omega} \times \mathbf{v} + (\rho' / \rho_0) \mathbf{g} - \nabla p / \rho_0 + (\mu / \rho_0) \nabla^2 \mathbf{v} .$$

El quocient (μ / ρ_0) s'anomena *viscositat cinemàtica* i s'indica amb la lletra ν . Tal com veiem a l'equació interior, és ν i no μ el que determina la dinàmica del fluid. Ambdues, però, actuen a escales petites i s'anomenen *viscositats moleculars*. A partir d'ara, per a referir-nos a una densitat de referència, no utilitzarem la notació ρ_0 , sinó que escriurem simplement ρ .

2. Característiques de la turbulència

2.1. El nombre de Reynolds i la viscositat turbulenta

A l'equació de Navier-Stokes el terme d'inèrcia, $\mathbf{v} \cdot \nabla \mathbf{v}$, és no lineal i, per tant, és el responsable del comportament caòtic (turbulent) del sistema. Per això, precisament el quocient entre aquest terme i el de les forces viscoses, les quals estableixen el sistema, és el que utilitzem per a avaluar si un sistema és turbulent o no, i és el que anomenem *nombre de Reynolds*, Re . Si suposem que L i V són la longitud i la velocitat característiques del flux mitjà, llavors podem escalar el terme $\mathbf{v} \cdot \nabla \mathbf{v}$ amb V^2/L i $(\mu/\rho)\nabla^2\mathbf{v}$ amb $\nu V/L^2$, de manera que $Re = VL/\nu$. Conseqüentment, com més gran sigui el quocient VL/ν , més importància tindran les forces d'inèrcia sobre les viscoses i més turbulent serà el flux.

Com que el nombre de Reynolds és un nombre adimensional, les unitats de ν i VL són les mateixes i, per tant, podríem interpretar ambdós termes de manera anàloga. Si abans hem comentat que ν era la viscositat cinemàtica i que actua a escales petites, ara podríem considerar el producte VL com una viscositat cinemàtica mitjana, ja que té en compte les escales característiques del flux. Aquestes escales, però, integren l'efecte de la turbulència i, per tant, també podríem anomenar aquesta viscositat cinemàtica mitjana *viscositat cinemàtica turbulenta*. Llavors, el nombre de Reynolds també es pot interpretar com el quocient entre la viscositat turbulenta (que el fluid té a causa de les característiques del flux turbulent) i la molecular (característica de cada fluid independentment de les característiques del flux). Així, doncs, com més gran és el nombre de Reynolds, més gran és la viscositat turbulenta (efectiva) d'un flux. Per tant, podem concloure que els fluxos turbulents són més viscosos que els laminars.

2.2. Turbulència activa i turbulència fòssil

Hem parlat de deformacions, però que una partícula es deformi no vol pas dir que rodi. Perquè una partícula rodi sobre un eix determinat (k , per exemple), les velocitats de rotació de dues línies perpendiculars entre si incloses sobre un pla de la partícula perpendicular a l'eix de rotació (pla ij) han de ser diferents o de sentit igual. Dit d'una altra manera, si $w_k = \partial v_j / \partial x_i - \partial v_i / \partial x_j$ és diferent de zero, llavors la partícula roda. Generalitzant aquesta expressió al cas tridimensional, tenim que $\omega = \nabla \times \mathbf{v}$, on \times indica el producte vectorial i ω és la vorticitat, que es defineix mitjançant aquesta equació. La vorticitat és, de fet, el doble de la velocitat angular mitjana d'una partícula; hom la defineix així per evitar arrossegar un factor dos quan es treballa amb les equacions del fluid.

El concepte de *vorticitat* és molt important en la turbulència, ja que aquesta és necessàriament vortical. El terme de vorticitat pot introduir-se fàcilment en l'equació de Navier-Stokes considerant la igualtat vectorial $(\nabla \times \mathbf{v}) \times \mathbf{v} = \mathbf{v} \cdot \nabla \mathbf{v} - \nabla v^2 / 2$.

Llavors, si substituïm el terme $\mathbf{v} \cdot \nabla \mathbf{v}$, responsable de la generació de turbulència, tenim que

$$\partial \mathbf{v} / \partial t = \mathbf{v} \times \omega + \mathbf{v} \times 2\Omega + (\rho' / \rho) \mathbf{g} - \nabla p / \rho - \nabla v^2 / 2 + \nu \nabla^2 \mathbf{v},$$

on ens apareix el terme $\mathbf{v} \times \omega$, que podem interpretar com una força. De fet, la intensitat relativa d'aquesta força s'ha proposat per a definir el caràcter turbulent d'un fluid: si supera la força viscosa, $\nu \nabla^2 \mathbf{v}$, el flux és turbulent; si no, no. Com que les dimensions de $\mathbf{v} \times \omega$ i $\mathbf{v} \cdot \nabla \mathbf{v}$ són les mateixes, el nombre de Reynolds també es pot interpretar com la relació entre la força vortical i la viscosa.

Tanmateix, en fluxos estratificats la força vortical acaba essent superada per la força gravitatòria (de flotabilitat), de manera que les forces de flotabilitat són les responsables principals de l'esmortiment de la turbulència. Hi ha autors que sostenen que, en aquesta situació, la turbulència es fossilitza, és a dir, que les inestabilitats que s'han generat fins a aquell punt perduraran encara durant un temps llarg malgrat que l'energia cinètica turbulenta s'hagi dissipat en gran part. Tindrem, per tant, una imatge fòssil del que ha estat un procés de barreja molt més energètic que no és en aquell moment. Una imatge intuïtiva per a il·lustrar aquesta idea és el deixant d'un avió de reacció.

La idea de la turbulència fòssil va ser proposada per Gibson el 1980 dintre d'un marc molt més ampli i ha generat molta controvèrsia entre la comunitat internacional d'oceanògrafs. Últimament, però, moltes de les seves propostes van essent acceptades i ha publicat recentment un article de revisió de la teoria de la turbulència fòssil (Gibson, 1999). Aquest autor manté que la majoria de les mesures de turbulència oceànica de què es disposa són de turbulència activofòssil o fòssil i que, per tant, la barreja en temps anteriors al mostreig havia de ser molt més gran que la que es dedueix de les característiques que té el clap turbulent en el moment de la mesura. Gibson, a més, proposa com s'ha d'inferir la velocitat de dissi-

pació de l'energia cinètica turbulenta al principi de la fossilització a partir de les dades de microestructura.

L'existència de turbulència fòssil explicaria el que s'ha anomenat la *paradoxa de la barreja negra* a l'oceà, que es denomina així per analogia amb la matèria negra, l'existència de la qual es dedueix a partir de consideracions cosmològiques, però que no es veu. En oceanografia, els fluxos de barreja que s'obtenen a partir de mesures del flux mitjà són els que s'utilitzen en els models numèrics de circulació que descriuen correctament la dinàmica general. Aquests valors, però, són superiors als que es dedueixen a partir de mesures directes de la turbulència. Naturalment, una causa d'aquesta discrepància podria ser que existissin processos de barreja oceànica importants que encara no s'hagin identificat, però també en podria ser la causa que els mostrejos que s'han fet siguin inadequats per a tenir-ne una bona estadística. De fet, segons la teoria de la turbulència fòssil hi ha una intermitència lognormal extrema de les dissipacions.

2.3. La turbulència és altament difusiva

Si apliquem el rotacional a l'equació de Navier-Stokes i considerem diverses relacions vectorials, obtenim l'equació $D\omega / Dt = (\omega \cdot \nabla) \mathbf{v} + \nu \nabla^2 \omega$, que descriu la dinàmica de la vorticitat en un fluid de densitat constata. El primer terme de la part esquerra de la igualtat és el que s'anomena un *terme font* i representa la creació/destrucció de vorticitat a causa del mecanisme d'estirament/inclinació dels vòrtexs. Aquest mecanisme es basa en la conservació del moment angular, i és el que el patinador sobre gel sap fer quan tanca o obre els braços per girar més o menys de pressa.

Si no considerem el terme font, l'equació de la vorticitat queda reduïda a $D\omega / Dt = \nu \nabla^2 \omega$, que formalment és igual a l'equació de transport de calor $DT / Dt = k \nabla^2 T$, o a la de massa, $DC / Dt = D \nabla^2 C$, quan tampoc no es considera cap terme de creació o destrucció; k i D són, respectivament, les difusivitats tèrmica i massica i tenen les mateixes dimensions que ν . Per tant, comparant les equacions de ω , T i C , podem interpretar la viscositat cinemàtica ν com la difusivitat de la vorticitat. Per tal causa, encara que relacionàvem la viscositat dinàmica μ amb el transport de la quantitat de moviment, el caràcter difusiu és determinat per ν i no pas per μ . Per això, tot i que l'aigua és més viscosa que l'aire, és prou més densa per a ser unes quinze vegades menys difusiva que l'aire.

En el cas d'un flux turbulent, abans hem vist que podem parlar d'una viscositat cinemàtica mitjana (turbulenta) que augmenta amb el nombre de Reynolds, és a dir, amb el grau de turbulència. Ara, però, veiem que la viscositat cinemàtica és, de fet, una difusivitat; per tant, difusivitats grans són també una de les característiques que s'associen amb turbulència. De fet, tots nosaltres sabem això quan en generem remenant amb la cullereta per barrejar més de pressa la llet i el cafè.

3. Aproximació de Reynolds a la turbulència

3.1. Descomposició de Reynolds

La irresolubilitat de la turbulència per a nombres de Reynolds alts afavoreix que se'n faci un tractament estadístic, i per això es diu de la turbulència que és un procés estocàstic. En estadística sovint parlem de propietats mitjanes, i en el nostre cas seran definides pel que s'anomena la *descomposició de Reynolds de les variables instantànies*. Les variables instantànies són les que hem considerat fins ara a les equacions que hem mencionat i corresponen al valor real que les variables prenen en cada moment. La descomposició de Reynolds es basa a considerar que el valor instantani d'una variable (a , per exemple) es pot expressar com la suma d'un valor mitjà (A) i una altra quantitat petita respecte a aquesta que s'anomena *fluctuació* (a'), de manera que $a = A + a'$; en el cas de la pressió, seria $p = P + p'$, i en el de la velocitat, $u_i = U_i + u_i'$.

Aquesta descomposició es pot aplicar a totes les variables de les equacions de Navier-Stokes, les quals llavors es poden amitjanar, tenint present les consideracions estadístiques següents: $\overline{A} = A$; $\overline{AB} = AB$; $\overline{Aa'} = 0$; $\overline{ab} = AB + \overline{a'b'}$. Com a resultat de fer tot això, tenim que les equacions de Reynolds per al flux mitjà són idèntiques a les de Navier-Stokes per a les variables instantànies, però amb un terme nou que es correspon amb la divergència de la correlació de les fluctuacions de les velocitats turbulentes, és a dir, $\partial(\overline{u_i' u_j'}) / \partial x_j$. Més concretament, la mitjana de Reynolds de la component i de l'equació de Navier-Stokes és

$$\partial U_i / \partial t + U_j \partial U_i / \partial x_j = -(1 / \rho) \partial P / \partial x_i - (\rho' / \rho) g_i + \partial (v \partial U_i / \partial x_j - \overline{u_i' u_j'}) / \partial x_j.$$

Considerant que $\partial (v \partial U_i / \partial x_j) / \partial x_j$ és, de fet, la divergència del tensor de deformacions dividit per la densitat, per analogia podem definir el tensor de Reynolds, τ_{Rij} , com a

$$\tau_{Rij} / \rho = - \overline{u_i' u_j'}.$$

El tensor de Reynolds, τ_{Rij} , descriu la transferència de quantitat de moviment associada amb les fluctuacions (turbulència) de la mateixa manera que el tensor de tensions, τ_{ij} , descriu la transferència de quantitat de moviment a causa de la viscositat. Fixem-nos que la traça del tensor de Reynolds és el doble de l'energia cinètica turbulenta, K :

$$\tau_{Rii} = \rho \overline{u_i' u_i'} = 2K.$$

Anàlogament, si fem la mitjana de Reynolds de l'equació de conservació de la temperatura, tenim

$$\partial T / \partial t + U_j \partial T / \partial x_j = \partial (k \partial T / \partial x_j - \overline{u_j' T'}) / \partial x_j,$$

on k és la difusivitat tèrmica. Aquesta equació és igual a la de conservació de la temperatura instantània, però amb un terme més en el qual apareix la correlació de les fluctuacions de velocitat i de temperatura. Si multipliquem l'equació anterior per ρC_p , tenim l'equació de l'energia i podem interpretar el terme de la dreta com la divergència del flux de calor Q_j , que escriurem com a

$$Q_j = \kappa \partial T / \partial x_j - \rho C_p \overline{u_j' T'}$$

El primer terme és el flux de calor molecular, on κ és la conductivitat tèrmica, i el segon és el flux turbulent de calor degut als remolins de la turbulència que interacciona a totes les escales i que també s'anomena *flux de calor de Reynolds*. C_p és la calor específica.

La descomposició de Reynolds permet, doncs, aïllar els processos turbulents i identificar-ne els fluxos, tal com hem vist per al cas de la quantitat de moviment (tensor de Reynolds) i de calor (flux de calor de Reynolds).

3.2. Equacions de l'energia cinètica turbulenta i de la variància de temperatura: un pas cap a la mesura

La velocitat de canvi de l'energia cinètica turbulenta, K , també es pot deduir a partir de les equacions de Navier-Stokes i de la descomposició de Reynolds. Més concretament, s'obté restant l'equació de Navier-Stokes, amitjanada per Reynolds, de l'equació amb variables instantànies però explicitant la descomposició de Reynolds. Després el resultat que s'obté es multiplica escalarment per la velocitat turbulenta i se'n treu la mitjana. El resultat és

$$\begin{aligned} \partial K / \partial t + \partial (U_k K) / \partial x_k + \partial (\overline{u_k' u_j' u_j'}) / \partial x_k = & -(1 / \rho) \partial (\overline{u_k' p'}) / \partial x_k + \\ & + \partial (2\nu u_j e_{kj}') / \partial x_k - \overline{u_i' u_j'} \partial U_i / \partial x_j - (g_i / \rho) \overline{\rho' u_i'} - 2\nu \overline{e_{ij}' e_{ij}'} \end{aligned}$$

on e_{ij}' és el tensor de deformacions a escales turbulentes i, per tant, conté els cisallaments turbulents. En aquesta equació s'identifiquen quatre termes divergents que són els responsables del transport de l'energia cinètica turbulenta. Aquest transport és degut al flux mitjà (segon terme de l'esquerra de la igualtat), al mateix flux turbulent a través dels remolins (últim terme de l'esquerra de la igualtat), a les fluctuacions de pressió (primer terme de la dreta de la igualtat) i, en menor mesura, a la viscositat molecular (segon terme de la dreta de la igualtat). Aquests termes de transport no contribueixen a la variació neta de l'energia cinètica turbulenta, sinó que només la redistribueixen dintre del mateix fluid. Per tant, si considerem el cas d'equilibri local, els podrem considerar zero. Anàlogament, si la turbulència és estacionària, dK/dt també val zero, i només ens queden els termes de producció i destrucció de l'energia cinètica turbulenta

$$-\overline{u_i' u_j'} \partial U_i / \partial x_j - (g_i / \rho) \overline{\rho' u_i'} - 2\nu \overline{e_{ij}' e_{ij}'} = 0 .$$

Més concretament, el terme $-\overline{u_i' u_j'} \partial U_i / \partial x_j$ correspon a la producció d'energia cinètica turbulenta a causa de l'acció dels esforços de Reynolds, $\overline{u_i' u_j'}$, sobre el cisallament del flux mitjà, $\partial U_i / \partial x_j$. La transferència d'energia cinètica entre el flux mitjà i el camp turbulent no es pot pas donar a la inversa, i, per tant, aquest terme sempre contribuirà positivament a l'evolució de K (malgrat el signe negatiu explícit a la fórmula).

El terme següent, $-(g_i/\rho)\overline{\rho' u_i'}$, és el de producció/destrucció per flotabilitat, i pot ser positiu o negatiu segons si transfereix l'energia potencial del sistema cap a energia turbulenta o a la inversa. En el primer cas, perquè augmenti l'energia cinètica turbulenta a costa de la potencial del sistema, cal que aquest sigui un sistema inestable; ens trobem, per tant, en el cas de la convecció. Contràriament, si el sistema és estable, llavors la turbulència pot desestabilitzar-lo (augmentar l'energia potencial) a costa precisament de l'energia cinètica turbulenta, i ens trobem en el cas que aquest terme és negatiu.

Finalment, el tercer terme, $-2\overline{v e_{ij}'} e_{ij}'$, és sempre negatiu i té a veure amb el treball fet per la viscositat a escales petites on l'energia es dissipa, i és exactament la velocitat de dissipació de l'energia turbulenta. Malgrat que la viscositat cinemàtica aparegui en aquest terme, cal tenir en compte que aquesta és una constant per a cada fluid; per tant, de l'equació de conservació de K es dedueix que e_{ij}' acabarà essent sempre prou gran per a poder dissipar tota l'energia. Així, doncs, ε no depèn de v .

Anàlogament, a partir de les equacions de la temperatura amitjanada per Reynolds es pot deduir l'equació per a la variància de la temperatura:

$$\begin{aligned} \overline{\partial T'^2} / \partial t + \partial(U_k \overline{T'^2}) / \partial x_k + \partial(\overline{u_k' T'^2}) / \partial x_k = + \partial(k \partial \overline{T'^2} / \partial x_k) / \partial x_k - \\ - 2 \overline{u_k' T'} \partial T / \partial x_k - 2k (\overline{\partial T' / \partial x_k})^2. \end{aligned}$$

En aquesta equació podem tornar a identificar uns termes divergents que són els responsables del transport de la variància que, en el supòsit d'equilibri local, es fan zero. Llavors, per a un cas d'equilibri local i estacionari, ens queden només $-2 \overline{u_k' T'} \partial T / \partial x_k$ i $-2k (\overline{\partial T' / \partial x_k})^2$, que descriuen la producció i la dissipació. L'últim terme és el de dissipació de la variància de temperatura i s'acostuma a indicar com a ε_T .

Fixem-nos, doncs, que la descripció de Reynolds ens permet també identificar els mecanismes generadors i aniquiladors de l'energia cinètica turbulenta, alhora que ens dóna les expressions formals de les velocitats de dissipació ε i ε_T , sobre les quals es basen, tal com veurem, algunes de les tècniques que en fan possible la determinació experimental. Naturalment, la mesura de la velocitat de dissipació ε , en un cas estacionari i amb equilibri local, també ens dóna la velocitat de producció, que és de més mal obtenir, ja que les velocitats turbulentes són difícils de mesurar. En canvi, el gradient de les velocitats turbulentes en la direcció vertical, dv'/dz , és una variable que, amb limitacions, és mesurable. Llavors, a partir de l'expressió $\varepsilon = 2v\overline{e_{ij}'} e_{ij}'$, si considerem isotropia local, tenim que $\varepsilon = 7,5 (dv'/dz)^2$, i, per tant, podem determinar el valor de la velocitat de dissipació a partir del cisallament turbulent. Paral·lelament, el terme de dissipació

de la variància de la temperatura $\varepsilon_T = 2k \overline{(\partial T' / \partial x_i)^2}$, per al cas d'isotropia local, val $6k \overline{(dT' / dz)^2}$ i també es pot mesurar.

3.3. Difusivitats i fluxos turbulents; mesures in situ

En aquest apartat reescriurem l'equació de l'energia cinètica turbulenta per al cas estacionari i d'equilibri local com a

$$Pc + F - \varepsilon = 0 ,$$

on $Pc = -\overline{u_i' u_j'} \partial U_i / \partial x_j$ és el terme de producció per cisallament; $F = -(\overline{g_i} / \rho) \overline{T' u_i'}$, el terme de flotabilitat, i $\varepsilon = -2\nu \overline{e_{ij}' e_{ij}'}$, el terme de dissipació, tal com ja l'havíem definit abans. Llavors, fent servir aquesta notació, definim el flux de Richardson com a

$$Ri_f = F / Ps .$$

D'aquesta definició veiem que, quan $F > 0$, el valor de Ri_f ens diu quina és la importància relativa del flux de flotabilitat respecte a la producció per cisallament. Per tant, Ri_f ens dóna una mesura de l'eficiència de barreja. Per a una turbulència plenament desenvolupada, es coneix que Ri_f assoleix un valor màxim i aproximadament igual a 0,15. Per tant, des d'un punt de vista de rendibilitat, veiem que la barreja turbulenta a partir del cisallament del flux mitjà és un procés poc efectiu.

Històricament, però, abans que el flux de Richardson es va definir un altre coeficient que s'anomena precisament *eficiència de barreja*, γ_m , i que és el quocient entre el flux de flotabilitat i la velocitat de dissipació:

$$\gamma_m = F / \varepsilon = -\overline{g' w'} / \rho \varepsilon ,$$

on w' és la component vertical de la velocitat turbulenta. En el cas d'equilibri local, γ_m i Ri_f es relacionen i $\gamma_m = Ri_f / (1 - Ri_f)$. Per tant, si $Ri_f = 0,15$, llavors $\gamma_m = 0,2$. Per a un cas qualsevol, γ_m se sol parametritzar d'acord amb el nombre de Richardson, que també s'anomena *gradient de Richardson*, el qual determina el grau d'inestabilitat d'un flux amb cisallament. Conèixer γ_m és importantíssim, ja que ens permet estimar la difusivitat turbulenta de massa, K_p , a partir de les mesures de dissipació.

Concretament, K_p es defineix com la constant que relaciona el flux turbulent de massa, $-\overline{\rho' w'}$, amb el gradient de densitat i, per tant,

$$K_p = -\overline{\rho' w'} / (\partial \rho / \partial z) .$$

Llavors, considerant la definició d'eficiència de barreja, tenim que

$$K_p = \gamma_m \varepsilon / (\partial \rho / \partial z) ;$$

i, per tant, considerant que $\gamma_m = 0,2$, podem estimar experimentalment K_p a partir de la mesura de ε i $\rho(z)$ i sense necessitat de tenir informació sobre les velocitats turbulentes, que, tal com hem comentat, són molt més difícils de mesurar.

D'altra banda, si considerem el cas d'un fluid estratificat amb turbulència isotròpica, podem escriure l'equació de conservació de la variància de temperatura com a $-2\overline{w'T'}\partial T/\partial z - 6k(\overline{\partial T'/\partial z})^2$, on w' és la component turbulenta vertical, i consegüentment tindrem que

$$-\overline{w'T'} = 3k(\overline{\partial T'/\partial z})^2 / \partial T / \partial z .$$

Normalment, però, no es treballa directament amb la variància del gradient de la temperatura turbulenta, sinó amb el nombre de Cox, que es defineix com a

$$C_x = 3(\overline{\partial T'/\partial z})^2 / (\partial T / \partial z)^2 .$$

I, doncs, $-\overline{w'T'} = k C_x \partial T / \partial z$. Fixem-nos que, si considerem la difusivitat tèrmica turbulenta $k_T = -\overline{w'T'} / \partial T / \partial z$, llavors de l'expressió anterior obtenim que

$$k_T = k C_x .$$

Per tant, el nombre de Cox ens diu precisament quant més difusiu, pel que fa a la temperatura, és el flux turbulent respecte al viscos.

Finalment, a partir de $-\overline{w'T'}$ podem obtenir el flux de calor turbulent o flux de calor de Reynolds $q = -\rho C_p \overline{w'T'} = -\rho C_p k C_x dT / dz$. A vegades, però, és més convenient treballar amb el flux de flotabilitat, q_{bi} , ja que és el que determina realment la dinàmica. Llavors, $q_{bi} = \alpha g \overline{u_i' T'} = (g/\rho) \overline{u_i' \rho'} = u_i' \overline{b}$, on $b = \rho'/\rho$ és la flotabilitat.

Tenim, doncs, un altre mètode alternatiu per al càlcul de fluxos turbulents (màssic o de flotabilitat) a partir del nombre de Cox i, per tant, a partir de la mesura del gradient de les temperatures turbulentes en comptes de ε i, consegüentment, la mesura del gradient de les velocitats turbulentes (cisallament turbulent). En l'un cas i en l'altre, però, cal resoldre experimentalment totes les escales turbulentes per a calcular la variància correctament, i això sovint no és possible. Normalment, les mesures de temperatura són de més bon fer, ja que no són sensibles a les petites pertorbacions que pateix l'aparell quan es desplaça a través de la columna d'aigua. Ara bé, d'altra banda, tenim fluctuacions de temperatura a escales més petites que l'escala inferior per a la velocitat i, per tant, necessitem més resolució espacial, la qual cosa no sempre és possible. Aquestes limitacions de les mesures, bé per falta de resolució de l'instrument o bé per

contaminació amb soroll, també es poden superar gràcies a unes altres aproximacions estadístiques a la turbulència que veurem seguidament.

4. Descripció espectral de la turbulència

4.1. El rang d'equilibri; definició i possibilitats de mesura

S'anomena *rang d'equilibri* el format per les escales en què s'ha perdut la informació del flux mitjà, on sovint la turbulència presenta complexes estructures espacials i, per tant, en el rang d'equilibri la turbulència és homogènia. El rang d'equilibri es divideix en dos: l'inercial i el dissipatiu. El rang inercial és format per les escales més grans, on les components del tensor de deformacions e_{ij}' són encara petites i es pot negligir el terme de dissipació; en conseqüència, es pot considerar que només hi ha transferència d'energia. En canvi, a escales més petites, e_{ij}' augmenta prou perquè el terme de dissipació viscosa sigui important, i llavors es parla de *rang dissipatiu*.

La descripció estadística de la turbulència a petita escala es basa en les hipòtesis de Kolmogorov, segons les quals la transmissió d'energia s'acompleix en cascada des de les escales grans fins a les més petites. Aquesta transferència és determinada per l'afluència mitjana d'energia des de les components de gran escala, que, en estat d'equilibri, és igual a la dissipació mitjana d'energia a les escales mínimes. El procés subjacent en aquesta transferència d'energia se suposa que és un procés de cascada basat en el mecanisme d'estirament de vòrtex. Gibson, però, defensa que aquesta transferència d'energia no està lligada a un procés de cascada de remolins des de les escales grans fins a les petites i creu que les inestabilitats que originen la turbulència es donen a escales petites i, mitjançant un procés d'aparellament de vòrtexs, afecten les escales més grans, de les quals s'extreu energia cinètica a partir de l'arrossegament del fluid no turbulent. De totes maneres, independentment de la visió que tinguem de quin és el mecanisme de transferència d'energia, en l'un cas i en l'altre s'arriba a la conclusió que a escales petites s'assoleix un estat d'equilibri on les hipòtesis de Kolmogorov són certes.

La primera d'aquestes hipòtesis estableix que l'escala espacial on s'esdevingui la dissipació, η , depèn només de dos paràmetres, ϵ i ν . Llavors, considerant que les unitats de ϵ són L^2/T^3 , i les de ν , L^2/T , per arguments dimensionals trobem que $\eta = (\nu^3/\epsilon)^{1/4}$, que és l'anomenada *escala de Kolmogorov*.

En el cas de l'aigua, però, si el fluid està estratificat, a escales en què ja no hi ha fluctuacions de velocitat encara hi ha fluctuacions de temperatura. Això es deu al fet que les fluctuacions de velocitat a escala molecular s'esmoreeixen més ràpidament que les de temperatura, ja que la viscositat cinemàtica de l'aigua, ν , és un ordre de magnitud més gran que la difusivitat tèrmica, k . L'equivalent de l'escala de Kolmogorov pel que fa a la temperatura és l'anomenada *escala de Batchelor*, η_b , la qual depèn de k , però també de ν , ja que les fluctuacions de velocitat i

de temperatura estan íntimament lligades. Llavors, també amb arguments dimensionals, es troba que $\eta_B = (kv^2/\nu)^{1/4}$. Per a unes altres escales també es defineix l'escala de Batchelor corresponent senzillament considerant el valor de la difusivitat molecular d'aquests en comptes de la de la temperatura.

Veiem que tant η com η_B depenen inversament del ritme de dissipació ϵ , el qual, en situacions estacionàries, és igual al de producció. A tall d'exemple, les dissipacions a la capa superficial de barreja poden arribar a ser de 10^{-5} W/kg; per tant, l'escala de Kolmogorov és de prop de 0,5 cm, i la de Batchelor (per a la temperatura), d'1 mm; per al cas de la salinitat, l'escala de Batchelor pot arribar a ser de poques dècimes de mil·límetre. D'altra banda, s'estima que a l'interior de l'oceà les velocitats de dissipació poden ser tan baixes com 10^{-11} W/kg, tot i que aquest valor ara per ara no és mesurable. En aquest cas, l'escala de Kolmogorov seria de prop de 10 cm i les escales de Batchelor per a la temperatura i la salinitat estarien per sobre i per sota d'1 cm respectivament.

Per a l'estudi experimental de la turbulència cal, doncs, tenir la resolució espacial corresponent a les escales de Batchelor i Kolmogorov, i els diferents sensors han de poder resoldre la magnitud de les fluctuacions en aquestes escales. Això, pel que fa a la temperatura, implica una resolució inferior a 10^{-3} °C; per a la conductivitat (a partir de la qual obtenim la informació de salinitat) es necessita 10^{-4} S m^{-1} , i per al cisallament turbulent, per sota de 10^{-2} s $^{-1}$. Aquest tipus d'instrumentació, que a partir dels anys vuitanta van començar a desenvolupar uns quants grups de recerca punta, ara és comercialitzada per tres o quatre companyies industrials implicades en la recerca, tot i que a preus elevats (d'uns setanta-cinc mil €). Els sensors corresponents solen estar muntats sobre un vehicle (anomenat *perfilador de microestructura*) que s'utilitza, normalment, en caiguda lliure, a velocitats prou baixes perquè, d'acord amb la freqüència de mostreig (entre 100 i 1.000 Hz), es puguin resoldre les escales espacials desitjades. Naturalment, el temps de resposta dels sensors també ha de ser l'adequat.

4.2. L'espectre de la turbulència en el rang d'equilibri

La densitat espectral de la turbulència, $S(k)$, és la contribució a la variància de la velocitat turbulenta per a un (rang de) nombre d'ones k , i, per tant, té a veure amb la contribució de les escales corresponents a l'energia cinètica turbulenta, K . Formalment, $\overline{u_i^2} = \int S(k) dk$ i $K = (\rho/2)\overline{u_i^2}$. Fixem-nos que $S(k)$ té unitats de L^3/T^2 . S'anomena *espectre* la representació (sovint logarítmica) de la densitat espectral segons el nombre d'ones.

La segona hipòtesi de Kolmogorov incideix en l'espectre de la turbulència en el rang inercial i diu que, en aquest rang, el paràmetre ν no té importància en la seva dinàmica, de manera que les seves característiques de $S(k)$ són determinades només per ϵ . Més amunt ja havíem comentat que el terme de dissipació $2\nu e_{ij}'e_{ij}'$ només és important quan les velocitats de deformació de les fluctuacions e_{ij}' són molt grans, i això només és així per a escales de prop de η . Llavors, a

partir d'aquesta suposició, i per consideracions dimensionals, s'obté l'anomenat *espectre de Kolmogorov*, que té la forma $S(k) = A \varepsilon^{2/3} k^{-5/3}$. Recordem que les dimensions de ε són L^2/T^3 .

Segons aquesta llei, quan el nombre de Reynolds és prou gran (només així existeix el rang inercial), si mesurem les velocitats turbulentes, en calculem la densitat espectral i representem $\ln S(k)$ respecte a $\ln k$, hem d'obtenir una recta amb pendent negatiu i igual a $-5/3$ i la seva situació a l'eix d'ordenades dependrà de la magnitud de ε . Aquest comportament, predit teòricament el 1941, només es va poder provar experimentalment el 1962 a partir de mesures oceanogràfiques amb $Re \approx 10^8$, ja que uns nivells de turbulència prou elevats perquè existeixi el rang inercial són pràcticament impossibles d'obtenir al laboratori.

En sistemes naturals, però, és més fàcil mesurar el cisallament turbulent que les velocitats. Tanmateix, com que el cisallament és la derivada de la velocitat i la densitat espectral té a veure amb la variància d'una magnitud, a partir de les propietats de la transformada de Fourier tenim que l'espectre del cisallament ha de ser k^2 vegades l'espectre de la velocitat i, per tant, ha de variar com a $k^{1/3}$. Així, doncs, l'espectre del cisallament en el rang inercial en representació logarítmica té un pendent positiu d' $1/3$. La seva localització en el nivell d'ordenades també és donada pel valor de la velocitat de dissipació de l'energia cinètica turbulenta.

A partir també d'arguments anàlegs als anteriors, i considerant ara que la velocitat de dissipació de la variància de temperatura, ε_T , té unitats de $^\circ C^2/T$, Obukhov va trobar que l'espectre de temperatura en el rang inercial (convectiu) també depèn de $k^{-5/3}$ i té la forma $S_T(k) = B \varepsilon_T \varepsilon^{-1/3} k^{-5/3}$, on ara $S_T(k)$ té unitats de $^\circ C^2 L/T$. Que S_T no depengui només de ε_T sinó també de ε es deu al fet que les fluctuacions de temperatura també estan relacionades amb les de velocitat. Evidentment, l'espectre del gradient de la temperatura turbulenta també varia com a $k^{1/3}$, igual que el del gradient de velocitat.

Més amunt hem dit, però, que, en el cas de l'aigua, les fluctuacions de temperatura abracen escales més petites que les de velocitat i, de fet, quan les fluctuacions de velocitat ja es dissipen a causa de la viscositat, les fluctuacions de temperatura encara no s'esmoreeixen, sinó que la convecció en aquestes escales encara és activa. En aquest rang, anomenat *convectivoviscós*, el procés de transferència de la variància de temperatura és diferent que en el rang inercial, on les fluctuacions de velocitat encara no s'han esmoreït. Per tant, en el rang convectivoviscós el comportament de l'espectre de la temperatura varia respecte al rang inercial i, segons Batchelor, depèn de k^{-1} . Finalment, per a escales inferiors a $10\eta_B$, l'espectre de temperatura decau exponencialment, i llavors es parla del *subrang difusiu*.

El coneixement d'aquests espectres universals no solament ens proporciona informació i comprensió sobre els processos turbulents, sinó que ens permet extrapolar les mesures que es tenen de turbulència a escales que no es poden resoldre, sia per limitació dels aparells, sia per contaminació de soroll. L'ajust de dades experimentals a espectres universals és, de fet, un dels mètodes emprats per a corregir la pèrdua de variància en les mesures de microestructura. La importància de conèixer les velocitats de dissipació per a determinar els fluxos turbulents i la producció d'energia cinètica turbulenta ja s'ha comentat.

5. Comentaris finals

Amb aquest document s'ha volgut donar una pinzellada dels pilars teòrics sobre els quals es fonamenten els treballs que s'han presentat en aquestes Jornades per tal d'il·lustrar una aproximació concreta a un sistema complex. Concretament, s'hi han comentat diferents estudis relacionats amb la turbulència en sistemes aquàtics naturals, línia en la qual es treballa dins del Grup de Física Ambiental de la Universitat de Girona. Els treballs presentats no es discuteixen en aquest document, ja que són documentats a Colomer *et al.* (2001), Muñiz (2000) i Roget *et al.* (2002). Pel que fa als treballs més metodològics relacionats amb aquesta línia, podem referir-nos, per exemple, a Piera *et al.* (2000).

Bibliografia

- COLOMER, Jordi; SERRA, Teresa; PIERA, Jaume; ROGET, Elena; CASAMITJANA, Xavier (2001). «Observations of a hydrothermal plume in a karstic lake». *Limnol. Oceanogr.*, núm. 46 (1), p. 197-203.
- GIBSON, Carl (1999). «Fossil turbulence revisited». *Journal of Marine Systems*, núm. 21 (1-4), p. 147-167.
- MUÑIZ, Miquel Àngel (2000). *Anàlisi de la microestructura en un plomall convectiu*. Girona: Universitat de Girona. 54 p. [Treball de recerca; doctorat de medi ambient]
- PIERA, Jaume; ROGET, Elena; CATALÁN, Jordi (2000). «Turbulent patch identification in microstructure profiles». A: LAWRENCE, G. A.; PIETERS, R.; YONEMITSU, N. [ed.]. *Stratified flows (II)*. P. 1233-1238.
- ROGET, Elena; FIGUEROA, Manuel; LOZOVATZKY, Iosif (2002). «The dissipation rate in the wind-induced Ekman boundary layer». Niça: European Geophysical Society, XXVII General Assembly.

Resum

La predicció numèrica és una de les branques més importants de la meteorologia moderna. A principis del segle passat Vilhelm Bjerknes ja va mostrar el procediment científic per a pronosticar l'estat futur de l'atmosfera. No obstant això, va ser Lewis Fry Richardson el primer meteoròleg que va intentar resoldre numèricament les equacions de la dinàmica i de la termodinàmica que expliquen el comportament de l'atmosfera. El seu llibre *Weather prediction by numerical process* (1922) constitueix el primer tractat de predicció numèrica del temps (PNT). Van haver de passar quasi trenta anys perquè Charney, Fjörtoft i Von Neumann obtinguessin el primer mapa del temps utilitzant aquesta tècnica. L'aplicació operativa i el desenvolupament consegüent de la PNT durant la dècada dels seixanta afavoriren l'interès pels temes relacionats amb la predictibilitat dels sistemes dinàmics i, en particular, per la determinació dels límits en la predicció del temps. En aquest sentit, els treballs de Lorenz demostren que la predicció determinista aplicada a l'atmosfera té un límit pròxim a les dues setmanes. Actualment, l'aplicació de noves tècniques com és la predicció per conjunts ha fet possible, d'alguna manera, superar aquests límits. A Catalunya l'aplicació amb finalitats operatives i de recerca de la PNT és relativament recent, encara que moltes de les iniciatives dutes a terme en els últims anys hagin estat pioneres a l'Estat espanyol, fet que permet mirar el futur amb un cert optimisme.

Abstract

Numerical prediction is one of the most important branches of modern meteorology. At the beginning of the last century, Wilhelm Bjerknes demonstrated the scientific procedure for forecasting the future state of the atmosphere. Nevertheless, Lewis Fry Richardson was the first meteorologist who tried to numerically solve the equations of Dynamics and Thermodynamics

that explain the behavior of the atmosphere. His book *Weather Prediction by Numerical Process* (1922) constituted the first treatise on the numerical prediction of the weather (NPW). Almost 30 years passed before Charney, Fjörtoft, and von Neumann obtained the first weather map using this technique. The operative application and consequent development of the NPW during the decade of the 60's favored interest in subjects related to the predictability of dynamic systems and, in particular, the determining of the limits of the prediction of the weather. In this sense, Lorenz's works demonstrated that deterministic forecasting applied to the atmosphere has an approximate limit of two weeks. Currently, the application of new techniques such as forecasting by groups has made it possible, in a way, to pass these thresholds. In Catalonia, its application of the NPW for operative purposes and research is relatively recent, although many of the initiatives carried out in recent years have been pioneers in Spain, which allows a certain optimism in looking at the future.

Introducció

La predicció numèrica del temps o, com suggereixen alguns experts, la predicció del temps per mètodes numèrics quedarà completament associada al passat segle xx. La seva història neix amb aquest segle i assoleix la maduresa completa al seu acabament. Encara que les lleis de la física que s'utilitzen en la PNT es coneixien des de molt temps abans, el primer meteoròleg que establí formalment les bases de la predicció del temps mitjançant la utilització de les equacions que expressen aquestes lleis és Vilhelm Bjerknes, l'autor de la teoria del front polar de l'escola noruega. En el seu treball (Bjerknes, 1904) explicà que, si es pogués trobar un mètode analític d'integració de les equacions de la dinàmica aplicades a l'atmosfera, la predicció del temps quedaria resolta tan sols determinant, a partir de les observacions meteorològiques, l'estat inicial de les set variables que en determinen l'estat (temperatura, pressió, densitat, humitat i les tres components del vent).

Com és conegut, el primer científic que va intentar resoldre el sistema d'equacions en derivades parcials va ser Lewis Fry Richardson utilitzant el mètode de les diferències finites que ell mateix havia aplicat a la resolució d'uns altres problemes físics. Richardson aplicà les dues equacions del moviment horitzontal (és a dir, la segona llei de la dinàmica aplicada al moviment horitzontal de l'aire), l'equació hidrostàtica (que substitueix l'equació del moviment vertical), l'equació que expressa el primer principi de la termodinàmica, l'equació de continuïtat aplicada a l'aire i a l'aigua que l'acompanya i l'equació d'estat. És el que es coneix com a *sistema d'equacions primitives*.

Per aplicar el mètode de les diferències finites a la resolució del sistema de equacions, Richardson va quadrillar una àmplia regió de l'atmosfera sobre el centre d'Europa assignant a cada quadre (d'uns 200 km de costat a la part central de la xarxa) uns valors determinats de les distintes variables meteorològiques. Per tal de tenir-ne en compte la variació vertical, divi-

dí l'atmosfera en cinc capes des del terra fins als 12 km d'altura aproximadament. El propòsit del càlcul era l'obtenció de la pressió atmosfèrica en el punt central de la quadrícula sis hores després. La predicció fou completament errònia, ja que pronosticà una variació de la pressió de quasi 150 hPa. L'assaig de Richardson, que fou publicat amb el títol *Weather prediction by numerical process* l'any 1922, constitueix el primer tractat de PNT.

És important, i a la vegada curiós, destacar que la causa principal del resultat incorrecte de la predicció de Richardson fou el fet de ser extremament rigorós i exacte en el tractament del problema, de les dades utilitzades i dels càlculs realitzats; va voler treballar amb tanta precisió que la mateixa complexitat del problema i la dificultat intrínseca amb la qual es determinen algunes variables meteorològiques feren que els resultats obtinguts no corresponguessin a les solucions meteorològiques cercades. Malgrat això, i encara que el resultat del seu assaig hagués estat més afortunat, era evident que amb els mitjans de càlcul de l'època el mètode proposat difícilment s'hauria pogut aplicar de manera pràctica. Richardson emprà sis setmanes per a calcular el pronòstic a sis hores de la pressió al centre de la xarxa. Van haver de passar quasi trenta anys perquè els meteoròlegs s'interessessin de nou per la predicció del temps mitjançant tècniques numèriques.

El model barotròpic

A finals de la dècada dels anys quaranta se celebrà una conferència de meteorologia a l'Institut d'Estudis Avançats de Princeton que reuní un grup de científics i meteoròlegs prestigiosos de l'època (Wiin-Nielsen, 1991) i l'objectiu principal de la qual era l'aplicació d'un model numèric de predicció del temps utilitzant una màquina de calcular electrònica. Després d'algunes setmanes de treball aconseguiren integrar numèricament l'equació de la vorticitat barotròpica aplicada al nivell de 500 hPa a una àmplia regió que contenia Amèrica del Nord. La màquina que van utilitzar fou l'ENIAC (Electronic Numerical Integrator and Computer), un dels primers ordinadors electrònics, que era propietat del Laboratori d'Investigacions Balístiques d'Aberdeen (Maryland). El treball es publicà a la revista *Tellus* el novembre de 1950 (Charney *et al.*) i constitueix el primer pronòstic del temps obtingut per mètodes numèrics mitjançant un ordinador.

D'aquesta primera predicció numèrica s'ha de destacar que, a diferència de l'intent de Richardson, no s'hi va utilitzar el sistema d'equacions primitives, sinó l'equació de la vorticitat barotròpica, que s'obté a partir d'aquelles després de simplificar dràsticament els termes que són inferiors en un ordre de magnitud o més als termes més grans, amb l'aproximació de vent geostrofic i aplicada al nivell de divergència nul·la. Dit d'una altra manera, l'èxit en la primera PNT s'aconsegüí gràcies a l'aplicació raonada, justificada i apropiada de simplificacions importants a les equacions que expliquen el comportament del sistema, les quals eviten, en molts casos, que les inevitables imperfeccions de les mateixes equacions i en la determinació de l'estat

inicial de l'atmosfera contaminin la part de la solució del sistema que té importància real en la descripció de l'estat futur del sistema que s'estudia.

La predictibilitat de l'atmosfera

El desenvolupament experimentat per la PNT durant la segona meitat del segle xx ha estat espectacular i, en gran part, paral·lel a l'augment vertiginós de la potència i la rapidesa dels ordinadors que utilitza. Així, del primer model d'un sol nivell i d'una xarxa solament de 200 punts i de 736 km de braç, s'ha passat a models globals de més de 40 capes o a la utilització operativa de models de mesoscala de 10 km de braç de malla i de 60 nivells a la vertical. A més, la qualitat i la bondat de les prediccions numèriques han experimentat una millora com a conseqüència de l'evolució dels models, de manera que els índexs que valoren l'exactitud dels pronòstics mostren, per exemple, que en menys de quinze anys s'ha aconseguit que els pronòstics per a setanta-dues hores tinguin la mateixa bondat que tenien els corresponents per a trenta-sis hores (Kalnay *et al.*, 1998).

La preocupació per aquestes qüestions apareix, no obstant això, molt abans que es millorés tan notablement l'encert de les prediccions. En realitat, els anys seixanta, quan ni tan sols les prediccions dels models no eren útils més enllà de dos dies, es comença a aprofundir en els temes relatius a la predictibilitat de l'atmosfera. En aquest camp, els treballs d'E. N. Lorenz establiren les bases per al desenvolupament de la teoria del caos (Lorenz, 1963, 1964 i 1965) i mostraren que el pronòstic meteorològic tenia un límit temporal més enllà del qual les solucions obtingudes en resoldre les equacions dinàmiques deixaven de ser útils per a pronosticar el temps en un lloc determinat. Els resultats obtinguts van demostrar ben aviat que més enllà de dues setmanes la PNT no podrà utilitzar-se per al pronòstic meteorològic a escala sinòptica com s'utilitza fins a deu dies. La raó d'aquestes conclusions es troba en la no-linealitat del sistema d'equacions que governa el comportament de l'atmosfera. Així, com explica Lorenz de manera magistralment senzilla (Lorenz, 1984), es podria pensar que, disminuint sistemàticament els errors que es cometen en obtenir les prediccions numèriques del temps, s'augmentaria progressivament el termini en què aquestes prediccions són útils. És a dir, millorant els sistemes d'observació meteorològica per a reduir els errors comesos en determinar l'estat inicial de l'atmosfera, utilitzant equacions que representin com més va més bé el comportament i l'evolució de les distintes variables meteorològiques i fent servir els mètodes numèrics i ordenadors més perfeccionats, ràpids i potents per a resoldre el sistema d'equacions de manera cada vegada més precisa, podríem augmentar indefinidament el límit temporal de validesa de la predicció. La realitat és ben diferent: el progrés en els aspectes esmentats sens dubte millorarà notablement la bondat de la predicció a curt termini i mitjà —és a dir, entre un i deu dies—, però més enllà de dues setmanes l'efecte serà totalment inapreciable. Seguint l'exemple proposat per Lorenz, i atès que les equacions que

expliquen el comportament de l'atmosfera no són lineals, podem triar una equació de segon grau de la forma

$$Y_{n+1} = a \times Y_n - Y_n^2 \tag{1}$$

i suposar que és aquesta l'equació de govern del nostre sistema atmosfèric. El valor de la constant a (3,750), en el terme lineal de l'equació, representarà el coneixement que tenim del sistema de govern i la dada y_1 (3,000) simbolitzarà l'estat inicial determinat a partir de les observacions. A la figura 1 s'han representat els valors y_i calculats a partir d'aquesta equació, que seran els estats futurs de la nostra atmosfera corresponents als instants i . Si repetim el procés partint d'un estat inicial lleugerament distint ($y_1 = 3,001$), amb el qual es vol fer palès l'error inherent a qualsevol observació meteorològica, l'equació [1] prediu (fig. 1) un estat futur del siste-

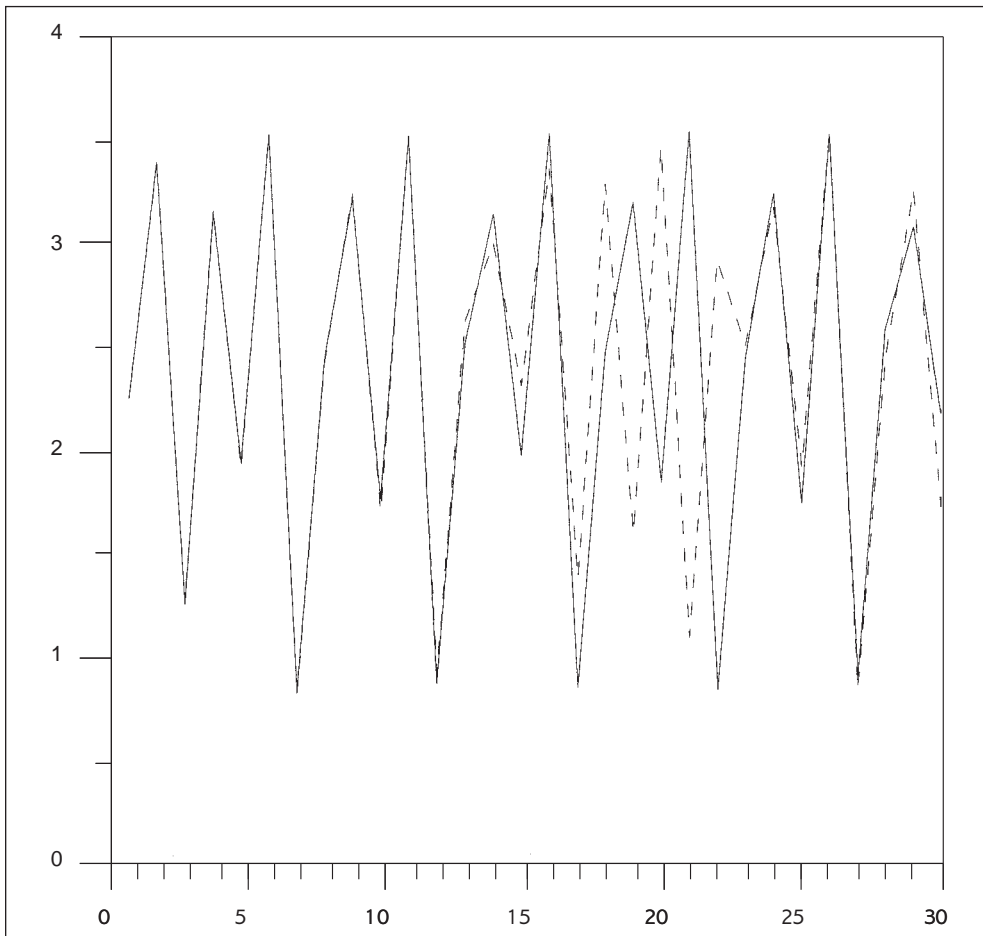


FIGURA 1. Solucions de l'equació [1] amb un valor inicial $y_1 = 3,000$ (línia contínua) i amb $y_1 = 3,001$ (línia discontinua).

ma que pràcticament coincideix amb el que s'havia obtingut per a les primeres catorze o setze iteracions; però, a partir d'aquest punt, les solucions difereixen de manera quasi aleatòria. Atès que l'equació no és lineal, la disminució de l'error en la dada inicial no assegura l'augment proporcional del nombre d'iteracions (termini de validesa de la predicció) en què les dues solucions són quasi coincidents.

Un resultat semblant es pot obtenir si es repeteix el càlcul prenent un valor de la constant a lleugerament diferent del veritable ($a = 3,751$); amb això es vol simbolitzar que el sistema d'equacions no és més que una representació matemàtica aproximada del sistema físic real. De nou (fig. 2) la solució correcta i l'obtinguda numèricament divergeixen a partir d'unes quantes iteracions.

Per al cas real de l'atmosfera, amb un sistema d'equacions primitives incomparablement més complex que l'il·lustrat aquí, hom ha trobat, com ja hem dit, que el límit de predictibilitat està al voltant de les dues setmanes.

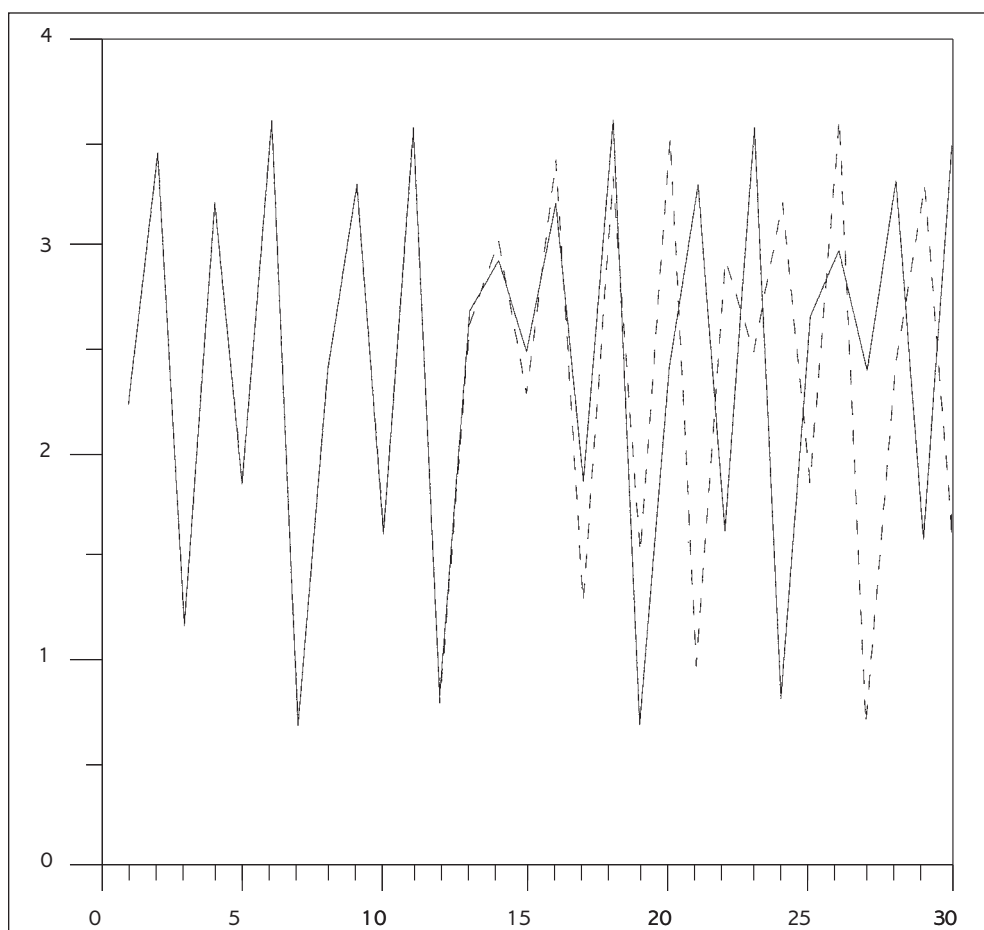


FIGURA 2. Solucions de l'equació [1] amb un valor de la constant $a = 3,750$ (línia contínua) i amb $a = 3,751$ (línia discontinua).

Prediccions estacionals

Ara, com en tantes ocasions, el coneixement de les limitacions ha servit perquè la ment enginyosa del científic aconseguís anar més enllà dels propis límits establerts. D'aquesta manera, el desenvolupament de la tècnica de predicció numèrica per a conjunts ha fet possible que en els últims cinc anys es vagin assajant prediccions estacionals (Palmer, 1999).

La PNT per a conjunts es basa en l'obtenció d'un cert nombre de pronòstics a partir d'unes condicions inicials lleugerament distintes, compatibles amb les observacions meteorològiques i obtingudes mitjançant la introducció de pertorbacions, convenientment escollides, a les dades inicials. Els primers pronòstics operacionals per conjunts comencen el 1992 i el mètode té dos objectius principals. El primer és proporcionar un pronòstic mitjà d'un conjunt que, en general, serà més aproximat que els pronòstics individuals més enllà dels primers tres dies, ja que els aspectes del pronòstic més incerts són eliminats per l'efecte del suavitzat en fer la mitjana. El segon objectiu, i més important, és proporcionar pronòstics amb una estimació de la seva confiança o probabilitat elaborant per a això, per exemple, mapes d'«espaguetis» que mostren la situació d'una mateixa isohipsa per a cadascun dels pronòstics obtinguts (fig. 3) o, també, pronòstics quantitativs probabilístics de precipitació acumulada. L'ús d'aquesta metodologia fa pos-

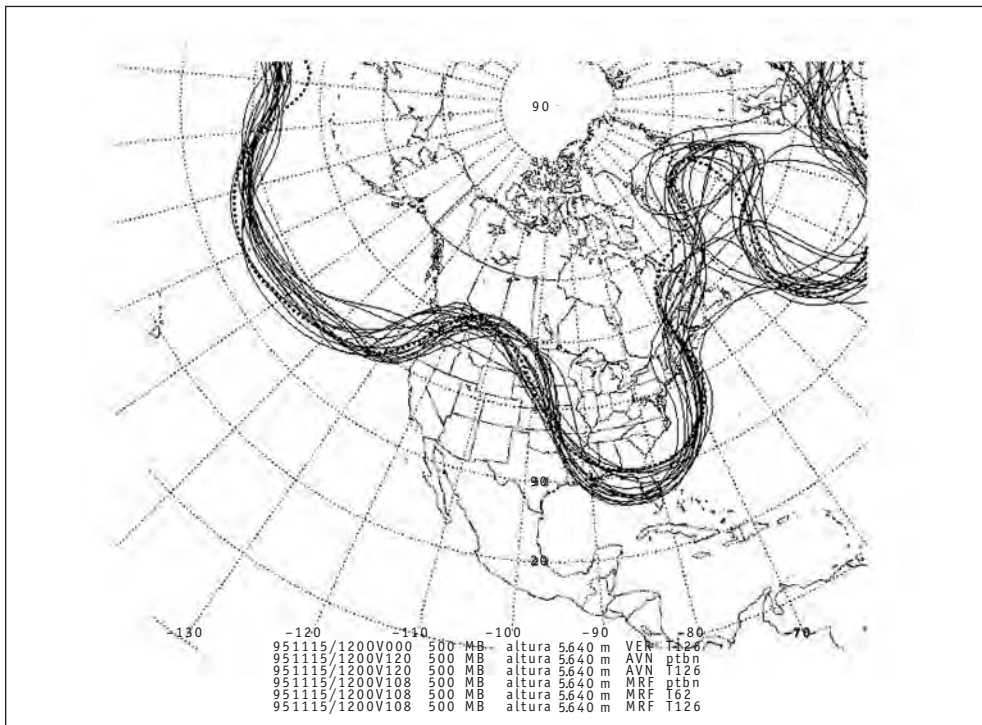


FIGURA 3. Pronòstic de les isohipses de 5,640 mgp a la superfície de 500 hPa obtinguda mitjançant la predicció per a conjunts (Kalnay, 1998).

sible, a més, la utilització d'un sistema d'observació adaptable que intensifiqui les observacions a les regions on la incertesa en les dades inicials produeix un major impacte sobre la variabilitat del pronòstic.

La predicció per a conjunts proporciona també l'eina fonamental per a estendre el pronòstic més enllà del límit de predictibilitat de Lorenz de dues setmanes. Normalment, els models numèrics de predicció utilitzen com a condicions de contorn dades que fan referència a propietats tant de la superfície terrestre com del mar. Si aquestes característiques són de variació lenta en comparació dels sistemes atmosfèrics que es volen predir, com és el cas, per exemple, de la temperatura superficial del mar o el tipus i l'extensió de la cobertura vegetal o del mantell de neu, aquestes dades es mantenen constants en els models de predicció durant un cert temps i són substituïdes periòdicament per uns altres valors que en tenen en compte la variació climàtica. No obstant això, a vegades ocorren anomalies que aparten aquestes variables geofísiques del seu valor mitjà climàtic de manera notable. Un dels fenòmens més importants per la magnitud i les conseqüències és sens dubte el corrent del Niño. La utilització de models de predicció numèrica acoblats atmosfera-ocèan ha posat de manifest la predictibilitat d'aquest fenomen amb una antelació superior a sis mesos, ja que es tracta d'un fenomen de gran període (de tres a set anys). L'obtenció d'un únic pronòstic del temps forçat amb una certa anomalia de la temperatura de la superfície del mar (TSM), com la corresponent al corrent del Niño, no seria útil més enllà de les dues setmanes, atès que la variabilitat imprevisible del temps podria ocultar les anomalies atmosfèriques originades per aquests forçaments de variació lenta. Ara bé, el valor mitjà d'un conjunt de nombrosos pronòstics obtinguts amb models atmosfèrics forçats per anomalies de la TSM (o per uns altres de variació lenta sobre la superfície terrestre, com poden ser l'extensió del mantell de neu o el contingut d'humitat del terra) permet filtrar les variacions imprevisibles del pronòstic i retenir la component forçada que es pot predir. Podem utilitzar l'equació [1] de Lorenz, lleugerament modificada, per a il·lustrar també l'essència del mètode descrit:

$$Y_{n+1} = a \times Y_n - b(n) \times Y_n^2. \quad [2]$$

Així, la consideració del forçament superficial de variació lenta (és a dir, l'anomalia de la TSM) seria representada per la variació temporal del coeficient del terme quadràtic de l'equació [2]. Si comparem els resultats anteriors amb els pronòstics obtinguts amb aquesta equació (fig. 4) per a un nombre considerable d'iteracions (cent en aquest exemple), hi podem observar que, eliminant la variabilitat pròpia del pronòstic diari, és possible quantificar la variació que produeix el forçament de variació lenta introduït a l'equació i que ha estat, en aquest cas, intencionadament sobredimensionat. La figura 5 mostra com a exemple el pronòstic de l'anomalia de la precipitació obtingut, a partir d'un conjunt de vint-i-set elements per a diferents períodes de tres mesos, pel Centre Europeu de Predicció a Mitjà Termini (Stockdale *et al.*, 1998).

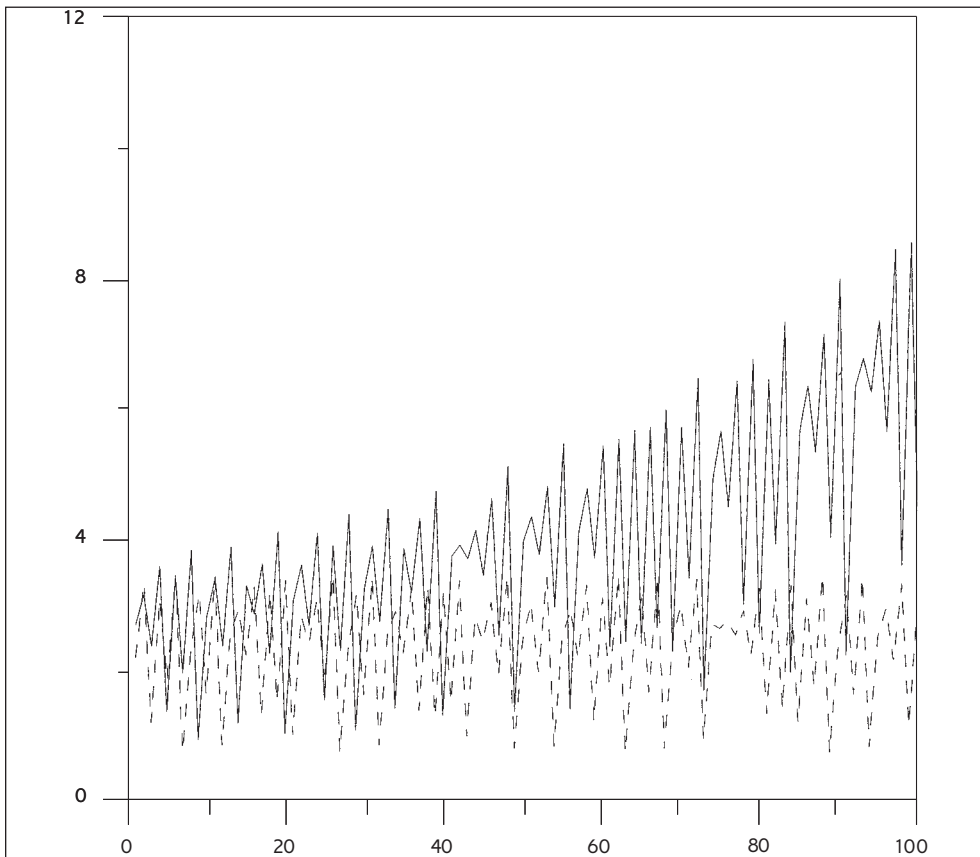


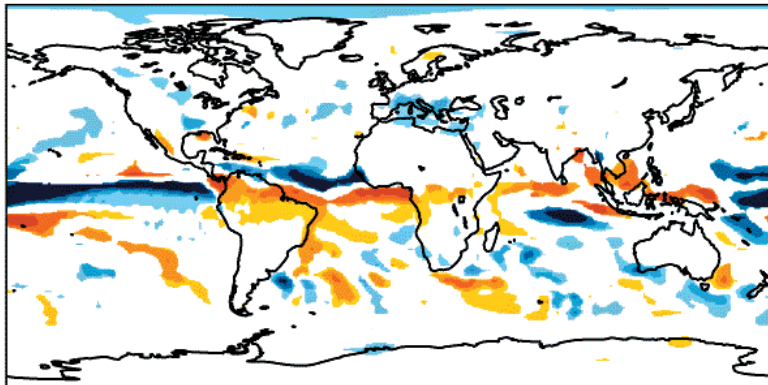
FIGURA 4. Solucions de l'equació [2] amb la introducció d'un terme (b) de forçament (línia contínua) i les corresponents a l'equació [1] (línia discontinua).

Hem vist, doncs, que les bases científiques de la PNT s'establiren durant els primers anys del segle passat, que el primer intent formal d'aplicació del nou mètode fou publicat a principis dels anys vint, que les primeres aplicacions operatives es van realitzar en la dècada dels anys cinquanta, que els treballs publicats durant les dècades dels setanta i dels vuitanta demostraren que la predictibilitat de l'atmosfera no podria superar les dues setmanes i que en els últims anys del segle xx s'apliquen models numèrics de predicció del temps per a pronosticar, amb mesos d'antelació, anomalies climàtiques d'algunes variables meteorològiques com la temperatura o la pluja.

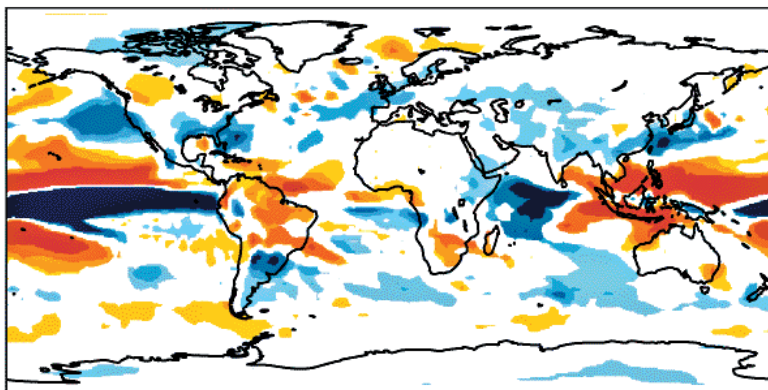
El futur de la predicció numèrica del temps

Sens dubte, la predicció numèrica del temps continuarà millorant per a fer cada vegada més precisos i incertats els pronòstics meteorològics. Com és lògic, això serà degut a l'augment de

a) Juny-agost de 1997



b) Desembre de 1997 - febrer de 1998



c) Març-maig de 1998

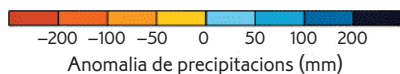
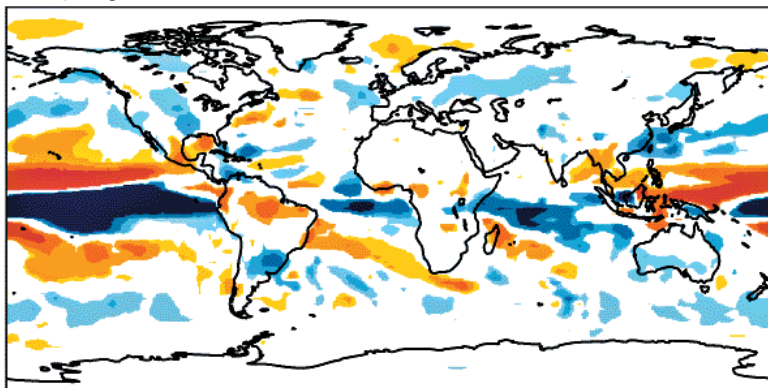


FIGURA 5. Pronòstic de l'anomalia de la precipitació per a diferents períodes obtingut a partir d'un conjunt de vint-i-set elements al Centre Europeu de Predicció a Mitjà Termini (Stockdale, 1998).

la capacitat, la potència i la rapidesa dels ordinadors, la millora dels sistemes d'observació i el perfeccionament dels models numèrics.

L'observació meteorològica millorarà amb la determinació del vent als oceans gràcies als radars dels satèl·lits. L'augment selectiu d'estacions de radiosondatge mòbils millorarà també la determinació del vent, la temperatura i la humitat a regions de la Terra on aquestes observacions encara són molt escasses. Les mesures realitzades amb GPS permetran determinar amb més gran resolució i exactitud el vapor d'aigua a l'atmosfera, de gran importància per al pronòstic a molt curt termini de la precipitació.

Els models numèrics representaran de manera cada vegada més acurada els fenòmens d'escala més petita, com són la turbulència i els processos de la microfísica dels núvols, la qual cosa millorarà especialment els pronòstics a curt termini. L'augment de les prestacions dels ordinadors s'utilitzarà potenciant la predicció per a conjunts, augmentant el nombre d'elements amb diferents condicions inicials i utilitzant conjunts obtinguts per diferents models numèrics; això beneficiarà fonamentalment el pronòstic meteorològic a mitjà termini i llarg.

La predicció numèrica a Catalunya

La predicció numèrica a Catalunya té una història molt més curta i recent, però, no obstant això, se'n poden destacar certs aspectes en els quals ha estat ja pionera. Poc temps després de la creació del Centre de Supercomputació de Catalunya (CESCA), els seus responsables van mostrar interès a utilitzar el recentment adquirit superordinador CRAY YMP en la PNT. A finals de 1993, la Fundació Catalana per a la Recerca, l'empresa CLABSA i el Departament d'Astronomia i Meteorologia de la Universitat de Barcelona signaren un conveni de col·laboració amb l'objectiu de posar a punt un model numèric de predicció del temps per a Catalunya.

A principis de 1994 es va instal·lar al CRAY del CESCA la versió 2.0 del model MASS (Mesoscale Atmospheric Simulation System), desenvolupat per l'empresa MESO dels Estats Units. Es tractava d'un model de predicció numèrica de mesoscala hidrostàtic que s'aplicava sobre una àrea sinòptica d'uns 3.000 km de base d'Europa occidental. El model utilitzava una malla de 55×55 punts, amb un braç d'uns 55 km aproximadament i 20 nivells a la vertical. Les prediccions a trenta-sis hores de les variables meteorològiques s'utilitzaven llavors com a condicions de contorn per a la simulació a escala més petita sobre Catalunya i el llevant peninsular (fig. 6). El model va mostrar aviat la seva utilitat en la predicció de situacions de mesoscala, ja que va reproduir amb èxit considerable les principals característiques dinàmiques d'un sistema convectiu de mesoscala que afectà Catalunya els dies 9 i 10 d'octubre de 1994 (Codiña *et al.*, 1997).

Un nou conveni de col·laboració entre aquestes institucions féu possible l'aplicació operativa del model MASS des de principis de 1996. Els mapes pronosticats pel model es difonien diàriament a través de la pàgina web InfoMet (<http://www.infomet.fcr.es/>) (creada també al De-

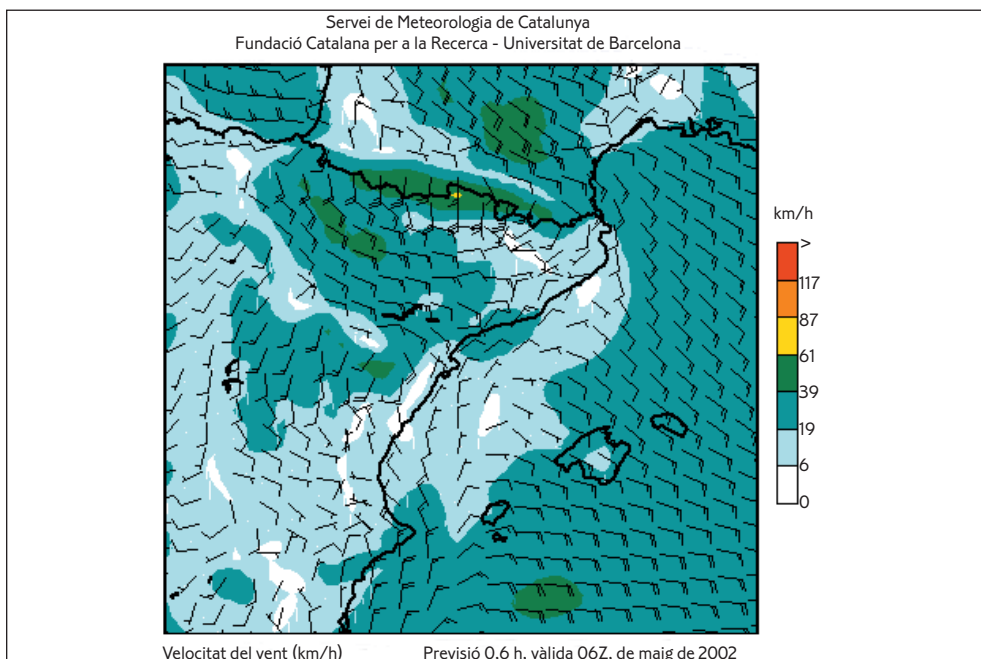


FIGURA 6. Pronòstic de la velocitat del vent obtingut pel model MASS.

partament), de manera que podien ser consultats lliurement per professionals, mitjans de comunicació i públic en general. El projecte fou pioner, ja que era la primera universitat europea que elaborava i difonia prediccions meteorològiques obtingudes amb un model numèric de mesoscala.

Actualment, el model MASS continua operatiu amb dues execucions diàries (a les 00 h i a les 12 h UTC) que duu a terme el Servei de Meteorologia de Catalunya. A més, els pronòstics de vent a la superfície obtinguts pel model s'utilitzen per a la inicialització del sistema de predicció d'onatge per a la costa de Catalunya que va posant a punt el Laboratori d'Enginyeria Marítima de la Universitat Politècnica de Catalunya (Ponce de León *et al.*, 2000). Pel que fa a la recerca, el model s'ha utilitzat per al desenvolupament de la metodologia necessària per a l'aprofitament de les dades GPS en la determinació del contingut de vapor d'aigua a l'atmosfera (Cucurull, 2001) i per a l'anàlisi de la cobertura radar a Catalunya amb la finalitat de millorar la interpretació i l'aprofitament de les dades dels radars meteorològics. Actualment, el model s'utilitza en el marc del projecte CARPE DIEM (*Critical assessment of available radar precipitation estimation techniques and development of innovative approaches for environmental management*) finançat per la Unió Europea. Així, doncs, la utilització en aquests últims anys de models numèrics per a la predicció del temps a Catalunya ha donat com a resultat la formació d'un grup de joves científics especialitzats en aquestes tècniques que permet mirar cap al futur amb un optimisme fonamentat.

Bibliografia

- BJERKNES, V. (1904). «Das Problem von der Wettervorhersage betrachtet vom Standpunkt der Mechanik und der Physik». *Meteor. Zeitsch.*, núm. 21, p. 1-7.
- CHARNEY, J. G.; FJÖRTOFT, R.; NEUMANN, J. von (1950). «Numerical integration of the barotropic vorticity equation». *Tellus*, vol. 2, núm. 4, p. 237-254.
- CODINA, B.; ARAN, M.; YOUNG, S.; REDAÑO, A. (1997). «Prediction of a mesoscale convective system over Catalonia (northeastern Spain) with a nested numerical model». *Meteorol. Atmos. Phys.*, núm. 62, p. 9-22.
- CUCURULL, L. (2001). *The use of global navigation satellite systems (GNSS) signals in numerical weather prediction (NWP)*. Barcelona: Universitat de Barcelona. [Tesi doctoral]
- KALNAY, E.; LORD, S. J.; MCPHERSON, R. D. (1998). «Maturity of operational numerical weather prediction: medium range». *Bull. Amer. Meteor. Soc.*, vol. 79, núm. 12, p. 2753-2769.
- LORENZ, E. N. (1963). «Deterministic nonperiodic flow». *J. Atmos. Sci.*, núm. 20, p. 130-141.
- (1964). «The problem of deducing the climate from the governing equation». *Tellus*, núm. 16, p. 1-11.
- (1965). «A study of the predictability of a 28-variable atmospheric model». *Tellus*, núm. 17, p. 321-333.
- (1984). «Some aspects of atmospheric predictability». A: *Problems and prospects in long and medium range weather forecasting*. Nova York: Springer-Verlag, p. 1-20.
- PALMER, T. N. (1999). «The scientific basis for seasonal prediction. Workshop proceedings». ECMRWF.
- PONCE DE LEÓN, S.; SÁNCHEZ-ARCILLA, A.; GÓMEZ, J.; SOSPEDRA, Q.; LORAIN, M. (2000). *Wind wave generation and prediction along the Catalan (Spanish Mediterranean) coast*. A: *27th International Conference on Coastal Engineering*. Sydney, 16-21, July.
- STOCKDALE, T. N.; ANDERSON, D. L. T.; ALVES, J. O. S.; BALMASEDA, M. A. (1998). «Global seasonal rainfall forecast using a coupled ocean-atmosphere model». *Nature*, núm. 392, p. 370-373.
- WIIN-NIELSEN, A. (1991). «The birth of numerical weather prediction». *Tellus*, núm. 43, sèrie AB, p. 36-52.

The physics of earthquakes, seismic occurrence and its prediction

Vladimir G. Kossobokov

International Institute of Earthquake Prediction Theory and Mathematical Geophysics,
Russian Academy of Sciences
Institut de Physique du Globe de Paris

Abstract

The observed physical parameters related to an earthquake result from a number of assumptions, heuristic extrapolations and other decisions made in course solution of a multitude of inverse problems and, therefore, an observable may be biased systematically or basically in error. Both experimental and mathematical modelings of earthquakes face the same problem, so that even impressive advances in these fields may be totally irrelevant to earthquakes. Despite the difficulty present day seismology delivers a number of confident predictions including intermediate-term medium-range earthquake forecasting, which have practical consequences for mitigation the effects of catastrophic earthquakes. Although of limited accuracy they may help reduce a substantial part of losses during a large catastrophic event by prompt escalation and de-escalation of realistic low-cost intermediate-term measures of civil defense type.

The high statistical significance of intermediate-term earthquake prediction methods based on inverse cascading of seismic activity achieved in course on-going real-time experiment confirms the following paradigms:

- Precursory seismicity patterns exist.
- The dimension of an area where precursory seismicity pattern appear is by far (about ten times) larger than that of source zone of the incipient large earthquake.
- Many precursory seismicity patterns are similar in regions of fundamentally different seismic and tectonic environment.
- Some precursory seismicity patterns are universal. Their analogues take place in advance extreme catastrophic events in other complex nonlinear systems.

Resum

Els paràmetres físics observats relatiu a un terratrèmol són el resultat de diverses assumpcions, extrapolacions heurístiques i algunes decisions més preses durant el procés de

solució de multitud de problemes inversos; així, doncs, un paràmetre observable pot quedar sistemàticament distorsionat o constituir bàsicament un error.

Tant el modelatge matemàtic de terratrèmols com el modelatge experimental s'enfronten amb el mateix problema, de manera que fins i tot els avenços més espectaculars en aquests camps poden resultar totalment irrelevants pel que fa als terratrèmols. Malgrat les dificultats, la sismologia actual proporciona prediccions fiables, incloent-ne d'abast mitjà i a mitjà termini, les quals tenen conseqüències pràctiques en la mitigació dels efectes dels terratrèmols catastròfics. Tot i la seva precisió limitada, poden ajudar a reduir una part substancial de les pèrdues en cas d'acomplir-se un esdeveniment catastròfic de gran abast, indicant l'increment i la minva de les mesures de protecció civil realistes a curt termini i de baix cost que cal adoptar.

La significació estadísticament elevada dels mètodes predictius de terratrèmols a mitjà termini basats en la cascada inversa de l'activitat sísmica i assolits durant l'experimentació en curs en temps real confirma els paradigmes següents:

- Existeixen patrons preliminars de sismicitat.
- Les dimensions d'una àrea on apareix un patró preliminar de sismicitat són molt més extenses (si fa no fa, deu cops) que les de la zona font d'un incipient gran terratrèmol.
- Hi ha nombrosos patrons preliminars de sismicitat que a regions amb un medi sísmic i tectònic fonamentalment distint resulten semblants.
- Alguns patrons preliminars de sismicitat són universals. Els seus anàlegs tenen lloc amb anterioritat a esdeveniments catastròfics extrems en uns altres sistemes complexos no lineals.

Introduction

Earthquakes are sudden fractures of the Earth's crust that radiate seismic waves and cause ground shaking. The extreme catastrophic nature of earthquakes is known for centuries due to resulted devastation in many of them. The abruptness along with apparent irregularity and infrequency of earthquake occurrences facilitate formation of a common perception that earthquakes are random unpredictable phenomena. The challenging questions remain pressing: What happens during an earthquake? How to size earthquakes? Why, Where and When do earthquakes occur? The basic difficulty in answering these questions comes from the fact that no earthquake has been ever observed directly. Therefore, the most important tasks in studying earthquakes at present is obtaining observational constrains from systematic analysis of the geologic, geodetic, gravity, heat-flow, petrochemical, geomagnetic, seismic and other data available and linking them together in plausible models and interpretations.

It would be misleading to pretend that the state-of-the-art Physics of earthquakes is a well-developed branch of Science, which deserves including into school programs and textbooks (although a basic knowledge about earthquakes in seismic regions, in particular, help avoiding losses of life and reduce physical damage). To the contrary most of seismologists clearly

understand the pioneering and, therefore, juvenile nature of the present day physical problems related to earthquakes (H. Kanamori & E. H. Brodsky, 2001. "The physics of earthquakes". *Physics Today* [June 2001], 34-40). The mature wisdom of any science is determined by its ability to predict phenomena in study. This paper intends to outline what can be predicted about a "ground shaking" caused by tectonic forces? and to discuss the uncertainties of such predictions either those obtained after the fact or in advance of an earthquake.

1. What are earthquakes and how to size them?

Fracturing of the Earth's crust that radiate seismic waves and cause ground shaking produces earthquakes. A ground shaking caused by other than tectonic sources (e.g., by an explosion) is disregarded here as an earthquake although these may be the basis for important studies and experiments providing conclusions about the Earth structure. Some historical records on earthquakes are known from 2100 B.C. However, most of them before the middle of the 18th century are generally lacking description or are not reliable.

In 1870's the English geologist John Milne designed a forerunner of modern seismographs: A simple pendulum and a needle suspended above a smoked-glass plate allowed to distinguish primary and secondary earthquake waves and, basing on their timing, to derive an accurate statement about location of an earthquake source. The Russian Prince Boris Golitzyn, who improved similar instruments of the 1890's, invented the modern seismograph in the early 20th century.

Figure 1 shows seismologists from Chinese Seismological Bureau visiting the epicenter zone of the November 14, 2001, Kokoxili Earthquake that happened along the Kunlun fault in Tibet.



FIGURE 1

The photograph (Xinhua/China News Agency) displays clearly the surface rupture of the earthquake, which appears to be a strike-slip with several meters of sinister slip and eventually more than 250 km long. The actual location, mechanism of faulting and size of rupturing are in a good agreement with quick post-the-fact predictions computed by many national and international seismological centers right after seismographs recorded the shaking.

Of course, a tiny proportion of registered earthquakes attract the interest of seismological community so that a special field trip or a temporary seismograph network be deployed to verify exact location and other parameters of an earthquake. Sometimes such a field trip outcomes with a failure after no traces of the earthquake are found in the predicted area either due to natural uncertainty of calculations or due to the depth of the quake. Most of the events from numerous catalogs of earthquakes have the parameters listed with uncertainties of seconds for the origin time, of kilometers for the hypocenter (this may vary much depending on locality and seismograph network), and of up to a decimal order for energy of an earthquake. Such an uncertainty comes from unavoidable natural obstacles of which the evident heterogeneity of the Earth's crust plays the main role.

In the 1930's Charles F. Richter, a California seismologist, introduced the concept of earthquake *magnitude*. His original definition held only for California earthquakes occurring within 600 km of a particular type of seismograph (the Woods-Anderson torsion instrument). Richter's original magnitude scale (M_L) was then extended to observations of earthquakes of any distance and of focal depths ranging between 0 and 700 km (Lay & Wallace, 1993). There are many magnitude scales, which are another source of a controversy about earthquakes. Reporters often mix the *Modified Mercalli Intensity Scale* (MM), which is a measure of the effects of an earthquake at a particular place and depends not only upon the size of the earthquake but also upon the distance from the epicenter and the local geology, with other *magnitudes*, each of which estimates a logarithmic value of the earthquake size from seismograms. Because earthquakes excite both body waves, which travel into and through the Earth, and surface waves, which are constrained to follow the Earth's uppermost layers, two basic magnitude scales evolved —the mb and M_S .

The standard body-wave magnitude formula is

$$mb = \log_{10}(A/T) + Q(D,h) ,$$

where A is the amplitude of ground motion; T is the corresponding period; and $Q(D,h)$ is an empirical function of distance, D , between epicenter and station and focal depth, h .

The standard surface-wave formula is

$$M_S = \log_{10} (A/T) + 1.66 \log_{10} (D) + 3.30 .$$

There is a belief, that a novel extension of the magnitude scale, known as *moment magnitude*, or M_W , is more uniformly applicable and gives, for very large earthquakes the most

reliable estimate of earthquake size. Indeed, the seismic moment is related to fundamental parameters of the faulting process:

$$M_0 = \mu S \langle d \rangle,$$

where μ is the shear strength of the faulted rock, S is the area of the fault, and $\langle d \rangle$ is the average displacement on the fault. However, these parameters are determined from waveform analysis of the seismograms, which parameters have the same uncertainties that are used in multiple station determinations of magnitudes. Therefore, the magnitude scale M_W is hardly a better estimate of the earthquake size than either of M_S or mb .

The standard seismic moment magnitude formula is:

$$M_W = 2/3 \log_{10} (M_0) - 10.7 .$$

The largest reported moments are 2.5×10^{30} dyn \times cm for the 1960 Chile earthquake ($M_S 8.5$; $M_W 9.6$) and 7.5×10^{29} dyn \times cm for the 1964 Alaska earthquake ($M_S 8.3$; $M_W 9.2$).

Let us illustrate the difficulty and controversy of magnitude determination of a particular earthquake with the most recent example (as on March 15). The red thumbnail on the official website of the U.S. National Earthquake Information Center reports the March 5, 2002, Magnitude 7.5 Earthquake on Mindanao, Philippines. This is M_W estimate. In this range M_W and M_S are supposed to match, although the NEIC average $M_S = 7.2$. Figure 2a displays the station M_S as a function of epicenter distance to the earthquake for 89 stations worldwide and

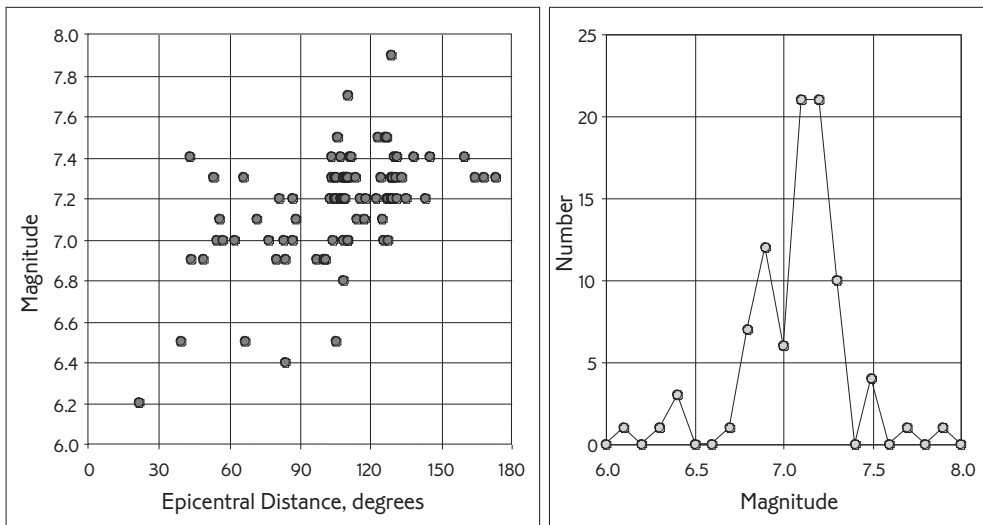


FIGURE 2

a clear tendency of these estimations to increase with the distance. It is not supposed to be so by definition of M_S scale, which is normalized theoretically by this very distance! Moreover, Figure 2b indicates two modes in the station magnitude distribution (attributing much sense to averaging of bimodal distributions are not welcomed in course statistical analysis), which fit well the epicenter and antipodal hemisphere estimates of 6.9 and 7.2. Finally, note that only six out of 89 stations reported 7.5 or more for the Mindanao earthquake, listed as 7.5 by NEIC.

Despite the uncertainties in determination of earthquake parameters some relations between them are beyond any doubt. One of such is so-called *Gutenberg-Richter frequency-magnitude relation* (Gutenberg & Richter, 1954): Averaged over a large territory and time the number of earthquakes equal or above certain magnitude, $N(M)$ scales as:

$$\log_{10} N(M) = a + b \times (8 - M) ,$$

where, for M being M_S scale, constant b is generally in the range $0.8 < b < 1.2$ (Frohlich & Davis, 1993) but varies from region to region. The constant a is a measure of the regional level of seismic activity. This law seems universal in the realm of multiple fracturing, where it is observed in a broad variety of conditions from laboratory samples of solid materials through geo-technical and engineering constructions to the lithosphere of the Earth (Barenblatt, 1982; Turcotte, 1995) and, perhaps, to extreme energies of “starquakes” (Kossobokov, Keilis-Borok & Cheng, 2000). Observations favor also the hypothesis that smaller earthquakes in moderate-sized regions occur at rates that are only weakly dependent on time. Thus, the rate of occurrence of smaller earthquakes can be extrapolated to assess the hazard of larger earthquakes in a region.

2. Where earthquakes occur?

Figure 3 shows the earthquake-prone pattern of the Earth as it is observed on a global scale. The plate-tectonics hypothesis explains its evident stability as basic concentration of earthquakes at plate boundaries, although a significant numbers of earthquakes occur within plate interiors. The Earth’s crust is extremely complex and faults and earthquakes in a region occur on a wide range of scale. There is considerable evidence that faults and earthquakes interact on a range of scales, from thousands of kilometers to millimeters or less. Evidence in support of this hypothesis comes the universal validity of scaling relations. Since the b -values for the two regions are about equal, this ratio is also applicable to other earthquake magnitudes. This extrapolation is the basis of our global seismic hazards assessment.

Figure 4 present a global map of the seismic intensity factor (Kossobokov *et al.*, 2000), which is based on the epicenters of earthquakes with magnitudes $m \times 4$, where the data has been obtained from the NEIC Global Hypocenters Data Base (1989) for the period 1964-1995. Dividing the surface of the earth into $1^\circ \times 1^\circ$ regions, the seismic intensity factor, I_4 , is defined to

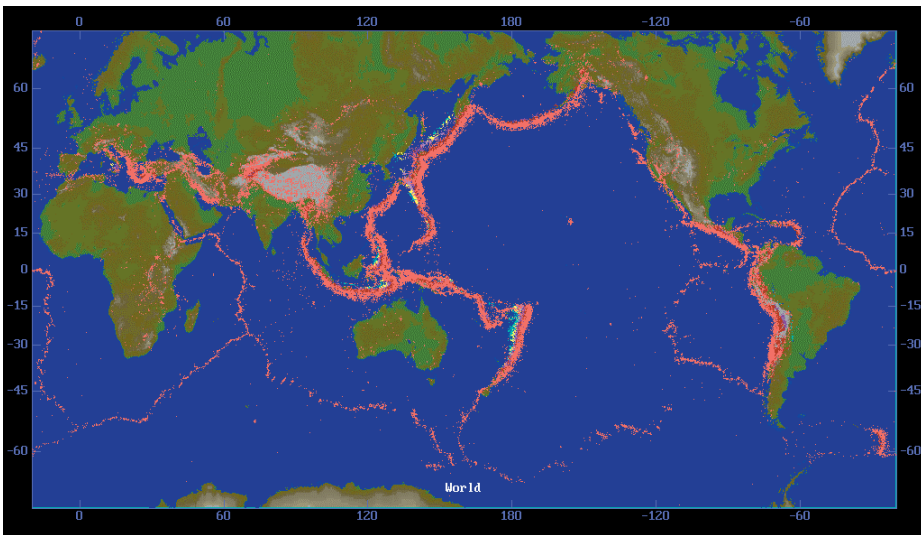


FIGURE 3

be the number of earthquakes per year with magnitudes $m \times 4$ in each region. A correction is made for the change in area with latitude.

The boundaries of the tectonic plates are clearly defined in Figures 3 and 4. Seismic activity is particularly intense in subduction zones (i.e., the ring of fire around the Pacific) as expected. A broad zone of seismic activity extends from southern Europe to southeastern Asia, which is associated with continent/continent collision zones between the Eurasian plate and the African, Arabian, and Indian plates. The minimum value of the seismic intensity factor in our compilation is $I_4 = 1/32 \text{ yr}^{-1}$, one magnitude 4 or larger earthquake in the 32 year period considered. The maximum value is about $I_4 = 40 \text{ yr}^{-1}$.

The range of I_4 is over three orders of magnitude. Using the Gutenberg-Richter relation the number of earthquakes per year in a $1^\circ \times 1^\circ$ area with magnitude greater than m is related

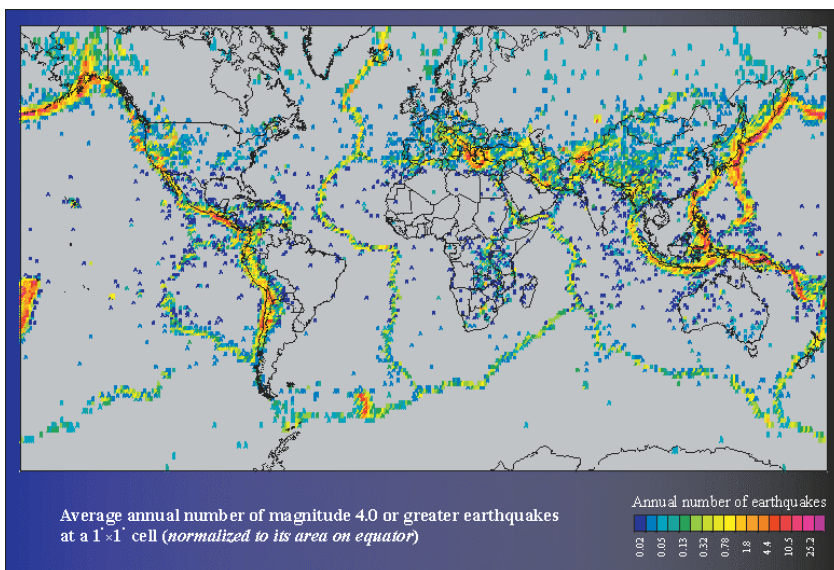


FIGURE 4

to I_4 . For example, in a block with $I_4 = 1 \text{ yr}^{-1}$, an earthquake with magnitude $m = 6$ has a return period of 63 years and one with $m = 8$ has a return period of 4,000 years.

Other seismic hazard maps have been prepared for many parts of the world. For the United States, the U. S. Geological Survey's June 1996 maps (Frankel *et al.*, 1996) are available on the Internet. Seismic hazard maps generally give the probability of exceeding specified peak ground acceleration during a specified time interval. The data used to formulate the maps include the historic seismicity, maps of active faults, and levels of seismic attenuation.

The hazard assessment maps are reasonably consistent with the seismic intensity maps, although the hazard maps apparently overestimate the seismic risk in aseismic zones due to the arbitrary choice of the *probabilistic space* and improper rescaling that ignore the local *fractal dimensionality of the earthquake-prone support* (Kossobokov & Mazhkenov, 1994). An objective knowledge about earthquake probabilistic space cannot be formed without bookkeeping of seismic record for extensive period of time and comprehensive statistical analysis of earthquake catalogs.

3. How to get information on earthquakes?

The known Internet type connections where original seismic data or seismic research information is available on-line could be found at "Surfing the Internet for Earthquake Data" provided by S. Malone at <http://www.geophys.washington.edu/seismosurfing.html>.

The complete SeismoSurfing index is also being mirrored for European users by ETH, Zurich at <http://seismo.ethz.ch/seismosurf/seismobig.html>.

Figure 5 shows the annual number of earthquakes from the NEIC Global Hypocenters Data Base and its updates through the 20th century. This frequency-magnitude graph demonstrates temporal variations of the global seismic activity. One can observe several "historic changes" of which the most dramatic reflects deployment of the World Wide Seismic Standard Seismograph Network in 1963. From a statistical viewpoint, since about that time the catalog appears to be surprisingly consistent in reporting magnitude 5.0 and above earthquakes: In logarithmic scale, the magnitude bands have almost the same width in agreement with the Gutenberg-Richter relationship. One can see also that the list of earthquakes above 7.0 in the NEIC GHDB is probably complete from the beginning of the century. Such a remarkable stability of the annual number of earthquakes suggests the global underlying processes rather stationary in the time scale of decades and encourages research aimed at intermediate-term prediction of earthquakes.

On the other hand, it is of common knowledge that all catalogs have some errors, which may render invalid conclusions derived in a study based on a catalog only. The errors in the data, as in general, may be neutralized in two ways:

- By postponing the analysis until the data are refined.
- By a robust analysis of existing data within the limits of its applicability.

The danger in pursuing the first way only was long ago indicated by a founder of non-linear mechanics and computational methods Professor N. Krylov: “Undue precision of computations is the first symptom of mathematical illiteracy.” People who earn their living by giving advice to the public on important economic questions do not have the luxury of waiting for decades before making a decision based on the available evidence. The forecasts and predictions made by seismologists already have a large economic impact. As the methods of forecasting improve along with seismological data these economic effects will increase.

What can we learn from a catalog of earthquakes? There are two extreme opinions on the subject:

Pessimistic (Habermann & Creamer, 1994): “[...] in the case of seismic data, most of the observed variations are, in fact, related to changes in the system for detecting and reporting earthquakes and not to actual changes in the Earth.”

Optimistic: Among existing data seismic catalogs remain the most reliable record on distribution of earthquakes in space and time.

4. Are earthquakes predictable?

Yes.

Figures 3 and 4 show a high level of the spatial predictability of earthquakes, while Figure 5 predict the annual number of magnitude 5 or larger earthquakes in the range from 90 to 200. The history, 1985-present, of the on-going real-time monitoring of the global seismic activity aimed at intermediate-term medium-range prediction of the largest earthquakes

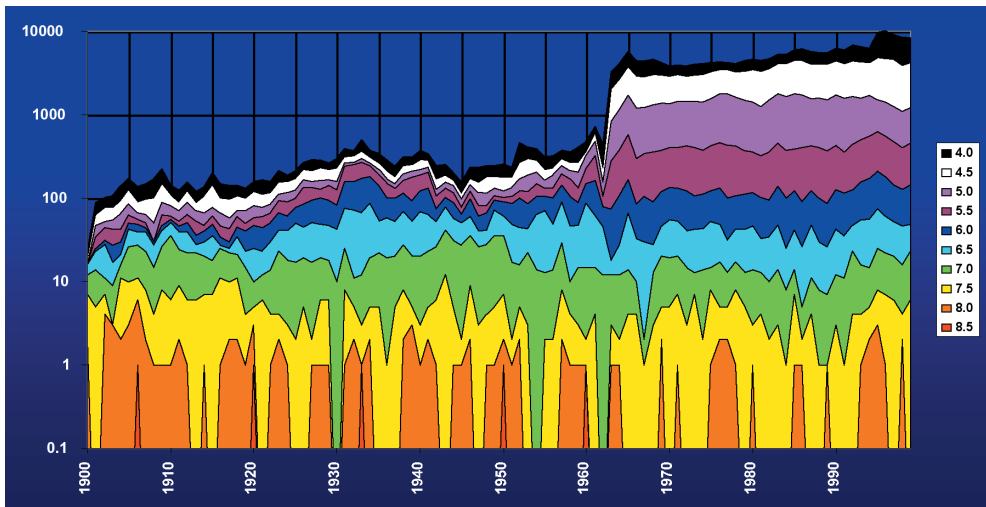


FIGURE 5

(<http://www.mitp.ru> or <http://www.phys.ualberta.ca/mirrors/mitp>) has proved the high statistical significance of the two methods, algorithms M8 (Keilis-Borok & Kossobokov, 1990) and MSc (Kossobokov, Keilis-Borok & Smith, 1990), based on inverse cascading of seismic activity to a catastrophe.

The temporal predictability of large earthquake occurrences requires a special comment on the recently revived discussions (Geller *et al.*, 1997). No current theory of dynamics of seismic activity can answer this question. Inevitably, a negative statement that asserts a non-trivial limitation on predictability is merely a conjecture. On the other hand, forward testing of a reproducible prediction method and, so far, in no other way, can unequivocally establish a certain degree of predictability of earthquakes. The results just mentioned and described below did confirm a positive statement on predictability of earthquakes at the intermediate-term scale (despite dramatic changes both in seismicity and in compilation of the earthquake catalog during the test period). Furthermore, it appears that premonitory activation evolves through long-, intermediate-, short-, and immediate-term or even nucleation (Ellsworth & Beroza, 1995) phases.

Let us first clarify what is earthquake prediction? The United States National Research Council, Panel on Earthquake Prediction of the Committee on Seismology suggested the following definition (1976, p. 7):

An earthquake prediction must specify the expected magnitude range, the geographical area within which it will occur, and the time interval within which it will happen with sufficient precision so that the ultimate success or failure of the prediction can readily be judged. Only by careful recording and analysis of failures as well as successes can the eventual success of the total effort be evaluated and future directions charted.

Rethinking Earthquake Prediction, Lynn Sykes *et al.* (1999), write:

The public perception in many countries and, in fact, that of many earth scientists is that earthquake prediction means short-term prediction, a warning of hours to days. They typically equate a successful prediction with one that is 100 % reliable. This is in the classical tradition of the oracle. Expectations and preparations to make a short-term prediction of a great earthquake in the Tokai region of Japan have this flavor. We ask instead are there any time, spatial and physical characteristics inherent in the earthquake process that might lead to other modes of prediction and what steps might be taken in response to such predictions to reduce losses?

Following common perception many investigators usually overlook as well spatial modes of predictions and concentrate their efforts on predicting the “exact” fault segment to rupture (e.g. Parkfield earthquake prediction experiment), which is by far a more difficult and might be

an unsolvable problem. Being related to the rupture size $L = L(M)$ of the incipient earthquake of magnitude M , such modes could be summarized in a classification of location of a source zone from a wider prediction ranges (Table 1).

<i>Temporal, in years</i>		<i>Spatial, in source zone size L</i>	
Long-term	10	Long-range	Up to 100
Intermediate-term	1	Medium-range	5-10
Short-term	0.01-0.1	Narrow	2-3
Immediate	0.001	Exact	1

TABLE 1. Classification of earthquake predictions.

From a viewpoint of such a classification, the earthquake prediction problem is naturally approached by a hierarchical, step-by-step prediction technique, which accounts for multi-scale escalation of seismic activity to the main rupture (Keilis-Borok, 1990).

Citing Christopher Scholz (1997):

Predicting earthquakes is as easy as one-two-three. Step 1: Deploy your precursor detection instruments at the site of the coming earthquake. Step 2: Detect and recognize the precursors. Step 3: Get all your colleagues to agree and then publicly predict the earthquake through approved channels.

No need to remind that some “precursor detection instruments” are already deployed worldwide, e.g., routine seismological data bases such as the US GS/NEIC Global Hypocenter Data Base, and their record available for general use. Some “precursors” are already detected, e.g. reproducible intermediate-term algorithms such as the M8 and MSc algorithms (Keilis-Borok & Kossobokov, 1990; Kossobokov, Keilis-Borok & Smith, 1990). And, finally, some earthquakes were already “publicly predicted” (Keilis-Borok *et al.*, 1990; Harris, 1998), although our colleagues continue to debate whether it is possible or not.

The on-going real-time monitoring of the global seismic activity aimed at intermediate-term medium-range prediction of the largest earthquakes has a long history now (Healy, Kossobokov & Dewey, 1992; Kossobokov *et al.*, 1999). Several largest earthquakes were predicted and some were missed. Table 2 gives the summary of the prediction outcomes for magnitude 8.0 or more earthquakes.

Test period	Large earthquakes			Percentage of alarms		Confidence level, %	
	Predicted by		Total	M8	M8-MSc	M8	M8-MSc
	M8	M8-MSc					
1985-2000	9	7	10	34.9	18.0	99.95	99.96
1992-2000	7	5	8	30.2	15.3	99.87	99.69

TABLE 2. Worldwide performance of earthquake prediction algorithms M8 and M8-MSc: Magnitude 8.0 or more.

Note: The confidence estimates use the most conservative measure of the alarm volume accounting for empirical distribution of epicenters (measure μ described below).

Are the results of the prediction experiment “good” or not? A statistical conclusion about that could be attributed in the following general way:

Let T and S be the total time and territory considered; A_t is the territory covered by the alarms at time t ; $\tau \times \mu$ is a measure on $T \times S$ (we consider here a direct product measure $\tau \times \mu$ reserving a general case of a time-space dependent measure ν for future more sophisticated null-hypotheses); N counts the total number of large earthquakes with $M \geq M_0$ within $T \times S$ and n counts how many of them are predicted. The time-space occupied by alarms, $A = \bigcup_{\tau} A_t$, in percentage to the total space-time considered equals

$$p = \int_A d(\tau \times \mu) / \int_{T \times S} d(\tau \times \mu).$$

By common definition the *statistical significance level* of the prediction results equals

$$\alpha = 1 - B(n-1, N, p),$$

where B is the cumulative binomial distribution function. The smaller is the significance level α , the larger is the *confidence level* $1-\alpha$ and the higher is the significance of the predictions under testing.

When testing temporal predictability of earthquakes it is natural to make the following choice of the product measure $\tau \times \mu$: Take the uniform measure τ , which corresponds to the Poisson, random recurrence of earthquakes and the measure μ proportional to spatial density of epicenters. Specifically, determine the measure μ of an area proportional to the number of epicenters of earthquakes from a sample catalog, for example, earthquakes above certain magnitude cutoff M_c . This empirical spatial measure of seismicity is much more adequate than the literal measure of territory in km^2 for estimating statistical significance of the prediction results (the last one equalizes the areas of high and low seismic activity, at the extreme, the areas where earthquake happen and do not happen).

The actual, empirical distribution of earthquake locations is the best present day knowledge estimate of where earthquakes may occur. The recipe of using α -measure and counting p is the following: Choose a sample catalog. Count how many events from the catalog are inside the territory considered; this will be your denominator. At a given time, count how many events from the catalog are inside the area of alarm; this will be your numerator. Integrate the ratio over the time of prediction experiment. This is the exact way of computing Percentage of alarms and Confidence level in Table 2. (All magnitude 4 or larger earthquakes, 1963-1984, from the NEIC Global Hypocenter Data Base formed the sample catalog.)

This simple recipe has a nice analogy that justifies using statistical tools available since Blaise Pascal (1623-1662):

Seismic Roulette: Consider a roulette wheel with as many sectors as the number of events in your sample catalog, a sector for each event. Make your bet according to prediction: determine, which events are inside area of alarm, and put one chip in each of the corresponding sectors. Nature turns the wheel.

If seismic roulette is not perfect, one can win systematically. The results of the on-going real-time global test of the intermediate-term prediction algorithms M8 and MSc did confirm it (Kossobokov *et al.*, 1999; Table 2).

Figure 6 shows the results of the analysis (as on January 1, 2002) aimed at prediction of the great magnitude 8.0 or larger earthquakes. The predictions apply to earthquake-prone

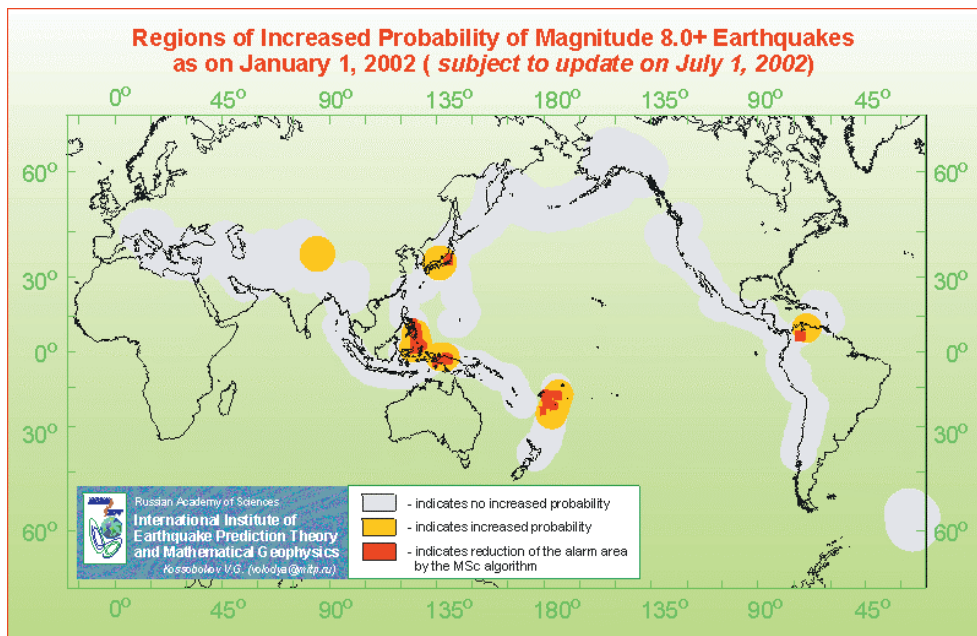


FIGURE 6

territories inside circles of investigation, each of the 667-km radii. Gray indicates the area under monitoring, where 80 % of the largest earthquakes on the Earth happen. The seismic activity in other areas is lower than that required in the prediction methods. Yellow indicates the areas of increased probability of magnitude 8 or more earthquakes and red indicates the areas of the highest probability of such earthquakes. The prediction is subject to update in July 2002. Similar analysis aimed at prediction of magnitude 7.5 or more is carried on in parallel since the formal declaration of the prediction experiment start in 1992 (Healy, Kossobokov & Dewey, 1992).

5. What stands behind M8 and MSc?

Both are reproducible earthquake prediction methods that make use of seismic activity reported in routine seismic catalogs. The M8 is applied first. It scans the territory in question for the areas in alarm, so-called Time of Increased Probability (TIP). The MSc is applied to reduce the area of alarm by analyzing dynamics at lower magnitude levels of seismic hierarchy. Sometimes, the data is enough to get a near-perfect outline of the incipient large earthquake. More often the catalog of earthquakes is exhausted already at the M8 analysis and the prediction remains in the medium range.

The M8 intermediate-term earthquake prediction algorithm was designed by retroactive analysis of dynamics of seismic activity preceding the greatest, magnitude 8.0 or more, earthquakes worldwide, hence its name. Its prototype (Keilis-Borok & Kossobokov, 1984) and the original version (Keilis-Borok & Kossobokov, 1987) were tested retroactively at recorded epicenters of earthquakes of magnitude 8.0 or greater from 1857-1983. The algorithm M8 uses traditional trailing description of a dynamical system adding to a common phase space of rate (i.e. number of mainshocks, N) and rate differential (i.e., deviation of N from a longer-term average, L) the dimensionless concentration (i.e., the average source size divided by the average distance between sources, Z) and a characteristic measure of clustering (i.e., maximum number of aftershocks, B).

The algorithm recognizes criterion, defined by extreme values of the phase space coordinates, as a vicinity of the system singularity. When a trajectory enters the criterion, probability of extreme event increases to the level sufficient for effective provision of a catastrophic event. The exact definitions and computer code of the M8 algorithm are published (Keilis-Borok & Kossobokov, 1990; Healy *et al.*, 1992; Kossobokov, 1997).

Retrospectively (Keilis-Borok & Kossobokov, 1990) the standard version of the algorithm was applied to predict earthquakes with magnitudes from above 8.0 to 4.9 in a number of regions worldwide. Its modified versions apply also in regions of seismic activity lower than required by the original version (Bhatia *et al.*, 1989; Kossobokov, Rastogi & Gaur, 1989; Latoussakis & Kossobokov, 1990; Gahalaut *et al.*, 1992; Romashkova *et al.*, 1998).

The second approximation prediction method MSc (Kossobokov, Keilis-Borok, Smith, 1990) was designed by retroactive analysis of the detailed regional seismic catalog prior to the Eureka earthquake (1980, $M = 7.2$) near Cape Mendocino in California, hence its name. Qualitatively, the MSc algorithm outlines such an area of the territory of the first approximation alarm where seismic activity is continuously high from the beginning of precursory inverse cascade and is infrequently interrupted for a short time. Such an interruption must have a sufficient temporal and/or spatial span. The phenomenon, which is used in the MSc algorithm, might reflect the second (possibly, shorter-term and, definitely, narrow-range) stage of the premonitory rise of seismic activity near the incipient source of main shock. MSc outperforms simple alternatives in reduction of territorial uncertainty of the M8 predictions, i.e. earthquake-prone cells and the most active of them that contain certain part of the recent seismic activity.

6. Seismic dynamics prior to and after the recent great earthquakes

The analysis of the great earthquakes since 1985 reveals the following (Romashkova & Kossobokov, 2001):

19/09/1985. Mexico Earthquake: Predicted both by M8 and MSc. Long lasting quiescence started in the middle of 1979 shortly interrupted with the 1981 $M_w = 7.4$ earthquake and smaller earthquake at the end of 1983. Sharp rise of activity started from the end of 1984 in a narrow vicinity of future epicenter.

20/10/1986. Kermadec Earthquake: Predicted both by M8 and MSc. Overall rise of seismic activity from the middle of 1984 extended over the whole area of M8 alarm, 667-km radius circle. Swarm of 24 $M_4 +$ earthquakes of which 11 were magnitude 5.0 or larger four month in advance the great shock.

23/05/1989. Macquarie Earthquake: Area of low seismic activity. Modified M8 predicts the occurrence of the great earthquake. Rather unusual earthquake of $M_{5.3}$ occurred two weeks in advance the great shock.

08/08/1993. Guam Earthquake: Predicted both by M8 and MSc. High activity from 1985 raised even more after a magnitude 7.5 earthquake with enormous number of aftershocks. Magnitude 6.6 event observed two months in advance and magnitude 5.1 event in the future epicenter zone 5 days in advance the great shock.

09/06/1994. Bolivia Deep Earthquake: The Great Deep Bolivia earthquake occurred under an aseismic area more than 600 km away from the Pacific coast. It was preceded by the January 10, 1994, magnitude 6.9, depth 595 km earthquake at distance of about 250 km. The previous earthquake that deep happened here in 1963. The epicenter is just out of test area and the M8 alarm.

04/10/1994. Shikotan Earthquake: Predicted both by M8 and MSc. The MSc reduced area fits almost exactly the aftershock zone of the earthquake. In August 1994, 5 weeks in advance

the great shock a swarm of 7 magnitude 6.0 + earthquakes occurred in about 200 km to the NE of its future epicenter.

07/04/1995. Samoa Earthquake: Predicted both by M8 and MSc. A swarm of 31 magnitude 4.0 + earthquakes, very unusual even in this very active area, happened at 250-km distance 3 months in advance the great shock, which itself did not produce notable aftershock series.

03/12/1995. Iturup Earthquake: Predicted both by M8 and MSc. Numerous foreshocks did occur prior to the great shock of December 3. This swarm-like series of 64 M4.0 + events started on November 24 with a magnitude 6.6 event and produced 19 magnitude 5.0 + and four magnitude 6.0 + events before the great rupture.

17/02/1996. New Guinea Earthquake: Predicted by M8 and missed by MSc reduced area. A remarkable rise of activity in 1994 some 400 km south of the future epicenter zone of the great earthquake did resolve temporarily with a series of 24 magnitude 4.0 + events, of which the largest two have magnitudes 6.1 and 7.1.

25/03/1998. Balleny Sea Earthquake: Area of low seismic activity. Modified M8 predicts the occurrence of the earthquake although its epicenter and aftershocks are outside seismic belt where no catalog reported seismic activity before 1998.

04/06/2000. South Sumatera Earthquake: Predicted both by M8 and MSc. Notable decline of activity in 1996-1997 changed with a significant rise in 1998-1999 in the future epicenter zone where the magnitude 4.7 foreshock happened less than a day before the great shock.

23/06/2001. Earthquake near coast of Peru: This earthquake is the only to-date failure-to-predict in M8-MSc testing aimed at magnitude 8.0 +. For a timely declaration of an alarm M8 algorithm was lacking rise and acceleration of activity.

14/11/2001. Kokoxili Earthquake (see figure 1): Predicted by M8, but not MSc, which was unable to come with a prediction of smaller area. No earthquake of such magnitude had been ever reported inside the area of the M8 alarm before 2001. The largest one in the 20th century has magnitude $M_s = 7.9$ and happened on November 08, 1997 four months after declaration of the M8 alarm in the test.

The analysis of earthquake sequences from the epicenter area of medium range size (i.e., 600-km radius) for each of the eleven earthquakes of magnitude 8 or above, 1985-2000, evidence consecutive stages of inverse cascading of seismic activity to the great shock although a power-law rise of the Benioff strain release (Varnes, 1989; Bowman *et al.*, 1998) in a timescale of years is not a common feature. Same with the aftershock sequences, which decay rate is hardly a uniform power-law for most of the series. Thus, the Omori Law (Omori, 1894; Utsu, Ogata & Matsu'ura, 1995) is by no means a characteristic behavior of the great earthquakes. The aftershock sequences indicate an apparent switching to higher level of seismic background activity. Does it witness "transient phase changes" in system of blocks-and-faults? Perhaps, but it's hard to make the final judgment without a systematic analysis of larger amount of data.

Discussion and conclusions

Seismic process is an essential part in dynamics of the lithosphere. Earthquakes evidently cascade into aftershocks that re-adjust the hierarchical system of blocks-and-faults in the locality of the main shock rupture. Systematic analysis shows less evident inverse cascade in seismic activity prior to the recent greatest earthquakes. Such a premonitory rise of activity first observed in intermediate-term at large distances, i.e. at the scale of up to 10 sizes of the earthquake source in preparation. In many cases the detailed analysis allows to distinguish a smaller territory, down to the size of the source, where the rise of activity shows intermittent switching from steadily high level to short periods of quiescence. On the background of such intermediate-term rise, episodes of activity, like swarms or sequences of foreshocks could be identified as short-term forerunners of a big earthquake. It is possible that premonitory rise of seismic activity evolves through long, intermediate, short, and immediate, or even nucleation, phases.

There is additional observational support for the hypothesis that the rate of occurrence of intermediate-sized earthquakes increases prior to a large earthquake (Varnes, 1989; Bufe *et al.*, 1994; Varnes & Bufe, 1996; Brehm & Braile, 1998, 1999; Bowman *et al.*, 1998), although it seems premature to claim a uniform power-law as a characteristic behavior. Both the Gutenberg-Richter scaling and the precursory activation are consistent with application of condensed-matter theory to regional seismicity. The scaling of small earthquakes that form seismic background in a region is equivalent to the thermal fluctuations in solids, liquids, and gases. The small earthquakes are the avalanches, which behavior falls under the general class of phenomena that exhibit self-organized criticality. Examples include the sand-pile, slider-block and forest-fire models. The distribution of avalanches can be explained in terms of an inverse cascade due to the coalescence of metastable clusters (Gabriellov *et al.*, 2000). Catalogs of earthquakes evidence consecutive self-organized stages of inverse cascading to main shock and direct cascading of aftershocks. The first may reflect coalescence of instabilities at the approach, while the second may indicate readjustment of a complex system of blocks-and-faults in a new state after a catastrophe.

It is generally accepted that more seismological and other data should be analyzed systematically to establish reliable correlations between the occurrence of extreme earthquakes and observable geophysical phenomena. At the same time a novel understanding of seismic occurrence process in terms of dynamics of a hierarchical system of blocks-and-faults, has led already to the design of reproducible intermediate-term medium-range earthquake prediction algorithms that passed statistical significance testing in forward application (Kossobokov *et al.*, 1999). Although of limited accuracy, the predictions by M8 and MSc algorithms can be used for prevention of damage and losses. The algorithms are neither optimal nor unique. Their accuracy could be improved in a systematic monitoring of the alarm areas and by designing a new generation of earthquake prediction technique. At the same time

the phenomena depicted by the algorithms provide empirical constraints for theoretical and computer modeling.

Acknowledgements

The author wishes to acknowledge long-lasting collaboration with Prof. V. I. Keilis-Borok, whose constant interest to mathematically precise analysis of geophysical data has led to design of a reliable earthquake prediction technique, which yet to be accepted by the majority of seismologists and geophysicists. This presentation was supported by the ISTC (grant #1538), by Russian Foundation for Basic Research (grant #00-15-98507), and by James S. McDonnell Foundation (the 21st Century Collaborative Activity Award for Studying Complex Systems, project "Understanding and Prediction of Critical Transitions in Complex Systems").

References

- BARENBLATT, G. I. (1982). *Similarity, selfsimilarity, and intermediate asymptotics*. Consultants Bureau, New York.
- BHATIA, S. C.; CHALAM, S. V.; GAUR, V. K.; KEILIS-BOROK, V. I.; KOSSOBOKOV, V. G. (1989). "On intermediate-term prediction of strong earthquakes in the Himalayan arc region using pattern recognition algorithm M8". *Proc. Indian Ac. Sci.. Earth and Planetary Sciences*, 98 (1), 111-123.
- BOWMAN, D. D.; OUIILLON, G.; SAMMIS, C. G.; SORNETTE, A.; SORNETTE, D. (1998). "An observational test of the critical earthquake concept". *J. Geophys. Res.*, 103, 24359-24372.
- BREHM, D. J.; BRAILE, L. W. (1998). "Intermediate-term earthquake prediction using precursory events in the New Madrid seismic zone". *Bull. Seis. Soc. Am.*, 88, 564-580.
- (1999). "Intermediate-term earthquake prediction using the modified time-to-failure method in Southern California". *Bull. Seis. Soc. Am.*, 89, 275-293.
- BUFE, C. G.; VARNES, D. J. (1993). "Predictive modeling of the seismic cycle of the greater San Francisco Bay region". *J. Geophys. Res.*, 98, 9871-9883.
- ELLSWORTH, W. L.; BEROZA, G. C. (1995). "Seismic evidence for a seismic nucleation phase". *Science*, 268, 851-855.
- FRANKEL, A.; MUELLER, C.; BARNHARD, T.; PERKINS, D.; LEYENDECKER, E. V.; DICKMAN, N.; HANSON, S.; HOPPER, M. (1996). *National Seismic Hazard Maps*, USGS Open-File Report, 96, 532.
- FROHLICH, C.; DAVIS, S. D. (1993). "Teleseismic b values; or, much ado about 1.0". *J. Geophys. Res.*, 98, 631-644.
- GABRIELOV, A.; KEILIS-BOROK, V. I.; ZALIAPIN, I. V.; NEWMAN, W. I. (2000). "Critical transitions in colliding cascades". *Phys. Rev. E*, 62, 1, 237-249.

- GAHALAUT, V. K.; KUZNETSOV, I. V.; KOSSOBOKOV, V. G.; GABRIELOV, A. M.; KEILIS-BOROK, V. I. (1992). "Application of pattern recognition algorithm in the seismic belts of the Indian convergent plate margins - M8 algorithm". *Proc. Indian Acad. Sci. (Earth Planet. Sci.)*, 101 (3), 239-254.
- GELLER, R. J.; JACKSON, D. D.; KAGAN, Y. Y.; MULARGIA, F. (1997). "Earthquakes cannot be predicted". *Science*, 275, 1616-1619.
- GUTENBERG, B.; RICHTER, C. F. (1954). "*Seismicity of the Earth and associated phenomenon*. 2nd ed. Princeton University Press, Princeton.
- HABERMANN, R. E.; CREAMER, F. (1994). "Catalog errors and the M8 earthquake prediction algorithm". *Bull. Seism. Soc. Am.*, 84, 1551-1559.
- HARRIS, R. A. (1998). "Forecasts of the 1989 Loma Prieta, California, earthquake". *Bull. Seism. Soc. Am.*, 88, 898-916.
- HEALY, J. H.; KOSSOBOKOV, V. G.; DEWEY, J. W. (1992). "A test to evaluate the earthquake prediction algorithm M8". *U. S. Geol. Surv. Open-File Report*, 92-401.
- KEILIS-BOROK, V. I.; ISMAIL-ZADEH, A.; KOSSOBOKOV, V.; SHEBALIN, P. (2001). "Non-linear dynamics of the lithosphere and intermediate-term earthquake prediction". *Tectonophysics*, 338, 247-280.
- KEILIS-BOROK, V. I.; KNOPOFF, L.; KOSSOBOKOV, V. G.; ROTVAIN, I. M. (1990). "Intermediate-term prediction in advance of the Loma Prieta earthquake". *Geophys. Res. Letters*, 17, 9, 1461-1464.
- KEILIS-BOROK, V. I.; KOSSOBOKOV, V. G. (1990). "Premonitory activation of seismic flow: algorithm M8". *Phys. Earth Planet. Inter.*, 61, 73-83.
- KOSSOBOKOV, V. G. (1997). "User manual for M8". In: HEALY, J. H.; KEILIS-BOROK, V. I.; LEE, W. H. K. [eds], "Algorithms for earthquake statistics and prediction". *IASPEI Software Library*, 6. Seismol. Soc. Am., El Cerrito CA.
- KOSSOBOKOV, V. G.; KEILIS-BOROK, V. I.; CHENG, B. (2000). "Similarities of multiple fracturing on a neutron star and on the Earth". *Phys. Rev. E.*, 61, 3529-3533.
- KOSSOBOKOV, V. G.; KEILIS-BOROK, V. I.; ROMASHKOVA, L. L.; HEALY, J. H. (1999). "Testing earthquake prediction algorithms: statistically significant real-time prediction of the largest earthquakes in the Circum-Pacific, 1992-1997". *Phys. Earth Planet. Inter.*, 111, 3-4, 187-196.
- KOSSOBOKOV, V. G.; KEILIS-BOROK, V. I.; SMITH, S.W. (1990). "Localization of intermediate-term earthquake prediction". *J. Geophys. Res.*, 95, B12, 19763-19772.
- KOSSOBOKOV, V. G.; KEILIS-BOROK, V. I.; TURCOTTE, D. L.; MALAMUD, B. D. (2000). "Implications of a statistical physics approach for earthquake hazard assessment and forecasting". *Pure Appl. Geophys.*, 157, 2323-2349.
- KOSSOBOKOV, V. G.; MAEDA, K.; UYEDA, S. (1999). "Precursory activation of seismicity in advance of the Kobe, 1995 earthquake". *Pure Appl. Geophys.*, 155, 409-423.
- KOSSOBOKOV, V. G.; MAZHKENOV, S. A. (1994). "Times of increased probability of large earthquakes in the Eastern Tien Shan diagnosed by the M8 algorithm". In: CHOWDHURY, D. K. [ed.]. *Computational seismology and geodynamics / Am. Geophys. Un.*, 1, Washington, D. C, 1, 16-19.

- KOSSOBOKOV, V. G.; RASTOGI, B. K.; GAUR, V. K. (1989). "On self-similarity of premonitory patterns in the regions of natural and induced seismicity". *Proc. Indian Ac. Sci. Earth and Planetary Sciences*, 98 (4), 309-318.
- LATOUSSAKIS, J.; KOSSOBOKOV, V. G. (1990). "Intermediate term earthquake prediction in the area of Greece: application of the algorithm M8". *Pure Appl. Geophys.*, 134, 2, 261-282.
- LAY, T.; WALLACE, T. C. (1995). *Modern global seismology*. Academic Press, San Diego.
- OMORI, F. (1894). "On the after-shocks of earthquakes". *J. Coll. Sci. Imp. Univ. Tokyo*, 7, 111-200.
- ROMASHKOVA, L. L.; KOSSOBOKOV, V. G. (2001). "Seismicity dynamics prior to and after the largest earthquakes worldwide, 1985-2000". In: *Problems of lithosphere dynamics and seismicity (computational seismology, 32)*, GEOS: Moscow, 162-189. [In Russian].
- ROMASHKOVA, L. L.; KOSSOBOKOV, V. G.; PANZA, G. F.; COSTA, G. (1998). "Intermediate-term prediction of earthquakes in Italy: algorithm M8". *Pure and Appl. Geophys.*, 152, 37-55.
- SCHOLZ, C. H. (1997). "Whatever happened to earthquake prediction". *Geotimes*, 42 (3), 16-19.
- SYKES, L. R.; SHAW, B. E.; SCHOLZ, C. H. (1999). "Rethinking earthquake prediction". *Pure Appl. Geophys.*, 155, 207-232.
- TURCOTTE, D. L. (1995). *Fractals and chaos in geophysics*. 2nd ed. Cambridge Univ. Press, Cambridge, U. K.
- UTSU, T.; OGATA, Y.; MATSU'URA, R. S. (1995). "The centenary of the Omori formula for a decay law of aftershock activity". *J. Phys. Earth*, 43, 1-33.
- VARNES, D. J. (1989). "Predicting earthquakes by analyzing accelerating precursory seismic activity". *Pure Appl. Geophys.*, 130, 661-686.
- VARNES, D. J.; BUFE, C. G. (1996). "The cyclic and fractal seismic series preceding an m_b 4,8 earthquake on 1980 February 14 near the Virgin Islands". *Geophys. J. Int.*, 124, 149-158.

Earthquake preparedness: the contribution of earthquake prediction and deterministic hazard research

Giuliano F. Panza

Department of Earth Sciences, University of Trieste
The Abdus Salam International Centre for Theoretical Physics-SAND Group ICTP

A. Peresan and F. Vaccari

Department of Earth Sciences, University of Trieste

L. Romashkova, V. Kossobokov, A. Gorshkov and I. Kuznetsov

International Institute of Earthquake Prediction Theory and Mathematical Geophysics,
Russian Academy of Sciences

Abstract

Several algorithms, structured according to a general pattern-recognition scheme, are currently applied for the space-time identification of strong earthquakes in the Italian territory. Two, namely CN and M8 algorithms, for intermediate-term medium-range earthquake prediction purposes and one for the recognition of earthquake-prone areas. These algorithms can be viewed as independent experts, hence they can be combined to better constrain the alerted seismogenic area.

We examine here the possibility to integrate the intermediate-term medium-range earthquake predictions with the procedures developed for the recognition of strong earthquake prone areas and for the deterministic assessment of seismic hazard. This permits to associate the alarms declared by the prediction algorithms to a set of possible seismic ground motion scenarios. We analyse in detail the case of CN predictions and we supply examples of scenarios corresponding to possible events with $M \geq 6.5$, that turn out to be the closest to several cities in Italy.

A practical advantage of the described procedure is the time information given by the intermediate-term earthquake prediction that is very useful to plan preparedness and rescue actions and to define priority criteria for the more detailed studies required by seismic microzonation.

Resum

Actualment s'apliquen alguns algorismes estructurats d'acord amb un esquema general de reconeixement de models amb vista a la identificació espaciotemporal de terratrèmols de gran intensitat a territori italià. N'hi ha dos, els algorismes CN i M8, de dedicats a la predicció a mitjà termini de terratrèmols. Aquests algorismes es poden considerar com a experts independents i, doncs, es poden combinar per a delimitar millor l'àrea sismogènica alertada.

Aquí examinem les possibilitats d'integrar les prediccions a mitjà termini de terratrèmols d'abast mitjà amb els procediments desenvolupats per al reconeixement de les àrees propenses a patir terratrèmols de gran intensitat i per a l'avaluació determinista del risc sísmic. Això permet d'associar les alarmes declarades pels algorismes de predicció a un conjunt de possibles escenaris de moviments sísmics. Analitzem en detall el cas de les prediccions del CN i oferim exemples d'escenaris corresponents a possibles esdeveniments amb $M \geq 6,5$ molt pròxims a algunes ciutats italianes.

El procediment descrit té l'avantatge pràctic del temps d'informació que ofereix la predicció de terratrèmols a mitjà termini, que resulta molt útil per als preparatius i per a les accions de rescat i per a definir criteris de prioritats per als estudis més detallats que requereix la microzonació sísmica.

Introduction

Earthquake preparedness requires a basic knowledge of the expected earthquake location and of the effects of local site conditions on shaking, since any strategy for seismic risk reduction should be outlined trying to answer the following basic questions:

- When, where and how big we have to expect a strong earthquake to strike a region?
- What should we expect when it occurs?

The answer to the first question is matter for earthquake prediction, while the second one is matter for seismic risk studies.

The complex nature of the seismic phenomenon does not seem to allow for deterministic predictions; therefore, we cannot know in advance location, magnitude, and time of occurrence of an earthquake, all within narrow limits, so that a planned evacuation can take place. Besides, the statistical predictions of destructive earthquakes cannot be performed, since there have been too few large events in any sufficiently small area to be able to define probabilities with the required accuracy. Currently, a realistic goal seems represented by the intermediate-term medium-range predictions, based on the observed variations of the background seismicity preceding large earthquakes. In this work we provide an example showing how, formally defined intermediate-term medium-range predictions, where alarms are declared for a time interval of a few years and with a space uncertainty of a few hundred kilometres, can be usefully integrated with the procedures for the deterministic hazard estimation (i.e. with the "prediction" of their effects).

The practical example of Italy is supplied, where three different algorithms are currently applied for the space-time identification of strong events. One of such algorithms is applied for the recognition of earthquake prone areas in the Italian territory and the others, namely CN and M8 algorithms, for earthquake prediction purposes. These procedures are independent experts, hence they can be combined to better constrain the alerted seismogenic area.

The algorithm for the recognition of Earthquake Prone Areas (EPAs) is used to identify the

sites where strong earthquakes are likely to occur, independently from seismicity information. This method is based on the assumption that strong events nucleate at the nodes (Gelfand *et al.*, 1972; Gabrielov *et al.*, 1996), specific structures that are formed at the intersections of lineaments. Lineaments are identified by the Morphostructural Zonation (MSZ) Method (Alexeevskaya *et al.*, 1977), that delineates a hierarchical block structure of the studied region, using tectonic and geological data, with special care to topography. In Italy, this study has allowed us to identify the sites where stronger events, with magnitude larger or equal to 6.0 or 6.5, may occur (Gorshkov *et al.*, 2002).

The algorithms CN (Keilis-Borok & Rotwain, 1990) and M8 (Keilis-Borok & Kossobokov, 1987) permit to identify the periods of time (TIP: Time of Increased Probability for the occurrence of a strong earthquake) when a strong event is likely to occur inside a given region, based on a quantitative analysis of the seismic sequence. For the application of the algorithm CN in Italy, a regionalization strictly based on the seismotectonic zoning and taking into account the main geodynamic features of the Italian area, is used (Peresan *et al.*, 1999). For M8 application to the Italian territory, a new spatially stabilized variant of the algorithm is considered, where the seismicity is analysed within a dense set of overlapping circles covering the monitored area (Kossobokov *et al.*, 2002).

In this work intermediate-term medium-range earthquake predictions and pattern recognition of earthquake-prone areas are combined with the deterministic hazard procedure, in order to associate the alarms (TIPs) declared by CN and M8 to a set of appropriate scenarios of hazard.

The procedure for deterministic seismic hazard assessment is based on the possibility to compute synthetic seismograms; the expected ground motion can be modelled at any desired point, starting from the available information about seismic sources and regional structural models (Panza *et al.*, 1999). Here, according to the flow chart shown in figure 1, the set of possible seismic sources defined within the alerted area (e.g. one of the CN regions) is used for the realistic modelling of the seismic ground motion (Panza *et al.*, 2001 and references therein). In such a way an alarm can be associated with maps describing the seismic ground motion (e.g. peak values of displacement, velocity and acceleration) caused by the potential sources in the alerted region. In a complementary step it is possible to define a set of scenarios, corresponding to the strong earthquake ($M > 6.0$) prone areas, the nodes, defined by MSZ and pattern recognition and contained within the alarmed region. We supply examples of scenarios corresponding to possible events with $M \geq 6.5$, that turn out to be the closest to several cities in Italy.

Intermediate-term medium-range earthquake predictions

To predict an earthquake means to indicate the main features (time, space and expected minimum magnitude) of an impending event, usually on the basis of the observation of a precursory signal. The prediction can be probabilistic, in the sense that the magnitude, time

and location might not be given exactly, but forecasts must demonstrate more predictability than a random guess, even accepting an intrinsic proportion of missed events or false alarms. To claim for a precursor, that is for a systematic relation between some quantifiable phenomenon and the occurrence of strong earthquakes, it is necessary to provide a significant number of case histories, globally distributed. Up to now, most of the proposed precursors have either failed or are, at best, unproven, mainly due to the lack of prolonged and systematic records. This problem is overcome, to some extent, when considering seismic precursors. Earthquake catalogues, in fact, represent the most widely available seismological data, containing systematically collected information for a prolonged period of time; this permits a widespread testing of the premonitory patterns which might be observed in the seismicity. Details about possible earthquake precursors, including the list of IASPEI significant precursors, is provided in Appendix I.

Nowadays, the most promising approach seems represented by the intermediate-term medium-range earthquake predictions (i.e. with a characteristic alarm-time from a few months to a few years and a space uncertainty of hundreds of kilometres) based on certain variations of the background seismicity preceding large earthquakes. In fact, the analysis of the seismic flow has evidenced that some specific seismicity patterns may prelude to an incumbent strong event. This observation is supported by the consideration that to initiate an earthquake a certain stress level is necessary that will maintain slow fracture growth; the latter tends to be discrete and manifests itself as background seismicity (Knopoff, 1996).

Several possible scenarios of precursory seismic activity have been proposed (see Appendix II); nevertheless, only a few formally defined algorithms allow for a systematic monitoring of seismicity, as well as for a widespread testing of their performances. In this study two algorithms are considered, namely CN and M8 algorithms (Keilis-Borok, 1996 and references therein). Both of them belong to a family of medium-range intermediate-term earthquake prediction algorithms structured according to a general pattern-recognition scheme and based on a quantitative analysis of the seismic flow. Such algorithms allow for a diagnosis of the Times of Increased Probability (TIPs) for the occurrence, inside a given region and time window, of the earthquakes with magnitude greater than a fixed threshold M_0 . The quantification of the seismicity patterns is obtained through a set of empirical functions of time, evaluated on the sequence of the main shocks occurred in the analysed region. Symptoms considered are “non-Earth-specific” (Keilis-Borok, 1996), and can be observed in many non-linear systems before a catastrophic event; in our case the non-linear system corresponds to the system of active faults and the small earthquakes are the source of perturbation of the system. The tests in advance predictions, performed on a global scale, allowed a first statistical evaluation of predictions; the significance level of the obtained results (larger than 95 %) seems to substantiate the predictive capability of these algorithms (Kossobokov *et al.*, 1999; Rotwain and Novikova, 1999). Algorithms have been developed to reduce the spatial uncertainty of predictions, making use of the information carried by lower

magnitude seismic activity (Kossobokov *et al.*, 1999), but their applicability is limited by the difficulty to keep a globally high level of detection.

The simple definition of alarm periods as “times of increased probability with respect to normal conditions”, which are not associated to a specific value of probability for the occurrence of a strong earthquake, is imposed by the fact that any attempt to quantify precisely the probability increase during TIPs would require several a priori assumptions (i.e. Poissonian recurrence, independence of TIPs, etc.), most of which would be poorly constrained by the available observations and hence below any critics.

The algorithm CN in Italy

According to CN rules, when a TIP is declared in a preselected region, the strong earthquake could occur in any point of the alerted area. Therefore, in order to reduce the space uncertainty of predictions, the region defined has to be as small as possible. On the other side it has been found empirically that the linear dimensions of the monitored area must be greater or equal to $5L-10L$, where L is the length (proportional to M_0) of the source of the events to be predicted. Therefore all the earthquakes with $M < M_0$ are point processes within the monitored area.

Considering the fractal character of the spatial distribution of events (e.g. Turcotte, 1992), the region must contain the major clusters of epicenters, characterised by a high density of epicentres and by large magnitudes. This choice affects clearly the frequency-magnitude distribution for events occurred within the region, because the log-linearity of the Gutenberg-Richter relation is preserved only on a global scale (i.e. for point processes) or, according to a multiscale approach (Molchan *et al.*, 1997), within an area of the appropriate hierarchical scale, depending on the maximum magnitude considered. When considering a limited area the frequency-magnitude distribution exhibits a good linearity only for the lower magnitudes, with increasing fluctuations due to the small number of events with larger magnitudes. In general, when applying the algorithm CN an upward bend can be observed in this pattern starting at a certain magnitude; depending both on the particular choice of the region, that must include the more seismically active zones, and on the possible occurrence of a characteristic earthquake for the major faults included in it (Schwartz & Coppersmith, 1984). According to Molchan *et al.* (1990), the magnitude threshold M_0 , for the selection of the events to be predicted, is chosen close to the minimum in the number of events, and this guarantees the stability of the results. Consequently, in any monitored region, CN makes use of the information given by small and moderate earthquakes, having a quite good statistic (i.e. following the GR law), to predict the stronger earthquakes, which are rare and possibly characteristic events of the same region.

The testing of the algorithm CN is ongoing in more than twenty regions world wide. From the global retrospective tests performed, it turns out that the algorithm CN is able to “predict”

the occurrence of about 80 % of the strong events, with TIPs occupying, on average, about 30 % of the total time (Keilis-Borok, 1996). The tests in advance predictions, carried on during the period 1983-1998, allowed to estimate the significance level of the results obtained by CN around 95 % (Rotwain & Novikova, 1999).

The algorithm CN is applied in Italy since 1990 (Keilis-Borok *et al.*, 1990). A regionalization (fig. 2), strictly following the seismotectonic zoning (Meletti *et al.*, 2000) and taking into account the main geodynamic features of the Italian area, is currently used for the application of the algorithm. The forward prediction is performed every two months, using the UCI2001 (Romashkova *et al.*, 2001) catalogue, which is composed by the Italian catalogue CCI1996 (Peresan *et al.*, 1997) for the period 1900-1985 and is updated with the NEIC Preliminary Determinations of Epicenters since 1986. The thresholds M_0 for the selection of the events to be predicted are fixed, taking into account their average return period, to $M_0 = 5.4$ for the Northern region and to $M_0 = 5.6$ for the Central and Southern regions. Details about CN application in Italy can be found in Peresan *et al.*, (1999). Results of predictions are routinely provided since January 1998; the results updated to January, 1 2002, are shown in figure 3.

The algorithm M8 in Italy

Algorithm M8 was designed for the prediction of the strongest earthquakes worldwide, with magnitude 8.0 and above (Keilis-Borok & Kossobokov, 1987), and was later adapted for the prediction of earthquakes with smaller magnitudes (Keilis-Borok & Kossobokov, 1990). The M8 algorithm analyses the seismic activity inside a set of Circles of Investigation, CIs, with radius normalized by the linear size of the events to be predicted, i.e. proportional to magnitude threshold M_0 . The choice of M_0 is determined by the condition that the average recurrence time of strong earthquakes is sufficiently long in the territory considered. A hierarchy of predictions is usually delivered for different magnitude ranges $M_0 +$, considering values of M_0 with an increment of 0.5 (i.e. $M_0 +$ indicates the magnitude range: $M_0 \leq M \leq M_0 + 0.5$).

For the monitoring of seismicity in the Italian territory, a new spatially stabilized scheme for the application of the algorithm M8 has been considered, where the seismicity is analysed within a dense set of overlapping CIs covering the monitored area (Kossobokov *et al.*, 2002). The steps of the analysis are the following. The territory is scanned with a set of small circles distributed over a fine grid, with the radius of the small circles approximately equal the grid spacing. The seismically active grid points are then selected by the condition that the average annual rate of seismic activity, within the small circle, is above a given threshold. The grid points where data are insufficient for the application of M8 algorithm and isolated grid points are excluded. The M8 algorithm is then applied with the circles of investigations, CIs, centred at each of the selected grid points. Finally, an alarm is declared for a CI only if the overwhelming majority (more than 75 %) of the CIs centred at the neighbouring grid points are also in state of

alarm. This scheme for the application of M8 is less subjective and less sensitive to the position of the CIs than the standard variant of the algorithm.

In Italy predictions are performed in the two different magnitude ranges defined by $M_0 = 6.5$ and 6.0 (i.e. $M6.5 +$ and $M6.0 +$), with CIs of radius $R = 192$ km and $R = 138$ km respectively, centered at the nodes of a grid with step $0.25^\circ \times 0.25^\circ$. The catalogue used for M8 application is the UCI2001 (Romashkova *et al.*, 2001), that is the same used by the algorithm CN; according to the worldwide tests the operating magnitude considered is M_{max} .

Figure 4 shows the current alarms in Italy as on January 1, 2002, displaying a rather large territory in state of alarm in the northern part of Italy for both magnitude ranges $M6.5 +$ (fig. 4a) and $M6.0 +$ (fig. 4b). The average space-time volume occupied by alarms has been estimated in the retrospective simulation of forward predictions as 38.6 % for $M6.5 +$ and 29.6 % for $M6.0 +$ (for details see Kossobokov *et al.*, 2002).

Pattern recognition of earthquake prone areas

The space uncertainty typical for the intermediate-term medium-range predictions is quite large. An attempt to better constrain the expected sources, however, seems possible through the combined use of seismological, geological and morphostructural information. In fact, the pattern-recognition technique can be used as well to identify the sites capable to generate the strongest events. The method for the pattern recognition of Earthquake-Prone Areas (EPAs) is based on the assumption that strong events nucleate at the nodes (Gelfand *et al.*, 1972), specific structures that are formed at the intersections of lineaments. Lineaments can be identified by the Morphostructural Zonation (MZS) Method (Alexeevskaya *et al.*, 1977), which delineates a hierarchical block structure of the studied region, using tectonic and geological data, with special care to present-day topography. Among the defined nodes, those prone to strong earthquakes are then identified by pattern recognition on the basis of the parameters characterising indirectly the intensity of neotectonic movements and fragmentation of the crust at nodes (e.g. elevation and its variations in mountain belts and watershed areas; orientation and density of linear topographic features; type and density of drainage pattern). The methodology developed for the recognition of EPAs has been successfully tested in many different regions of the world (e.g. Gorshkov *et al.*, 2000 and references therein).

This methodology has been applied to peninsular Italy and Sicily by Gorshkov *et al.* (2002), revising and updating, on a more detailed scale, the analysis performed by Caputo *et al.* (1980). The morphostructural map has been compiled at the scale of 1:1,000,000 by the combined analysis of topographic, tectonic, geological maps and satellite photos. Under the assumption that the future strong events will occur at the nodes, the seismic potential of each node has been evaluated, by the pattern recognition technique, for two magnitude thresholds, $M \geq 6.0$ and $M \geq 6.5$. For recognition purposes, the nodes have been defined as circles of radius $R = 25$

km surrounding each point of intersection of lineaments (Gorshkov *et al.*, 2002). Such node dimension is comparable with the size of the earthquake source for the magnitude range considered in this work (Wells & Coppersmith, 1994) and with the smoothing window used in the deterministic hazard computation (Costa *et al.*, 1993). The results of the classification of the nodes for both magnitude thresholds are in good agreement with the recorded seismicity, in fact almost all (more than 90 %) of the past strong earthquakes occurred at the recognised nodes. The nodes prone to $M \geq 6.5$ events are shown in figure 5.

The identification of EPAs is especially important in areas where historical and instrumental information is scarce. In fact EPAs can be used as input sources in the deterministic seismic hazard assessment, thus allowing an effective estimate of the seismic risk, more realistic than that based on unavoidably incomplete seismological observations.

Seismic hazard scenarios

Earthquake hazard analysis can be performed in various ways, e.g. by means of “ground-shaking scenarios”, that is a description of the ground-shaking intensity due to an earthquake of a given magnitude, or using probabilistic maps of relevant parameters to describe the ground motion.

The probabilistic analysis (Cornell, 1968) supplies indications that could be useful, but are not sufficiently reliable to characterise the seismic hazard, since it is basically conditioned by the definition of the seismogenic zones, inside which the seismogenic process is assumed to be rather uniform. For a given zone, the estimation of the recurrence time of seismic events is based on the classical Poisson hypothesis, where the seismic events are assumed to be time independent, and the related values of probability are derived from empirical data that may fail to describe adequately the reality. The assumption of homogeneity can introduce severe errors in the estimate of the seismic hazard at a specific site, especially in view of the limited seismological data and of the Multiscale Seismicity (MS) model (Molchan *et al.*, 1997). In fact, according to the MS model, a log-linear Frequency-Magnitude (FM) relation can describe the ensemble of events only if they are geometrically small, compared with the elements of the seismotectonic regionalization. This condition, largely fulfilled by the early global investigation by Gutenberg and Richter, has been subsequently violated in many investigations performed at regional scale, leading to the Characteristic Earthquake (CE) model (Schwartz and Coppersmith, 1984). The MS model implies that, when assessing the seismic hazard with the probabilistic approach, the seismic zonation should be made at several scales, depending upon the self-similarity conditions of the seismic events and the linearity of the log FM relation, in the magnitude range of interest.

Many different renewal process models (e.g., the “Brownian Passage Time” model by Ellsworth *et al.* [1998] or the “Time-Predictable” model by Shimazaki and Nakata [1980])

have been proposed for probabilistic estimates, as more adequate alternatives to the classical Poisson hypothesis (“memoryless” random occurrence of the events). Such models make use of mixed functions of the density of probability to describe the occurrence of the events and these functions may differ from place to place. As a consequence, the available observations appear often not sufficient to constrain significantly the distribution functions. Another shortcoming of the renewal process models is the evaluation of the time elapsed from the last event: such an evaluation can be impossible if the length of the catalogue is shorter than the storage and release time interval and paleoseismological data are not available (Decanini *et al.*, 2001).

Due to the aforementioned limitations of the probabilistic methods in this work a deterministic approach is considered, which is completely different and complementary to the probabilistic approach as in general proposed (Costa *et al.*, 1993; Panza *et al.*, 2001). The deterministic approach capably addresses aspects largely overlooked in the probabilistic one such as:

- the direct evaluation of resulting maps in terms of design parameters, without requiring the adaptation of probabilistic maps to design ground motions;
- the effect of crustal properties on attenuation;
- the derivation of ground motion parameters from synthetic time series, instead of using overly simplified attenuation functions;
- the generalization of design parameters to locations where there is little seismic history.

The numerical modelling of ground motion allows for a first-order seismic hazard mapping, based on the seismograms computed, for example, at the nodes of a grid with the desired spacing (Panza, Radulian & Trifu, 2000). Moreover, for the urban areas where synthetic seismograms have been successfully compared with strong motion records, the theoretical computations may permit a detailed microzoning (Panza *et al.*, 1996), taking into account a set of possible scenario earthquakes. The seismic input, defined by means of complete waveforms modelling, can be used to perform detailed studies of the effects of the expected ground motion on the relevant man made structures. In fact, complete seismograms are necessary for engineering dynamic analysis, in order to compute the full non-linear response of the structures (Field *et al.*, 2000).

Integrated deterministic seismic hazard

The deterministic hazard modelling by Costa *et al.* (1993) and its developements by Panza *et al.* (1999), employed in this integrated approach, are based on the computation of complete synthetic seismograms with parameters defined from a wide geophysical and geological data set. The different structural and seismic source models, necessary for the computation, are defined with regional polygons.

The steps of the procedure can be summarised as follows:

1. the seismic sources are grouped into homogeneous seismogenic zones and for each group the representative focal mechanism is assigned;
2. the scalar seismic moment associated with each source is estimated considering the maximum observed magnitudes in the epicentral area, eventually integrated with other available information on the seismic potential of active faults;
3. a database of synthetic seismograms is computed by modal summation (Panza, 1985; Florsch *et al.*, 1991), to model ground motion, taking into account the available knowledge on the earthquake generation and wave propagation processes;
4. seismic hazard maps are compiled, considering maximum displacement, velocity, design ground acceleration (DGA), or any other parameter that can be extracted from the complete synthetic seismograms, which blend information from geology, historical seismicity and observational seismology.

Thus, given a set of expected sources and the average properties of the structural model, a zoning can be performed at a regional scale, identifying areas prone to the heaviest seismic input.

In the present study we make use of such methodology, in order to generate scenarios of expected ground motion that can be considered, in case of a declared TIP, to increase earthquake preparedness. The space information provided by the prediction algorithms can be used directly with the deterministic hazard procedure, to associate to the alerted area an appropriate hazard scenario (space: level 1). Local scenarios (space: level 2) can be computed for each earthquake prone node capable of the strongest events included in the alerted region. To give an idea of the maps (space: level 1 and 2) that can be obtained following the proposed procedure, we supply the scenarios associated with the CN regionalization in Italy (Peresan *et al.*, 1999) as well as examples of scenarios corresponding to possible events with $M \geq 6.5$ (Gorshkov *et al.*, 2002), that turn out to be the closest to several cities in Central Italy.

CN for deterministic hazard

The algorithm CN indicates if an event, with magnitude larger or equal than a fixed threshold M_0 , is likely to occur within a given region and time interval. In order to describe what should be expected, in terms of ground shaking, when a TIP is declared, we apply the procedure for the deterministic evaluation of seismic hazard following closely Panza *et al.* (1999), considering only the set of possible sources included in the CN region.

To derive the distribution of the maximum observed magnitudes, we consider the catalogue NT 4.1 (Camassi & Stucchi, 1996). Since each CN region represents an individual seismic domain, as assumed by the regionalization based on the seismotectonic model, only the sub-catalogue of events occurred inside the region is considered. The seismicity is then discretized into $0.2^\circ \times 0.2^\circ$ cells, assigning to each cell the maximum magnitude M recorded within it; a smoothing procedure

(Panza *et al.*, 1999) is then applied to account for the spatial uncertainty and for source dimensions. Only the cells located within a seismogenic zone and within the CN region are retained, and a double-couple point source, with a representative focal mechanism, is placed at the centre of each cell. If the assigned magnitude M is lower than the CN threshold M_0 , we put $M = M_0$.

Synthetic seismograms are then computed with receivers placed at the nodes of a grid with step $0.2^\circ \times 0.2^\circ$, covering the studied area; at each receiver the sum vector of the radial, vertical and transverse components of ground motion is computed. To reduce the amount of computations, the maximum source-receiver distance is set equal to 25, 50 and 90 km, depending on the magnitude associated with the source ($M < 6,6 \leq m < 7$, and $M \geq 7$, respectively). More details about the procedure can be found in Panza *et al.* (1999).

Each receiver is thus associated to recordings of many different sources and any parameter of interest can be extracted from such complete time series, therefore different maps can be produced. To provide an example of the possible scenarios associated with CN TIPs, the maps of horizontal velocities produced for the three Italian regions are shown in figure 6. The corresponding maximum intensities at some Italian cities along the monitored regions, are then estimated (Table 1), considering the relations among the parameters of ground motion and the ISG maximum observed intensities (Molin *et al.*, 1996), derived by Panza *et al.* (1997).

The integrated CN-deterministic hazard approach allows us to observe, for example, that a TIP in the Northern region mainly concerns the north-eastern part of Italy. In the meanwhile, comparing fig. 6a and fig. 6b, it is possible to notice that along the coast of the Adriatic sea, between latitudes 40°N and 42°N , the highest hazard is associated to alarms in the Central

City	CN region	I_{max} (observed)	I_{max} (computed)		
			Displacement	Velocity	DGA
Trieste	north	VII	VI	VII	VI
Bologna	north, centre	VIII	VIII	VIII	IX
Assisi	centre	VIII	VIII	IX	IX
Firenze	centre	VIII	VII	VIII	VIII
Roma	centre	VIII	IX	VIII	IX
Napoli	centre, south	VIII	X	IX	IX
Messina	south	X	X	X	X

TABLE 1. Maximum intensities at some cities along the CN monitored area (fig. 6) estimated, using the relations derived by Panza *et al.* (1997) for the parameters of ground motion and the ISG observed intensities (Molin *et al.*, 1996). The CN region providing the maximum ground shaking is evidenced in bold. Design Ground Acceleration (DGA) has been obtained using the design spectra of EC8 for soil A (EUROCODICE 8, 1993).

region. This is mainly due to the lower level of the observed seismic activity in such part of the compressive belt composing the Northern region, with respect to the higher magnitude seismicity associated with the extensional belt in the Central region.

Earthquake-Prone Areas for deterministic hazard

Scenarios associated with CN regions provide information about the whole area which may be concerned by a declared TIP; nevertheless a single earthquake does not generate such an extended damage, nor every point within the alerted region is capable to generate the strongest events ($M \geq 6.0$ or $M \geq 6.5$). Since the space uncertainty of CN predictions is intrinsically quite large, an attempt to better constrain the expected sources is done through the pattern recognition of earthquake-prone areas. When a TIP is declared, the strongest events should nucleate at the identified nodes inside the alerted region; hence, it is possible to associate each of these nodes (fig. 5), corresponding to circles with radius $R = 25$ km (Gorshkov *et al.*, 2002), to a scenario with dimensions comparable to the area which may be realistically affected by a single strong event.

To provide an example of such pattern recognition and deterministic hazard integrated approach, we consider the particular case of Central region and the nodes prone to earthquakes with $M \geq 6.5$, that have been identified by Gorshkov *et al.* (2002) for peninsular Italy (fig. 5). The procedure is the same we followed in the previous section, except that only the subset of sources (discretized into $0.2^\circ \times 0.2^\circ$ cells) included in a node are retained now, considering each circle separately. Moreover if the magnitude M assigned to a representative source, based on historical data, is lower than the magnitude expected accordingly to the seismogenic potential indicated by the morphostructural analysis, then we set $M = 6.5$. The pattern recognition of earthquake prone areas, however, does not provide an upper limit for the magnitude to be expected at a given node. Hence, special attention should be paid to those areas identified as prone to strong earthquakes, but characterised only by moderate seismic activity in historical time, in order to assess their effective maximum seismogenic potential.

A set of scenarios is then obtained through the realistic modelling of ground motion, for the earthquake-prone areas within the CN Central region (fig. 7d). From these scenarios it is possible to select, for example, those associated with places of special interest (e.g. large cities or some special industrial areas). Here we provide the maps of horizontal velocities for the nodes giving the maximum ground motion in the cities of Assisi, Roma and Napoli (fig. 7). The maximum intensities, as obtained from the regression (Panza *et al.*, 1997) of estimated velocity, displacement and DGA, are provided in Table 2. The observed maximum intensities drawn from ISG data are reported for comparison.

Apart from Roma and Napoli, which are the largest cities in the study area, we consider the case of Assisi since the corresponding node hosted at least three strong earthquakes with $M \geq 6.0$ in the past millennium ($M = 6.3$ in 1279 and $M = 6.1$ in 1832, according to Boschi *et al.*, 1995)

City	Number of nodes	I_{max} (observed)	I_{max} (computed)		
			Displacement	Velocity	DGA
Assisi	9	VIII	VIII	IX	IX
Roma	5	VIII	IX	IX	IX
Napoli	3	VIII	X	IX	IX

TABLE 2. Maximum intensities at some cities in Central Italy (fig. 7), corresponding to the nodes prone to earthquakes with $M \geq 6.5$, and maximum observed intensity from ISG (Molin *et al.*, 1996). Only the scenarios for the nodes providing the maximum ground shaking are considered and the number of nodes, which may interest each city, is indicated.

and the most recent one occurred in 1997 ($M = 6.0$). The synthetic scenario appears reasonably comparable with the available macroseismic observations and experimental strong motion records (Aoudia *et al.*, 2000); from Table 2 we can observe that the computed intensity can be slightly higher than the observed one, and that the city is surrounded by a large number of nodes.

The scenarios associated with the cities of Roma and Napoli (fig. 7b, 7c) involve a wider area than that associated to Assisi, according to the large magnitudes reported for this part of the Italian peninsula.

Conclusions

An example has been provided showing how different methodologies can be integrated, blending together the available information in a set of realistic scenarios, to supply a useful tool for decision-makers in order to increase earthquake preparedness.

The retrospective analysis of the case of the Umbria-Marche earthquake, occurred in Central Italy on September 26, 1997, supports the adequacy of the proposed procedure. According to the CN algorithm, the $M = 6.0$ event was preceded by a TIP declared for the Central region (fig. 4b). Furthermore the epicentre is localised within one of the nodes recognised to be prone to earthquakes with both $M \geq 6.0$ and $M \geq 6.5$. Finally, the scenario associated to the corresponding node, by the deterministic hazard procedure, appears reasonably comparable with the experimental strong motion records and the available macroseismic observations (Aoudia *et al.*, 2000).

Deterministic hazard and recognition of earthquake prone areas procedures are especially useful as a mean of prevention in areas that have not yet been struck, but are potentially prone to earthquakes. In fact, they are based on a wide geophysical and geological data set, as well as on the current knowledge of the physical process of earthquake generation and wave propagation in realistic anelastic media, and not only on the available macroseismic observations. Moreover, the

procedure for seismic hazard assessment based on the computation of synthetic seismograms provide a realistic modelling of ground motion instead of a less specific upper bound for the maximum possible ground shaking. According to Field *et al.* (2000), “waveform modelling represents our best hope for making more accurate estimates of ground motion at a site” and “is also in line with the trend toward dynamic analysis in the engineering community”.

A practical advantage of the described procedure is the time information given by the intermediate-term earthquake prediction that is very useful to plan preparedness and rescue actions (e.g. placement and survey of the first-aid resources) and to define priority criteria for the more detailed studies required by seismic microzonation. The quantitative approach proposed for the evaluation of seismic risk can be helpful in reorienting the current strategies, shifting the focus from the highly expensive post-disaster rescue and relief operations (prevalent in many countries) to cost effective advance actions aimed at creating knowledge based hazard resilient public assets.

Acknowledgements

We are grateful to A. A. Soloviev for the useful discussions and to F. Vaccari and G. Costa for their precious help. This research has been developed in the framework of the UNESCO-IGCP project 414 and has been supported by INTAS funds (n° 94-0232), by MIUR-COFINANZIAMENTO (2000 and 2001), by CNR (Agenzia 2000) and by NATO SFP 972266.

References

- ALEKSEEVSKAYA, M. A.; GABRIELOV, A. M.; GVISHIANI, A. D.; GELFAND, I. M.; RANZMAN, E. Ya (1977). “Formal morphostructural zoning of mountain territories”. *J. Geophys.*, 43, 227-233.
- AUDIA, A.; CHIMERA, G.; COSTA, G.; NUNZIATA, C.; PANZA, G. F.; ROMANELLI, F.; SARAO', A.; SUHADOLC, P.; VACCARI, F. (2000). “Modelling of the seismic ground motion of the Umbria-Manche sequence (September 1997). Proceedings of the 12th World Conference on Earthquake Engineering”. CdRom paperId-2500, Auckland - New Zealand.
- BOSCHI, E.; FERRARI, G.; GASPERINI, P.; GUIDOBONI, E.; SMRIGLIO, G.; VALENSISE, G. (1995). *Catalogo dei forti terremoti in Italia dal 461 a. C. al 1980*. Istituto Nazionale di Geofisica e Storia Geofisica Ambiente.
- BUFE, C. G.; NISHENKO, S. P.; VARNES, D. J. (1994). “Seismicity trends and potential for large earthquakes in the Alaska-Aleutian region”. *Pure and Appl. Geophys.*, 142, 83-99.
- CAMASSI, R.; STUCCHI, M. (1996). *NT 4.1, un catalogo parametrico di terremoti di area Italiana al di sopra della soglia di danno. Open data file*. Consiglio Nazionale delle Ricerche GNDT.

- CAPUTO, M.; KEILIS-BOROK, V. I.; OFICEROVA, E.; RANZMAN, E.; ROTWAIN, I.; SOLOVIEFF, A. (1980). "Pattern recognition of earthquake-prone areas in Italy". *Phys. Earth Planet. Inter.*, 2-1, 305-320.
- CORNELL, C. A. (1968). "Engineering seismic risk analysis". *Bull. Seism. Soc. Am.*, 58.
- COSTA, G.; PANZA, G. F.; SUHADOLC, P.; VACCARI, F. (1993). "Zoning of the Italian territory in terms of expected peak ground acceleration derived from complete synthetic seismograms". In: CASINIS, R.; HELBIG, K.; PANZA, G. F. [eds.]. "Geophysical exploration in areas of complex geology". *J. Appl. Geophys.*, 30, 149-160.
- DECANINI, L.; MOLLAJOLI, F.; PANZA, G. F.; ROMANELLI, F.; VACCARI, F. (2001). "Probabilistic vs deterministic evaluation of seismic hazard and damage earthquake scenarios: a general problem, particularly relevant for seismic isolation". In: *Proceedings of the 7th International Seminar on Seismic Isolation, Passive Energy Dissipation and Active Control of Vibrations of Structures: 2-5 October 2001: Assisi*.
- ELLSWORTH, W. L.; MATTHEWS, M. V.; NADEAU, R. M.; NISHENKO, S. P.; REASENBERG, P. A.; SIMPSON, R. W. (1998). "A physically-based earthquake recurrence model for estimation of long-term earthquake probabilities". In: *Proceedings of the Second Joint Meeting of the UJNR Panel on Earthquake Research*. 135-149.
- EUROCODICE 8 (1993). Eurocode 8 structures in seismic regions — design — part 1 general and building, Doc TC250/SC8/N57A.
- FIELD, E. H.; THE SCEC PHASE III WORKING GROUP (2000). "Accounting for site effects in probabilistic seismic hazard analyses of Southern California: overview of the SCEC Phase III Report". *Bull. Seism. Soc. Am.*, 90, S1-S31.
- FLORSCH, N.; FAEH, D.; SUHADOLC, P.; PANZA G. F. (1991). "Complete synthetic seismograms for high frequency multimode love waves". *Pure Appl. Geophys.*, 136, 529-560.
- GABRIELOV, I.; KEILIS-BOROK, V.; JACKSON, D. D. (1996). "Geometric incompatibility in a fault system". *Proc. Natl. Acad. Sci. USA*, 93 (9): 3838-3842.
- GELFAND, I.; GUBERMAN, SH.; IZVEKOVA, M.; KEILIS-BOROK, V. I.; RANTSMAN, E. (1972). "Criteria of high seismicity, determined by pattern recognition". *Tectonophysics*, 13, 415-422.
- GELLER, R. J. (1997). "Earthquakes cannot be predicted". *Geophys. J. Int.*, 131, 425-450.
- GORSHKOV, A.; KUZNETSOV, I.; PANZA, G.; SOLOVIEV, A. (2000). "Identification of future earthquake sources in the Carpatho-Balkan orogenic belt using morphostructural criteria". *Pure and Appl. Geophys.*, 157, 79-95.
- GORSHKOV, A.; PANZA, G. F.; SOLOVIEV, A. A.; AODIA, A. (2002). "Morphostructural zonation and preliminary recognition of seismogenic nodes around the Adria margin in peninsular Italy and Sicily". *JSEE. J. of Seismology and Earthquake Engineering*. [Submitted]
- KAGAN, Y. Y. (1997). "Are earthquakes predictable?" *Geophys. J. Int.*, 131, 505-525.
- KAGAN, Y. Y.; JACKSON, D. D. (1995). "New seismic gap hypothesis: five years after". *J. Geophys. Res.*, 99, 3943-3959.
- KEILIS-BOROK, V. I. (1996). "Intermediate term earthquake prediction". *Proc. Natl. Acad. Sci. USA*, 93, 3748-3755.

- KEILIS-BOROK, V. I.; KNOPOFF, L.; ROTWAIN, I. M. (1980). "Bursts of aftershocks, long term precursors of strong earthquakes". *Nature*, 283, 259-263.
- KEILIS-BOROK, V. I.; KOSSOBOKOV, V. G. (1987). "Periods of high probability of occurrence of the world's strongest earthquakes". *Computational Seismology* [Allerton Press Inc.], 19, 45-53.
- KEILIS-BOROK, V. I.; KOSSOBOKOV, V. G. (1990). "Premonitory activation of seismic flow: algorithm M8". *Phys. Earth Planet. Inter.*, 61, 73-83.
- KEILIS-BOROK, V. I.; KUTZNETSOV, I. V.; PANZA, G. F.; ROTWAIN, I. M.; COSTA, G. (1990). "On intermediate-term earthquake prediction in Central Italy". *Pure and Appl. Geophys.*, 134, 79-92.
- KEILIS-BOROK, V. I.; ROTWAIN I. M. (1990). "Diagnosis of time of increased probability of strong earthquakes in different regions of the world: algorithm CN". *Phys. Earth Planet. Inter.*, 61, 57-72.
- KNOPOFF, L. (1996). "Earthquake prediction: the scientific challenge". *Proc. Natl. Acad. Sci. USA*, 93.
- KOSSOBOKOV, V. G.; KEILIS-BOROK, V. I.; SMITH, S. W. (1990). "Reduction of territorial uncertainty of earthquake forecasting". *Phys. Earth Planet. Inter.*, 61, R1.
- KOSSOBOKOV, V. G.; ROMASHKOVA, L. L.; KEILIS-BOROK, V. I.; HEALY, J. H. (1999). "Testing earthquake prediction algorithms: statistically significant advance prediction of the largest earthquakes in the Circum-Pacific, 1992-1997". *Phys. Earth Planet. Inter.*, 111, 187-196.
- KOSSOBOKOV, V. G.; ROMASHKOVA, L. L.; PANZA, G. F.; PERESAN, A. (2002). "Stabilizing intermediate-term medium-range earthquake predictions". *J. of Seismology and Earthquake Engineering*. [Submitted]
- LOMNIZ, C. (1994). *Fundamentals of earthquake prediction*. J.Wiley and Sons Inc., New York.
- MELETTI, C.; PATACCA, E.; SCANDONE, P. (2000). "Construction of a seismotectonic model: the case of Italy". *Pure and Appl. Geophys.*, 157, 11-35.
- MOGI, K. (1992). "Earthquake prediction research in Japan". In: BOSCHI, E.; DRAGONI, M. [eds.]. *Earthquake prediction: Proceedings of the International School of Solid Earth Geophysics, 5th course*. "Ettore Majorana" Centre for Scientific Culture, Istituto Nazionale di Geofisica, 11-31.
- MOLCHAN, G. M.; DMITRIEVA, O. E.; ROTWAIN, I. M.; DEWEY, J. (1990). "Statistical analysis of the results of earthquake prediction, based on bursts of aftershocks". *Phys. of the Earth and Planet. Int.*, 61, 128-139.
- MOLCHAN, G. M.; KRONROD, T. L.; PANZA, G. F. (1997). "Multiscale seismicity model for seismic risk". *Bull. Seismol. Soc. Am.*, 87, 5, 1220-1229.
- MOLIN, D.; STUCCHI, M.; VALENSISE, G. (1996). *Massime intensità macrosismiche osservate nei comuni italiani. Elaborato per il Dipartimento della Protezione Civile*. GNDT, ING, SSN, Roma.
- NARKUNSKAYA, G. S.; SHNIRMAN, M. G. (1994). "On an algorithm of earthquake prediction". *Computational Seismology and Geodynamics* (AGU, Washington, D. C.), 1, 20-24.
- NISHENKO, S. P. (1991). "Circum-Pacific seismic potential: 1989-1999". *Pure and Appl. Geophys.*, 135, 169-259.

- PANZA, G. F. (1985). "Synthetic seismograms: the Rayleigh waves model summation". *J. Geophys.*, 58, 125-145.
- PANZA, G. F.; CAZZARO, R.; VACCARI, F. (1997). "Correlation between macroseismic intensities and seismic ground motion parameters". *Annali di Geofisica*, 5, 1371-1381.
- PANZA, G. F.; RADULIAN, M.; TRIFU, C. [eds.] (2000). "Seismic hazard in the circum pannonian region", PAGEOPH Topical Volume, Birkhauser Verlag, 280.
- PANZA, G. F.; VACCARI, F.; CAZZARO, R. (1999). "Deterministic seismic hazard assessment". In: WENZEL, F. [et al.] [eds.]. *Vrancea Earthquakes: Tectonics, Hazard and Risk Mitigation*, 269-286. Kluwer Academic Publishers, The Netherland.
- PANZA, G. F.; VACCARI, F.; COSTA, G.; SUHADOLC, P.; FAH, D. (1996). "Seismic input modelling for zoning and microzoning". *Earthquake Spectra*, 12 (3), 529-566.
- PANZA, G. F.; VACCARI, F.; ROMANELLI, F. (2001). "Seismic wave propagation in laterally heterogeneous anelastic media: theory and applications to the seismic zonation. Advances in Geophysics". *Academic Press*, 43, 1-95.
- PERESAN, A.; COSTA, G.; PANZA, G. F. (1999). "Seismotectonic model and CN earthquake prediction in Italy". *Pure and Appl. Geophys.*, 154, 281-306.
- PERESAN, A.; COSTA, G.; VACCARI, F. (1997). *CCI1996: The current catalogue of Italy*. The Abdus Salam International Centre for Theoretical Physics, Trieste. [Internal report]
- REID, H. F. (1910). "The California earthquake of April 19, 1906". In: *The mechanics of the earthquake*. Vol. 2. Carnegie Institution of Washington, Washington.
- ROELOFFS, E.; QUILTY, E. (1997). "Water level and Strain changes preceding and following the August 4, 1985 Kettleman Hills, California, earthquake". *Pure and Appl. Geophys.*, 149, 21-61.
- ROMASHKOVA, L. L.; KOSSOBOKOV, V. G.; PERESAN, A.; PANZA, G. F. (2001). *Intermediate-term earthquake prediction algorithm M8: a new spatially stabilized application in Italy*. The Abdus Salam International Centre for Theoretical Physics. ICTP, Trieste. [Internal report]
- ROTWAIN, I. M.; NOVIKOVA, O. (1999). "Performance of the earthquake prediction algorithm CN in 22 regions of the world". *Phys. Earth Planet. Inter.*, 111, 207-213.
- SCHWARTZ, D. P.; COPPERSMITH, K. J. (1984). "Fault behaviour and characteristic earthquakes: examples from the Wasatch and San Andreas fault zones". *J. Geophys. Res.*, 89, 5681-5698.
- SHIMAZAKI, K.; NAKATA, T. (1980). "Time-predictable recurrence model for large earthquakes". *Geophysical Research Letters*, 7, 279-282.
- WAKITA, H.; NAKAMURA, Y.; SANO, Y. (1988). "Short-term and intermediate-term geochemical precursors". *Pure and Appl. Geophys.*, 126, 267-278.
- WALLACE, R. E.; DAVID, J. F.; McNALLY, K. C. (1984). "Terms for expressing earthquake potential, prediction and probability". *Bull. Seism. Soc. Am.*, 74, 1819-1825.
- WELLS, D. L.; COPPERSMITH, K. J. (1994). "New empirical relationships among magnitude, rupture length, rupture width, and surface displacement". *Bull. Seism. Soc. Am.*, 84, 974-1002.

- WYSS, M. (1997). "Second round of evaluation of proposed earthquake precursors". *Pure and Appl. Geophys.*, 149, 3-16.
- WYSS, M.; HABERMANN, R. E.; BODIN, P. (1992). "Seismic quiescence: a test of the hypothesis and a precursor to the next Parkfield, California, earthquake". *Geophys. J. Int.*, 110, 518-536.

Appendix I: Earthquake precursors

Scientific earthquake prediction consists in the “specification of the expected magnitude, geographic location and time of occurrence of a future event, with sufficient precision that the ultimate success or failure of a prediction can be evaluated” (Wallace *et al.*, 1984). Hence, predictions based on some premonitory phenomena, which can be observed at or near the surface of the Earth, must be quantifiable and statistically significant. According to the IASPEI (International Association of Seismology and Physics of the Earth’s Interior) Subcommittee on Earthquake Prediction (Wyss, 1997), the validation criteria for precursor candidates are, briefly, the following: 1) the observed anomaly should be related to some mechanism leading to earthquakes; 2) the anomaly should be simultaneously observed at more than one site or instrument; 3) the definition of the anomaly and of the rules for its association with subsequent earthquakes should be precise; 4) both anomaly and rules should be derived from an independent set of data, than the one for which the precursory anomaly is claimed.

An earthquake occurs due to the fracturing or, more frequently, the slipping of lithospheric blocks along a pre-existing fault, under the action of the tectonic stresses generated by plate motions. Consequently, possible precursors could be those phenomena that may take place in the lithosphere during the accumulation of stresses. One idea is, for example, that high stresses around the impending source area may induce recognisable inelastic processes, such as creep or stress weakening. Indeed, earthquakes respond to stress changes, but the response is complex, as demonstrated by the fact that often aftershocks occur on planes for which the shear stress should have been reduced by the main shock. Moreover, according to Mogi (1992), the strongest events ($M \geq 8$) could be characterised by quite clearly visible related phenomena, even different from time to time, while for the smaller events the focal area is not as big and consequently precursors may be very local and not discernible (e.g. a possible quiescence in the nucleation area).

Among the several “signals” which have been indicated as possible precursors relevant to earthquake prediction, there are (Lomnitz, 1994; Geller, 1997; Kagan, 1997):

- variations in the seismic activity;
- changes in the velocity and in the spectral content of the seismic waves and of the earthquakes sources;
- crustal deformations in a wide area;
- anomalous variations of the stress in the crust;
- changes in the gravitational and geomagnetic fields;
- geoelectrical precursors (e.g. resistivity variations);
- anomalous changes in the underground water level and in its content of different chemical components (Rn , F , CO_2 , NO_3);
- anomalies in the atmospheric pressure, temperature and heat flow.

Some of these phenomena are consistent with the dilatancy model, nevertheless most of them have been claimed as precursors only by retrospective analysis (i.e. they have been associated to a given earthquake only after it occurred). Special attention must be paid to validate precursors, because earthquakes are too infrequent and each phenomenon has its own non-seismic natural variations and to establish a precursory connection requires observations of many cases. The IASPEI Preliminary List of Significant Precursors issued on March 1994 (Wyss, 1997), for example, includes only five possible precursors, out of the forty precursory observations submitted until now: one is based on ground water chemistry (Wakita, 1988), one is a measurement of crustal deformation by ground water levels (Roeloffs & Quilty, 1997) and three are seismicity patterns. The seismic precursors are based on *foreshocks* (several days before the event), on *presocks swarms* (several months before the earthquake) and on the *seismic quiescences* that can be observed before major aftershocks. Four out of these five precursors, however, are based on a single example; only the last premonitory pattern, which should allow for the prediction of the strong aftershocks, has been observed in several cases. These five precursory observations seem to deserve further study in the framework of earthquake prediction, nevertheless none of them can be considered yet as a validated precursor.

Appendix II: Seismicity patterns

The seismicity patterns candidates to be precursors of an impending strong earthquake generally reflect some underlying hypothesis on the dynamics of earthquake generation. For example, based on Reid's (1910) theory of elastic rebound, it should be possible to study the time intervals between the earthquakes in a given region and magnitude range (*earthquake-recurrence hypothesis*). Paleoseismicity studies, however, evidenced that the intervals of time between the strongest earthquakes on a given fault may have large variability. Besides, the statistical distribution of the inter-event times provided by paleoseismicity is poorly constrained and the extrapolation from a brief set of recent data is still insufficient to provide adequate information. Therefore, a long time elapsed since the last occurrence of a strong earthquake does not guarantee that the next event is imminent, because it depends on the actual inter-event time distribution, which is unknown. Several prediction methods, however, rely on the earthquake-recurrence hypothesis. The best known is maybe the seismic-gap model (Nishenko, 1991), which is based on the following assumptions: *a*) earthquakes occur on recognised faults, thus locations and mechanisms are known; *b*) large events are quasi-periodic (i.e. after a strong earthquake the probability of a strong shock is low); *c*) the frequency-magnitude distribution of events on a single fault is described by the characteristic earthquake model. Tests in forward predictions are still ongoing for this prediction method; nevertheless, it has been shown by Kagan and Jackson (1995) that, up to now, the seismic gap pattern can be

rejected with 95 % of confidence level. In fact, the strong earthquakes have occurred preferentially near the sites of previous, recent large events.

The linear elastic Reid's approach has been recently replaced by the perception that the Earth's crust brittle part behaves as a dissipative non-linear system, possibly characterised by self-organization and chaotic behaviour. In such kind of systems the evolution is very sensitive to small perturbations and to the initial conditions; thus, the symptoms of an impending earthquake may be different from time to time and for different regions. Nevertheless, some grossly averaged regularities have been found in the earthquake sequences preceding the strong events. The observed premonitory patterns possess some basic features (Keilis-Borok, 1996), which are common to many non-linear systems before collapse; in particular, the response to a perturbation 1) increases, 2) becomes more chaotic, and 3) acts at large distances. In our case, the non-linear system corresponds to the system of seismically active faults, while the small earthquakes are the sources of the perturbation to the system. The premonitory seismicity patterns, which have been observed in a time interval of 2-3 years before strong earthquakes, are the following:

1. increase of the seismic activity, clustering of the earthquakes in time and space, and spatial concentration of sources; in other words, the increase of the response to the perturbation;
2. increase of the variation of seismicity and its clustering, which reflects the chaotic response to the perturbation;
3. long-range interaction of earthquakes, which can be interpreted as an increase of the range of influence of the perturbation.

Among the formally defined premonitory patterns, we recall the following:

- the *burst of aftershocks* (Keilis-Borok *et al.*, 1980; Molchan *et al.*, 1990), which is associated to moderate magnitude events characterised by a large number of aftershocks;
- the *seismic quiescence* (Wyss *et al.*, 1992);
- the relative increase of the *b-value* for the moderate events, with respect to smaller events (Narkunskaya & Shnirman, 1994);
- the increase of the *spatial correlation* in the earthquake flow and the log-periodic variations of the earthquake flow on the background of its exponential rise (Bufe *et al.*, 1994).

Considering that the process of preparation of a large earthquake may involve a system of faults rather than a single fault, the seismic precursors must be searched within regions having linear dimensions of several hundred kilometres, in order to take into account possible long-range correlations. In fact, non-local precursors, including migrations of seismicity, may reflect some underlying large-scale process (e.g. the movement of tectonic plates), which cannot be explained by the simple post-seismic stress redistribution in an elastic medium, but are typical in self-organized critical systems. During the stage of earthquake preparation, it is not clear

where and when the earthquake will occur inside the considered region; then, as the time approaches, the precursors tend to cluster around the impending source zone. This observation allows for a second order approximation in predictions (Kossobokov *et al.*, 1990), evaluating premonitory seismicity patterns in a lower space and magnitude range, compatibly with the completeness of the available data.

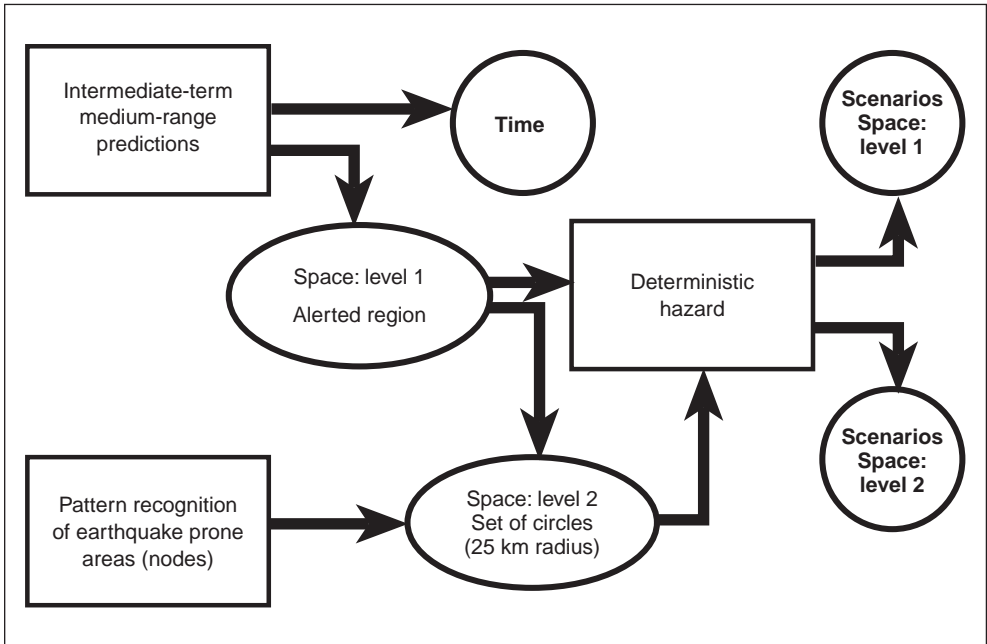


FIGURE 1. Flow chart of the proposed procedure: earthquake predictions + deterministic hazard → scenarios (space level 1); earthquake predictions + pattern recognition + deterministic hazard → scenarios (space level 2).

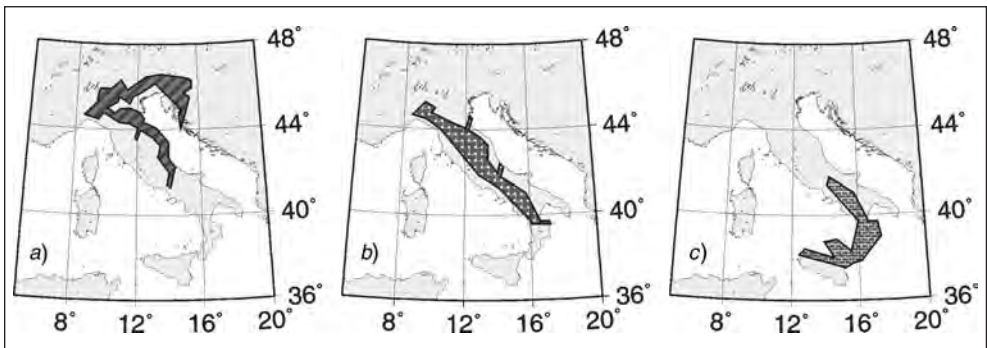


FIGURE 2. Regionalization defined, on the basis of the seismotectonic model, for CN application to the Italian territory: a) Northern Region; b) Central Region; c) Southern Region (Peresan *et al.*, 1999).

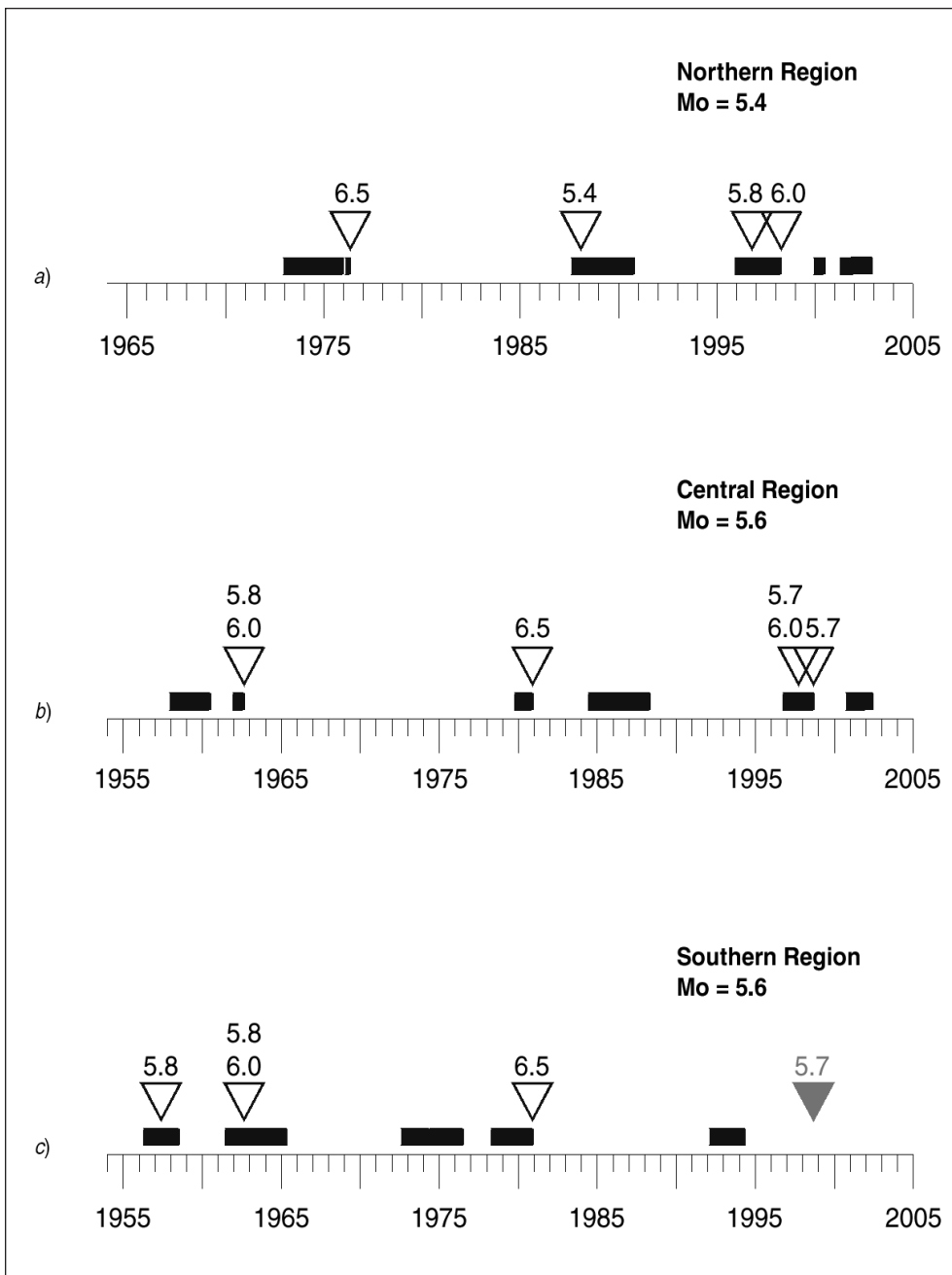


FIGURE 3. Diagrams of the Time of Increased Probability (TIPs) obtained for the three Italian regions in the monitoring of seismicity (updated: January 1, 2002). Black boxes represent the periods of alarm, while a triangle with a number above indicates the occurrence of a strong event together with its magnitude. The catalogue used for the monitoring is the UCI2001 (Romashkova *et al.*, 2001).

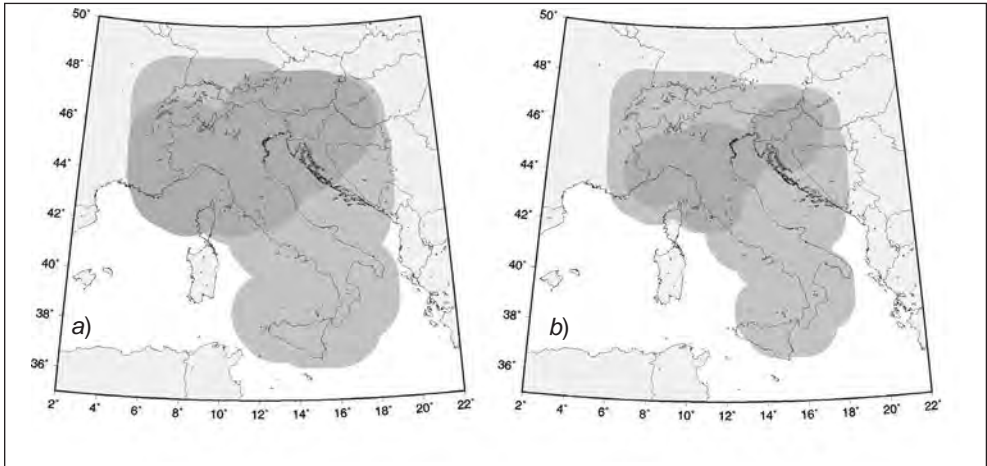


FIGURE 4. Results of the application of the M8 algorithm in Italy for the prediction earthquakes in two magnitude ranges: M6.5 + (fig. 4a) and M6.0 + (fig. 4b). The grey circles, both light and dark, outline the territory where the algorithm M8 has been applied; the dark ones display the alarm area (updated: January 1, 2002).

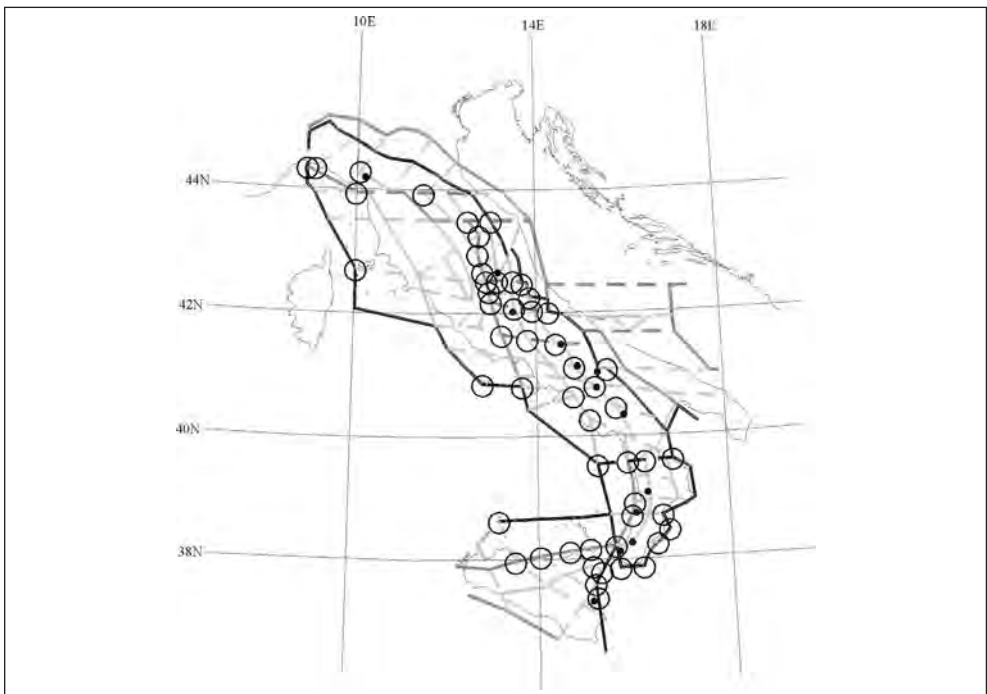


FIGURE 5. Morphostructural map of peninsular Italy and Sicily. Circles indicate the nodes identified as prone to earthquakes with $M \geq 6.5$; the black dots denote the epicentres of the events with $M \geq 6.5$ reported in the considered catalogue (from Gorshkov *et al.*, 2002).

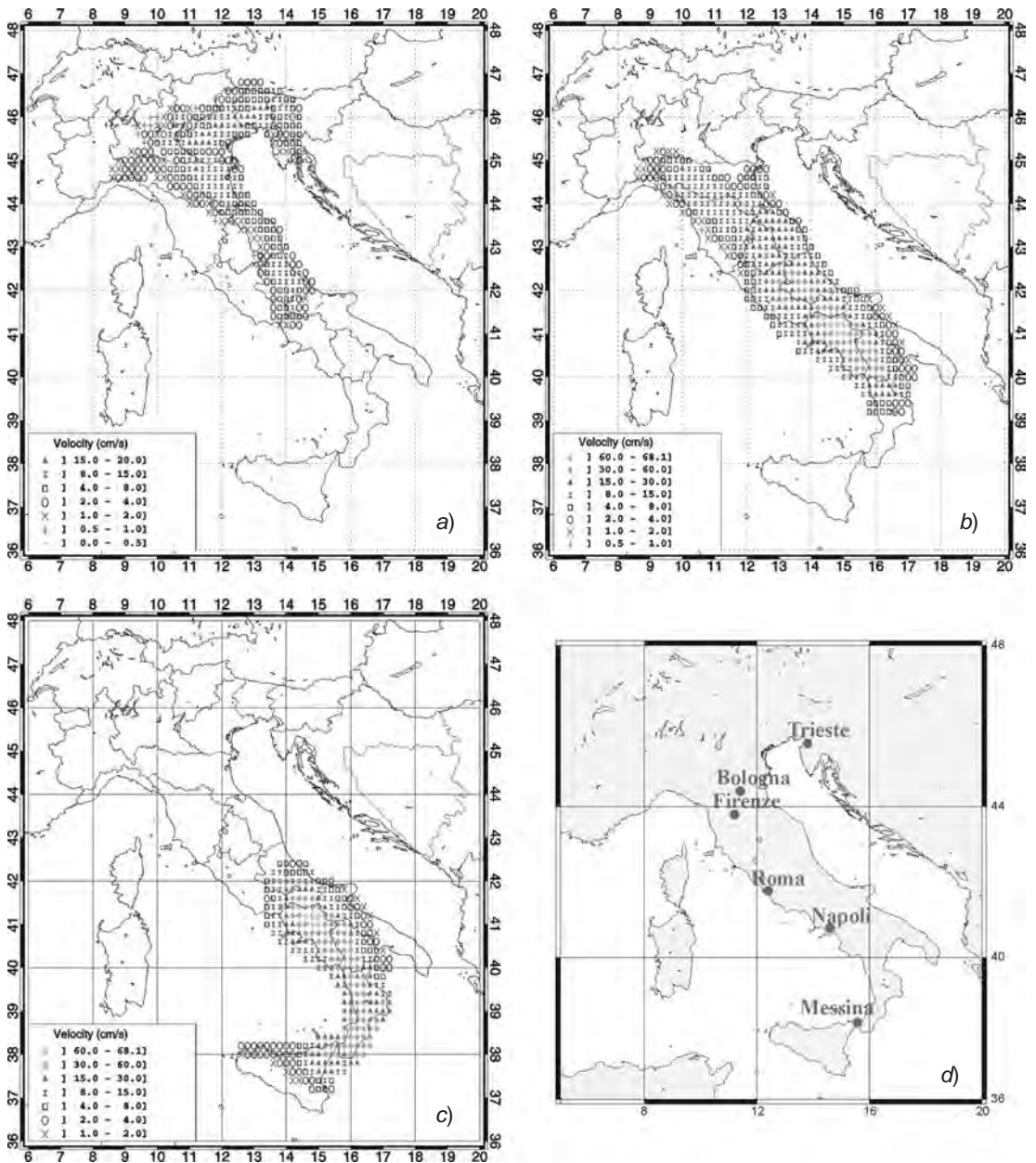


FIGURE 6. Maps of horizontal velocities computed, according to Panza *et al.* (1999), for the a) Northern, b) Central and c) Southern regions; d) map of the cities considered in Table 1.

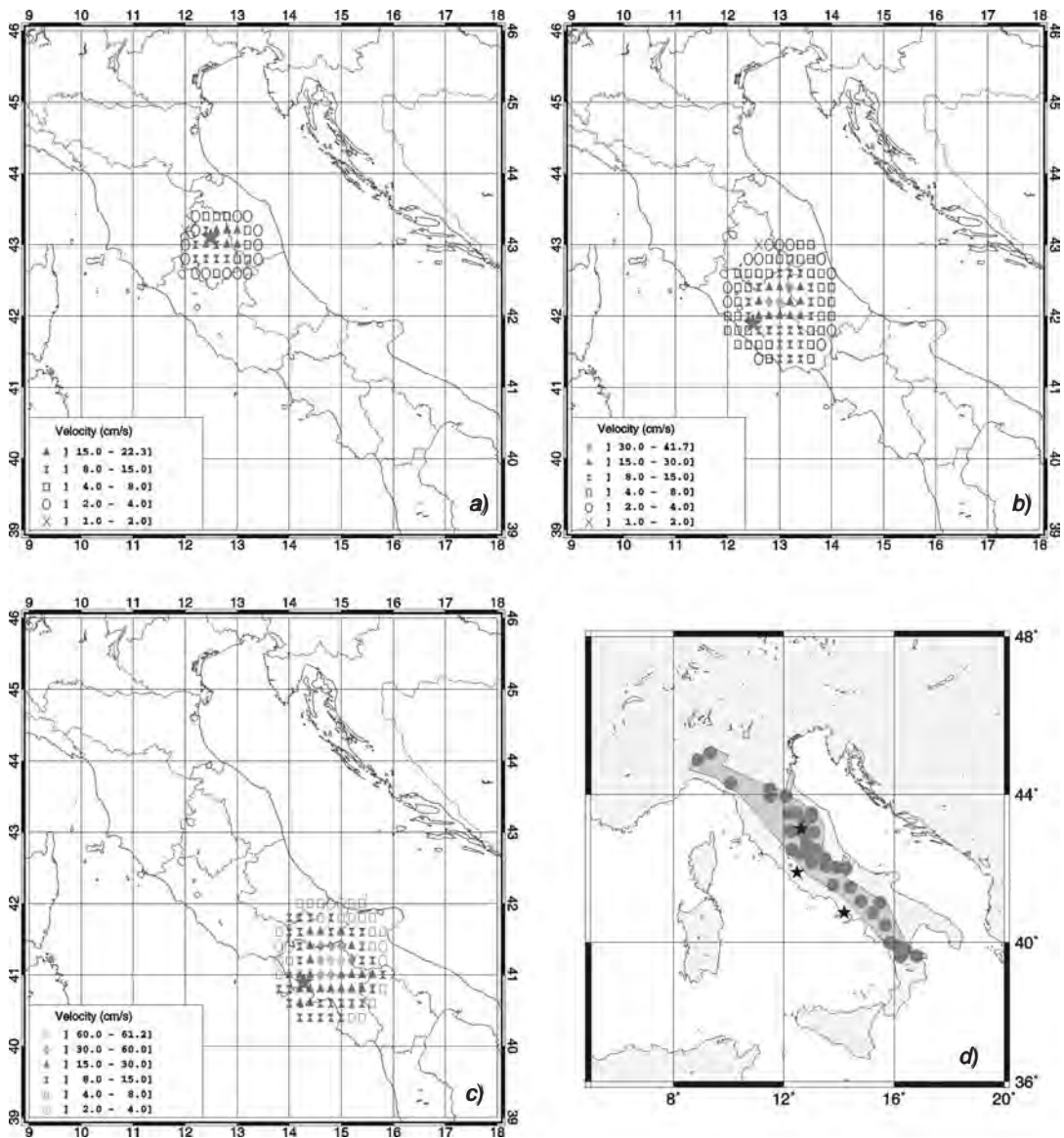


FIGURE 7. Maps of horizontal velocities generated for the nodes closest to the cities of *a*) Assisi, *b*) Rome and *c*) Naples and *d*) map of the nodes prone to earthquakes with $M \geq 6.0$ (light gray) or $M \geq 6.5$ (dark gray), within the Central region. Cities are indicated by stars.

Abstract

Complexity describes the transition from order to chaos, where nonlinear interactions inextricably link space and time, where the whole is greater than the sum of its parts. Complexity may be seen in patterns that can reproduce at different scales embracing a hierarchy of interactions, where fractal structures are often produced, yet preserving a subtle sensitivity in its detail to its starting point.

We begin our discussion using a simple example from geophysics germane to the earth's interior and to its atmosphere, the emergence of convection by heating a fluid from below. We will discuss the bifurcation of such convective cells and the concept of a cascade in turbulence, and ideas relevant to vortex motion in the atmosphere and the Earth's dynamo. Two of the above are "infinite degree of freedom" systems—we will consider Burgers' equation as a model for an inverse cascade and explore the role of nonlinearity. We then consider how to reduce infinite degree of freedom systems to low degree of freedom systems, i.e., ordinary differential equations and mappings. We will begin with the Bullard and Rikitake models for the Earth's dynamo, and follow this discussion with the Lorenz model for turbulence.

We then move our discussion to the issue of mappings, beginning with the Feigenbaum map. We then graduate to percolation models, focusing on the emergence of criticality and relevance to modeling forest fires and earthquake events. Slider-block models are considered as well as their non-inertial equivalents, so-called "sand-box" models. We present fiber bundle models as a description of hierarchical structure, and then progress to colliding-cascades model for seismicity which have been successful recently in describing all observed scalings present in seismicity but have led as well to the discovery of new scalings.

Resum

La complexitat descriu la transició de l'ordre al caos, on les interaccions no lineals enllacen inextricablement l'espai i el temps i on el tot és més gran que la suma de les parts. La complexitat es pot veure en models que es poden reproduir a escales diferents i que abasten una jerarquia d'interaccions, en què sovint s'esdevenen estructures fractals, i que, tanmateix, mantenen una sensibilitat subtil en el seu detall respecte al punt inicial.

Comencem la nostra argumentació amb un exemple senzill procedent de la geofísica i relatiu a l'interior de la Terra i a la seva atmosfera: l'emergència de convecció mitjançant l'escalfament des de sota d'un fluid. Abordarem la bifurcació d'aquestes cèl·lules convectives i el concepte d'una cascada en turbulència, així com també algunes altres idees referents al moviment de vòrtex a l'atmosfera i la dinamo de la Terra. Tot això suava esmentat són sistemes amb infinits graus de llibertat: considerarem l'equació de Burgers com a model per a una cascada a la inversa i hi explorarem el paper de la no-linealitat. Després veurem la manera de reduir els sistemes d'un grau de llibertat infinit a sistemes amb un grau de llibertat baix; per exemple, equacions diferencials i mapatges normals. Començarem pels models de Bullard i Rikitake per a la dinamo de la Terra i continuarem amb el model de Lorenz per a la turbulència.

Aleshores durem el debat cap al tema dels mapatges, començant per l'aplicació de Feigenbaum. Després passarem als models de percolació centrant-nos en l'aparició de la criticitat i la rellevància per a crear models d'incendis forestals i terratrèmols. També hi abordarem els models de bloc lliscant, així com els seus equivalents no inerts, els anomenats *models de caixa de sorra*. Hi presentem models de feix de fibres com a descripció d'una estructura jeràrquica i, en acabat, passem al model de cascades en col·lisió per a la sismicitat, el qual recentment ha funcionat bé en la descripció de tots els escalatges observats presents en la sismicitat i que també ha dut al descobriment de nous escalatges.

THE MATHEMATICS AND SIMULATION OF COMPLEX SYSTEMS. I

William I. Newman, UCLA

Outline

- Introduction
- FDEs: Feigenbaum Map
- ODEs: Competition; Convection and Lorenz Model
- PDEs: Turing Model and Solitary Waves
- Turbulent Cascades, Renormalization, and Fractals

Complexity describes ...

- Transition from order to chaos (sensitivity to initial conditions)
- Nonlinear interactions that link space and time
- Pattern formation on many scales, including hierarchy of interactions, cascades, and emergence of fractal structures
- Situations where the whole is greater than sum of its parts
- Complexity lacks universally accepted definition
- Physical problems generally correspond to a high (∞) degree of freedom (DOF) system
- *Challenge*: How can we render complex problems mathematically and computationally tractable?

— or —

How can we simplify complexity?

Conceptual breakthroughs ...

- Ruelle and Takens (1971): many ∞ -DOF systems relax, via dissipation, to finite DOF behavior; turbulence
- Implication: if we're clever enough, we can reduce partial differential equations (PDEs) to finite set of ODEs
- May and Oster (1976) and Feigenbaum (1978) showed that complex behavior (cf. complexity) can emerge from finite difference equations (FDEs) or maps generated by ODEs in biological examples
- Complexity is more involved—necessarily has high DOF but could be describable by coupled ODEs or FDEs (cellular automata) instead of PDEs
- Introduce topic via biological examples, because of simplicity, before turning to geophysical ones

From Feigenbaum map to predator prey

- Growth law for population ρ undergoing Malthusian growth with rate γ with $\rho(t=0) = \rho_0$ and saturating at ρ_s

$$\frac{d\rho}{dt} = \gamma\rho \left(1 - \frac{\rho}{\rho_s}\right) \Rightarrow \rho(t) = \frac{\rho_0\rho_s}{\rho_0 + (\rho_s - \rho_0)\exp(-\gamma t)}$$

- Discretize with $t_n \equiv t_0 + n\Delta t$, $\rho_n \equiv \rho(t_n)$; first order explicit integrator (Euler)

$$\rho_{n+1} = \rho_n + \gamma\Delta t\rho_n \left(1 - \frac{\rho_n}{\rho_s}\right) \Rightarrow x_n \equiv \frac{\gamma\Delta t}{1 + \gamma\Delta t} \cdot \frac{\rho_n}{\rho_s}$$

to obtain Feigenbaum or quadratic map, $\alpha \equiv 1 + \gamma\Delta t > 1$,

$$x_{n+1} = \alpha x_n (1 - x_n)$$

- Paragon of stability transformed into a paradigm for chaos

From Feigenbaum map to predator prey (cont)

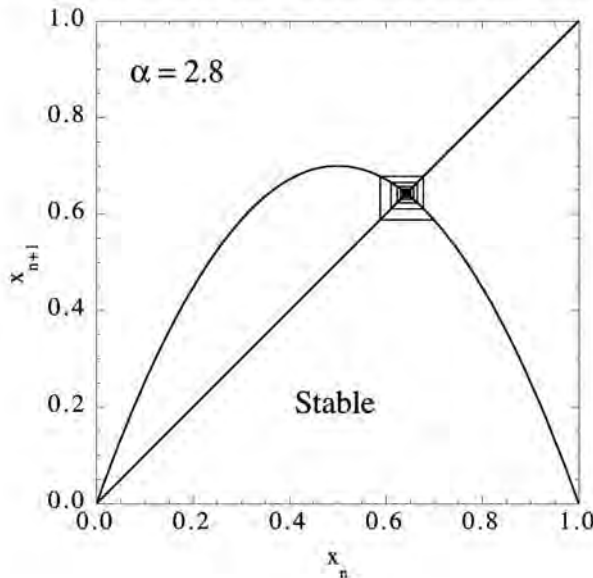
- Mapping has form $x_{n+1} = f(x_n)$, $f(x) = \alpha x(1-x)$
- “Fixed points” satisfy $x = f(x)$, namely 0 and $1 - 1/\alpha$
- Stability of mapping: let \hat{x} be a fixed point;
 $x_{n+1} - \hat{x} = f(x_n) - f(\hat{x}) \approx f'(\hat{x})(x_n - \hat{x})$; therefore,

$$\left| \frac{x_{n+1} - \hat{x}}{x_n - \hat{x}} \right| \approx |f'(\hat{x})| < 1$$

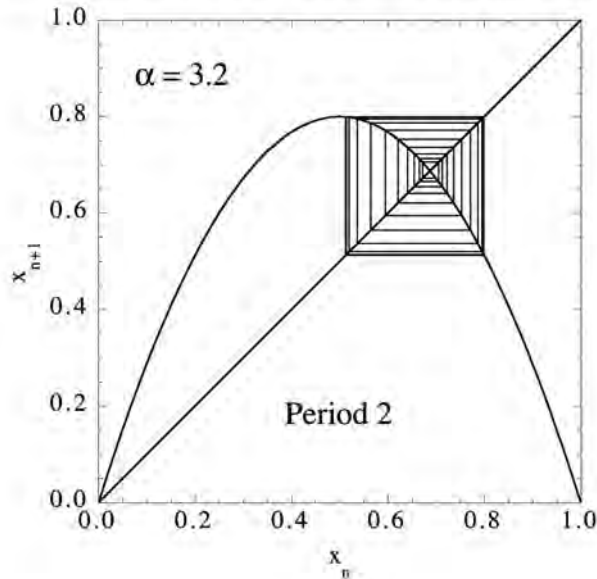
for convergence; strange attractors in higher dimension

- Fixed point at origin, where $f' = \alpha > 1$, always unstable!
- Other fixed point has $|f'| \rightarrow |2 - \alpha|$ which is greater than 1 if $\alpha > 3$; e.g., $\alpha = 2.8$ vs. $\alpha = 3.2$; “deterministic chaos” and pitchfork bifurcation (sequence)
- *Caveat emptor*: numerical integrations must be checked!

From Feigenbaum map to predator prey (cont)



From Feigenbaum map to predator prey (cont)



From Feigenbaum map to predator prey (cont)

- Turn to differential equation models—“predator-prey” or 2-dimensional competition between two species (rabbits R and wolves W), including stationary or fixed points,

$$\frac{dR}{dt} = \gamma_R R (1 - W/W_s); \quad \frac{dW}{dt} = -\gamma_W W (1 - R/R_s)$$

- Overall growth rate of each population depends on the other; (R_s, W_s) is a fixed point (FP); linearization

$$R = R_s + r; \quad W = W_s + w$$

$$\Rightarrow \frac{dr}{dt} = -\frac{\gamma_R R_s}{W_s} w; \quad \frac{dw}{dt} = +\frac{\gamma_W W_s}{R_s} r \Rightarrow \frac{d^2 r}{dt^2} = -(\gamma_W \gamma_R) r$$

- Similar result for $d^2 w/dt^2$; population cycles with frequency $\sqrt{\gamma_W \gamma_R}$

From Feigenbaum map to predator prey (*cont*)

- Amazingly, this nonlinear system has a (nonphysical) constant of motion (reminiscent but unlike Hamiltonian),

$$c = \gamma_W \left[\ln(R) - \frac{R}{R_s} \right] + \gamma_R \left[\ln(W) - \frac{W}{W_s} \right]$$

- Expanding the latter close to the stationary point recovers the nature of the “phase portrait” recovered earlier
- Bullard, Rikitake, and others in late 1960’s developed competition model for Earth’s dynamo, linking mechanical and electrical behavior; failed to get field reversals, etc.
- Generally, Poincaré-Bendixson thm. applies in 2D; system either (a) goes to FP, (b) has periodic orbit around FP, or (c) approaches periodic orbit via “limit cycle”
- Malkus and Robbins realized higher dimension needed

A simple paradigm ...

- Consider Earth’s surface heated by sun and resulting temperature distribution in troposphere
- Ideal gas law and hydrostatic equation govern;

$$P = \rho \frac{k_B T}{\mu}; \quad \frac{dP}{dz} = -\rho g$$

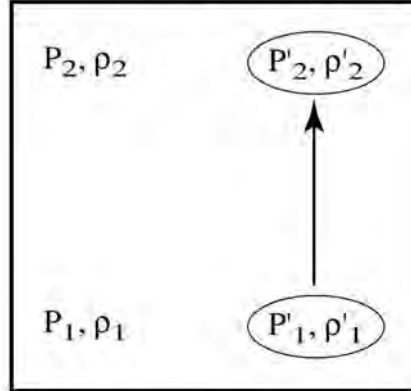
P pressure, ρ density, k_B Boltzmann constant, T temperature, μ mean molecular weight, z altitude, g grav. accel.

- Adiabatic (second law) $P \propto \rho^\gamma$, $\gamma \approx 1.4$ for diatomic atmos.
- “Superadiabatic” temperature gradient (“lapse rate”) produces instability; sustained convection reduces gradient to

$$\left. \frac{dT}{dz} \right|_{\text{ad}} = -\frac{\mu g}{k_B} \cdot \frac{\gamma - 1}{\gamma}$$

A simple paradigm ... (cont)

- “Air parcel” in balloon (primed) perturbed adiabatically
- $\rho'_1 = \rho_1, P'_1 = P_1, P'_2 = P_2; \rho'_2 < \rho_2$ unstable; unavoidable



- Adiabatic lapse rate is marginally stable; quantitative analysis possible via full fluid PDEs

A simple paradigm ... (cont)

- Lorenz (1963) simplification for supercritical behavior; Turcotte (1997), Lichtenberg & Lieberman (1992) derivations

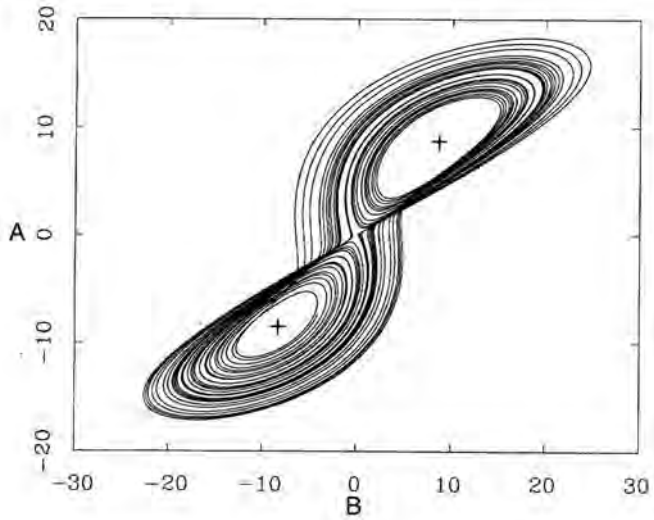
$$\frac{dA}{dt} = \sigma(B - A); \frac{dB}{dt} = rA - B - AC; \frac{dC}{dt} = -bC + AB$$

where t is scaled time, σ is Prandtl number (10 typical), r is normalized Rayleigh number (around 28), and b is geometrical factor (typically 8/3); r “tunable” in experiment

- Nontrivial stationary points at $B = A = \pm\sqrt{b(r-1)}$ (symmetry) and $C = r - 1$ for $r > 1$ (physical constraint)
- Stability criterion emerges from linearized analysis

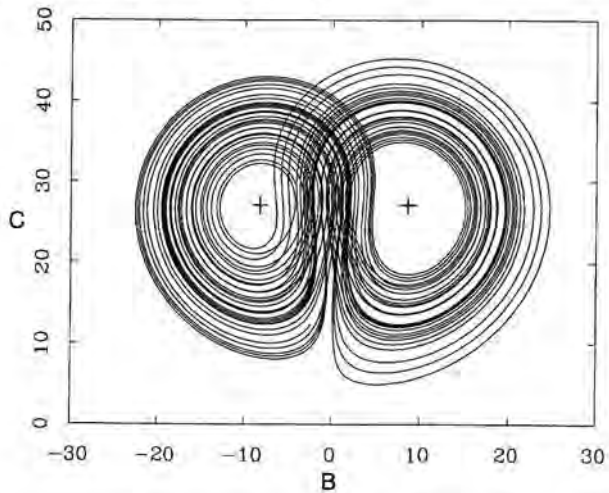
$$r < \frac{\sigma(\sigma + b + 3)}{\sigma - b - 1} = \frac{470}{19} \approx 24.74$$

A simple paradigm ... (cont)



- Appearance of homoclinicity; deterministic chaos

A simple paradigm ... (cont)



- Apparent “self-similar” structure; fractals as attractors

Turing Model and Solitary Waves

- Formation of coherent pattern is symptomatic of many ∞ -DOF systems; linearized analysis often shows sensitivity to initial conditions, hence spatio-temporal chaos
- Original model for spatio-temporal chaos emerged from model for two chemical species A and B governed by

$$\frac{\partial A}{\partial t} = F(A, B) + D_A \nabla^2 A; \quad \frac{\partial B}{\partial t} = G(A, B) + D_B \nabla^2 B$$

- F and G describe nonlinear kinetic interaction; linear analysis ($D_A = D_B = 0$) indicates stability around fixed point
- If $0 \neq D_A \neq D_B$, linearization plus Fourier analysis, i.e., $A \propto B \propto \exp(i\vec{k} \cdot \vec{x})$, around stationary solutions yields instability for a range of \vec{k} ; infrequent in geophysics

Turing Model and Solitary Waves (cont)

- Burgers equation simplest model for shock and turbulence

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2}$$

- Exact solution available from Cole-Hopf transformation

$$u = -2\nu \frac{\theta_x}{\theta} = -2\nu \frac{\partial}{\partial x} [\ln(\theta)] \Rightarrow \frac{\partial \theta}{\partial t} = \nu \frac{\partial^2 \theta}{\partial x^2}$$

- Solution method: transform initial-boundary conditions from u to θ at initial time, solve θ equation via linear methods, transform back to u
- Nonlinear term $u \frac{\partial u}{\partial x}$ produces “mode coupling”; evident from Fourier analysis; allows for “cascades”—inverse (largest to smallest scale; turbulence) and direct (smallest to largest)

Turing Model and Solitary Waves (cont)

- Wave motion generally considered in terms of dispersion relation and Fourier (linear) analysis; $\exp(i\vec{k} \cdot \vec{x} - i\omega t)$, $\omega = \omega(k)$; e.g., seismic or sound waves $\omega = c_s k$
- Wave is dispersive if $\frac{\partial \omega}{\partial k}$ varies with \vec{k} ; wave is diffusive if $\Im[\omega(\vec{k})] < 0$; wave is unstable (growing) if $\Im[\omega(\vec{k})] > 0$
- Solitary wave, e.g., shock structure, permitted by Burgers equation; modes dispersive and interact and grow, balancing dissipation due to diffusion (requires energy source)
- Nonlinear diffusion can also produce solitary waves, but nonlinearity imparts special properties (including robust propagation speeds)
- Potential for application to crack propagation, etc.

Turing Model and Solitary Waves (cont)

- Special subclass of solitary waves have conservative properties, e.g., many—possibly ∞ —moments of a distribution can be preserved \Rightarrow *solitons*
- J. Scott Russell observed shallow water waves that, when well-separated, traveled at their own characteristic speeds, collided, and then re-emerged unscathed in form and moving at their former speed
- Korteweg-de Vries (KdV) equation is simplest nonlinear, dispersive equation, expressed (for algebraic convenience)

$$\frac{\partial u}{\partial t} - 6u \frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3} = 0$$

- Has traveling wave solutions $-\frac{c}{2} \operatorname{sech}^2\left[\frac{\sqrt{c}}{2}(x - ct - x_0)\right]$

Turing Model and Solitary Waves (*cont*)

- Similarity of KdV equation to Burgers advanced by Cole-Hopf transformation which becomes “Inverse Scattering” transformation and yields nonlinear Schrödinger equation
- Quantity $u(x, t)$ has infinite number of conserved quantities, related to special class of moments, discovered by Kruskal and his students at Princeton in the early 1970’s
- Solitons are singular in their possessing an infinite number of conservation laws, yet they have great importance in geophysics, ranging from weakly ionized plasmas to fracture propagation
- Almost all geophysical problems do not share this amazing panoply of properties, just as Hamiltonians seldom occur
- Move focus of lecture from mathematically formal to physically intuitive

Turbulent Cascades, Renormalization, and Fractals

- Stirring of vessel (production range) produces cells which break down into smaller ones (bifurcation sequence), etc. via direct cascade, until dissipation controls smallest scales
- Kolmogorov (1941) scaling: turbule of size ℓ stirred at velocity v_ℓ over time $t_\ell \approx \ell/v_\ell$; kinetic energy per mass of turbule $\approx v_\ell^2 \Rightarrow$ energy dissipation rate per mass

$$\epsilon_\ell \approx v_\ell^2/t_\ell \approx v_\ell^3/\ell \Rightarrow v_\ell \approx (\epsilon_\ell \ell)^{1/3}$$

(Kolmogorov-Obukhov Ansatz); $\epsilon_\ell = \epsilon$ independent of ℓ

- Assumes isotropy, time and spatial averaging—statistical theory; direct cascade (direction unimportant); wavenumber

$$k \propto 1/\ell, \quad v(k) \propto \epsilon^{1/3} k^{-1/3}$$

Turbulent Cascades, Renormalization, and Fractals (*cont*)

- Spectrum in N -dim. requires care in normalization

$$S(k) k^{N-1} dk \approx v^2(k) k^{-1} dk \Rightarrow S(k) \propto k^{-N-2/3}$$

- Previous is example of “renormalization,” notion that physics is scale invariant over “inertial subrange” (apart from production and dissipation subranges)
- Renormalization first developed in quantum mechanics for summing series from integral equations; Ising model is exact as are many others; sometimes, approximate
- Fast Fourier Transform (FFT) and decimation:

$$f_k = \sum_{\ell=0}^{N-1} g_{\ell} \exp\left(2\pi i \frac{k\ell}{N}\right), \quad k = 0, \dots, N-1$$

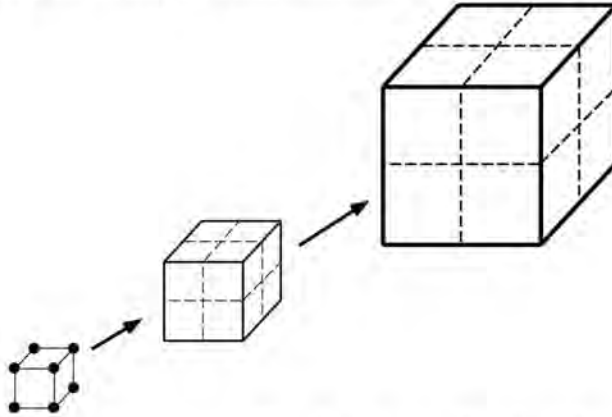
Turbulent Cascades, Renormalization, and Fractals (*cont*)

- Normally, we have $N \times N$ operations
- RHS is sum over $N/2$ even and odd numbers; $\ell = \ell_0 + 2\ell_1$ where $\ell_0 = 0, 1, \ell_1 = 0, \dots, N/2 - 1 \Rightarrow g_{\ell_0, \ell_1} \equiv g_{\ell}$
- Series over N terms can be written over $N/2$ terms,

$$f_k = \sum_{\ell_1=0}^{N/2-1} [g_{0, \ell_1} + g_{1, \ell_1} \exp(2\pi i k/N)] \exp\left(2\pi i \frac{k\ell_1}{N/2}\right)$$

- Former has appearance of FFT over $N/2$; repeat process $\log_2 N$ times, requiring $N \times \log_2 N \ll N^2$ operations
- Role of “self-similarity” is fundamental mathematically and conceptually, the essence of fractals
- Allègre et al. (1982) construction describing failure of a lattice; similar ideas by Turcotte and Newman & Knopoff

Turbulent Cascades, Renormalization, and Fractals (*cont*)



- Eight vertices form small cube, 8 small cubes form a larger one, etc.
- What does this have to do with fractal dimension?

Turbulent Cascades, Renormalization, and Fractals (*cont*)

- What is a fractal? The outcome of an iterative and/or hierarchical *process* is often fractal; renormalization exploits the existence of the iterative hierarchy.
- Richardson (1926) recognized that the “length of the coastline of Britain” depended on the length of the ruler used; if coastline is Euclidean and ruler reduced by factor r and number of ruler segments needed increased by N , then Hausdorff or fractal dimension D satisfies

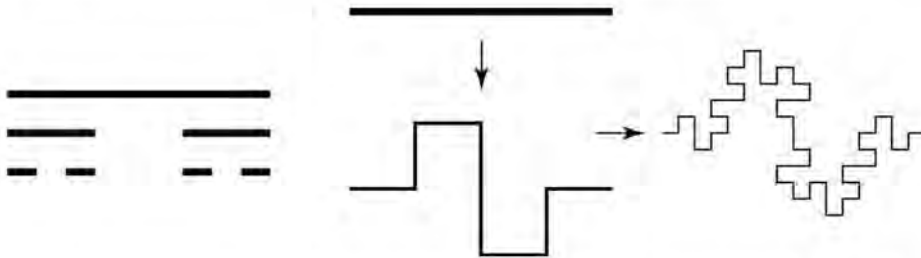
$$Nr^D = 1 \Rightarrow D = \frac{\log(N)}{\log(1/r)}$$

- Allègre et al. construction has $N = 8$, $r = 1/2 \Rightarrow D = 3$, making it “space filling”
- Other “deterministic fractals” shown below

Turbulent Cascades, Renormalization, and Fractals (*cont*)

Cantor Dust $D = 0.6309$

Koch Snowflake $D = 1.5$



- Stochastic fractals randomize aspects of the iteration, e.g., r is drawn from a probabilistic distribution
- Above are self-similar; self-affine structures have different scalings in different dimensions (e.g., Brownian motion)
- Fractal dimension algorithms exist (but lacunarity, multi-fractals); many aspects of the natural world are fractal

From classical to contemporary complexity...

- We have reviewed aspects of mappings, as well as ordinary and partial differential equations germane to complexity
- Contemporary problems in complexity are primarily discrete in structure—in contrast with continuous—and forms the basis of the second lecture

THE MATHEMATICS AND SIMULATION OF COMPLEX SYSTEMS. II

William I. Newman, UCLA

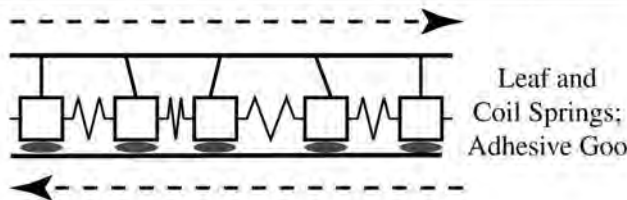
Outline

- Slider-block models
- Dynamical systems and the seismic cycle
- Percolation and lattice models
- Fiber bundle models
- Clustering, universality, and self-organization
- Cascades: inverse, direct, and their collision
- Will our future be complex?

Slider-block models

- Knopoff and Burridge (1967) developed conceptual paradigm for fault behavior based on a system of identical masses connected by identical coil springs (representing elasticity) and identical leaf springs to one fault surface and via adhesive (representing friction) to the opposing fault surface

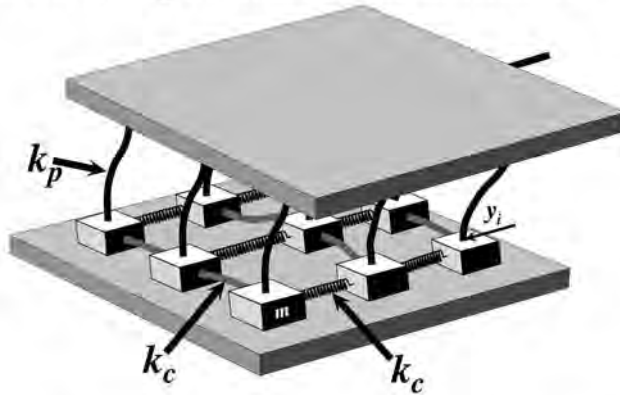
$$m\ddot{u}_i = k(u_{i+1} - 2u_i + u_{i-1}) - \beta(u_i - u_{i0}) - \alpha\dot{u}_i + f_i$$



- Model reproduced Gutenberg-Richter power-law statistics; critical issue—GR statistics apply to fault network

Slider-block models (cont)

- Model revitalized by Carlson and Langer in 1990's
- Extensions to higher dimension now available



- Power-law frequency-magnitude statistics widely regarded as evidence for scale-invariance and self-organization

Dynamical systems and the seismic cycle

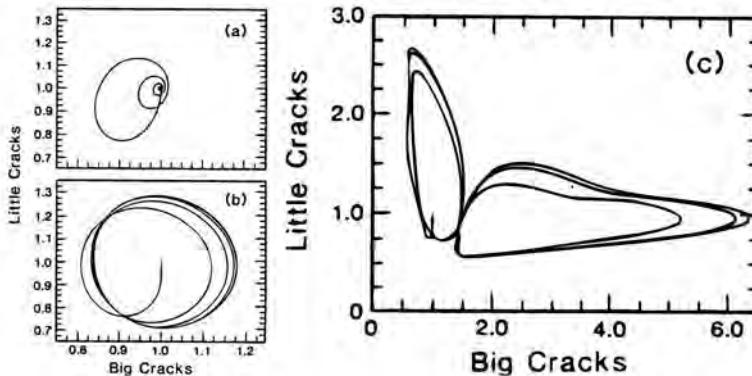
- Newman & Knopoff (1982, 1983) sought a low degree of freedom *phenomenological* model to describe seismic cycle
- Crack fusion model focused on process of fusion of small cracks into successively larger ones, accommodating the influence of main and aftershocks on crack populations, juvenile crack genesis from tectonic stress, crack healing, and anelastic-creep induced time delays, plus other effects
- N & K discovered that critical ingredient of model was time delay due to cascade of crack fusion events from smallest to largest sizes, and stress release
- Thermodynamic model utilizing dynamic renormalization group later developed to help explain universality
- Tertiary phase of brittle failure characterized by accelerating process that culminates in catastrophic failure

Dynamical systems and the seismic cycle (cont)



- Essentially *all* materials display this scaling independent of structure; *Zhurkov's law*
- Slope depends on temperature and humidity, only—water has dramatic catalytic role
- Figure gives outcome of crack fusion model showing role of state-dependent time-delay on relative populations of big and little cracks; strange attractor & chaos?
- Thermodynamic model utilizing dynamic renormalization group later developed to help explain universality and became central to exploring richer problem
- Importantly, time delays render dynamics ∞ -DOF

Dynamical systems and the seismic cycle (cont)

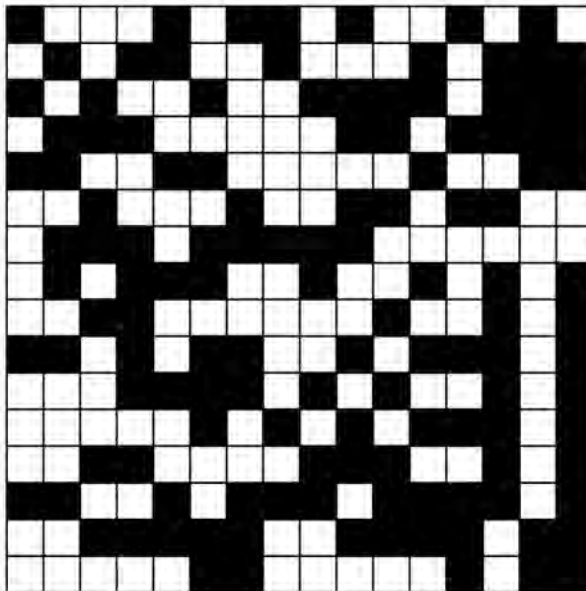


- Remarkably little work has been done on earthquake-related problems using ODEs or continuum models
- Remainder of lecture will focus on discrete systems and, often, discrete time

Percolation and Lattice Models

- Concept of percolation on a lattice emerged in application to epidemiology, conductivity, and forest fires
- Suppose probability that a site on a 2D square lattice is occupied is p ; Monte Carlo results below employs $p = 0.5$
- Critical probability $p_c \approx 0.59275$ corresponds to situation where percolation will occur from one face of the lattice to its opposing face (in infinite lattice limit)
- Significance to above applications clear; can also describe “phase transition” in physical materials
- Relevant to earthquake models where sites correspond to broken elements and percolation corresponds to fracture of the system, but physics for individual site failure required
- Return later to forest fires and percolation

Percolation and Lattice Models (*cont*)



Percolation and Lattice Models (*cont*)

- Bak, Tang, and Wiesenfeld (1988) developed “sandpile model” and claimed new phenomena “self-organized criticality” was present; subtleties negate conclusion but ideas very important
- Consider lattice where discrete units of load are being randomly deposited on sites; sites bearing excess of 4 units load redistribute them uniformly on 4 nearest neighbors
- Avalanche statistics are power-law; now recognized that these models are intimately related to slider-block models in the zero-mass or inertia-free limit (stress transfer)
- Pervasiveness of power-law behavior plus conceptual link provides compelling reason for further investigation

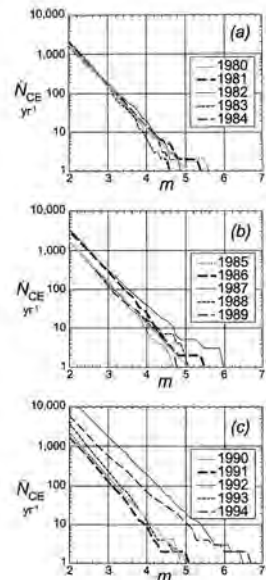
Percolation and Lattice Models (*cont*)

- Cumulative number of earthquakes per year, \dot{N}_{CE} , occurring in southern California with magnitudes $> m$; SCSN Catalog (1995)
- Larger numbers of earthquakes in 1987, 1992, and 1994 attributed to aftershocks of the Whittier, Landers, and Northridge earthquakes
- Solid straight line is Gutenberg-Richter relation

$$\log N = -bm + \log a$$

with $b = 1.05$ and $a = 2.06 \times 10^5 \text{ yr}^{-1}$

- Slider-block models show power laws



Percolation and Lattice Models (*cont*)

Why do we obtain power laws?

- While developing a model that produces power-law behavior is encouraging, since it mimics a feature observed in nature, it is no substitute for an explanation
- Widespread confusion exists in that many believe that power-laws are a product of “self-organized criticality”
- We will show *analytically* (fiber bundle and clustering models) that self-organization to limiting power-laws is to be expected, but that there is nothing “critical” here (apart from the critical probability at the percolation threshold)
- We will begin by considering static and dynamic hierarchical fiber bundle models

Fiber Bundle Models

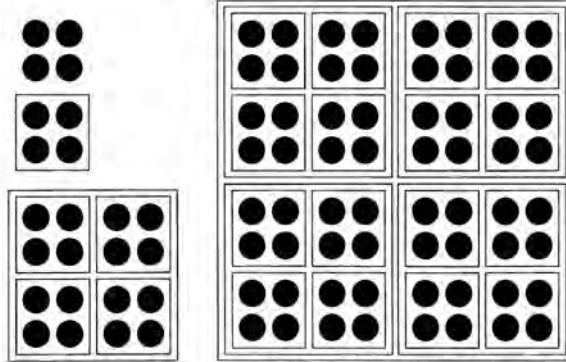
- Smalley, Turcotte, and Sola (1985) imprinted structure of fractal network onto static fiber bundle model to explore “fault weakening”; refined by Newman and Gabrielov (1991)
- Static binary fiber model assumed that fibers organized hierarchically, i.e., in pairs, pairs-of-pairs, etc.; a given unit would fail only if all associated subunits would fail
- Suppose stress σ *per fiber* is initially applied to a binary hierarchical ensemble of fibers where the cumulative probability distribution for failure of a single fiber is $P_0(\sigma)$
- Bundle of order n has 2^n fibers; probability of failure of a pair of order $n + 1$ is probability that both fail or that one fails (2 ways) while survivor fails under load between $\sigma - 2\sigma$, i.e., $P_{n+1}(\sigma) = P_n^2(\sigma) + 2P_n(\sigma)[P_n(2\sigma) - P_n(\sigma)]$

Fiber Bundle Models (*cont*)

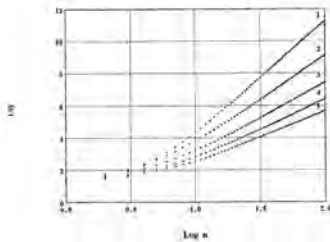
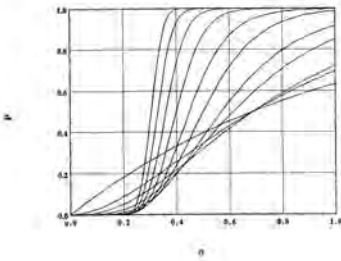
- Iteration simplifies to

$$P_{n+1}(\sigma) = P_n(\sigma) [2P_n(2\sigma) - P_n(\sigma)]$$

- 2D structure in hierarchical fiber bundle with “coordination number” $m = 4$



Fiber Bundle Models (*cont*)



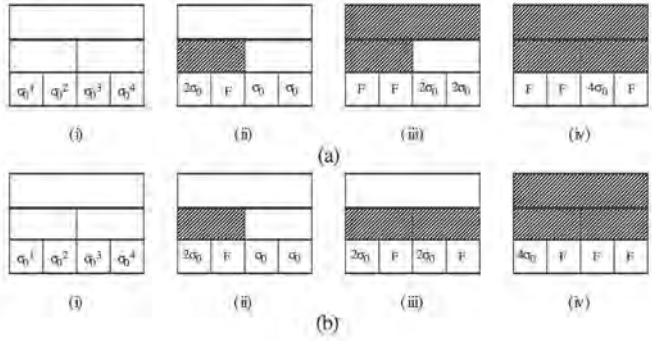
- Figure shows, as a function of stress σ in a bundle with 2^n fibers, the cumulative probability distribution $P_n(\sigma)$ as a function of n and trend to weakness with increasing system size—but *no* critical point for failure (step function)
- Median stress σ_n for failure as a function of n for different Weibull model parameterizations

- Theorem shows $\nu = \frac{\ln 2}{\ln [m / (m - 1)]}$

$$\sigma_n^{-\nu} \propto \log [\log (\text{mass})]$$

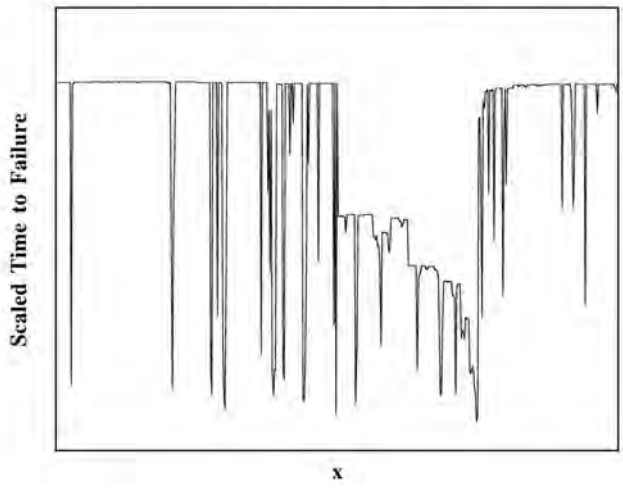
Fiber Bundle Models (cont)

- We have now identified remarkable *universal* scaling features for static fiber bundle model, with power-law index depending in modest way upon geometry
- Turn to dynamic (time-dependent) problem; consider two independent scenarios (allowing for equivalence under interchange): critical time demonstrated analytically



Fiber Bundle Models (cont)

- Inverse cascade shown below; location of failures against logarithm of (remaining) time to failure



Fiber Bundle Models (*cont*)

- Bufe and Varnes (1993), among others, proposed that possible critical or failure time t_f could be established using Benioff strain E given by

$$E = C (t_f - t)^\alpha$$

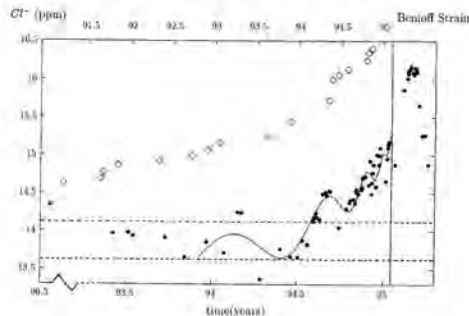
- Perturbation theory allows for α complex, i.e., $\alpha = \xi + i\eta$ in expansion $E = \Re \left\{ C (t_f - t)^{\xi+i\eta} \right\}$ so that

$$E = \Re \left\{ C (t_f - t)^\xi \exp [i\eta \ln (t_f - t)] \right\} = C (t_f - t)^\xi \cos [\eta \ln (t_f - t)]$$

- Benioff strain—plus, possibly, gas release due to dilatancy—might show this power-law approach to failure with superimposed form of intermittency—known as “log periodicity”—emerging from accelerating, oscillating term

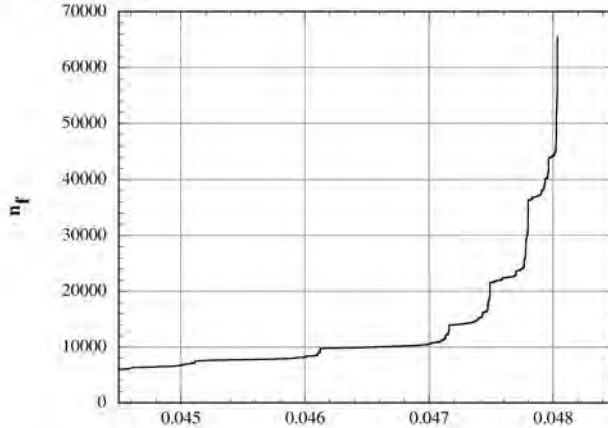
Fiber Bundle Models (*cont*)

- Chlorine concentration (black dots; lower and left scales) from around 1993.5 to 1995.3; cumulative Benioff strain (diamonds; upper and right scales) around epicenter. Time of Kobe earthquake indicated by vertical line showing large jump in ion concentration. Johansen et al. (1996).



Fiber Bundle Models (cont)

- “Benioff strain” from Newman, Turcotte, and Gabriellov (1995) fiber bundle model showing hint of log-periodicity as well as power-law; not always seen in simulations

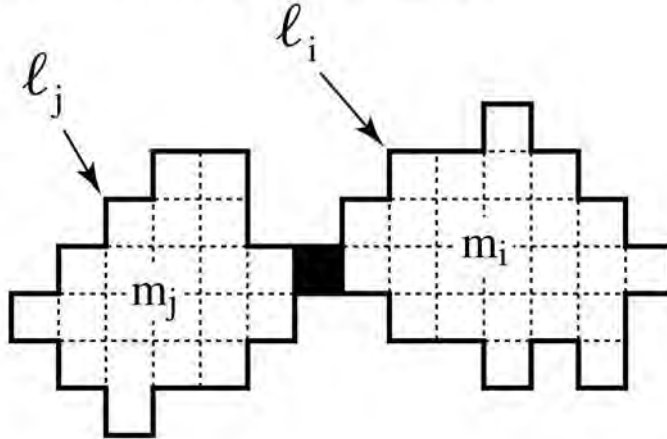


Clustering, universality, and self-organization

- Model germane to forest fires, mergers of cracks, sandpile, and slider-block models
- Define a single element that is added to a system to be a cluster of rank 1; valid in 2 or 3 dimensions, need not be on regular lattice; see figure below
- If a new element is added adjacent to a pre-existing cluster, rank of cluster unchanged (unless it is itself a single element—then rank would become 2)
- If new element connects two existing clusters of ranks $i \geq j$, then new cluster remains rank i unless $i = j$ whereupon new cluster becomes rank $i + 1$
- Ignore 4 body, etc. effects; cut-off at largest size scales; Gabriellov, Newman, & Turcotte (1999)

Clustering, universality, and self-organization (cont)

- Figure below shows how model works (Cartesian lattice used for clarity, but is not necessary)



Clustering, universality, and self-organization (cont)

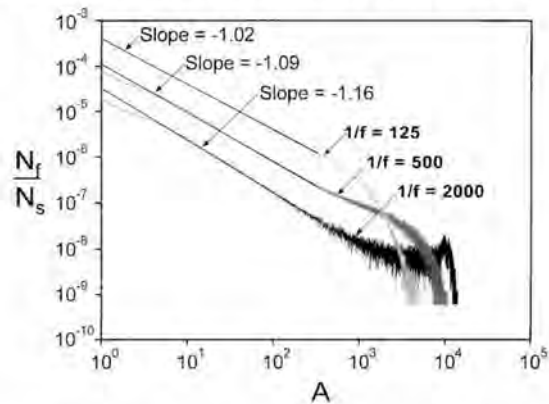
- Look at evolution of N_i as function of time, where N_i is population of rank i objects
- Similarly, look at evolution of m_i , where m_i is mass
- Presume that rate of clustering of clusters with rank i and j depends on products of the form $N_i l_i \times N_j l_j$ where l_i is the length (or area) of the boundary of rank i clusters in two (or three) spatial dimensions
- In Euclidean geometry in 2-D, expect that $l_i \propto \sqrt{m_i}$, and obtain (analytic!) self-organized “critical” (?) scaling

$$N_i \propto \frac{1}{m_i}$$

- Theory extends directly to branching networks, Tokunaga trees (geomorphology and biology), etc.

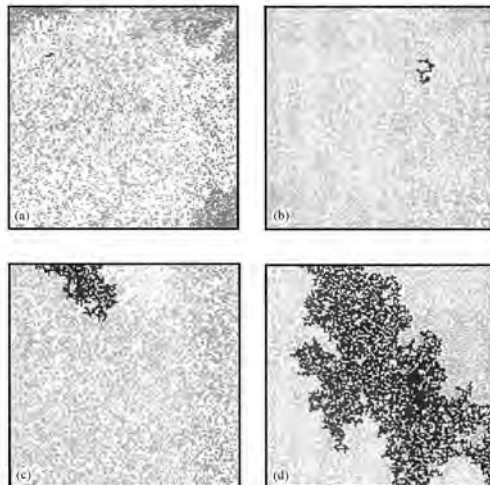
Clustering, universality, and self-organization (cont)

- Corresponds directly to observed statistics of forest fires



- Simulation on a 128×128 grid; sparking frequency f is ratio of “matches” dropped to “trees” planted; $f = 1/125, 1/500, 1/2000$ shown

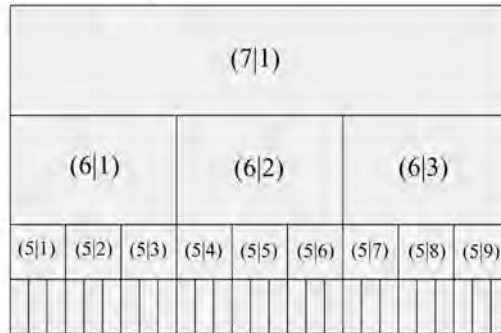
Clustering, universality, and self-organization (cont)



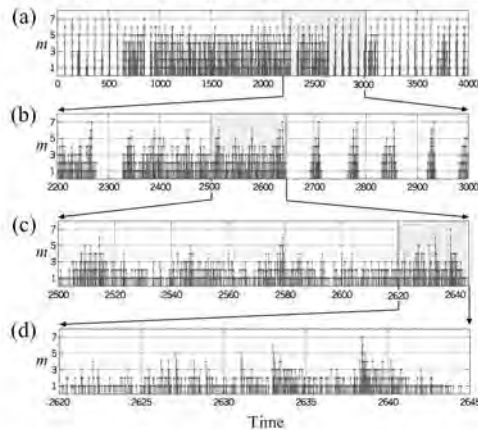
- Four examples with $f = 1/2000$; dark areas in flames, gray areas are forested, white areas are barren

Cascades: inverse, direct, and their collision

- Colliding Cascades (CC) model reproduces premonitory behavior, aftershocks, frequency-magnitude relation, seismic cycle, and clustering—all seismic phenomena (apart from migration) built on fractal network—simplest case \Rightarrow ternary hierarchy

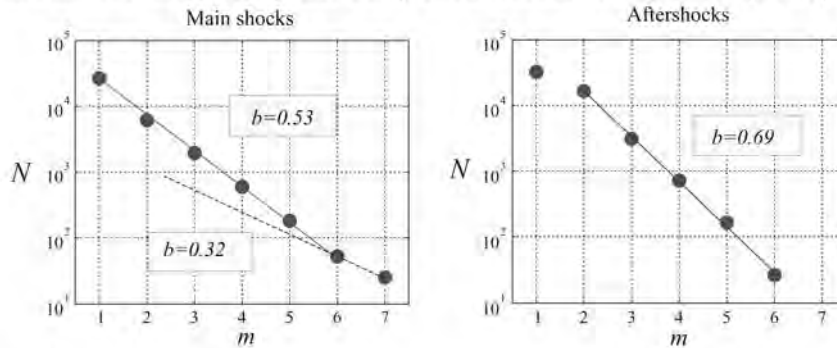


Cascades: inverse, direct, and their collision (*cont*)



- CC model displays expected temporal self-similarity within context of seismic cycle; inverse cascade of failure events with direct cascade of stress transfer within network

Cascades: inverse, direct, and their collision (*cont*)



- Model reproduces relevant frequency-magnitude and aftershock statistics plus spatio-temporal clustering
- Importantly, model showed a new scaling later found to be present in real seismic data (Gabriellov et al., *Geophys. J. Int.*, **143**, 427–43, 2000)

Will our future be complex?

- Absolutely! Many natural phenomena are inextricably complicated yet offer several seemingly simple scaling relationships
- Complexity reminiscent of turbulence but intrinsically more complicated due to interplay of “messy” physics
- First principles calculations are very useful, but limited due to inability to deal with underlying diversity
- Carefully-constructed phenomenological models that pay attention to the natural setting can help provide essential insights into what otherwise might be intractable problems
- Role of mathematics and computation in understanding complexity remains secure

Some References (complexity is still too new!)

- (1) Drazin, P. G. and R. S. Johnson (1989). *Solitons: An Introduction*, Cambridge: Cambridge University Press.
- (2) Lichtenberg, A. J. and M. A. Lieberman (1991). *Regular and Chaotic Dynamics*, New York: Springer.
- (3) Murray, J.D. (1993). *Mathematical Biology*, Berlin: Springer.
- (4) Nayfeh, A. H. and B. Balachandran (1995). *Applied Non-linear Dynamics: Analytical, Computational, and Experimental Methods*, New York: Wiley.
- (5) Peitgen, H.-O. and D. Saupe, eds. (1988). *The Science of Fractal Images*, New York: Springer.
- (6) Turcotte, D. L. (1997). *Fractals and Chaos in Geology and Geophysics*, Cambridge: Cambridge University Press.
- (7) Whitham, G. B. (1974). *Linear and Nonlinear Waves*, New York: Wiley.

Antoni M. Correig

Laboratori d'Estudis Geofísics Eduard Fontserè, Institut d'Estudis Catalans
Departament d'Astronomia i Meteorologia, Universitat de Barcelona

Resum

L'únic coneixement directe que tenim dels fenòmens naturals consisteix en l'observació i la mesura de les seves manifestacions. El seguit de mesures d'un paràmetre determinat (com, per exemple, la pressió atmosfèrica o el moviment del sòl causat per un terratrèmol) es coneix com a *sèrie temporal*. El pas següent, i l'objecte del present treball, consisteix en la descodificació de la informació continguda a les sèries temporals a poder conèixer les característiques del sistema dinàmic en estudi. Sense entrar en metodologies numèriques concretes, ens centrem en la validesa de les diferents aproximacions pensant sempre que les aplicarem a l'estudi de fenòmens naturals. Comencem presentant l'anàlisi de Fourier, aplicable a sistemes lineals estacionaris i per als quals el principi de superposició és vàlid, i seguidament generalitzem a l'espectre de potència. Reconeixent a continuació que els sistemes naturals en general no obeeixen a una distribució gaussiana i que són no lineals, introduïm l'atractor (sovint fractal), caracteritzat per la seva dimensió, l'exponent de Lyapunov i l'entropia de Kolmogorov. Finalitzem discutint la influència de les fluctuacions (generades pel mateix sistema) i dels possibles canvis en la dinàmica del sistema.

Abstract

The only direct knowledge that we have of natural phenomena consists of the observation and measurement of their manifestations. The series of measurements of a specific parameter (such as, for example, atmospheric pressure or the movement of the ground caused by an earthquake) is known as a temporal series. The next step, the object of the present work, consists of the decoding of the information contained in temporal series, in order to be able to find out the characteristics of the dynamic system under study. Without entering into specific numerical methodologies, we focus on the validity of the different approaches, always thinking

that we will apply them to the study of natural phenomena. We begin by presenting Fourier's analysis, applicable to stationary linear systems, and for which the principle of overlapping is valid, and then we generalize the power spectrum. Recognizing, then, that natural systems in general do not obey a Gaussian distribution and that they are not linear, we introduce the attractor (often fractal) characterized by its dimension, Lyapunov exponent, and Kolmogorov entropy. We end by discussing the influence of fluctuations (generated by the system itself) and possible changes in the system's dynamic.

1. Introducció

Les notes que presentem, bàsicament conceptuals, pretenen donar compte dels principis en els quals es basa l'anàlisi de les sèries temporals. El motiu d'escriure-les rau en el fet que, massa sovint, la metodologia vàlida per a l'estudi de sistemes lineals estacionaris és cegament aplicada a sistemes no lineals, amb què s'obtenen resultats aparentment correctes però sense cap ni peus. L'anàlisi de les sèries temporals, especialment quan provenen de processos no lineals, no deixa de ser un art. D'acord amb Hegger *et al.* (2000), podem dir que l'anàlisi de les sèries temporals «és encara una tasca difícil. [...] Si hom utilitza els programes com una capsula negra, amb tota seguretat obtindrà resultats sense sentit. Fins i tot sabent què tenim entre mans, hem d'interpretar els resultats obtinguts amb la màxima circumspicció i sense acceptar-los, d'entrada, tal com ens apareixen».

L'estudi dels fenòmens naturals es basa en l'anàlisi de les mesures d'un observable, típicament un escalar. Les mesures obtingudes $s(n)$ ens defineixen un arranjamant de nombres, la *sèrie temporal*, la qual conté la informació proporcionada pel sistema dinàmic subjacent. L'estudi de les sèries temporals és l'única cosa que ens proveirà, doncs, d'alguna informació respecte al procés físic en estudi.

Actualment existeix una àmplia literatura sobre el tema, i no és pas la meua intenció repassar-la. No tractaré, tampoc, de les tècniques numèriques d'anàlisi. De manera temptativa, dono algunes referències representatives, escollides bàsicament per la claredat de la seva exposició. Aquestes referències són:

- Fonaments teòrics de l'anàlisi de sèrie temporals: Bendat i Piersol (1986).
- Tècniques numèriques: Press *et al.* (1992).
- Anàlisi no lineal de sèries temporals: Abarbanel *et al.* (1993) i Kantz i Schreiber (1997).
- Tècniques numèriques per a l'anàlisi no lineal: Hegger *et al.* (2000).

El treball s'organitza de la manera següent: en primer lloc, a la secció 2 introduïm l'anàlisi lineal a partir de les sèries de Fourier i comentem l'estacionarietat i el principi de superposició, i, seguidament, generalitzem els resultats previs en termes de la integral de Fourier, l'espectre de potència i les funcions de distribució de probabilitat. A la secció 3 introduïm els sistemes no lineals, en concret l'estadística fractal i els sistemes caòtics i com hem de caracteritzar-los a

partir de la dimensió de l'atractor, l'exponent de Lyapunov i l'entropia de Kolmogorov. Finalment, en un apèndix presentem alguns exemples numèrics que ajuden a aclarir el text.

No tractem en aquest treball del tema del soroll, sempre present en els fenòmens naturals, ja que necessitaria un tractament específic amb un altre enfocament.

Una primera versió d'aquest escrit fou presentada al «Sixth Workshop on Non-Linear Dynamics and Earthquake Prediction», a Abdus Salam International Centre for Theoretical Physics, Trieste, 15-27 d'octubre de 2001.

2. Anàlisi lineal

Les tècniques d'anàlisi lineal han tingut un paper clau en l'anàlisi de sèries temporals, deterministes o estocàstiques fins als anys setanta del passat segle XX, quan hom començà a analitzar resultats experimentals i fenòmens naturals, bàsicament el temps atmosfèric i la climatologia, des del punt de vista de la dinàmica dels sistemes caòtics.

2.1. Sèries de Fourier

L'anàlisi de sèries temporals es basa en el descobriment del matemàtic francès Joseph Fourier, referent al fet que qualsevol funció $f(t)$, si satisfà algunes condicions restrictives, es pot expressar com la suma d'un nombre infinit de termes sinusoidals (Fourier, 1882). La hipòtesi fonamental és que $f(t)$ ha de ser una funció periòdica. Així, si $x(t)$ és una funció periòdica, o, pel que fa al nostre cas, una sèrie temporal periòdica amb període T ,

$$x(t) = x(t \pm nT), \quad n = 1, 2, 3, \dots, \quad [1]$$

podem representar $x(t)$ en termes d'una sèrie de Fourier:

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos n\omega t + b_n \sin n\omega t), \quad \omega = 2\pi f, \quad [2]$$

en què

$$\begin{aligned} f &= \frac{1}{T} \\ a_n &= \frac{2}{T} \int_0^T x(t) \cos n\omega t dt, \quad n = 1, 2, \dots \\ b_n &= \frac{2}{T} \int_0^T x(t) \sin n\omega t dt, \quad n = 1, 2, \dots \end{aligned} \quad [3]$$

En forma complexa podem escriure:

$$x(t) = c_n \sum_{-\infty}^{+\infty} e^{in\omega t}, \quad [4]$$

en què

$$c_n = \frac{1}{2T} \int_{-\infty}^{+\infty} x(t) e^{-in\omega t} dt. \quad [5]$$

Les equacions [2] o [4] constitueixen el *teorema de les sèries de Fourier*.

Veiem que una funció $x(t)$ definida en termes d'una sèrie de Fourier és una superposició lineal d'un nombre infinit de termes sinusoidals definida a l'interval $(0, \infty)$, o bé $(-\infty, \infty)$ si representem la sèrie en forma complexa. A partir de la seva definició, se segueix que $x(t)$ és estacionària.

2.2. Estacionarietat

Una sèrie temporal és estacionària si les seves propietats estadístiques són independents del temps. Des d'un punt de vista matemàtic, hom distingeix dos tipus d'estacionarietat: la forta i la feble. L'*estacionarietat forta* requereix que tots els moments de la sèrie temporal siguin constants, mentre que l'*estacionarietat feble* requereix només la constància dels dos primers moments (el valor mitjà i la variància).

El *valor mitjà* d'un flux continu de dades es defineix com a

$$\bar{x} = \frac{1}{T} \int_0^T x(t) dt; \quad [6]$$

i per a un conjunt discret x_i de N mostres,

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i. \quad [7]$$

El *valor quadràtic mitjà* és la mitjana del quadrat dels valors de la sèrie temporal. Per a un flux continu és

$$\bar{x}^2 = \frac{1}{T} \int_0^T x^2(t) dt; \quad [8]$$

i per a un conjunt discret,

$$\bar{x}^2 = \frac{1}{N} \sum_{i=1}^N (x_i)^2. \quad [9]$$

La *variància* és una mesura de la dispersió de les dades respecte al seu valor mitjà. Per a un flux continu tenim

$$V(x) = \frac{1}{T} \int_0^T [x(t) - \bar{x}]^2 dt ; \quad [10]$$

i per a un conjunt discret,

$$V(x) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 . \quad [11]$$

2.3. Superposició

Hem comentat ja que la representació d'una funció $x(t)$ en termes d'una sèrie de Fourier és una superposició de distints components harmònics, cada un amb un coeficient d'amplitud donat. Hom pot visualitzar aquesta superposició com la contribució d'un gran nombre d'oscil·ladors harmònics, cada un definit com a

$$m\ddot{x} + kx = 0 . \quad [12]$$

Tanmateix, aquest tipus d'oscil·lador tan simple no es dóna a la naturalesa. És més realista considerar el moviment d'un oscil·lador amortit subjecte a una força externa $F(t)$. La força externa pot consistir en la suma de distintes forces $F_1(t)$, $F_2(t)$..., i el moviment resultant $x(t)$, com la suma de moviments $x_1(t)$, $x_2(t)$... que tindria cada oscil·lador si les distintes forces actuessin separatament. Aquest és el contingut del *principi de superposició*, que podem expressar de la manera següent: sigui el conjunt de funcions $x_n(t)$, $n = 1, 2, \dots$, una solució del conjunt d'equacions,

$$m\ddot{x}_n + b\dot{x}_n + kx_n = F_n(t) , \quad [13]$$

on la massa m i la constant de recuperació k poden diferir d'un oscil·lador a un altre, i sigui

$$F(t) = \sum_n F_n(t) . \quad [14]$$

Llavors, la funció

$$x(t) = \sum_n x_n(t) \quad [15]$$

satisfà l'equació

$$m\ddot{x} + b\dot{x} + kx = F(t) . \quad [16]$$

En particular, si $F(t)$ es pot escriure com una suma de termes oscil·latoris sinusoidals

$$F(t) = \sum_n C_n \cos(\omega_n t + \theta_n) \quad [17]$$

i si $F(t)$ és periòdica amb període T

$$F(t + T) = F(t) \quad , \quad [18]$$

aleshores $F(t)$ es pot escriure com una sèrie de Fourier tal com l'hem definit a [2]. Cal recordar que, diferentment de la solució de l'equació [12], si qualsevol de les freqüències $f = 2\pi n / T$ coincideix amb la freqüència natural ω_0 de l'oscil·lador, llavors el terme corresponent de la solució és relativament més gran que la resta; aquest fenomen és conegut com a *ressonància*. Per a més detalls vegeu Symon (1963).

2.4. L'anàlisi de Fourier

Una manera útil d'apropar-nos a la comprensió de la dinàmica del sistema consisteix a construir models de les observacions en termes de les variables usuals (velocitat, pressió, temperatura, voltatge, etc.). La tasca del físic és explorar l'existència de *quantitats invariants* del model de la dinàmica del sistema. Per al cas de sistemes lineals, l'anàlisi de Fourier ha mostrat ser una eina molt potent, ja que la posició de les estretes línies de l'espectre de Fourier són característiques del sistema físic específic. Ni que variem l'amplitud de la força externa, ni que iniciem el sistema en temps diferents, la posició dels pics de Fourier no quedarà alterada. Així, les freqüències característiques, conegudes també com a *freqüències ressonants del sistema lineal*, són *invariants del sistema* i ens seran útils per a classificar-lo.

Sigui $\vec{x}(t) = x_1(t), x_2(t), \dots, x_n(t)$ una sèrie temporal. Si considerem que les condicions [17] i [18] són vàlides, l'anàlisi de Fourier, conegut també com a *anàlisi harmònica*, consisteix en l'obtenció de les freqüències $\omega_n = 2\pi n f$ i dels coeficients a_0, a_n i b_n de la sèrie de Fourier que considerem representativa de les observacions. Molt sovint, l'obtenció d'aquests paràmetres es porta a terme mitjançant l'algorisme FFT (*fast Fourier transform*). Per a detalls, vegeu, per exemple, Brigham (1974).

2.5. La integral de Fourier i l'anàlisi espectral

El teorema de les sèries de Fourier, equació [2], es pot generalitzar fent el límit $T \rightarrow \infty$ per incloure-hi forces no periòdiques, així com per poder incloure-hi funcions estocàstiques. Aquesta generalització, coneguda com a *teorema integral de Fourier*, ens permet de representar qualse-

vol funció lineal contínua (o contínua a trossos), subjecta a certes limitacions, com una superposició de forces harmòniques i la força resultant pot ser no periòdica. O sigui, prescindim del constrenyiment [18] (força periòdica) alhora que mantenim la hipòtesi [17] (superposició).

Abans de prosseguir, recordem que l'anàlisi de Fourier ens proporciona les freqüències fonamentals presents en el procés físic original juntament amb les seves amplituds corresponents. La freqüència més baixa és coneguda com a *freqüència fonamental*, i les freqüències superiors, múltiples enters de la freqüència fonamental, són conegudes com a *harmònics d'ordre superior*. En el cas de la integral de Fourier, hom considera el període del senyal infinit (encara que la longitud del senyal que cal analitzar sigui òbviament finita). Aquest és el contingut de la *transformada integral de Fourier*, en la qual totes les freqüències són presents i no solament els múltiples enters de la freqüència fonamental.

Un cop portat a terme el procés de generalització de les sèries de Fourier a la integral de Fourier, el resultat final (l'equivalent a les equacions [2] o [4]) consisteix en el parell de transformades

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt, \quad [19]$$

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega. \quad [20]$$

$F(\omega)$ és sovint coneguda com la *transformada de Fourier* de $f(t)$ (transformada directa), i $f(t)$, com la *transformada inversa de Fourier* de $F(\omega)$. Escrivint les dues transformacions successives com una integral repetida, obtenim el *teorema integral de Fourier*:

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \right] e^{i\omega t} d\omega. \quad [21]$$

Amb vista a l'anàlisi de dades digitalitzades, ens cal considerar una versió finita de les sèries de Fourier, així com una forma discreta de la transformada de Fourier. Des d'un punt de vista computacional, les sèries de Fourier clàssiques són idèntiques a la transformada discreta de Fourier, si bé els fonaments teòrics són molt diferents. De manera similar a l'anàlisi de Fourier, l'anàlisi espectral s'acostuma a portar a terme mitjançant l'algorisme FFT.

En un context més general, utilitzem les sèries i les integrals de Fourier per a representar i/o aproximar funcions, i són casos particulars de desenvolupaments en termes de funcions ortogonals, convenientment tractades en el context dels espais de Hilbert. Es poden deduir directament del teorema de Weierstrass per a dues variables fent el canvi a coordenades polars, restringint el domini de definició al cercle unitari i utilitzant el teorema d'Euler. Per a més detalls vegeu, per exemple, Korn i Korn (1968) i Byron i Fuller (1969). En aquest context es veu de seguida que la transformada de Fourier, un tipus de transformació integral, és un procés lineal.

2.6. Espectre de potència

Sigui x_j una sèrie temporal discreta. Hom defineix la *funció d'autocorrelació* del senyal x_j com a

$$\psi_m = \frac{1}{n} \sum_{j=1}^N x_j x_{j+m}, \quad [22]$$

on $\psi_m = \psi(m\delta t)$. Des d'un punt de vista físic, aquesta funció representa la mitjana del producte dels valors del senyal a un temps donat t i a un temps posterior $m\delta t$. Per tant, a partir de ψ_m hom pot deduir si, i per quina durada, el valor instantani del senyal depèn dels seus valors previs. Hom pot també considerar la funció d'autocorrelació com una mesura del grau de semblança d'un senyal amb si mateix en el decurs del temps.

Es pot veure (Berge *et al.*, 1984) que la funció d'autocorrelació ψ_k és la transformada de Fourier de $|X_k|^2$, on X_k és la transformada de Fourier de x_k . El gràfic que representa $|X_k|^2$ com a funció de la freqüència f ($f = k\delta f$) és conegut com a *espectre de potència*, per analogia amb el cas que les ordenades representin potència (energia per unitat de temps). Considerem, per exemple, un sismograma (l'enregistrament d'un camp ondulatori com ara el desplaçament o la velocitat del sòl). Si el detector (el sismògraf) és lineal i no introdueix cap distorsió a l'interval de freqüències considerat, la quantitat mesurada és proporcional a la variació del camp ondulatori. La teoria ondulatoria ens mostra que la potència transportada per una ona és proporcional al quadrat de la seva amplitud mitjanada en el temps. El teorema de Parseval ens diu que podem reemplaçar la mitjana temporal per la mitjana freqüencial, cosa que justifica el nom de *espectre de potència*.

Des d'un punt de vista pràctic, computacional, l'espectre de potència no és més que el quadrat de l'amplitud de la transformada de Fourier del senyal. Cal que ens adonem que, en considerar $|X_k|^2$, hom perd la informació relativa a la fase del component (X_k).

La representació de l'espectre de potència depèn clarament de com el senyal $x(t)$ evoluciona en el decurs del temps. Per al cas d'un senyal periòdic pur amb període T , l'espectre de potència consisteix en un simple pic (ratlla espectral) localitzat en una freqüència $1/T$; i, si tractem amb funcions periòdiques que contenen harmònics, l'espectre de potència consisteix en una sèrie d'harmònics de la freqüència fonamental localitzats a les freqüències n/T , $n = 1, 2, \dots$. Una mica més complicat és el cas de senyals quasiperiòdics, per als quals la raó dels seus components freqüencials és un nombre irracional. Aquesta complicació es deriva del fet que l'espectre de Fourier d'una funció quasiperiòdica $x(t)$, que conté, diguem-ne, f_r components freqüencials, els quals depenen no linealment de la funció periòdica de les variables ω, t , conté components de totes les freqüències de la forma $|m_1 f_1 + m_2 f_2 + \dots + m_r f_r|$, on m_i són enters arbitraris.

2.7. Funcions de distribució de probabilitat

En analitzar l'estacionarietat feble d'una sèrie temporal, n'hem determinat les propietats estadístiques a partir de dos descriptors: el valor mitjà i la variància. Aquests dos descriptors es refereixen a la descripció estadística de fenòmens que ja han ocorregut. Des d'un altre punt de vista, en molts casos cada un dels valors x_i , $i = 1, \dots, N$ d'una sèrie temporal té una probabilitat d'ocurrència p_i . Per definició:

$$\sum_{i=1}^N p_i = 1 . \quad [23]$$

La funció de distribució de probabilitat (PDF, *probability density function*), $p(x)$, és la probabilitat que cada mesura tingui un valor comprès entre x i $x + dx$. En termes de PDF, el valor mitjà μ i la variància σ^2 d'una variable aleatòria x són donats per

$$\mu = \int_{-\infty}^{\infty} xp(x)dx \quad [24]$$

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 p(x)dx . \quad [25]$$

Tal com ja hem dit, una sèrie temporal x és estacionària si els seus paràmetres estadístics no depenen del temps. Des del punt de vista de la PDF, una sèrie és estacionària feble (o estacionària d'ordre 2) si només existeixen els moments fins a ordre 2 de les distribucions de probabilitat anteriors i són idèntics. Això vol dir que els valors esperats de x i x^2

$$\begin{aligned} E[x] &= \mu \\ E[x^2] &= \mu' \end{aligned} \quad [26]$$

són constants; i, per tant, tenint en compte [25],

$$\text{Var}[x] = \sigma^2 , \quad [27]$$

és una constant, en què en conseqüència se satisfà l'estacionarietat feble.

3. Anàlisi no lineal

A la secció anterior hem donat una breu ullada als mètodes d'anàlisi de sistemes que considerem evolucionen linealment; o sigui, de sistemes per als quals la «sortida» és proporcional a l'«entrada». Tanmateix, molt sovint descrivim l'evolució del sistema mitjançant equacions que hem linealitzat. Com a exemple, tenim l'equació no lineal del pèndol simple, amb freqüència pròpia ω_0 i angle x :

$$\ddot{x} + \omega_0^2 \sin x = 0 . \quad [28]$$

Freqüentment hom fa l'aproximació lineal per desplaçaments infinitesimals ($\sin x \sim x$) i obté com a solució una oscil·lació harmònica. Per al cas dels fenòmens naturals, però, amb desplaçaments finits, aquesta aproximació no és en general vàlida. Com a conseqüència, el principi de superposició perd la seva validesa: si $x_1(t)$ i $x_2(t)$ són solucions de l'equació [28], la suma de $x_1(t)$ i $x_2(t)$ no és solució perquè $\sin x_1 + \sin x_2 \neq \sin(x_1 + x_2)$. El principi de superposició queda substituït ara per la competència entre freqüències; per exemple, entre la freqüència pròpia del sistema i la freqüència d'una força externa.

Un altre punt que cal tenir en compte és que, implícitament, hom considera que els processos naturals són gaussians, o sigui, que obeeixen a una distribució normal, la qual està completament especificada pel seu valor mitjà i la variància. L'únic problema que ens presenta la hipòtesi d'una distribució gaussiana és que, en general, no és vàlida per als processos naturals: per exemple, hi ha un nombre reduït de terratrèmols grans i molts de petits, tal com passa amb un riu i els seus afluents. Liebovitch i Scheurle (2000) presenten una discussió interessant del tema en termes de jocs d'atzar.

3.1. Estadística fractal

És un fet que molts fenòmens de la naturalesa són *autosimilars*, això és, són d'escala invariant: presenten el mateix aspecte independentment de l'escala utilitzada (rigorosament parlant, a causa del fet que la naturalesa és finita, tenim uns límits superior i inferior). L'autosimilitud implica que en el fenomen no hi intervé cap escala característica, atès que hi són totes. Com a exemple de fenòmens autosimilars, tenim la distribució de la freqüència de terratrèmols d'acord amb la magnitud o bé la superfície dels núvols i la pluja respecte al perímetre.

Des del punt de vista de l'estudi dels fenòmens naturals, podem definir un conjunt fractal com a

$$N_i = \frac{C}{r_i^D}, \quad [29]$$

on N_i és el nombre d'objectes amb dimensió lineal característica r_i , C és una constant de proporcionalitat i D és la dimensió fractal. Si D és un enter, la dimensió fractal coincideix amb la dimensió euclidiana.

Podem aplicar també els conceptes fractals a una distribució estadística d'objectes. Si el nombre d'objectes N amb una dimensió característica *major que* r satisfà la relació

$$N = \frac{C}{r^D}, \quad [30]$$

llavors és una distribució fractal. Notem que, a diferència de [29], la relació [30] és una distribució acumulativa, més útil que la relació [29] per al cas de les distribucions contínues. De [30] podem definir la *dimensió fractal* com a

$$D = \frac{\log(N)}{\log(\frac{1}{r})}. \quad [31]$$

Podem interpretar la relació [30] també com una distribució de probabilitat

$$p(x) = Ax^{-\alpha},$$

on x és definida a l'interval $(0, \infty)$ i $\alpha > 0$. Aquesta distribució fractal, una llei potencial, és *l'única que satisfà la propietat d'invariància d'escala*.

Tenint en compte les definicions [24] i [25], podem comprovar que, per a una distribució fractal, ni el valor mitjà ni la variància no són definits: divergeixen per $x \rightarrow 0$ i s'esvaeixen per $x \rightarrow \infty$. Això es tradueix en el fet que, per a una distribució fractal, hi ha un petit nombre d'esdeveniments grans i un gran nombre de petits, de manera que el valor mitjà d'un conjunt de dades depèn de la mida del conjunt, així com de si a la mostra hi ha inclosos més o menys esdeveniments grans. Com a conseqüència, la sèrie temporal no és estacionària, cosa que ens presenta un problema seriós: la major part dels mètodes d'anàlisi consideren l'estacionarietat de la sèrie temporal. Cal fer notar també que, excepte per al cas $\alpha = -1$, les lleis potencials són no lineals, la qual cosa ens porta a la utilització de tècniques d'anàlisi no lineals.

Hem comentat ja que el valor mitjà de la sèrie temporal depèn del conjunt de dades analitzat. El valor mitjà no és, per tant, un invariant del sistema (tal com ho és, per exemple, en el cas d'una distribució gaussiana). El que realment importa és el nombre relatiu de valors «petits» de les dades respecte als valors «grans». Altrement dit, i des d'un punt de vista més general d'un conjunt fractal, de la seva *dimensió fractal*. Per a un objecte fractal, la dimensió fractal D , definida per l'equació [31], ens proporciona precisament el nombre d'elements N que hom pot detectar en mesurar l'objecte amb una resolució r donada.

3.2. Sistemes caòtics

La mecànica clàssica es basa en el determinisme newtonià: donats un model (un conjunt de regles o equacions) i un conjunt de condicions inicials, hom pot predir exactament tot moviment passat i futur. Malauradament, aquest determinisme és només vàlid per al (molt reduït) conjunt de sistemes integrables (en el sentit de la mecànica hamiltoniana: un sistema és integrable si totes les equacions es poden reduir a quadratures).

En un article ja clàssic, Henon i Heiles (1964), que era un estudi numèric sobre sistemes conservatius no integrables, descobriren que, depenent del valor de l'energia (una constant del moviment), la solució resultant podia ser *periòdica* o *estocàstica* (cal recordar que un sistema hamiltonià és totalment determinista!). Rigorosament parlant, aquest descobriment es deu a Poincaré a començaments del segle xx, si bé passà desapercbut excepte per a uns quants matemàtics.

Al mateix temps, Lorenz (1963) féu un altre descobriment decisiu: la *sensitivitat a les condicions inicials*. En un estudi sobre la predictibilitat d'un sistema convectiu, un model simplificat de convecció atmosfèrica, descobrí que l'evolució del sistema era fortament sensible a les condicions inicials, alhora que seguia una trajectòria molt irregular, caòtica. A diferència del sistema conservatiu estudiant per Henon i Heiles (1964), el sistema de Lorenz és un sistema obert, dissipatiu.

L'anàlisi de sèries temporals generades pel sistema de Lorenz mostrà que les trajectòries són *aperiòdiques*. Recordem que l'espectre de potència d'un senyal aperiòdic és continu (en el sentit, per exemple, del valor mitjà d'un interval) i no ens mostra cap freqüència predominant. Tanmateix, si bé l'espectre d'un senyal aperiòdic és continu, no tot espectre continu ha de ser aperiòdic forçosament (per exemple, aquest és el cas de l'espectre d'un senyal quasi periòdic amb un gran nombre de freqüències). D'altra banda, en cas que el senyal sigui realment aperiòdic, és d'importància fonamental esbrinar el nombre de graus de llibertat del sistema dinàmic subjacent: si el nombre de graus de llibertat és petit (diguem inferior a deu), podem abordar una descripció determinista; mentre que, si el nombre de graus de llibertat és gran, l'única aproximació al problema és una descripció probabilista.

D'una manera similar a la mecànica hamiltoniana, estudiarem l'evolució del sistema a l'espai de fase. En tractar amb sistemes dinàmics, un exemple dels quals és el sistema de Lorenz, definim l'espai de fase de manera més genèrica en termes de *posició* i *velocitat*, on per *posició* entenem el paràmetre observat, i per *velocitat*, el seu ritme de variació. En tractar amb sèries temporals observades, per a les quals efectuar-ne la derivada numèrica pot ser delicat, hom acostuma a utilitzar el *mètode dels retards* (Packard *et al.*, 1980), consistent a representar la sèrie temporal respecte a la mateixa sèrie retardada.

Un estudi de l'evolució del sistema de Lorenz a l'espai de fase ens revela una propietat important, compartida per tots els sistemes caòtics dissipatius: el moviment tendeix, asimptòticament, a un *atractor*, un subdomini de l'espai de fase.

Un sistema caòtic, conservatiu o dissipatiu, necessàriament ha de ser no lineal, cosa que fa que analitzar-lo sigui difícil, no trivial. Per una banda, tenim la sensibilitat a les condicions inicials, que fa que un experiment estudiat a partir d'una sèrie temporal no sigui repetible i sigui impredecible; i, per una altra banda, la no-linealitat invalida el principi de superposició, que és substituït per la competitivitat entre els diversos components del sistema. Finalment, ja hem comentat que l'espectre de potència d'un sistema caòtic és aperiòdic, sense ressonàncies. Ens trobem mancats dels invariants coneguts.

Necessitem, per tant, definir nous invariants, uns paràmetres que es mantinguin constants a l'espai de fase al llarg de les trajectòries i que siguin també insensibles a petites variacions de les condicions inicials. Ja hem dit que, a l'espai de fase, asimptòticament el moviment queda atrapat a l'atractor. Hom pot comprovar que, si bé la sèrie temporal varia d'un interval a un altre i d'un experiment a un altre, l'atractor és sempre el mateix. Per tant, l'atractor és un invariant del sistema. Es tracta, doncs, d'esbrinar com cal caracteritzar-lo.

3.3. Caracterització de l'atractor

En primer lloc, vegem què és un atractor. Seguint Schuster (1995), considerem un sistema caòtic dissipatiu, descrit sia per un flux (sistema continu), sia per una aplicació (sistema discret). Considerem, per simplicitat, el cas d'un flux, descrit per un conjunt d'equacions diferencials autònomes de primer ordre:

$$\vec{\dot{x}} = \vec{F}(\vec{x}), \quad \vec{x} = (x_1, x_2, \dots, x_d). \quad [33]$$

El terme *dissipatiu* significa que un element de volum V delimitat per una superfície S a l'espai de fase $|\vec{x}|$ es contreu en el decurs del temps. Determinarem l'evolució de la superfície S a partir de l'evolució de cada un dels punts que la defineixen d'acord amb l'equació [33]. Aplicant-hi el teorema de la divergència, obtenim

$$\frac{dV}{dt} = \int_V d^d x \left(\sum_{i=1}^d \frac{\partial F_i}{\partial x_i} \right). \quad [34]$$

Un sistema dinàmic és dissipatiu si $dV/dt < 0$. Hom troba que, per a un sistema dinàmic dissipatiu, les trajectòries generades per l'equació [33] es caracteritzen pel que segueix: *a)* la tendència a una regió delimitada de l'espai de fase (l'atractor), *b)* el moviment és erràtic, i *c)* les trajectòries són sensibles a petites variacions de les condicions inicials. Com que per hipòtesi el sistema (de dimensió d) és dissipatiu, el volum de l'atractor, definit com un subdomini de dimensió menor de d , és nul a l'espai de dimensió d . Sovint l'atractor és un objecte fractal, de dimensió no entera, i rep el nom de *atractor estrany*.

Resumint, podem dir que un atractor estrany té les propietats següents:

a) És un atractor, o sigui, una regió delimitada de l'espai de fase $|\vec{x}|$ a què tendeixen asimptòticament totes les trajectòries prou properes. El conjunt de trajectòries que van a parar a l'atractor defineixen una regió coneguda com a *conca d'atracció*. Endemés, l'atractor és indiscomponible, cosa que vol dir que, en el transcurs del temps, una trajectòria en visitarà tots els punts.

b) La propietat segons la qual un atractor és estrany és la dependència del sistema a les condicions inicials: malgrat la contracció del volum, les longituds no s'encongeixen en totes les direccions i punts inicialment propers dins l'atractor, per uns temps prou grans, se separen exponencialment.

Les dues propietats anteriors ens mostren que, bàsicament, podem caracteritzar l'atractor amb dos invariants: un invariant geomètric, la dimensió de l'atractor (propietat *a*), i un invariant dinàmic, l'exponent de Lyapunov (propietat *b*), relatiu a la separació de trajectòries properes. Ambdós invariants es poden obtenir a partir de les sèries temporals i són l'equivalent, en els sistemes caòtics, de les ressonàncies de Fourier dels sistemes lineals. Seguint Abarbanel *et al.* (1993), a les properes seccions comentarem més àmpliament els dos invariants.

3.3.1. Dimensió

Més enllà del simple concepte de dimensió com el nombre de coordenades necessàries per a especificar un estat, tenim el concepte geomètric relacionat de *com un (hiper)volum s'escala com a funció d'una longitud característica*

$$V \propto L^D . \quad [35]$$

Les àrees planars s'escalen quadràticament amb la longitud d'un costat i els volums com el cub. Partint del concepte de volum, de l'equació [35] podem definir una *dimensió* com a

$$D = \frac{\log V}{\log L} . \quad [36]$$

Les tècniques antigues per a estimar la dimensió recobrien l'atractor per tal de calcular el volum V . Considerem una partició de l'espai de dimensió d_E de l'espai de fase amb una malla de dimensió ϵ . Comptem quants elements de la partició contenen com a mínim un punt de les dades mostrejades i utilitzem aquest valor com una mesura del volum V amb una resolució ϵ . Seguidament refinem la resolució a l'espai de fase disminuint el valor de ϵ i tornem a comptar quantes d'aquestes noves particions, més petites, contenen algun element. Continuem repetint aquest procés fins que ϵ hagi variat alguns ordres de magnitud. Teòricament, en el límit $\epsilon \rightarrow 0$, el quocient a l'equació [36] ens descriu les propietats de cobertura de l'espai del conjunt de punts que cal analitzar. Aquest procediment es coneix com a *mètode de comptatge de caps* (*box counting method*). Intuïtivament, l'equació [35] considera la «quantitat d'un objecte» com a equivalent al seu volum, però aquesta no és l'única manera de mesurar «quant hi ha» d'alguna cosa determinada. L'equació [35] es pot escriure d'una manera més genèrica com a

$$mida \propto longitud^{dimensió} . \quad [37]$$

Considerem ara el cas que mesurem la «mida» en termes d'algun moment d'una distribució invariant donada. La distribució (o mesura) invariant a l'atractor és la densitat natural de punts dins del volum de l'atractor. La definició d'aquesta densitat és

$$\rho(\vec{x}) = \frac{1}{N} \sum_{k=1}^N \delta^d(\vec{x} - \vec{y}(k)) \quad [38]$$

a l'espai de fase de dimensió d .

Sigui $g(\vec{x})$ una funció qualsevol (o bé una sèrie temporal) a l'espai de fase. En definim el moment $\langle g \rangle$ com a

$$\langle g \rangle = \int d^d x g(\vec{x}) \rho(\vec{x}) = \frac{1}{N} \sum_{k=1}^N g(\vec{y}(k)) . \quad [39]$$

Considerem, a l'equació [38], $g(\bar{x}) = \rho^p(\bar{x})$ i utilitzem $\langle g \rangle^{1/p}$ per a mesurar la «mida» de l'equació [37]. El resultat final és

$$D = \lim_{r \rightarrow 0} \frac{1}{q-1} \frac{\log \sum \rho^q}{\log r} = D_q . \quad [40]$$

Aquest resultat constitueix la definició de *dimensió generalitzada* D_q , la qual ens proporciona un espectre complet de quantitats invariants per a $-\infty < q < \infty$.

Són d'un interès especial les dimensions corresponents a $q = 0, 1$ i 2 :

q = 0: Substituint-ho, obtenim

$$D_0 = \frac{\log N_0}{\log r} , \quad [41]$$

que no és més que la *dimensió de capacitat* estimada pel mètode del comptatge de capsos.

q = 1: Fent el límit $q \rightarrow 1$ a l'equació [40], obtenim

$$D_1 = \lim_{r \rightarrow 0} \frac{\sum_i p_i \log p_i}{\log r} . \quad [42]$$

Notem que, si $p_i = 1/N$ per a tot i , llavors $D_1 = D_0$; però, si els p_i no són tots iguals, aleshores $D_1 \leq D_0$.

La diferència entre les dues mesures es pot explicar en termes de com utilitzem les particions a l'espai de fase. Per a D_0 definim la mida pel nombre d'elements de la partició plens, mentre que per a D_1 calculem la mida tenint en compte que la contribució de cada partició és proporcional al nombre de punts que conté. Així, en D_0 la mida es refereix al volum, mentre que per a D_1 la mida és equivalent a la massa. A causa de la forma funcional del numerador de l'equació [42], D_1 és coneguda com a *dimensió d'informació*. Aquest numerador expressa la quantitat d'informació necessària per a especificar un estat amb una precisió r , i el quocient que ens defineix la dimensió ens diu com creix ràpidament la informació en decreïxer r .

q = 2: Obtenim la dimensió de correlació D_2 :

$$D_2 = \lim_{r \rightarrow 0} \frac{-\log \sum_i p_i^2}{\log r} . \quad [43]$$

En aquest cas, el numerador no és més que una funció de correlació de dos punts, la qual ens mesura la probabilitat de trobar un parell de punts aleatoris dins un element de partició donat, d'una manera semblant a com a la definició de D_1 el numerador ens mesura la probabilitat de trobar un punt en un element donat. Podem estimar el valor numèric del numerador

de [43] comptant quants parells de punts tenen una distància de separació menor que un valor r donat. Grassberger i Procaccia (1983) suggeriren la utilització de la *integral de correlació*

$$C_2(r) = \int d^d x \rho(\vec{x}) n(r, \vec{x}) = \frac{2}{N(N-1)} \sum_{i \neq j}^N \theta(r - |\vec{y}(j) - \vec{y}(i)|) \quad [44]$$

per tal d'estimar $\sum_i \rho_i^2$ de manera simple i eficient.

3.3.2. L'espectre de Lyapunov

Mentre que la dimensió ens caracteritza la distribució de punts a l'espai de fase, els exponents de Lyapunov ens diuen com la dinàmica del sistema afecta les trajectòries a l'espai de fase.

A fi d'entendre el significat de l'espectre de Lyapunov, considerem com evoluciona en el temps un hipervolum esfèric en un subespai donat de l'espai de fase. Una dinàmica caòtica provoca que, al llarg de la seva trajectòria, l'esfera inicial adopti formes extremament complexes. Tanmateix, per a escales de longitud prou petites i escales temporals prou curtes, la dinàmica del sistema distorsiona l'esfera inicial a una forma el·lipsoïdal en la qual alguns dels eixos s'allarguen i uns altres s'encongeixen. L'eix principal (el més llarg) de l'el·lipsoide correspon a la direcció més inestable del flux, i és precisament el ritme asimptòtic d'extensió d'aquest eix el que ens mesura el major exponent de Lyapunov. Més concretament, si el radi infinitesimal i -èssim a un temps t és $l_i(t)$, definim l'exponent de Lyapunov i -èssim com a

$$\lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \log \frac{l_i(t)}{r(0)} . \quad [45]$$

Per conveni, hom presenta els exponents de Lyapunov de manera que $\lambda_1 > \lambda_2 > \lambda_3 \dots$. En el cas d'una anàlisi de sèries temporals, normalment només es pot calcular amb precisió l'exponent de Lyapunov màxim.

3.4. L'entropia de Kolmogorov

En calcular els exponents de Kolmogorov, hom no té en compte la possible dependència respecte al punt inicial. Aquesta dependència queda inclosa a la definició de l'entropia de Kolmogorov, coneguda també com a *K-entropia*:

$$K = \int_M \rho(x) \sum \lambda_i(x) dx , \quad [46]$$

on $\rho(x)$ representa una funció de distribució invariant (normalitzada) a l'atractor M i la prima sobre el sumatori indica que només tenim en compte els exponents de Lyapunov positius.

Sovint hom tracta amb sistemes per als quals els exponents de Lyapunov són independents del punt de l'atractor; en aquest cas, com que la funció de distribució està normalitzada, la K-entropia és donada per

$$K = \sum \lambda_i. \quad [47]$$

La K-entropia representa el ritme de producció d'informació i d'augment d'incertesa. En un sistema caòtic, per al qual $K > 0$, el nombre d'estats possibles del sistema en evolució (des d'unes condicions inicials donades) s'incrementa en el temps. Per tant, la informació també s'incrementa. Des d'un altre punt de vista, considerem una mesura d'un punt del sistema a l'espai de fase. Considerem un sistema amb $K > 0$ i la mesura d'un punt del sistema a $t > 0$ amb una incertesa mínima; això implica que la incertesa del sistema a $t = 0$ deu ser menor que el mínim. Conseqüentment, hi ha hagut un augment d'informació. En evolucionar el sistema, el seu coneixement subsegüent (determinat a partir de l'estat inicial) es deteriora. El sistema va generant informació i, si no realitzem cap nova mesura, el coneixement que tenim del sistema va disminuint i la incertesa, per tant, augmenta.

Seguint Schuster (1995), podem utilitzar la K-entropia per a determinar el temps mitjà de predicció de l'estat del sistema. Considerem, per exemple, una aplicació unidimensional, localitzada en un quadrat unitat, per al qual després de n passos un interval l s'incrementa a $L = le^{\lambda n}$, essent λ l'exponent de Lyapunov. Si L esdevé major que 1 (o sigui, un exponent de Lyapunov positiu), no podem ja localitzar la trajectòria a l'interval $[0, 1]$ i tot el que podem dir és que el sistema té una probabilitat

$$p(x) = \rho_0 dx \quad [48]$$

de raure en un interval $[x, x + dx] \in [0, 1]$, on $\rho_0(x)$ és la densitat de probabilitat invariant del sistema. Altrament dit, una predicció precisa de l'estat del sistema solament és possible per a un temps n menor que un temps T_0 :

$$le^{\lambda T_m} = 1 \rightarrow T_m = \frac{1}{\lambda} \log \left(\frac{1}{l} \right). \quad [49]$$

Per a temps superiors a T_m només podem fer prediccions estadístiques. L'equació [48] es pot generalitzar a sistemes dinàmics de dimensió superior reemplaçant λ per la K-entropia

$$T_m \propto \frac{1}{K} \log \left(\frac{1}{l} \right). \quad [50]$$

Notem, per acabar, que la precisió l amb la qual localitzem l'estat inicial només influencia T_m d'una manera logarítmica.

Bibliografia

- ABARBANEL, H. D. I.; BROWN, R.; SIDIROWICH, J. J.; TSIMIRING, L. SH. (1993). «The analysis of observed chaotic data in physical systems». *Rev. Mod. Phys.*, núm. 65, p. 1331-1392.
- BENDAT, J. S.; PIERSOL, A. G. (1986). *Random data: Analysis and measurement procedures*. Nova York: Wiley.
- BERGE, P.; POMEAU, Y.; VIDAL, C. (1984). *Order within chaos*. Nova York: Wiley.
- BRIGHAM, E. O. (1974). *The fast Fourier transform*. Englewood Cliffs: Prentice-Hall.
- BYRON, F. W.; FULLER, R. W. (1969). *Mathematics of classical and quantum physics*. Reading: Addison-Wesley.
- FOURIER, J. (1822). *Théorie analytique de la chaleur*. París: Didod.
- GRASSBERGER, P.; PROCACCIA, I. (1983). «Characterization of strange attractors». *Phys. Rev. Lett.*, núm. 50, 346-349.
- HEGGER, R.; KANTZ, H.; SCHREIBER, T. (2000). *TISEAN-2.1, Nonlinear Time Series Analysis* [en línia]. <<http://www.mpipks-dresden.mpg.de/tisean>>
- HENON, M.; HEILES, C. (1964). «The applicability of the third integral of the motion: some numerical results». *Astron. J.*, núm. 69, p. 73.
- KANTZ, H.; SCHREIBER, T. (1997). *Nonlinear time series analysis*. Cambridge.
- KORN, G. A.; KORN, T. M. (1968). *Mathematical handbook for scientists and engineers*. Nova York: McGraw-Hill.
- LIEBOVITCH, L. S.; SCHERLUE, D. (2000). «Two lessons from fractals and chaos». *Complexity*, núm. 5, p. 34-43.
- LORENZ, E. N. (1963). «Deterministic nonperiodic flow». *J. Atmos. Sci.*, núm. 20, p. 130.
- PACKARD, M. H.; CRUTCHFIELD, J. F.; FARMER, J. D.; SHAW, R. S. (1980). «Geometry from a time series». *Phys. Rev. Lett.*, núm. 45, p. 712.
- PRESS, W. H.; TEUKOLSKY, S. A.; VETTERLING, W. T.; FLANNERY, B. P. (1992). «Numerical recipes in Fortran. The art of scientific computing». Cambridge.
- SCHUSTER, H. G. (1995). «Deterministic chaos». Weinheim: VCH.
- SYMON, K. R. (1963). *Mechanics*. Reading: Addison-Wesley.

Apèndix

En aquest apèndix mostrem alguns exemples útils per a clarificar els conceptes exposats. No hi pretenem, en absolut, cobrir tots els punts, sinó només els que hem considerat prioritaris.

FIGURA 1. Exemple d'una sèrie de Fourier en el domini del temps i de les freqüències. A la part superior hem presentat la funció

$$f(x) = |\sin x|, \quad -\pi < x < \pi, \quad [51]$$

per a la qual la sèrie corresponent de Fourier és

$$F(x) = \frac{2}{\pi} - \frac{4}{\pi} \sum_n \frac{\cos(2nx)}{(2n-1)(2n+1)}, \quad n = 1, 2, \dots \quad [52]$$

Aquesta funció és periòdica amb període $T = 2\pi/x = 1s$. A la part inferior hem representat l'espectre de potència, on clarament hi ha representades les freqüències $f = n/T$, $n = 1, 2, \dots$. $n = 1$ correspon a la freqüència fonamental i $n > 1$ correspon als harmònics d'ordre superior.

FIGURA 2. Superposició de tres ones harmòniques (part superior):

$$f(x) = \sum_{i=1}^3 a_i \sin(2\pi f_i t) \quad [53]$$

per a $f_1 = 0,04$ Hz, $f_2 = 0,2$ Hz i $f_3 = 0,4$ Hz. A la part central n'hem representat l'evolució a l'espai de fase: hi podem veure que la trajectòria és tancada, fet que ens indica que la funció és periòdica. A la part inferior hem representat l'espectre de potència, on podem veure clarament les tres freqüències indicades. Com que partim de la superposició de tres ones harmòniques, no obtindrem harmònics d'ordre superior.

FIGURA 3. Exemple d'oscil·lacions no lineals generades a partir de l'oscil·lador de Duffing:

$$\ddot{x} + \delta\dot{x} - \omega_0 x + \beta x^3 = \gamma \cos(\omega t). \quad [54]$$

Per a dibuixar el gràfic hem utilitzat els valors $\delta = 0,05$, $\omega_0 = 0$, $\beta = 1$, $\gamma = 7,5$ i $\omega = 1$. A la figura hi ha superposades dues trajectòries corresponents a dues condicions inicials (posició i velocitat) lleugerament diferents $(0, 0)$ i $(0,0001, 0)$. Hi podem veure que fins a $t \approx 35$ s les trajectòries coincideixen i que a partir d'aquest valor comencen a separar-se i a canviar la seva forma. Aquest comportament és el que es coneix com a *sensitivitat a les condicions inicials*.

FIGURA 4. És igual que la figura 3 treballant, però, a l'espai de fase. A la part superior hem dibuixat les trajectòries per 1.000 s, suficients per a delimitar-nos l'atractor (oscil·lacions en un potencial biestable) i a la part inferior hem dibuixat els primers 80 s: hi podem veure clarament com en el decurs del temps les trajectòries van divergint.

FIGURA 5. Exemple d'aplicació del principi de superposició per a un sistema lineal (equació [54] amb $\beta=0$) a la part superior; i a la part inferior, la competitivitat entre freqüències (equació [54] amb $\beta=0,01$). En tots els casos la freqüència pròpia del sistema és $f_0=0,4$ Hz i hem variat la freqüència de la força externa donant-los els valors $f=0,75$ (línia contínua), $f=0,45$ Hz (línia de traços) i $f=0,41$ Hz (línia puntejada). En la *simulació lineal* podem adonar-nos que l'espectre resultant és la suma d'espectres. Hi podem veure també que, en apropar-se f cap a f_0 , la ressonància hi va apareixent. En la *simulació no lineal* els resultats són força més complicats. En primer lloc, hi podem veure, per al cas de la línia contínua, que la posició del pic ressonant s'ha desplaçat lleugerament cap a les altes freqüències (de fet, en sistemes no lineals el coeficient ω_0 perd el seu sentit com a freqüència pròpia del sistema). En segon lloc, veiem que, en acostar-nos a la freqüència de ressonància, hi apareixen *subharmònics* (pics a freqüències ω/n) a més d'*harmònics d'ordre superior*. Els subharmònics no apareixen mai en sistemes lineals. Per al cas d'un coeficient β més gran, veuríem que, en apropar-nos a la ressonància, l'amplada del pic ressonant augmenta. Aquests tres fets són característics de les oscil·lacions no lineals.

FIGURA 6. Exemple d'una distribució normal i d'una de fractal, tretes de Liebovitch i Scherlue (2000). Els valors dels conjunts de dades hi són representats per la seva mida. Per al cas d'una distribució normal, les mides són d'un ordre semblant, mentre que per a una distribució fractal n'hi ha poques de grans i moltes de petites seguint una llei potencial. Clarament, per al cas de la distribució normal, com més dades utilitzem, més ens acostem al valor real del valor mitjà; mentre que, per al cas de la distribució fractal, el valor mitjà depèn del nombre de mostres grans o petites incloses al conjunt analitzat.

FIGURA 7. Representació de l'atractor de Henon, definit per l'equació

$$\begin{aligned} X_{k+1} &= X_k + 1 - \alpha X_k^2 \\ Y_{k+1} &= \beta X_k \end{aligned} \tag{55}$$

havent-hi utilitzat els valors $\alpha=1,4$ i $\beta=0,3$. De manera intuïtiva hi podem veure que la dimensió de l'atractor és fractal: major que 1 (cas d'una línia) i menor que 2, ja que cobreix una petita porció de superfície.

FIGURA 8. Càlcul de la dimensió de correlació de l'atractor de Henon, definida com a

$$D_2 = \lim_{r \rightarrow 0} \frac{C_2(r)}{\log r}, \quad [56]$$

on $C_2(r)$ és la integral de correlació definida per l'equació [44]. A la part superior hem representat la integral de correlació com a funció del radi per a diversos valors de la dimensió d'inclusió (l'espai euclidià en què reconstruïm l'atractor) des de $d=2$ (línea inferior) fins a $d=7$ (línia superior). Hi podem veure que aproximadament fins a $r=1$ les línies són rectes paral·leles que se saturen per a $r > 1$. A la part inferior simplement hem dibuixat els pendents de les línies de la figura superior, als quals hem ajustat una recta (paral·lela a l'eix de les abscisses) dins l'interval $r = 0,004 - 0,4$ (dues dècades). El valor de la dimensió de correlació és $d_c = 1,2059$. Si la part plana dels pendents és inferior a una dècada, la dimensió obtinguda pot no tenir sentit.

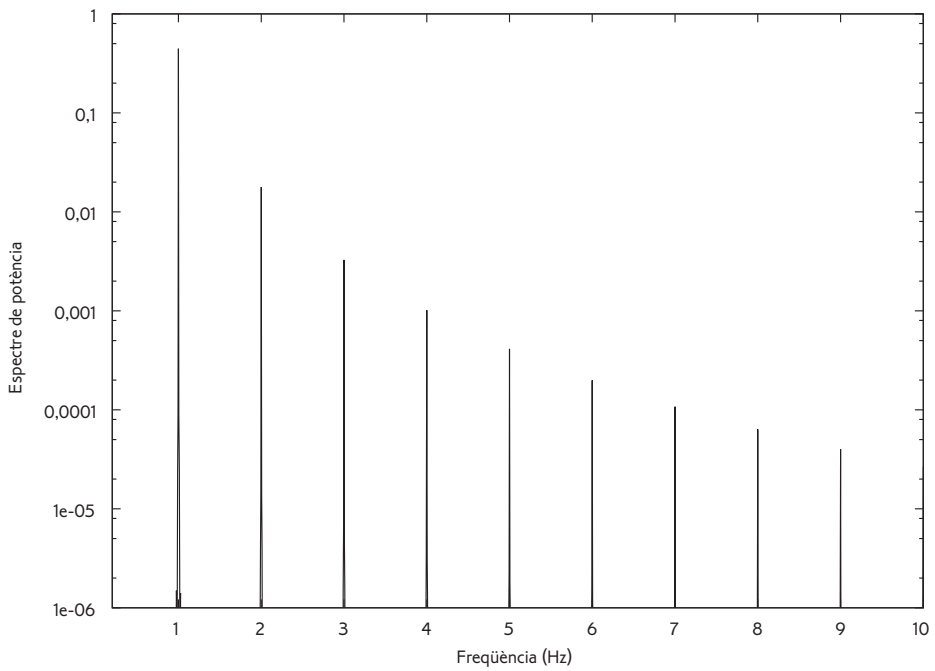
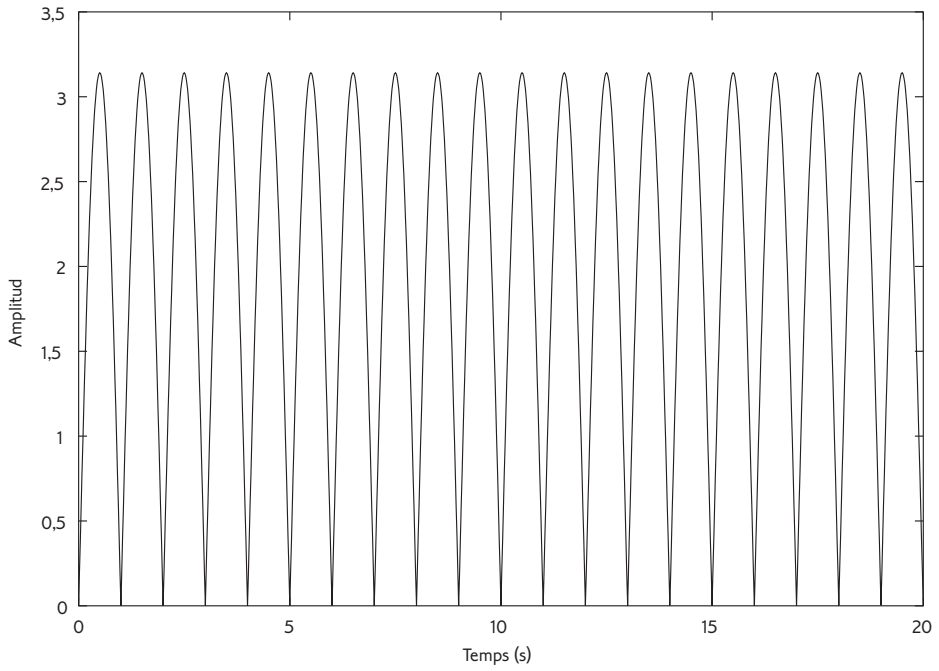


FIGURA 1

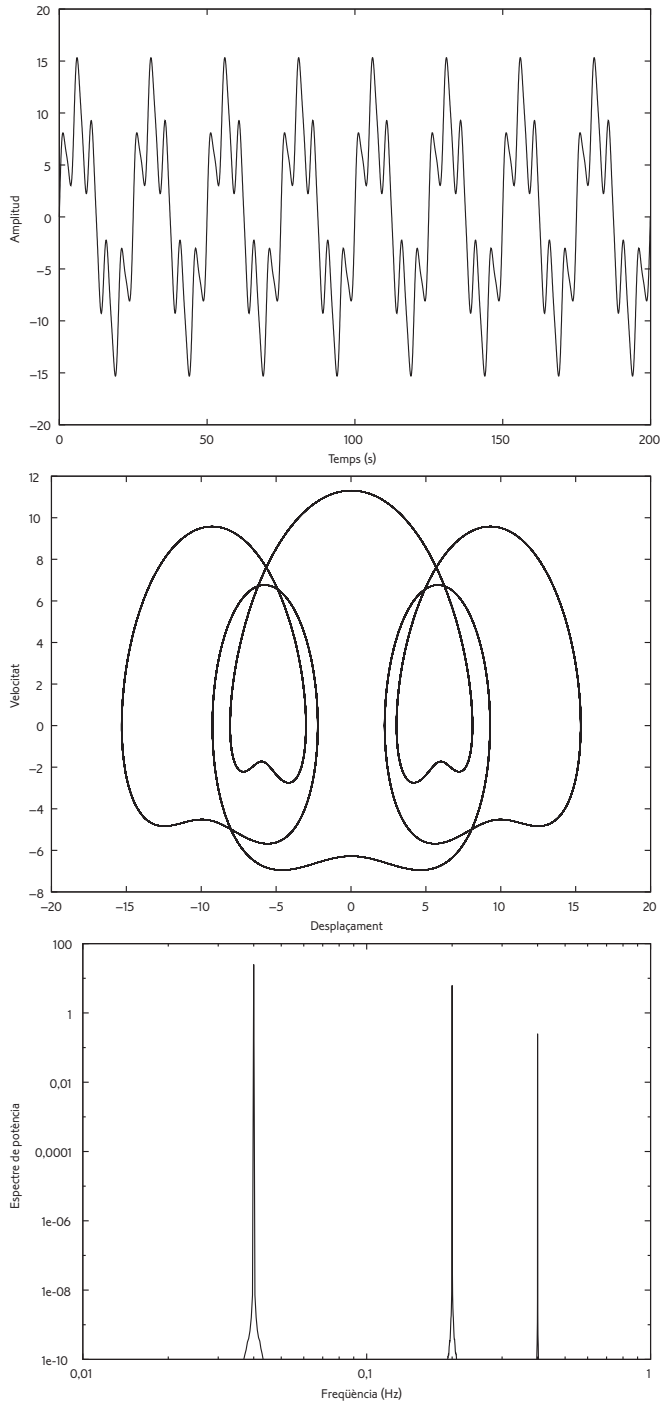


FIGURA 2

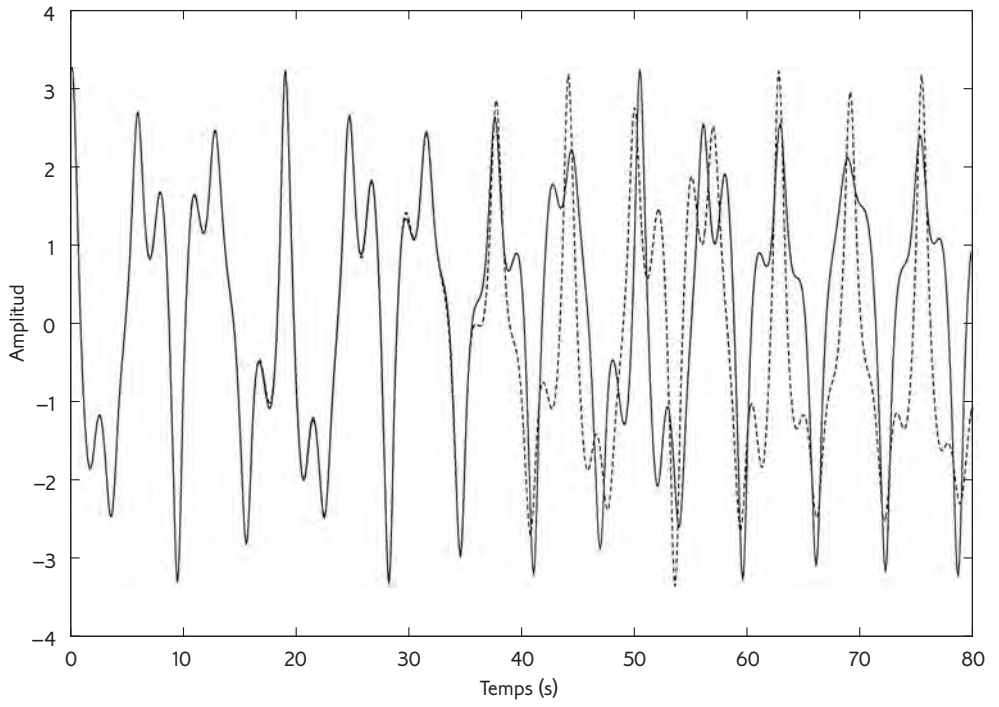


FIGURA 3

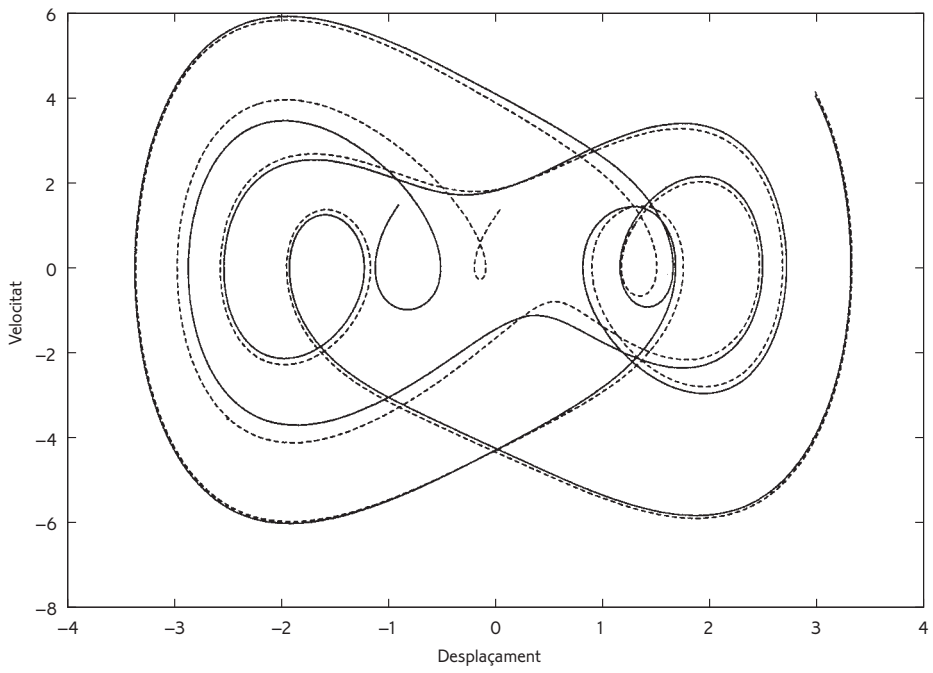
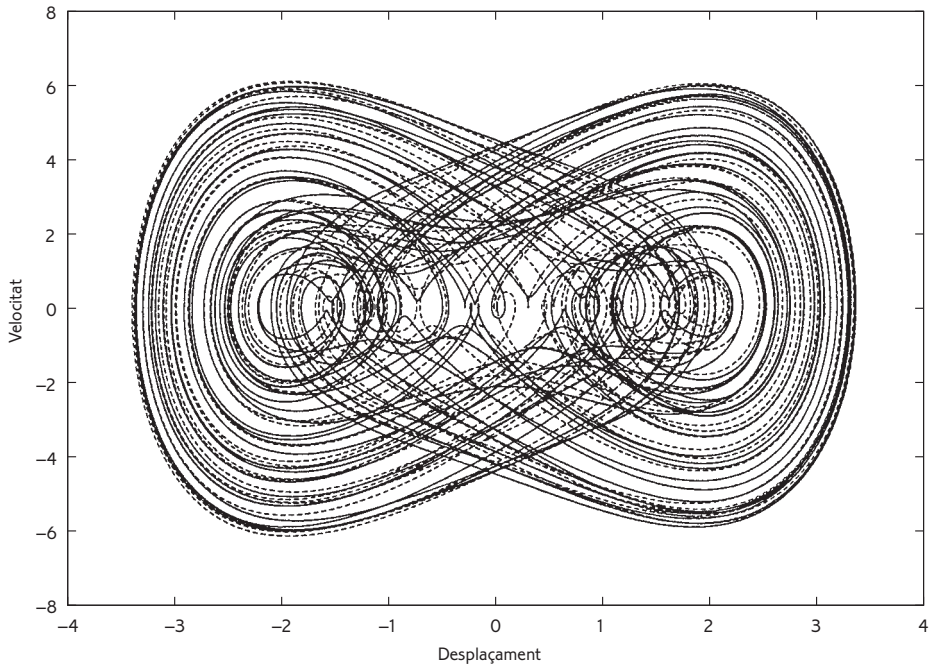


FIGURA 4

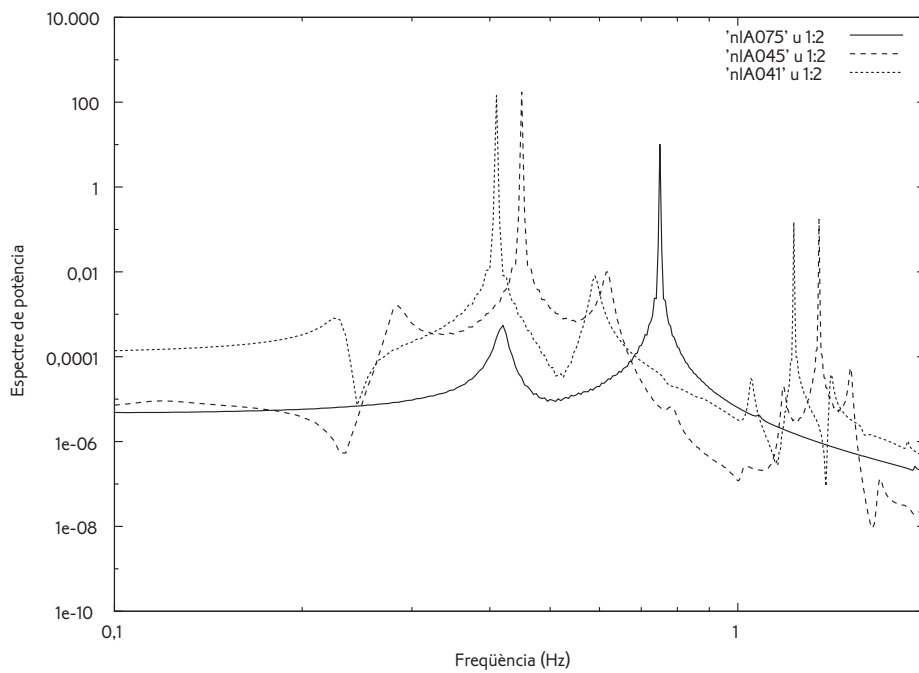
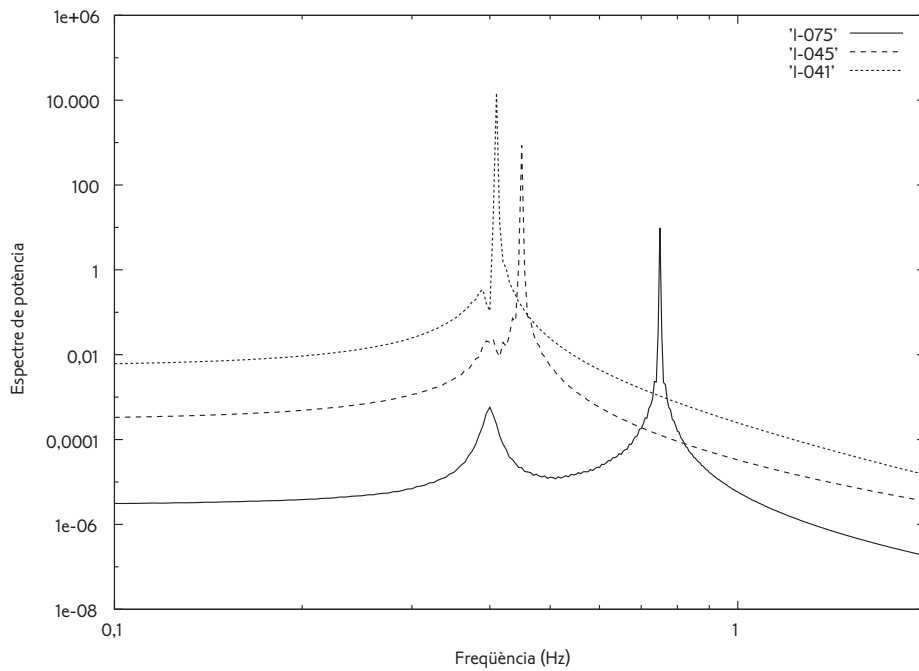
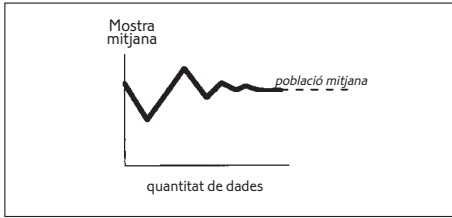
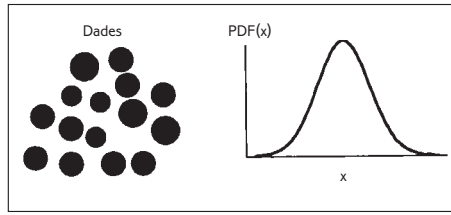


FIGURA 5

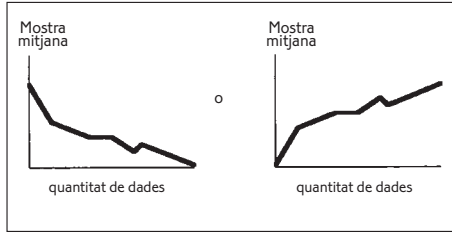
«Normal»



«Normal»



Fractal



Fractal

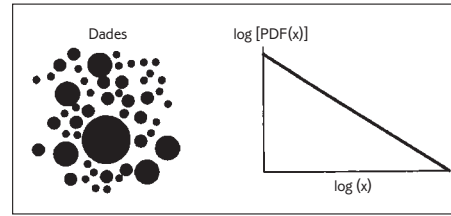


FIGURA 6

Mapa de Henon

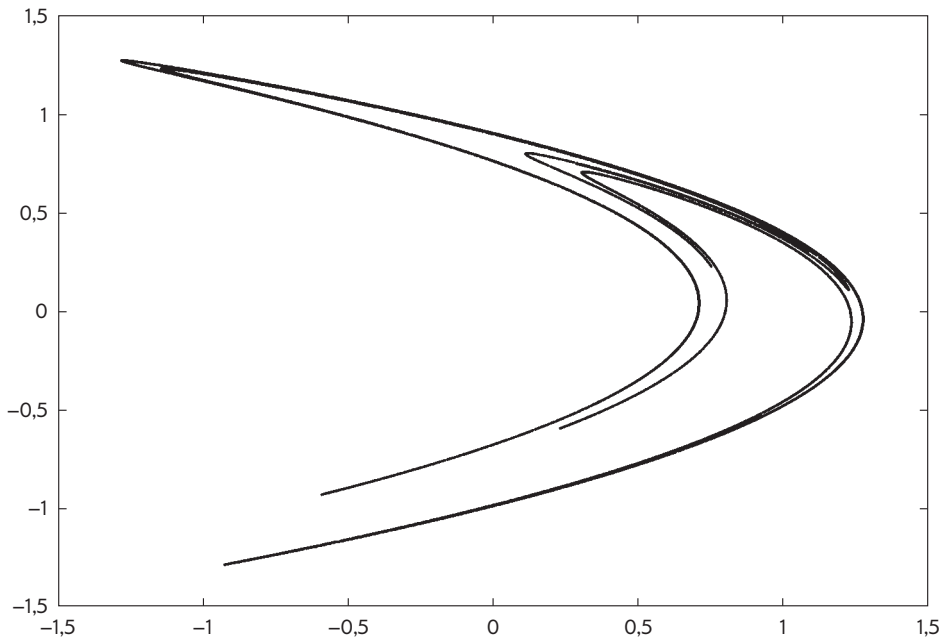


FIGURA 7

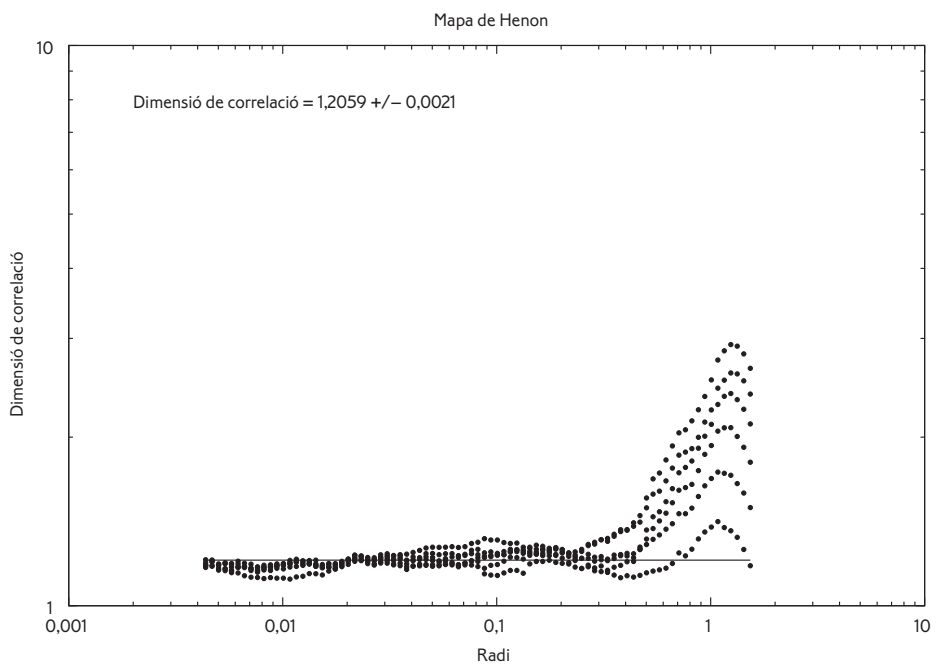
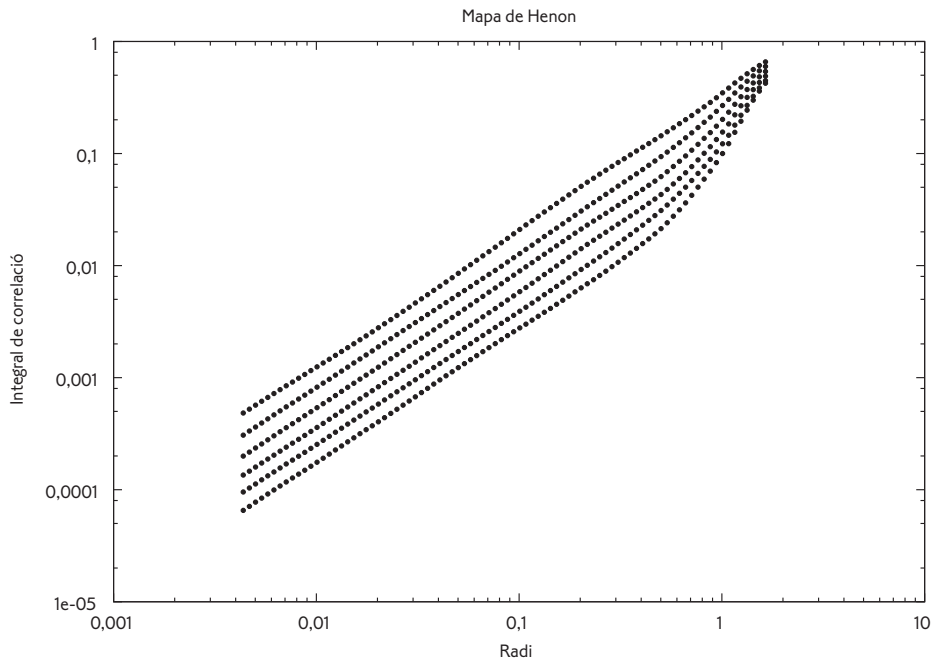


FIGURA 8

Prediction of critical transitions in complex systems: colliding cascades model and real disasters

Vladimir I. Keilis-Borok

International Institute of Earthquake Prediction Theory and Mathematical Geophysics,
Russian Academy of Sciences

Institute of Geophysics and Planetary Physics

Department of Earth and Space Sciences, University of California

The reactions and the attitudes with respect to such complexity are linked between two extremes: on the one hand is he who identifies as the sole possible solution to the problem a meticulous treatment of every process operating on the ... system, on the other hand is he who considers as the only hope that of "guessing" the right equations.

G. Puppi & A. Speranza

Abstract

This is a review of the recent studies in predictability of critical transitions (abrupt overall changes) in the hierarchical complex systems (i.e. in the aggregates of interacting elements).

1. A critical transition is preceded by certain patterns of the permanent background activity (static) of the system. A recently developed *lattice model of colliding cascades* (direct cascades of "loading" and inverse cascades of "failures") reproduces four basic types of such patterns, used in the rigorously defined robust prediction algorithms.

2. Most of the applications concern earthquake prediction; pilot applications to socio-economic predictions yield promising results.

3. Exhaustive sensitivity analysis demonstrated stability of predictions to inevitably non-unique choice of adjustable parameters and raw data. Experiments in advance prediction have been launched.

4. A prediction algorithm may be used in two ways: (i) as an unambiguous, reproducible description of phenomena, premonitory to a critical transition; that is an important heuristic constraint for development of fundamental theory of the process considered. (ii) On the practical side —for disasters' prediction and preparedness. Such is the usual twofold goal of a prediction research.

Resum

Aquesta és una ressenya dels estudis recents sobre la predictibilitat de transicions crítiques (canvis generals abruptes) en sistemes jeràrquicament complexos (per exemple, en agregats d'elements interactuants).

1. Una transició crítica va precedida per uns patrons determinats de l'activitat de fons permanent (estàtica) del sistema. Hi ha un *model d'enreixat de cascades en col·lisió* (cascades directes de «càrrega» i cascades inverses d'«errors») que reproduïx quatre tipus bàsics de patrons utilitzats en els algorismes de predicció robusta rigorosament definits.

2. La major part de les aplicacions tenen a veure amb la predicció de terratrèmols; les aplicacions pilot a la predicció socioeconòmica també ofereixen resultats prometedors.

3. L'anàlisi de sensibilitat exhaustiva ha manifestat una estabilitat de prediccions per a la tria no única de paràmetres ajustables i de dades sense processar. S'han iniciat experiments en el camp de la predicció prèvia.

4. Un algorisme de predicció es pot utilitzar de dues maneres: *a)* com a descripció de fenòmens reproduïble i sense ambigüitats, premonitòria d'una transició crítica; es tracta, doncs, d'una important limitació heurística al desenvolupament d'una teoria fonamental del fenomen considerat; *b)* pel cantó pràctic —per a la predicció i la prevenció de desastres. Aquest és el doble objectiu habitual de la recerca enfocada a la predicció.

Introduction

One of the most important and least understood features of complex systems is *persistent reoccurrence of the abrupt overall changes*, called “critical transitions” or “critical phenomena”. In applications they are referred to as crises, catastrophes, and disasters. Here, we consider hierarchical dissipative complex systems, which play important role in the global village. Among such systems is *the Earth's crust*, prone of geological disasters, which trigger in turn ecological and socio-economic catastrophes; *economy*, prone of depressions; *society*, prone of the bursts of large-scale violence; *a megacity*, on a way to self-inflicted collapse; etc.

Like the study of gravity had been between T. Brahe and J. Kepler, the study of such systems is in the “pre-equation stage”: the heuristic search of major regularities necessary to develop a fundamental theory. At that stage prediction is necessary for fundamental understanding of the system, as well as for practical goal of disasters preparedness.

The problem. Complex systems are not predictable with absolute precision, in the Laplacean sense. However, after a coarse-graining (averaging) the regular behavior patterns emerge; among them are scenarios of development of critical transitions. The problem of prediction —the focus of this review— is posed then as a consecutive, step-by-step reduction of the time-space domain, where a critical transition has to be expected (Keilis-Borok, 1996, 1990, 1999). At each step we consider the system on different levels of averaging, in a “holistic” approach —from the whole to details. Coarse-graining limits the accuracy of prediction, but it remains practically important (Molchan, 1990, 1991, 1997).

The division into consecutive approximations is dictated by the stage-by-stage development

of critical transitions (Kossobokov *et al.*, 1999). At the same time this division corresponds to the needs of disaster preparedness (Kantorovich & Keilis-Borok, 1991).

A more specific version of that problem is reviewed here: Consider dynamics of a system. Let t be the current moment in time, m —the scale (“magnitude”) of a critical transition. Given the behavior of the system prior to t , our problem is *to decide whether a critical transition with magnitude $m > m_0$ will occur or not during subsequent time interval $(t, t + \Delta)$* . In other words, we have to localize in time-space a specific singular trait of the process. This formulation is considerably different from a more traditional one —extrapolation of a process in time. Decision rule is called *the prediction algorithm* (an example is described in 2.3.2).

Dual nature of critical transitions. Phenomena precursory to critical transitions have been first encountered in seismicity, where critical transitions are the strong earthquakes (Keilis-Borok, 1990). Mathematical modeling and analysis of real data show that these phenomena are partly “universal”, common to the complex systems of different origin. On that background the system-specific precursors emerge.

Relevance to statistical physics. Premonitory seismicity patterns discussed below are qualitatively reminiscent of asymptotic behavior of a system near the point of phase transition of second kind. However, our problem is unusual for statistical physics: we consider not the equilibrium state, but the growing disequilibrium, culminated by a critical transition.

Raw data include observable background activity (“static”) of the system and external factors affecting it. Among examples of the static are the small earthquakes; variations of macroeconomics indicators; flow of misdemeanors; etc. The static in hierarchical systems may include critical transitions with smaller magnitude, forming their own hierarchy (Holland, 1995); the same phenomenon may be regarded as a part of the static for the whole system and as a critical transition for some part of the system.

Pattern recognition. Having to be heuristic, the data analysis is based on the “pattern recognition of infrequent events” —the methodology, developed by the school of I. Gelfand (Gelfand *et al.*, 1976). It allows to overcome the complexity and imprecision of the data in the situations when the data are insufficient for more traditional statistical analysis.

Performance of a prediction algorithm is quantitatively characterized by the rate of false alarms, rate of failures to predict, and total volume of alarms (see 2.3.2). The tradeoff between these characteristics is summed up by the *error diagrams*. They also provide an interface between prediction and preparedness (Molchan, 1997; Kantorovich & Keilis-Borok, 1991).

Validation: (“with four exponents I can fit the elephant”, *E. Fermi*). Inevitably in the absence of adequate fundamental equations, there is some freedom in the design of a prediction algorithm. The validation of the algorithms, taking a lion share of the efforts, comprises the following three stages:¹

1. Each algorithm described below was validated at the first two stages and put to test by advance prediction. We do not repeat this henceforth, to avoid monotonous repetitions.

(i) “*Sensitivity analysis*”: Retrospective check whether performance is insensitive to variations in the raw data and in adjustable numerical parameters.

(ii) “*Out of sample*” retrospective evaluation of performance on independent data, not used in the design of the algorithm.

(iii) *Advance prediction* —the only decisive test of a prediction algorithm.

Content. Section 2 describes the recently developed *model of colliding cascades* propagating in hierarchical chaotic systems (Gabrielov *et al.*, 2000a, 2000b, 2001; Zaliapin *et al.*, 2001a, 2001b). This is one of the lattice models of statistical physics type intensely used in the study of critical transitions. We consider that specific model because it exhibits a wide set of known precursors to a critical transition (Keilis-Borok, 1990, 1999) and some unknown precursors, to be tested on observations. Next two sections demonstrate prediction of the real (observed) critical transitions. *Section 3* concerns the earthquakes, *Section 4* —economic recessions and unemployment.

1. Colliding cascades

The model of colliding cascades (Gabrielov *et al.*, 2000a, 2000b, 2001; Zaliapin *et al.*, 2001a, 2001b) synthesizes three major factors responsible for generation of critical transitions in the complex systems.

(i) *Hierarchical structure.* Specifically we consider the ternary tree shown in the fig. 1a.

(ii) “*Loading*” by external sources. The load is applied to the largest element and transferred downwards, thus forming direct cascades.

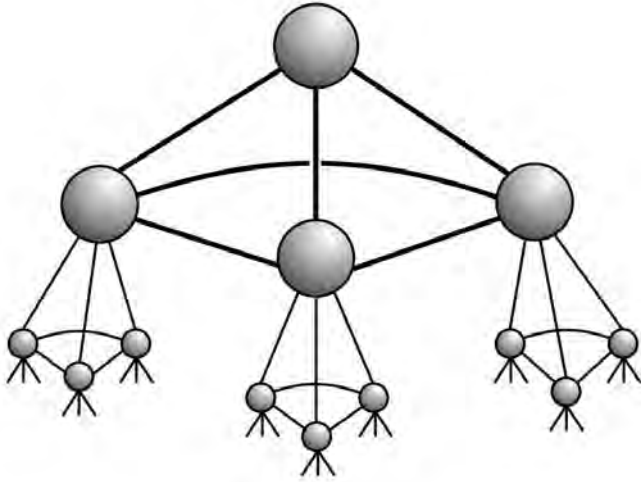
(iii) “*Failures*” of the elements under the load. The failures start from the smallest elements and gradually expand upwards the hierarchy, thus forming inverse cascades. A broken element eventually “heals”, which ensures continuous operation of the system.

Cascades are realized through interaction of adjacent elements, as shown in fig. 1b. Direct and inverse cascades collide and interact: loading triggers the failures, while failures release and redistribute the load. Dynamics of the model is described in (Gabrielov *et al.*, 2000a).

Fig. 2 shows an example of the sequence of failures. What is the relevance of this sequence (and of the whole model) to reality? The answer is that it fits the major heuristic constraints, derived from the earthquakes studies, and exhibits yet unknown regularities, which can be tested on observations.

Separately, the cascades of each type have been intensely studied in many fields. Among classical examples are direct cascades of eddies in the three-dimensional turbulent flow (Kolmogoroff, 1941a, 1941b; Frish, 1995), and inverse cascades in percolation (Stauffer & Aharony, 1992). Pioneering lattice models of seismicity (Burrige & Knopoff, 1967; Bak *et al.*, 1988; Allégre *et al.*, 1982) have been focused on the inverse cascades. This review concerns a much less explored phenomenon —*interaction* of direct and inverse cascades. It was considered also in a model of magnetic field reversals (Blanter *et al.*, 1999).

a)



b)

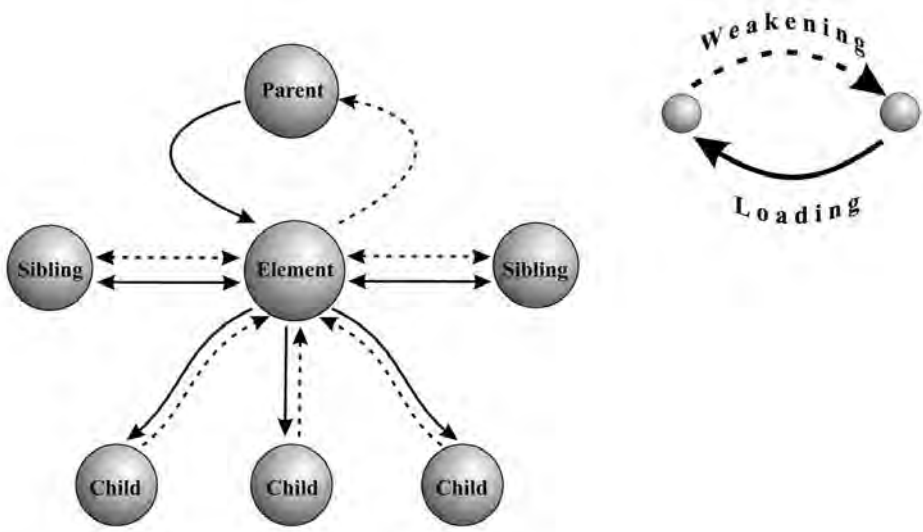


FIGURE 1. Structure of the colliding cascades model. a) Three highest levels of the hierarchy. b) Interaction with the nearest neighbors. [After Gabrielov *et al.*, 2000b]

2. Earthquakes

2.1. Definitions

To interpret colliding cascades in terms of seismicity, we associate the loading with the impact of tectonic forces, and the failures with earthquakes. The simplest standard representation of an observed earthquake sequence is

$$\{t_k, m_k, g_k\}, k=1, 2, \dots \quad [1]$$

Here, t_k is the occurrence time of an earthquake; m_k is its magnitude, i.e. logarithmic measure of energy released by the earthquake; g_k indicates coordinates of the hypocenter; and k is the sequence number of an earthquake, $t_k \leq t_{k+1}$.

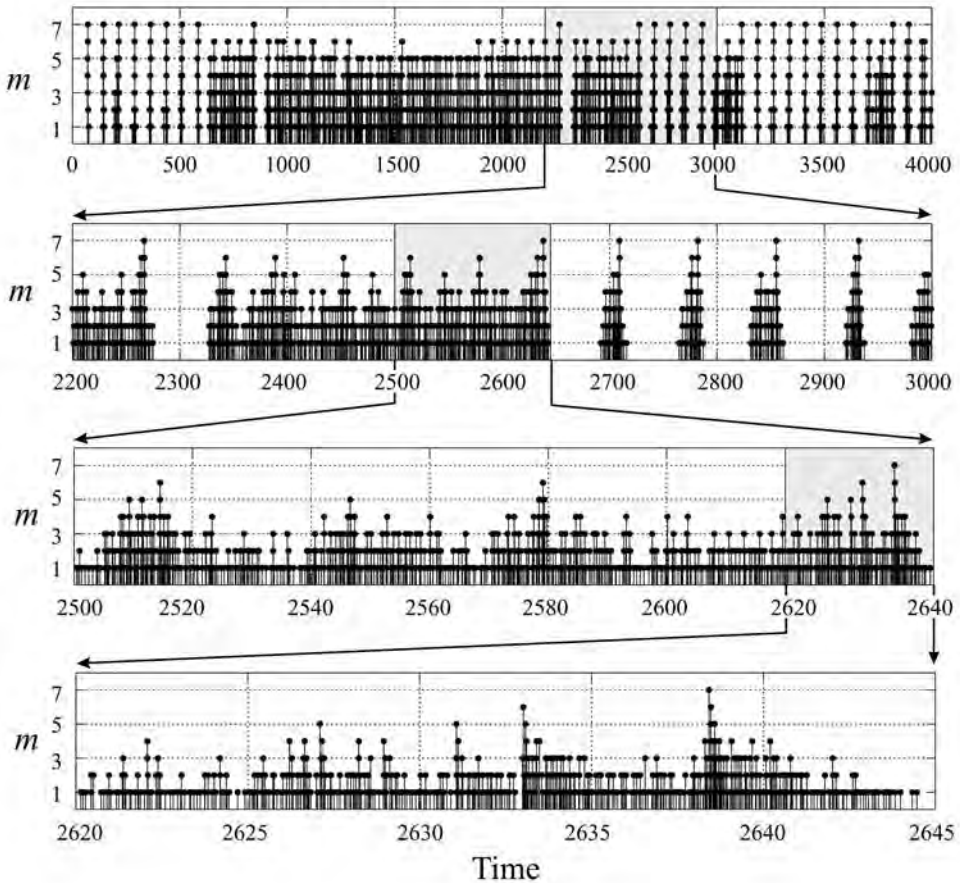


FIGURE 2. Synthetic earthquake sequence generated by the colliding cascades model. The complete sequence is shown in the top panel; exploded views —in the following three panels. [After Gabrielov *et al.*, 2000b]

Identifying the failures in the model with the real (observed) earthquakes, we regard m as the magnitude; this is a natural analogy, since the size of the rupture that originates a real earthquake is strongly correlated with magnitude. The position of an element in the tree is regarded as a hypocenter; this is an admittedly coarse analogy, since, strictly speaking, the model has no Euclidean hypocenter space.

2.2. Heuristic constrains

Synthetic sequence of failures, shown in fig. 2, exhibits, upon averaging, the major regular features of observed seismicity: *seismic cycles*, *intermittency*, *scale invariance*, and *a specific kind of clustering* (Gabrielov *et al.*, 2000a, 2000b, 2001). (The reader will recognize also a qualitative similarity with many other processes, e.g. dynamics of economy.) In the next section we describe one more constraint, central for the problem of prediction: the set of the spatio-temporal patterns of seismicity, preceding a strong earthquake.

(i) *Seismic cycles*. Our synthetic sequence is dominated by easily identifiable cycles, each culminating in a major earthquake. A single cycle is shown in the bottom panel of fig. 2. It comprises three consecutive phases (Scholz, 1990): raise of seismic activity, culminating in one or several major earthquakes; gradual decline of activity; and period of low activity, eventually followed by a next cycle.

(ii) *Intermittency of seismic regime*. From time to time the sequence of cycles abruptly changes its basic features: maximal magnitude, degree of periodicity, and duration of separate cycles (compare for example three intervals: 100-500, 1000-1200 and 1500-1700 in the top panel of fig. 2).

(iii) *Scale invariance*. Typically for complex systems with persistent critical transitions, the distribution of earthquakes' magnitude follows the power law, known in seismology as Gutenberg-Richter relation:

$$dN(m) \sim 10^{-bm} dm , \quad [2]$$

with the value of b being constant in considerable magnitude ranges (Gutenberg & Richter, 1954; Molchan & Dmitrieva, 1990; Turcotte, 1997; Kagan, 1999). This relation emerges after a sufficient averaging over territory and time. Fig. 3 shows that synthetic seismicity does follow this law (perhaps too perfectly); this is partly predetermined by the design of the model.

(iv) *Clustering*. Real earthquakes are clustered in time and space. The clustering is hierarchical, taking place in different scales. A most prominent type of clusters consists of a *main shock*, closely followed by a decaying sequence of weaker *aftershocks*. Also, about 30 % of main shocks are closely preceded by few weaker *foreshocks* (Molchan & Dmitrieva, 1990). The modeled sequence does exhibit the aftershocks as well as some other forms of clusters (Gabrielov *et al.*, 2000a). As in reality, the slope b in the distribution [2] is steeper for the aftershocks than for the main shocks (fig. 3).

2.3. Premonitory seismicity patterns

2.3.1. Three types of premonitory phenomena

Studies of observed and modeled seismicity show that an earthquake of magnitude m_0 is preceded by certain patterns of seismicity in an area and magnitude range normalized by m_0 . Specifically, these patterns reflect the following changes in seismicity (Keilis-Borok, 1994, 1996; Turcotte, 1997; Sornette & Sammis, 1995):

- (i) Rise of the earthquakes' clustering in space and time.
- (ii) Rise of intensity of earthquakes' flow.
- (iii) Rise of the range of correlations between the earthquakes.

Premonitory patterns of the first two types belong to heuristic constrains for the model. They have been found mainly by the analysis of observations and are used in the intermediate-term

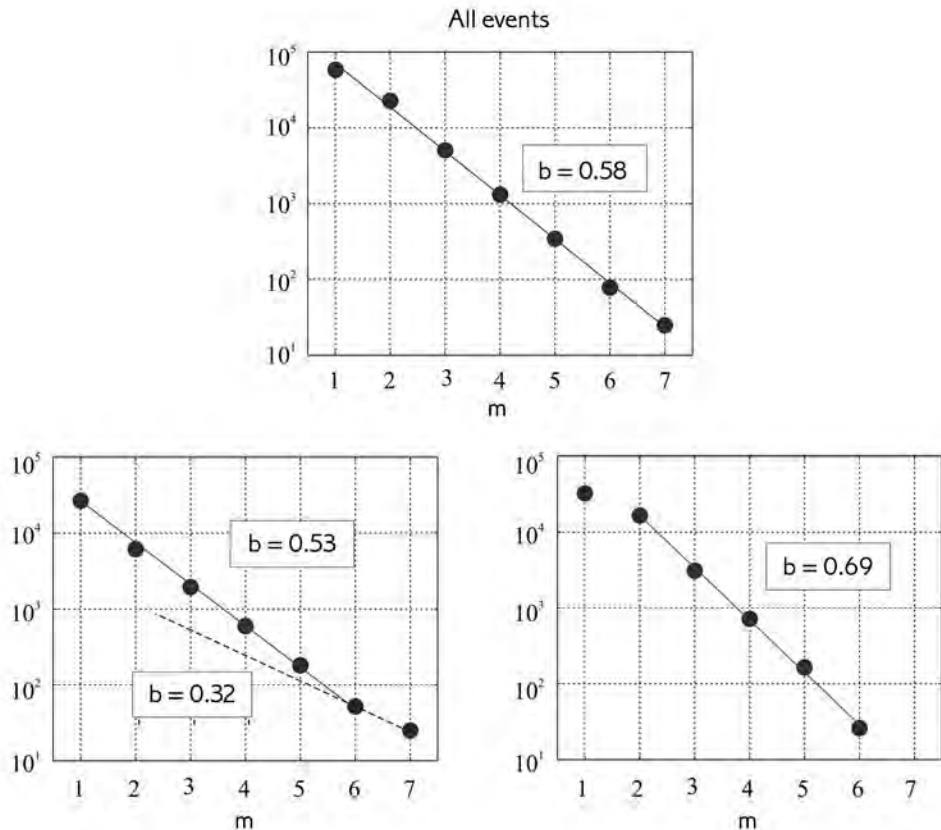


FIGURE 3. Magnitude distribution, $\log N(m) = a - bm$, $N(m)$ is the number of events with magnitude m . Note that the magnitude has discrete integer values from 1 to 7. a) All events. b) Main shocks. c) Aftershocks. [After Gabrielov *et al.*, 2000b]

earthquake prediction algorithms, with characteristic duration of alarms years (Keilis-Borok, 1990, 1999). The third type of patterns has been found just recently in the CC model (Gabrielov *et al.*, 2000a, 2000b, 2001), although it was previously hypothesized in (Keilis-Borok, 1994, 1996).

2.3.2. General scheme of prediction (Gabrielov *et al.*, 1986; Keilis-Borok, 1990; Keilis-Borok & Shebalin, 1999)

(i) *Areas.* The territory considered is scanned by overlapping areas; their size is normalised by the magnitude m_0 of the earthquakes targeted by prediction.

(ii) *Functionals.* An earthquake sequence in each area is described by a time function $F(t)$, depicting the premonitory changes of seismicity (see 2.3.1). Each type of changes can be depicted by different functionals (for example, clustering of the earthquakes may be measured by generation of aftershocks and by the swarms of main shocks).

The functionals are also normalized by m_0 .

(iii) *Premonitory patterns.* The emergence of a premonitory pattern in the area considered is recognized by the condition

$$F(t) \geq C_F \quad [3]$$

The threshold C_F is an adjustable parameter. It is usually defined to be a certain percentile of the functional F .

(iv) *Prediction algorithms.* An algorithm based on a single pattern is formulated as follows. Whenever $F(t) > C_F$, an alarm is declared for a time period Δ_F in the area considered. Prediction is *correct*, if an earthquake with magnitude in the range $(m_0, m_0 + c)$ occurs in that area and time interval; the opposite case is a *false alarm*. A *failure to predict* is the case when a major earthquake occurs outside of an alarm area and period.

Performance of an algorithm is quantitatively defined by three parameters: the rate of false alarms, the rate of failures to predict, and the total space-time, occupied by the alarms.

Most of the algorithms of that kind are based not on one but on several premonitory patterns. An alarm is declared when certain combinations of the patterns emerge (Gabrielov *et al.*, 1986; Keilis-Borok, 1996; Keilis-Borok & Shebalin, 1999).

(v) *Robustness.* A highly complex process (an earthquake sequence) is described in this analysis by the few averaged functionals, which in turn are defined at the lowest (binary) level of resolution —above or below certain threshold. Predictions themselves are also binary, of “yes or no” kind, with unambiguous definition of the area and duration of alarm. Probabilistic aspect of prediction is reflected in the errors diagram, showing the tradeoff between the parameters, characterizing performance of a prediction method (see item [iv] above).

Such robustness, usual in the exploratory data analysis (Gelfand *et al.*, 1976; Tukey, 1977), allows to overcome high complexity of the process considered, and chronic incompleteness of the data. This is achieved at the price, however: the limited accuracy of prediction.

2.3.3. Application to synthetic seismicity

The colliding cascades model reproduces the whole set of premonitory patterns described above. Prediction algorithms, based on each pattern, have been applied to synthetic earthquake sequence, shown in fig. 2 (Gabrielov *et al.*, 2000b). Prediction was targeted at the 25 “major earthquakes”, with $m = 7$. For the patterns of the first two types, depicting clustering and level of seismic activity, we used the *a priori* definitions, developed in analysis of observations (Keilis-Borok, 1990). For the patterns of the new type, depicting correlation range, we used definitions developed in (Gabrielov *et al.*, 2000a).

Fig. 4 summarizes performance of these patterns. The top panel shows, in separate boxes, the emergence of the patterns before each major earthquake. The bottom panel shows the false alarms. Evidently, most of them are close to the earthquakes with magnitude $m = 6$.

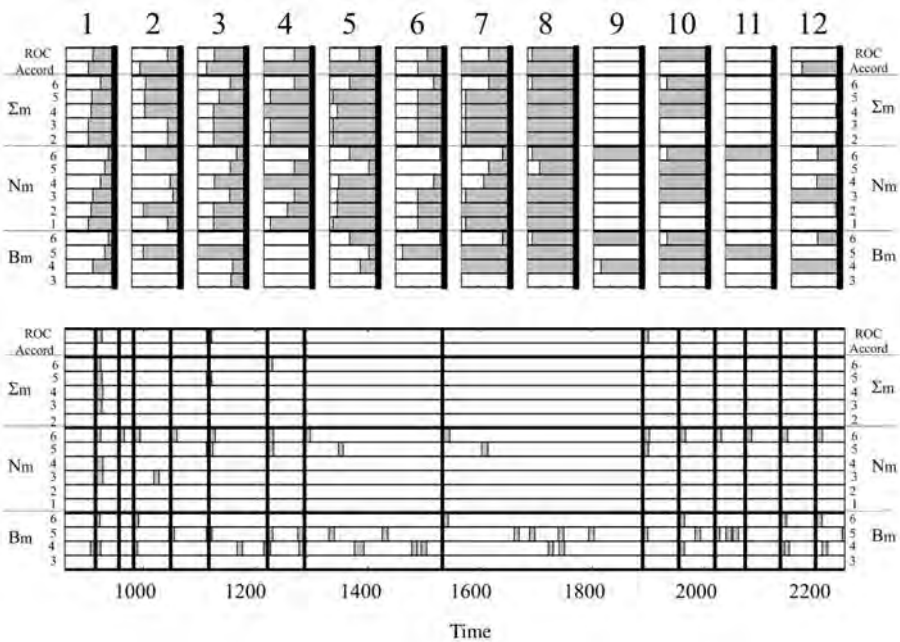


FIGURE 4. Collective performance of premonitory patterns. Prediction is targeted at the first 12 major events ($m = 7$) from synthetic sequence shown in fig. 2. Figure juxtaposes alarms generated by all 17 precursors considered in Gabrielov *et al.*, 2000b. The patterns *ROC* and *Accord* reflect correlation range, the patterns Σ and *N* —seismic activity, the pattern *B*— clustering (Gabrielov *et al.*, 2000a, 2000b). The top panel shows, in separate boxes, emergence of precursors before each of major earthquakes in the synthetic sequence. The right edge of each box is the moment of a major earthquake. The time interval of three units is considered before each one. The bottom panel shows the alarms determined during the time periods when strong earthquakes did not occur; those are the false alarms. Vertical lines show the moments of events with magnitude $m = 6$. Evidently, $m = 6$ events are associated with most of the false alarms. Each row shows the track record of a single precursor. Shaded areas show the alarms determined by the precursor. Values of m indicate the magnitude range in which a precursor is determined. [After Gabrielov *et al.*, 2000b]

2.3.4. Applications to real seismicity

Premonitory patterns of the first two kinds have been put to tests by advance prediction worldwide, by the scientists from Russia, US, France, Italy, and New Zealand; the tests, unprecedented in rigor and volume, established the high statistical significance of predictions (Molchan, 1990; Kossobokov *et al.*, 1999; Vorobieva, 1999; Harte *et al.*, 2000 and references therein), though the probability gain is low so far, between 3 and 10. Among predicted are 7 out of last 8 strongest earthquakes, of magnitude 8 or more.²

Patterns of the third type have yet to be validated by observations. First applications to observed seismicity, in S. California (Zaliapin *et al.*, 2000) and Lesser Antilles (Shebalin *et al.*, 2000), are encouraging.

Figs. 5-7 illustrate the predictions described above. Figure 5 shows successful prediction of Sumatera earthquake $M=8$. Figure 6 shows prediction of second strong earthquake —Northridge— that follows Landers earthquake within two years. Figure 7 depict similarity of premonitory seismicity patterns preceding Aquaba earthquake and major starquake —flash of energy radiated by a neutron star in the form of the soft γ -rays repeaters (Kossobokov *et al.*, 2000).

3. Recessions and unemployment

The lattice models used in the studies of seismicity (Burridge & Knopoff, 1967; Turcotte, 1997; Rundle *et al.*, 2000; Gabrielov & Newman, 1994; Shnirman & Blanter, 1999; Huang *et al.*, 1998), colliding cascade model included, are not necessarily specific to seismicity only. The notions central in such models —*hierarchy, loading, failures, and healing*— might be interpreted in terms of different complex processes.³ Here, we describe application of similar approach to socio-economic predictions (Keilis-Borok *et al.*, 2000a, 2000b; Lichtman and Keilis-Borok, 1999; Lichtman, 1996).

3.1. American economic recessions

Each of the five recessions, which occurred in the USA since 1962, has been preceded by a specific pattern of 6 macroeconomic indicators, determined in (Keilis-Borok *et al.*, 2000a). This pattern emerged within 5 to 13 months before each recession and at no other time, suggesting

2. Results of that test are routinely available on the website: <http://www.mitp.ru>.

3. Hypothetically, we encounter in the studies reviewed here an “arithmetic” of critical transitions —the basic precursory phenomena, common for a wide class of complex systems. However, if that is true, the universality is not unlimited: on the background of “universal” precursors the system-specific ones do emerge (Keilis-Borok, 1990, 1999).

a hypothetical prediction algorithm. It is put to test by advance prediction, starting from 1996. Since then no more recessions did occur, and by April 2001, when this is written, the algorithm generated no alarms.

3.1.1. Raw data. The following routinely available monthly macroeconomic indicators have been analyzed

1. Difference between the interest rate on 10 years U.S. Treasury bonds, and federal funds' interest rate, on annual basis.

2. The "Stock-Watson index" of overall monthly economic activity. It is a weighted average of four measures, depicting employment, manufacturing output, and retail sales, which emphasize services.

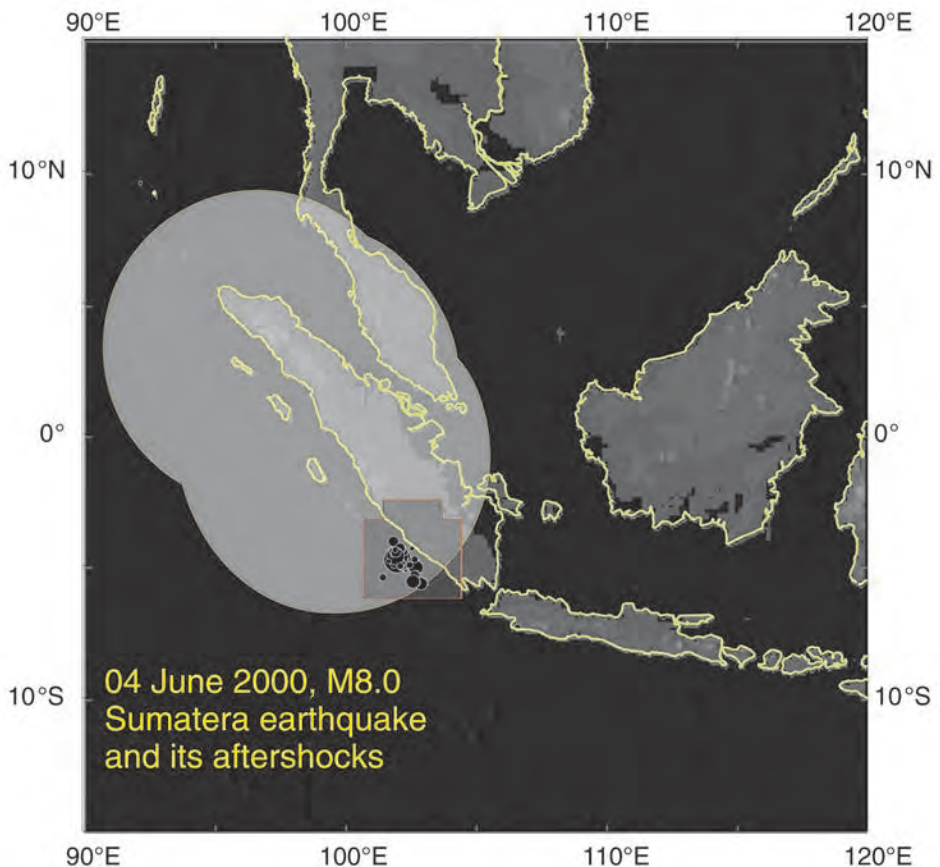


FIGURE 5. Successful advance prediction of the Southern Sumatra earthquake (4 June 2000, magnitude $M_s = 8.0$). The alarm areas in the first (M8 algorithm) and the second (MSc algorithm) approximations are highlighted by yellow and red respectively. [From the web site: <http://www.mitp.ru>]

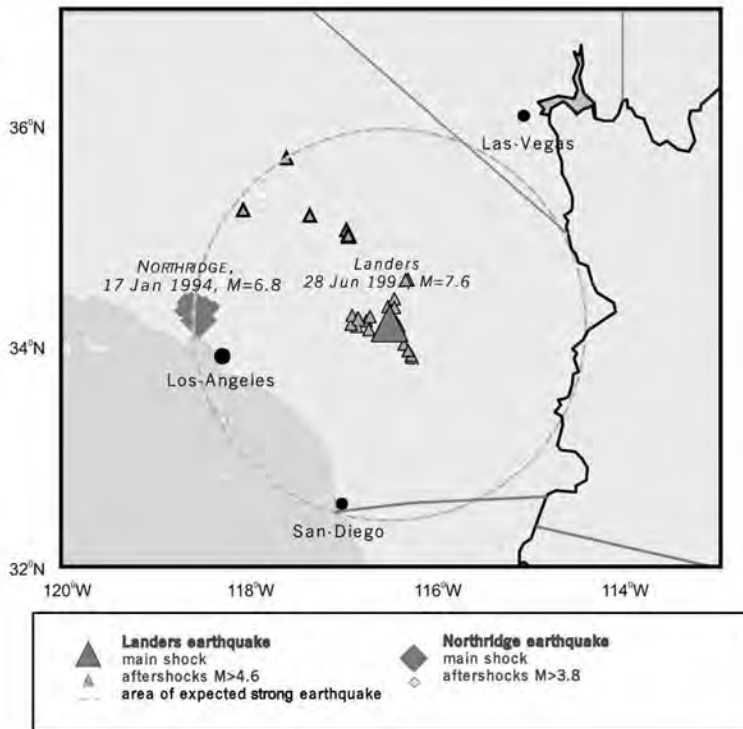


FIGURE 6. Prediction of the second strong earthquake. The Landers earthquake (28 June, 1992 M 7.6) was followed by the Northridge earthquake (17 January, 1994, M 6.8). The circle in the figure shows the territory, pointed by the algorithm, where occurrence of the second strong earthquake should be expected. [After Vorobieva, 1999]

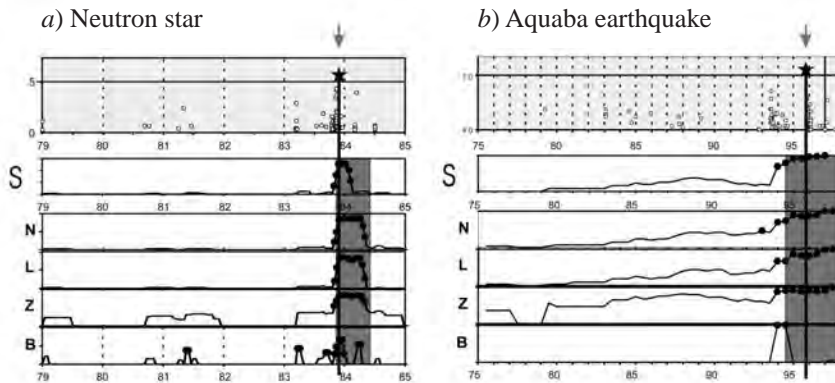


FIGURE 7. Similarity of premonitory seismicity patterns in Aquaba gulf (panel *b*) and on neutron star (panel *a*). Top - sequence of events. Major ones are shown by a star: Aquaba earthquake, 1995, energy $\sim 10^{23}$ erg and the starquake recorded in 1983, energy $\sim 10^{41}$ erg. Other boxes show premonitory patterns used in earthquake prediction. **S** depicts the change in energy distribution, **N** - the rate of earthquake occurrence, **L** - its deviation from long term trend, **Z** - concentration of events in space, **B** - clustering in space and time. Dots show the values, exceeding a standard threshold. [After Kossobokov *et al.*, 2000]

3. Index of “help wanted” advertising. This is put together by a private publishing company that measures the amount of job advertising (column-inches) in a number of major newspapers.
4. Average weekly number of people claiming unemployment insurance.
5. Total inventories in manufacturing and trade, in real dollars. Includes intermediate inventories (for example held by manufacturers, ready to be sent to retailers) and final goods inventories (goods on shelves in stores).
6. Interest rate on 90 days U. S. treasury bills at an annual rate.

These indicators happen to be sufficient for prediction; other potentially relevant indicators are not yet considered in this approach.

3.1.2. Hypothetical premonitory patterns

Premonitory behavior of each indicator was defined (see 2.3.2) by a robust binary condition (3). An example is shown in fig. 8a. It was found that the following patterns of the above indicators are premonitory, emerging more frequently, as a recession approaches: low value of indicator 1; strong downward trend of indicators 2 and 5; and strong upward trend of 3 other indicators. Qualitatively, this would be expected from the nature of each indicator.

A hypothetical prediction algorithm based on these patterns generated retrospective alarms, juxtaposed with recessions in fig. 8b. In the advance prediction such performance would be quite satisfactory.

Comparison with the more traditional multiregression analysis is summarized in Sect. 4 below.

3.2. *Unemployment*

Here, we summarize the study (Keilis-Borok *et al.*, 2000b) in prediction of unemployment in Western Europe and USA. The targets of prediction are the formally defined episodes of a sharp increase in the rate of unemployment, named “FAUs”, for “Fast Acceleration of Unemployment”. The FAUs in France, Germany, Italy, and USA since early sixties have been considered. Most of them are preceded by a uniform pattern of three macroeconomic indicators. A hypothetical prediction algorithm based on these patterns is put to test by advance prediction.

3.2.1. Raw data

The algorithm is based on the following three indicators, selected for a start among many relevant ones:

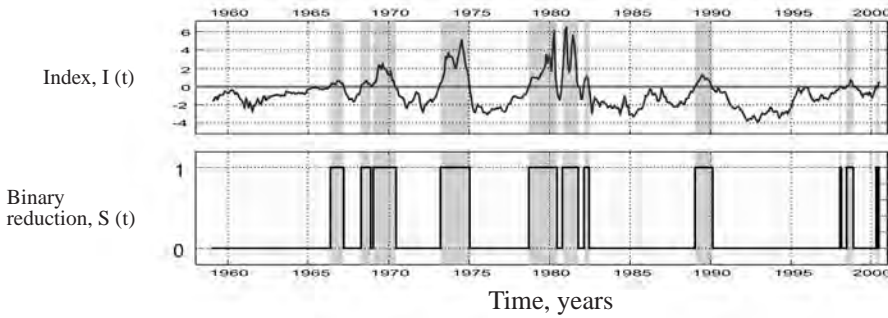
1. Industrial production index, composed of weighted production levels in numerous sectors of the economy, in % relative to the index for 1990.

2. Long-term interest rate on 10-year government bonds.
3. Short-term interest rate on 3-month bills.

3.2.2. Hypothetical premonitory pattern

Most of the FAUs considered are preceded by steep increase of all three indicators. The corresponding (hypothetical) prediction algorithm has been developed for France and applied then to Germany, Italy and USA. The alarms determined by that algorithm are juxtaposed with FAUs in fig. 9. Such results would be quite satisfactory in real time prediction, and encourage the further tests of an algorithm. So far, only one prediction, for USA, was made in advance (fig. 10). An alarm was determined for the period from February to November 2000 (a shaded area in Fig. 10 and unemployment did start to rise in July 2000).

a)



b)

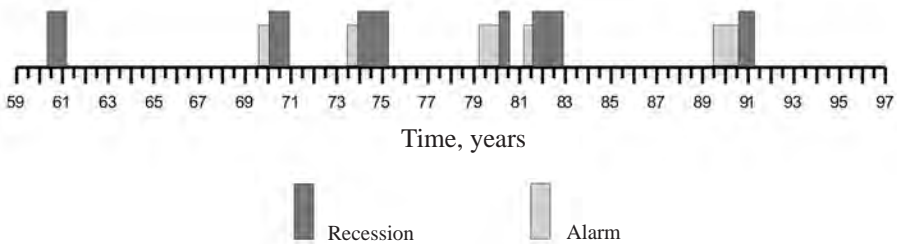


FIGURE 8. Prediction of US recession by premonitory patterns of 6 macroeconomic indicators. a) Each index $I(t)$ is replaced by its binary representation $S(t)$. Here I is the difference between the interest rate on short-term and long-term bonds. When $I(t) > 0$ $S(t) = 1$ (typical before recessions). When $I(t) < 0$, $S(t) = 0$. b) Performance of hypothetical prediction algorithm: an alarm is declared for 9 months when 4 or more indicators signal approach of a recession.

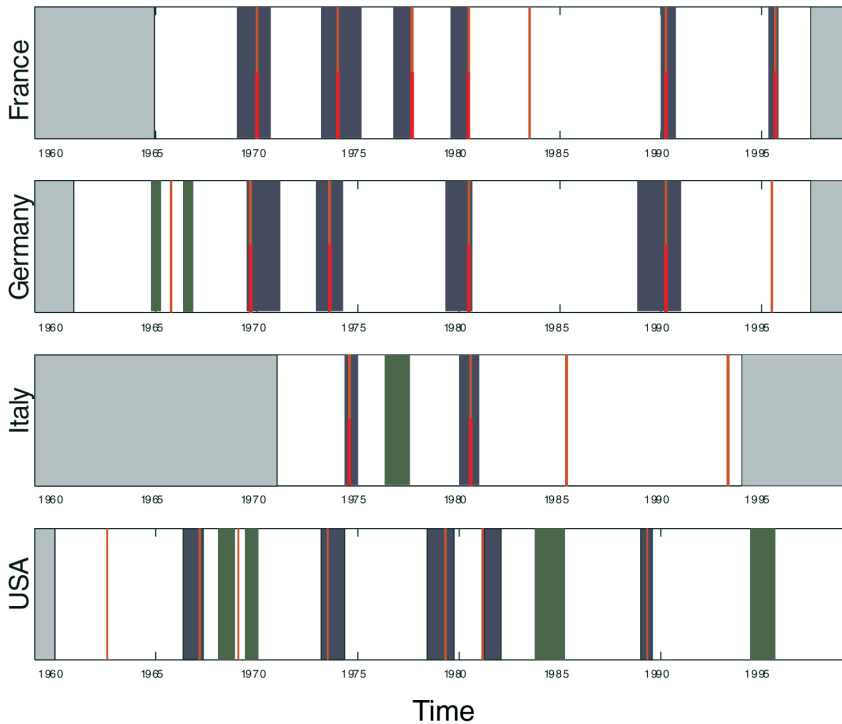


FIGURE 9. Prediction of fast acceleration of unemployment (FAU). Precursor is the steep rise of three national macro-economic indicators. Red vertical lines - moments of FAU. Blue bars - periods of alarms. Green bars - false alarms. Gray areas on both sides - periods, for which the economic indicators were unavailable.

4. Summary

1. The following findings seem promising.

— In each case, prediction rule was uniform, transcending the immense complexity of the process considered, diversity of prediction targets, and change of circumstances in time. For the earthquakes, premonitory seismicity patterns happen to be similar for microcracks in laboratory samples, largest earthquakes of the world, in the energy range from 10^{-1} erg to 10^{26} erg; possibly, also, for ruptures in a neutron star, 10^{41} erg. For recessions and unemployment similarity overcomes the changes of economy since 1962; for FAUs, also the differences between the countries.

— Prediction was based on the routinely available data —global catalogs of earthquakes, major macroeconomic indicators, etc.

— In each case predictability was achieved by extremely robust analysis. It is interesting to quote how an economist with excellent track record in recessions' prediction describes his impressions on such a robustness:

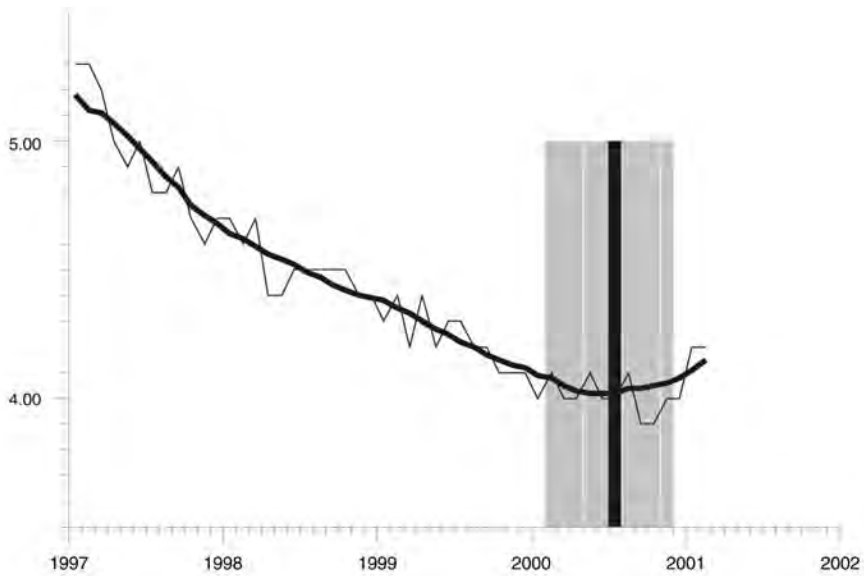


FIGURE 10. Unemployment rate in the U. S., July 2000: thin curve shows original data; thick curve show the rates with seasonal variations smoothed out. The gray bar shows the alarm period, defined by analysis of microeconomic indicators. The black bar - actual start of the unemployment rise, defined by analysis of monthly unemployment rates for the U. S. civilian labor force.

Prediction of recessions [...] requires fitting non-linear, high-dimensional models to a handful of observations generated by a possibly non-stationary economic environment [...]. The evidence presented here suggests that these simple binary transformations of economic indicators have significant predictive content comparable to or, in many cases, better than that of more conventional models (J. Stock, from Keilis-Borok *et al.*, 2000a).

The accuracy of prediction is limited. However, only a small part of relevant models and available data was used, so that a wealth of possibilities for better prediction remains unexplored.

2. Similar approach was successfully applied to prediction of president and mid-term senatorial elections in the USA (Lichtman & Keilis-Borok, 1989; Lichtman, 1996).

3. Although, obviously, neither a panacea, nor an easy ride are implied, the approaches, discussed here, open a reasonable hope to break the current stalemate in disasters' prediction.

4. The studies summarized here comprise cooperation of about 20 institutions in 12 countries and several international projects. Still, this review is not intended to be encyclopedic and covers a specific part of the much broader effort.

Acknowledgements

Preparation of this review was partly supported by The 21st Century Collaborative Activity Award for Studying Complex Systems (James S. McDonnell Foundation).

References

1. ALLÉGRE, C. J.; LE MOUËL, J. L.; PROVOST, A. (1982). "Scaling rules in rock fracture and possible implications for earthquake prediction". *Nature*, 297, 47-49.
2. BAK, P.; TANG, C.; WIESENFELD, K. (1988). "Self-organized criticality". *Phys. Rev. A*, 38, 364-374.
3. BLANTER, E. M.; NARTEAU, C.; SHNIRMAN, M. G.; LE MOUËL, J. L. (1999). "Up and down cascade in a dynamo model: spontaneous symmetry breaking". *Phys. Rev. E*, 59, 5112-5123.
4. BURRIDGE, R.; KNOPOFF, L. (1967). "Model and theoretical seismicity". *Bull. Seism. Soc. Am.*, 57, 341-371.
5. FRISH, U. (1995). *Turbulence: The legacy of Kolmogorov*. Cambridge University Press, Cambridge.
6. GABRIELOV, A.; DMITRIEVA, O. E.; KEILIS-BOROK, V. I.; KOSSOBOKOV, V. G.; KUZNETSOV, I. V.; LEVSHINA, T. A.; MIRZOEV, K. M.; MOLCHAN, G. M.; NEGMATULLAEV, S. Kh.; PISARENKO, V. F.; PROZOROFF, A. G.; RINEHART, W.; ROTWAIN, I. M.; SHEBALIN, P. N.; SHNIRMAN, M. G.; SHREIDER, S. Yu (1986). *Algorithm of long-term earthquakes' prediction*. Centro Regional de Sismología para América del Sur, Lima.
7. GABRIELOV, A.; NEWMAN, W. I. (1994). "Seismicity modeling and earthquake prediction: a review". In: NEWMAN, W. I.; GABRIELOV, A.; TURCOTTE, D. L. [eds.]. *Nonlinear dynamics and predictability of geophysical phenomena*. AGU., Int. Un. of Geodesy and Geophys., 7-13. Geophysical Monograph 83, IUGG, 18.
8. GABRIELOV, A.; KEILIS-BOROK, V. I.; ZALIAPIN, I. V.; NEWMAN, W. I. (2000a). "Critical transitions in colliding cascades". *Phys. Rev. E*, 62, 237-249.
9. GABRIELOV, A. M.; KEILIS-BOROK, V. I.; ZALIAPIN, I. V.; NEWMAN, W. I. (2000b). "Colliding cascades model for earthquake prediction". *JGI*, 143, 427-437.
10. GABRIELOV, A.; ZALIAPIN, I.; KEILIS-BOROK, V. I. (2001). *Premonitory seismicity patterns in cellular automata with colliding cascades*. [Manuscript]
11. GELFAND, I. M.; GUBERMAN, Sh. A.; KEILIS-BOROK, V. I.; KNOPOFF, L.; PRESS, F.; RANZMAN, E. Ya; ROTWAIN, I. M.; SADOVSKY, A. M. (1976). "Pattern recognition applied to earthquake epicenters in California". *Phys. Earth and Planet. Inter.*, 11, 227-283.
12. GUTENBERG, B.; RICHTER, C. F. (1954). *Seismicity of the Earth and associated phenomena*. Princeton University Press, Princeton.
13. HARTE, D.; LI, D. F.; VREEDE, M.; VERE-JONES, D. (2000). *A model for earthquake probabilities based on the M8 algorithm*. [Manuscript]

14. HOLLAND, J. H. (1995). "Hidden order: how adaptation builds complexity". Addison-Wesley, Reading, MA.
15. HUANG, Y.; SALEUR, H.; SAMMIS, C.; SORNETTE, D. (1998). "Precursors, aftershocks, criticality and self-organized criticality". *Europhys.*, lett. 41, 43-48.
16. KAGAN, Y. (1999). "Universality of the seismic moment-frequency relation". *Pure Appl. Geoph. (PAGEOPH)*, 155, 537-573.
17. KANTOROVICH, L. V.; KEILIS-BOROK, V. I. (1991). "Earthquake prediction and decision-making: social, economic and civil protection aspects". International Conference on Earthquake Prediction: State-of-the-Art, Strasbourg, France. *Scientific-Technical Contributions, CSEM-EMSC*, 586-593.
18. KEILIS-BOROK, V. I. (1994). "Symptoms of instability in a system of earthquake-prone faults". *Physica D*, 77, 193-199.
19. KEILIS-BOROK, V. I. (1996). "Intermediate-term earthquake prediction". *Proc. Nat. Acad. Sci. USA*, 93, 3748-3755.
20. KEILIS-BOROK, V. I. [ed.] (1990). "Intermediate-term earthquake prediction: models, phenomenology, worldwide tests". *Phys. Earth Planet. Inter.*, 61, Special Issue, 1-2, 144.
21. KEILIS-BOROK, V. I.; SHEBALIN, P. N. [eds.] (1999). "Dynamics of lithosphere and earthquake prediction". *PEPI*, Special Issue, III, 179-330.
22. KEILIS-BOROK, V. I.; STOCK, J. H.; SOLOVIEV, A. A.; MIKHALEV, P. (2000a). "Pre-recession pattern of six economic indicators in the USA". *Journal of Forecasting*, 19, 1, 65-80.
23. KEILIS-BOROK, V. I.; SOLOVIEV, A. A.; ALLÈGRE, C. B.; SOBOLEVSKII, A. N.; INTRILIGATOR, M. D. (2000b). *Patterns of macroeconomic indicators preceding the fast acceleration of unemployment in France*. [Manuscript]
24. KOLMOGOROFF, A. N. (1941a). "The local structure of turbulence in an incompressible viscous fluid for very large Reynolds number". *Doklady Akademii Nauk SSSR*, 30, 299-303.
25. KOLMOGOROFF, A. N. (1941b). "On degeneration (decay) of isotropic turbulence in an incompressible viscous fluid". *Doklady Akademii Nauk SSSR*, 31, 538-540.
26. KOSSOBOKOV, V. G.; ROMASHKOVA, L. L.; KEILIS-BOROK, V. I.; HEALY, J. H. (1999). "Testing earthquake prediction algorithms: statistically significant advance prediction of the largest earthquakes in the Circum-Pacific, 1992-1997". *PEPI*, 111, 187-196.
27. KOSSOBOKOV, V. G.; KEILIS-BOROK, V. I.; CHENG, B. (2000). "Similarities of multiple fracturing on a neutron star and on the Earth". *Phys. Rev. E*, 61, 4, 3529-3533.
28. LICHTMAN, A. J.; KEILIS-BOROK, V. I. (1989). "Aggregate-level analysis and prediction of midterm senatorial elections in the United States, 1974-1986". *Proc. Natl. Acad. Sci. USA*, 86, 10176-10180.
29. LICHTMAN, A. J. (1996). "The keys to the White House". Lanham [ed.]: Madison Books.
30. MOLCHAN, G. M.; DMITRIEVA, O. E.; ROTWAIN, I. M.; DEWEY, J. (1990). "Statistical analysis of the results of earthquake prediction, based on burst of aftershocks". *Phys. Earth Planet. Inter.*, 61, 128-139.

31. MOLCHAN, G. M. (1990). "Strategies in strong earthquake prediction". *Phys. Earth Planet. Inter.*, 61, 84-98.
32. MOLCHAN G. M.; DMITRIEVA, O. (1990). "Dynamics of the magnitude-frequency relation for foreshocks". *Phys. Earth Planet. Inter.*, 61, 99-112.
33. MOLCHAN, G. M. (1991). "Structure of optimal strategies in earthquake prediction". *Tectonophysics*, 193, 267-276.
34. MOLCHAN, G. M. (1997). "Earthquake prediction as a decision-making problem". *Pure and Appl. Geophys.*, 149, 233-247.
35. RUNDLE, J. B.; TURCOTTE, D. L.; KLEIN, W. [eds.] (2000). *Geocomplexity and the physics of earthquakes*. AGU, Washington, DC.
36. SCHOLZ, C. H. (1990). *The mechanics of earthquakes and faulting*. Cambridge University Press, Cambridge.
37. SHEBALIN, P.; ZALIAPIN, I. V.; KEILIS-BOROK, V. I. (2000). "Premonitory raise of the earthquakes' correlation range: Lesser Antilles". *PEPI*, 122, 3-4, 241-249.
38. SHNIRMAN, M. G.; BLANTER, E. M. (1999). "Mixed hierarchical model of seismicity: scaling and prediction". *PEPI*, 111, 295-304.
39. SORNETTE, D.; Sammis, C. G. (1995). "Complex critical exponents from renormalization group theory of earthquakes: implications for earthquake predictions". *J. Phys. I. France*, 5, 607-619.
40. STAUFFER, D.; AHARONY, A. (1992). *Introduction to percolation theory*. 2nd ed. Taylor and Francis, London. 181 p.
41. TUKEY, J. W. (1977). *Exploratory data analysis*. Addison-Wesley, Reading, MA.
42. TURCOTTE, D. L. (1997). *Fractals and chaos in geology and geophysics*. 2nd ed. Cambridge University Press, Cambridge.
43. ZALIAPIN, I. V.; KEILIS-BOROK, V. I.; GHIL, M. (2001a). *A boolean delay model of colliding cascades. I: Multiple seismic regimes*. Preprint ICTP, Trieste.
44. ZALIAPIN, I. V.; KEILIS-BOROK, V. I.; GHIL, M. (2001b). *A boolean delay model of colliding cascades. II: Prediction of critical transitions*. Preprint ICTP, Trieste.
45. ZALIAPIN, I. V.; KEILIS-BOROK, V. I.; AXEN, G. (2000). *Premonitory spreading of seismicity over the fault network in S. California: precursor accord*. 20 p. [Submitted to *J. Geophys. Res.*]
46. VOROBIEVA, I. A. (1999). "Prediction of a subsequent large earthquake". *Phys. Earth and Planet. Inter.*, 111, 3-4, 197-206.



INSTITUT
D'ESTUDIS
CATALANS

