

**Resoconto sulla “Third Mediterranean School on
Mesoscale Meteorology: Fundamentals and
applications of probabilistic forecasts and Ensemble
Prediction Systems”
Alghero 25–30 Maggio 2008**

Agostino Manzato

ARPA – Agenzia Regionale per la Protezione dell’Ambiente del FVG,
settore OSMER – Osservatorio Meteorologico Regionale, Palmanova (UD), Italy

1 Introduzione

Quest’anno si è svolta la terza edizione della Mediterranean School on Mesoscale Meteorology (MSMM) organizzata dal SAR (Servizio Agrometeorologico Regionale per la Sardegna), dall’ARPA Sardegna e dall’ISAC (Istituto di Scienze dell’Atmosfera e del Clima del CNR). Gli organizzatori sono stati in particolare il dr. Andrea Buzzi dell’ISAC e il dr. Piero Chessa, che ha lavorato come ricercatore al SAR fino all’anno scorso, mentre poi ha cambiato tipologia di lavoro. Per questo motivo è stato sostituito all’ultimo momento dal dr. Alessandro Delitala, sempre del SAR (che sta transitando in ARPA Sardegna, come avvenuto tempo fa per il nostro CSA). Il sito di riferimento della scuola MSMM è:

<http://msmm.sar.sardegna.it>

da dove è possibile scaricare il programma e le lezioni, che sono anche disponibili in un CD-ROM depositato presso la nostra biblioteca.

Se la prima edizione della scuola brillava per la presenza di insegnanti come Kerry Emanuel (MIT) e Chuck Doswell (CIMMS), la seconda edizione ha portato ad Alghero il prof. Brian Hoskins (Università di Reading), mentre in questa terza edizione tutti i tre docenti principali erano molto famosi e hanno richiamato ad Alghero una partecipazione di studenti superiore alle edizioni precedenti. Infatti i partecipanti sono stati circa una sessantina, con una folta partecipazione di Italiani e Spagnoli (una decina per gruppo), seguiti da Tedeschi e Francesi, ma non mancavano studenti americani e cinesi.

Gli argomenti trattati sono stati abbastanza diversi da quelli di cui mi sono occupato finora e per questo ho avuto delle difficoltà a seguire i dettagli più tecnici.

In particolare sono stati approfonditi i seguenti aspetti: *data assimilation* per creare l'analisi dei modelli (tecniche variazionali 3D-var e 4D-var vs. Kalman Filters), *perturbazioni iniziali* per creare un ensemble di previsioni diverse (Perturbed Observations, Bred Vectors, Singular Vectors, Ensemble Transform Kalman Filters), *verifica* delle previsioni probabilistiche e tecniche di calibrazione (incluse le *reforecast*). Di seguito potete trovare un riassunto delle cose che più mi hanno colpito delle lezioni. L'ultimo giorno ci sono stati dei seminari tenuti da altre persone, di cui riferirò brevemente.

2 Eugenia Kalnay

ekalnay@atmos.umd.edu ,
Department of Atmospheric and Oceanic Science. University of Maryland, College Park (MD), USA
<http://www.atmos.umd.edu/~ekalnay/>



Questa famosa professoressa del Maryland ha ormai collezionato un numero elevatissimo di pubblicazioni, tra le quali va citato almeno il recente libro intitolato *Atmospheric Modeling, Data Assimilation and Predictability*, pubblicato nel 2002 dalla Cambridge University Press. Di origine argentina, ha studiato meteorologia all'Università di Buenos Aires e successivamente ha fatto il PhD al MIT con Julie G. Charney. Dopo aver lavorato alla NASA, è diventata direttrice per 10 anni del gruppo di sviluppo dell'NCEP, dirigendo tra l'altro il progetto per i 50 anni di NCEP/NCAR reanalysis (Kalnay, Kanamitsu et al. 1996: *The NCEP/NCAR 40-Year Reanalysis Project*, BAMS, **77**, 437–431.) Assieme a Zoltan Toth ha sviluppato il modello probabilistico di NCEP (Toth e Kalnay 1993: *Ensemble forecasting at NMC: the generation of perturbations*, BAMS, **74**, 2317–2330), introducendo il formalismo dei *Bred Vectors* per sviluppare le perturbazioni iniziali.

2.1 Il tempo è un sistema caotico

All'inizio delle sue lezioni, Kalnay ha fatto un'introduzione al *chaos* nei sistemi dinamici, definito da Ed Lorenz come “Quando il presente determina il futuro, ma l'approssimazione del presente non determina l'approssimazione del futuro”. Questo per dire che ci sono situazioni in cui piccole perturbazioni dello stato iniziale portano a soluzioni molto divergenti nel futuro. Da qui nasce l'esigenza di testare l'*instabilità dinamica* di un sistema, ovvero di stimare l'incertezza sulla sua evoluzione futura. Il sito del loro gruppo che approfondisce queste tematiche è: <http://www.weatherchaos.umd.edu>

2.2 L'assimilazione delle osservazioni per l'analisi iniziale

Il problema del *data assimilation* nasce dal fatto che i modelli meteorologici (NWP) attuali hanno un numero di variabili usate per *definire* lo stato dell'atmosfera dell'ordine di $10^7 - 10^8$ (vettore \mathbf{x}), mentre tutte le osservazioni usate per *misurare* lo stato dell'atmosfera sono dell'ordine di 10^5 (vettore \mathbf{y}) e quindi sono molte di meno dei gradi di libertà del modello¹. Per questo motivo bisogna inizializzare i modelli con i campi previsti dai modelli precedenti (background field o first guess) modificati in base alle osservazioni acquisite. Il processo con cui il campo iniziale (fatto dalle ultime previsioni disponibili) viene interpolato in base alle osservazioni si chiama "data assimilation". In termini generali si può parlare di stima bayesiana, in cui uno stato noto *a priori* (il background) viene combinato con la verosimiglianza delle osservazioni per produrre lo stato *a posteriori* (l'analisi aggiornata).

Per far ciò, all'inizio si cerca un operatore matriciale \mathbf{H} che proietta il vettore che definisce lo stato dell'atmosfera previsto (\mathbf{x}_0) sullo spazio delle osservazioni, che, come detto, ha una dimensione minore dello spazio "del modello". Si ottiene quindi un vettore chiamato *innovation* definito come: $[\mathbf{y}_{\text{obs}} - \mathbf{H}(\mathbf{x}_0)]$, che misura le discrepanze tra il background e le misure effettuate, cercando di bilanciare l'*incertezza della previsione* precedente con l'*incertezza associata alla misura* dei dati osservati. Ci sono diversi metodi possibili per trovare il modo di pesare (tramite la matrice "rettangolare" \mathbf{W} , dell'ordine di $10^5 \times 10^8$) le diverse differenze ma alla fine ciò che viene prodotto è l'*analisi iniziale*, definita come:

$$\mathbf{x}_a(t_0) = \mathbf{x}_0 + \mathbf{W} \cdot [\mathbf{y}_{\text{obs}} - \mathbf{H}(\mathbf{x}_0)] \quad (1)$$

Una volta ottenuta l'analisi si può finalmente procedere a stimare la previsione futura calcolando $\mathbf{x}(t) = \mathbf{M}[\mathbf{x}_a(t_0)]$, dove \mathbf{M} rappresenta il modello meteorologico non-lineare.

Uno dei primi metodi per l'assimilazione è stato l'*Optimal Interpolation*, introdotto dal grande Lev Gandin (Gandin 1963: *Objective analysis of meteorological field*, Gidrometeorologicheskoe Izdate'stvo., Leningrad, URSS., 286pp.), che ha sviluppato le idee abbozzate da Kolmogorov, trovando \mathbf{W} come minimizzazione dell'errore su ogni grid-point e quindi ottenendo la matrice di covarianza tra previsione e osservazione.

I metodi variazionali tipo *3D-var* sono invece basati sull'idea di cambiare direttamente il modello invece che di trovare la matrice dei pesi \mathbf{W} e ciò viene perseguito minimizzando un'apposita funzione di costo J . Si ottiene la stessa soluzione di prima (ma con tecniche diverse) se $\mathbf{W} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}^{-1})^{-1}$, dove \mathbf{R} è la matrice di covarianza delle osservazioni e \mathbf{B} la covarianza delle previsioni (background field).

I metodi *4D-var* estendono l'idea *3D-var* ad un intervallo temporale invece che ad un istante. Ovvero, per produrre l'analisi delle 12 UTC non si assimilano solo i

¹Inoltre molte variabili misurate hanno una *natura* diversa da quelle usate per definire il modello, ad es. i dati di radiazione derivati da satellite.

dati delle 12 UTC, ma tutti i dati raccolti tra le 3 UTC e le 15 UTC. In questo senso il sistema è uno “smoother”, che cambia il background interpolando nello spazio e nel tempo le nuove osservazioni.

Senza scendere nei dettagli matematici, il problema principale dei metodi variazionali e in particolare delle 4D-var, è che per applicare questo metodo bisogna calcolare l'operatore *aggiunto* del modello linearizzato (*linear tangent model*) e questo è computazionalmente molto oneroso, viste le dimensioni delle matrici coinvolte. Inoltre, l'operatore aggiunto \mathbf{K} è definito operativamente in base al prodotto scalare definito sullo spazio tangente in modo tale che $\langle \mathbf{x}, \mathbf{K}\mathbf{y} \rangle = \langle \mathbf{K}^T \mathbf{x}, \mathbf{y} \rangle$, e quindi dipende dalla scelta del prodotto scalare (*norma*) che viene fatta.

L'uso di questo operatore ci permette anche di fare il *targeted observation* se lo si usa per integrare all'indietro e trovare le osservazioni che più influenzano la previsione, altrimenti i metodi variazionali non ci danno una stima della *sensitivity* delle osservazioni (Langland e Baker 2004: *Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system*, Tellus A, **56**, 189–201).

Infine ci sono i metodi basati sui filtri di Kalman, ovvero su un loro ensemble (EnKF) che risolve un problema approssimato, poiché una stima esatta del filtro di Kalman relativo allo stato dell'atmosfera è praticamente irrealizzabile. Un filtro di Kalman è un processo *sequenziale* che dato uno stato attuale e un nuovo dato ad un tempo successivo, modifica lo stato attuale assimilando il nuovo dato e quindi producendo lo stato futuro. Anche in questo caso la matematica coinvolta va oltre lo scopo di questo riassunto, basti sottolineare che i vantaggi della famiglia degli EnKF sono il fatto di dare immediatamente una stima dell'errore nell'analisi (senza dover fare il *targeted observation*) e che sono computazionalmente più “leggeri” del 4D-var.

La variante che sembra essere più promettente è quella del *Local Ensemble Kalman Filter* (LEKF), in cui l'assimilazione viene fatta indipendentemente su ogni grid-point del dominio (vedi Ott et al. 2004: *A Local Ensemble Kalman Filter for Atmospheric Data Assimilation*, Tellus A, **56**, 415–428) e in particolare i *Local Ensemble Transform Kalman Filter* (LETKF) descritti in Hunt et al. 2004: *Four-Dimensional Ensemble Kalman Filtering*, Tellus A, **56**, 273–277, e Hunt, Kostelich e Szunyogh 2007: *Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter*, Physica D, **230**, 112–126. Comunque queste tecniche sono ancora “giovani” e il gruppo della Kalnay ne sta sviluppando nuove “varianti” per migliorarne le prestazioni e raggiungere (o battere) il metodo 4D-var con costi computazionali minori.

Un confronto tra diverse tecniche di assimilazione è stato fatto dal gruppo di Kalnay (ad esempio Kalnay et al. 2007: *4D-Var or Ensemble Kalman Filter?*, Tellus A, **59**, 758–773), ma anche da un gruppo giapponese che ha implementato i LETKF operativamente, vedi <http://www3.es.jamstec.go.jp/alera/about.html> o Miyoshi e Yamane 2007: *Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution*, Mon. Wea. Rev., **135**, 3841–3861. Da questi studi pa-

re che LETKF vada meglio di 4D-var (che –ovviamente– va meglio di 3D-var) nell’emisfero nord, mentre in quello sud 4D-var ha ancora un leggero vantaggio.

Per approfondimenti sul data assimilation si consiglia di consultare Bouttier e Courtier 1999: *Data assimilation concepts and methods*, ECMWF training course lecture series.

2.3 La previsione probabilistica e con condizioni iniziali diverse

Il metodo del *Breeding* consiste nel far ri-girare il modello deterministico da condizioni iniziali perturbate e verificare, punto per punto, dove si sviluppano le differenze maggiori, su una variabile prestabilita, ad esempio la vorticità potenziale. A quel punto si riscalda questa previsione perturbata in modo da rendere minime le differenze piccole e da lasciare traccia solo di quelle più grandi. Quindi si riparte da questa evoluzione “perturbata–ridotta” come condizione iniziale per lo step successivo e di nuovo si va a fare il confronto con l’evoluzione deterministica iniziale (relativamente al nuovo tempo). In questo modo solo le zone con i “breeding vectors” più intensi vengono identificate e corrispondono a quelle dove l’incertezza sulla previsione è dominante. Il breeding growth rate si misura come:

$$g(t) = \frac{1}{n\Delta t} \ln \frac{|\delta\mathbf{x}(t)|}{|\delta\mathbf{x}_0|} \quad (2)$$

dove Δt è l’intervallo d’integrazione e n è il numero di step necessari per raggiungere quel tempo, mentre i $\delta\mathbf{x}$ sono le perturbazioni. I Bred Vectors sono una generalizzazione a tempi finiti dei vettori di Lyapunov (“Leading Lyapunov Vectors”) del sistema dinamico, e tendono a questi ultimi per tempi lunghi. In generale le perturbazioni iniziali, che sono casuali e quindi tutte diverse, tendono a collassare su poche “direzioni” prevalenti (i modi non–lineari del sistema). I Bred Vectors sono il metodo utilizzato da NCEP per inizializzare il suo modello probabilistico con 20 condizioni iniziali diverse.

In base alla scelta della variabile che si perturba e della scala temporale che si sceglie i Bred Vectors si adattano a risolvere problemi diversi. Ad esempio, tempi molto brevi possono essere indicati per studiare la convezione, mentre tempi molto lunghi vengono usati per studiare fenomeni a scala climatica come l’ENSO (per il quale si usa $\Delta t \cong 15$ gg, vedi Yang et al. 2006: *ENSO Bred Vectors in Coupled Ocean–Atmosphere General Circulation Models*, J. of Climate, **19**, 1422–1436).

È interessante notare anche come ci sia una scuola italiana che si occupa di queste questioni a livello teorico, capeggiata da nomi come Francesco Uboldi, Anna Trevisan e Alberto Carrassi (vedi ad esempio Carrassi, Trevisan e Uboldi 2007: *Adaptive observations and assimilation in the unstable subspace by breeding on the data-assimilation system*, Tellus, **59A**, 101–113).

I Singular Vectors sono invece un formalismo matematico per trovare una generalizzazione degli autovettori (*Singular Value Decomposition*) nel caso di una metrica non–simmetrica, definita dalla norma del linear tangent model. La matematica

coinvolta va oltre lo scopo di questo resoconto e quindi si rimanda per approfondimenti ad esempio Buizza e Palmer, 1995: *The singular–vector structure of the atmospheric general circulation*, J. Atmos. Sci., **52**, 1434–1456, o alla pagina del dr. Roberto Buizza:

http://www.ecmwf.int/staff/roberto.buizza/Work/Selected_articles/Singular_Vectors/index.html

Qui basta sottolineare che i Singular Vector utilizzati da ECMWF per inizializzare il suo modello probabilistico con 50 membri sono tarati usando una metrica basata sull’“energia totale” massimizzata nelle prime 48 ore e per questo motivo non ha molto senso guardare i risultati prodotti nei primi due giorni². Inoltre, i Singular Vectors vengono trovati in un contesto di tipo 4D–var, visto che serve il linear tangent model.

3 Olivier Talagrand

talagran@lmd.ens.fr,

Laboratoire de Météorologie Dynamique. École Normale Supérieure (LMD/ENS), Paris, France

<http://www.lmd.jussieu.fr>



Anche questo noto professore francese sta scrivendo un libro sull’assimilazione dei dati, il cui titolo sarà *Data Assimilation in Meteorology And Oceanography* (previsto per febbraio 2010). Pur non avendo mai lavorato direttamente a Reading, Talagrand è comunque legato allo sviluppo del modello ECMWF, avendo scritto dei lavori come: Talagrand e Courtier, 1987: *Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory*, QJRMS, **113**, 1311–1328 o Candille e Talagrand 2003: *Some comparisons between ECMWF and NCEP EPS*, Geophysical Research Abstracts, Vol. 5, 11106.

Ha cominciato la sua lezione con una descrizione dei modelli probabilistici (tipo ECMWF EPS) spiegando ad esempio la differenza tra il *control run* degli EPS e il modello *operativo*: entrambi sono run deterministici basati sull’analisi “non-perturbata”, ma il control viene fatto girare alla risoluzione (T399L62) degli altri membri dell’EPS (quelli con condizioni iniziali perturbate), che è più bassa di quella del modello operativo (a risoluzione T799L90).

Dal punto di vista teorico, lo scopo delle previsioni probabilistiche è secondo lui quello di valutare la probabilità che si verifichi un certo stato futuro dell’atmosfera e quindi l’idea empirica è quella di fare più previsioni per poter campionare al meglio la Probability Density Function (PDF). Ha fatto anche degli studi (vedi più

²Se ho capito bene, con l’ultimo aggiornamento di EPS di febbraio 2006 le perturbazioni iniziali vengono generate solo per metà coi Singular Vectors e per l’altra metà si usano i Bred Vectors, e quindi adesso anche i primi due giorni di previsione dovrebbero avere maggior skill.

avanti) per dimostrare come un numero limitato (tra 30 e 50) di possibili membri sia sufficiente per sondare tale PDF.

A questo punto il prof. Talagrand si è lanciato in un volo pindarico relativamente alla struttura dell'attrattore strano di Lorenz. Vale la pena cercare di seguire il suo esempio di *simple chaos motion*. Si prenda un quadrato di lato 1 che rappresenta lo spazio delle fasi e si isoli dentro un cerchietto che rappresenta le condizioni iniziali "vicine". A questo punto si allunghi la coordinata x del doppio e si accorci la coordinata y di metà. Si otterra' in tal modo un rettangolo lungo 2 e alto 0.5, mentre il cerchietto diventerà un ovale schiacciato. Quindi si tagli il rettangolo a metà e si ponga una metà sopra l'altra, in modo da riformare un quadrato grande come quello iniziale. Il processo descritto si chiama *stretch and fold conservativo*, perché l'area totale del quadrato iniziale viene conservata.

Se si itera all'infinito questo processo l'intorno scelto all'inizio (che rappresenta le condizioni iniziali) verrà sparpagliato all'interno del quadrato sotto forma di segmenti orizzontali, tra loro anche molto distanti. C'è un modo molto elegante di descrivere matematicamente quanto appena fatto. Infatti se si scrivono le coordinate x e y di un punto nello spazio delle fasi con uno sviluppo binario si può dimostrare che ogni iterazione precedente corrisponde nell'affiancare a destra del numero binario della y il numero binario della x e a spostare il punto della virgola di una cifra a destra (*shift di Bernulli*).

Dopo un numero finito di iterazioni si sarà persa ogni possibile informazione sulla x , indipendentemente dalla precisione con la quale era stata misurata all'inizio. Si può dimostrare che la predicibilità di questa coordinata varia col logaritmo dell'accuratezza nella sua misura iniziale.

Ma il bello viene se si rifà l'esperimento nel caso *dissipativo*. Si prenda un quadrato 1×1 come prima e si moltiplichino per 2 la x come prima, ma si divida per 3 la y . Si tagli il rettangolo $2 \times 0.\bar{3}$ che si ottiene in tre rettangolini sovrapponendoli uno all'altro e ottenendo in tal modo un rettangolo di base $2/3$ e altezza 1. Il sistema è dissipativo perché non conserva l'area iniziale e ad ogni iterazione di questo processo l'area diminuisce.

Se si descrivono le coordinate x in binario e y nel sistema ternario (fatto con 0, 1 e 2) allora si ottiene nuovamente uno shift di Bernulli tra questi due numeri e a lungo andare la x tende all'*insieme ternario di Cantor*. Il sottoinsieme ternario di Cantor si ottiene prendendo l'intervallo continuo $[0, 1]$, dividendolo in 3 ed eliminando il terzo centrale. Nei due terzi che restano si ripete il processo, ovvero si divide ognuno dei due in 3 e si elimina il terzo centrale. Andando avanti all'infinito si ottiene un insieme di punti con volume nullo (non ci sono dentro segmenti di lunghezza non-nulla) ma con cardinalità pari a quella del continuo (\mathbb{R}). Questa è esattamente la struttura dell'attrattore del sistema descritto da Lorenz '63 (Ed Lorenz 1963: *Deterministic Nonperiodic Flow*, J. Atmos. Sci., **20**, 130–141)!

Uno dei modi più utilizzati per descrivere la teoria del chaos introdotta da Lo-

renz è quello dell'*effetto farfalla*, secondo il quale anche una piccola perturbazione prodotta dal battito della farfalla da qualche parte potrebbe portare a conseguenze disastrose (es. un uragano) da un'altra parte se tali perturbazioni crescono in modo esponenziale. Ci sono alcuni detrattori di questa teoria secondo i quali la dissipazione presente nel fluido non permetterebbe mai a perturbazioni così piccole di crescere così tanto, creando di fatto un'interruzione di scala alla propagazione delle perturbazioni. Una possibile spiegazione di ciò potrebbe essere che nel fluido non c'è la potenza del continuo, ma in qualche modo la presenza di molecole discretizza il sistema e impedisce che si verifichino quelle conseguenze che la matematica teoricamente permette (una specie di paradosso di Achille all'inverso).

Passando poi a commentare i diversi modi per creare le perturbazioni iniziali, secondo Talagrand i Singular Vector sono i modi del sistema dinamico che crescono più rapidamente, indipendentemente dall'incertezza sullo stato iniziale (analisi) e dalla norma scelta per misurare l'errore sui dati (ad esempio come confrontare un errore di 1 hPa nella pressione con quello di 1 grado Celsius nella temperatura). Invece i Bred Vectors nascono dall'incertezza iniziale e la amplificano. Infine, gli EnKF sono un processo sequenziale (a differenza del 4D-var che assimila contemporaneamente un periodo di tempo) che si basa sull'ipotesi che gli errori nelle osservazioni e nelle previsioni siano *non-correlati nel tempo* e questo secondo lui è un vincolo molto stringente (difficile da rispettare).

Per quanto riguarda la verifica delle previsioni probabilistiche, secondo lui bisogna guardare due proprietà: la *reliability* e la *resolution*. Secondo la sua definizione la prima è la proprietà che quando prevedo un certo intervallo con una probabilità f si osservi con la stessa probabilità, $p(o|f)$, un valore in quel intervallo, ovvero: $p(o|f) = f$ per ogni intervallo di verifica f ³. Graficamente si costruisce un diagramma in cui per ogni intervallo previsto f si va a calcolare la probabilità condizionata di avere l'osservazione in quell'intervallo, ovvero $p(o|f)$ e si plotta il punto $[f, p(o|f)]$. Più la linea che congiunge questi punti è vicina alla diagonale è più la previsione è ben calibrata.

Invece la *resolution* è definita da Talagrand come la capacità di fare una previsione con una PDF relativamente stretta (rispetto alla varianza del parametro in questione) pur riuscendo a far cadere al suo interno il valore osservato.

Come metriche per valutare la bontà di queste proprietà suggerisce il *rank histogram* e la componente di *reliability* nella decomposizione del Brier Score (BSS) per la reliability e la curva di ROC (con l'area sottesa), il Brier Skill Score (e varianti come il Ranked Probability Score) o il RMSE (che però non è una misura propria) come misure di resolution. Secondo lui il rank histogram ha il limite che se le previsioni stanno "fuori" dallo spread previsto assume una forma a U, indi-

³Nel caso di previsioni binarie questo concetto si riduce ad avere una quantità di eventi previsti uguale al *prior* (climatologia) dell'evento osservato, che in termini di tabella di contingenza significa avere un BIAS unitario.

pendentemente se sono fuori “di tanto” o “di poco”. Per questo preferisce misure *continue* come il BSS.

Va sottolineato che di solito si può migliorare a posteriori la reliability con un’opportuna *calibrazione* e quindi ha senso concentrare l’attenzione sulle misure di resolution, in particolari su quelle come l’area del ROC (o il maximum Pierce Skill Score) che sono invarianti alla calibrazione⁴. Ulteriori approfondimenti si possono trovare in Candille e Talagrand 2005: *Evaluation of Probabilistic Prediction Systems for a Scalar Variable*, QJRMS, 1–20.

Invece in Talagrand et al. 1999: *Evaluation of probabilistic prediction systems*, Proc. ECMWF 1997 Workshop on Predictability, Reading UK, 1–25, si mostra come un numero di EPS inferiore a 50 membri sia sufficiente a raggiungere tutto lo spread che il sistema può generare⁵, tranne nel caso di eventi molto rari. Quello che però è più interessante è che la dipendenza della dimensione del database usato per fare la verifica è molto forte e, ovviamente, lo è ancor di più per gli eventi rari. Ciò introduce il problema di avere un database lungo di previsioni *omogenee* (senza cambi di modello dopo pochi anni) e quindi all’idea delle *reforecast* introdotte da Hamill. Ad una mia domanda esplicita, ha risposto che ECMWF non intende seguire NCEP sulla via delle reforecast per il modello deterministico, mentre sta sviluppando un reforecast di EPS a 5 (cinque) membri.

La conclusione delle sue lezioni è stata inaspettata. Infatti, come si può leggere in Descamps e Talagrand 2007: *On some aspects of the definition of initial conditions for ensemble prediction*, Mon. Wea. Rev., **135**, 3260–3272, lui e Descamps (presente anch’egli al corso) testano diversi metodi per creare le perturbazioni iniziali nel sistema di Lorenz 1996, nel sistema quasi-geostrofico di Molteni–Palmer 1993 (Molteni e Palmer 1993: *Predictability and finite-time instability of the northern winter circulation*, QJRMS, **119**, 269–298) e infine nel modello operativo ARPEGE, trovando sempre che gli EnKF vanno meglio dei Bred Vectors, che a loro volta vanno meglio dei Singular Vectors. Stranamente, dopo questa conclusione che pare contraria ai Singular Vectors implementati a ECMWF, Talagrand si è impegnato in una loro generalizzazione appena pubblicata: Rivière, Lapeyre e Talagrand giugno 2008: *Nonlinear Generalization of Singular Vectors: Behavior in a Baroclinic Unstable Flow*, J. Atmos. Scie., **65**, 1896–1911.

⁴Più precisamente il ROC è invariante a trasformazioni monotoniche, come spiegato anche in Manzato 2007: *A Note on the Maximum Peirce Skill Score*, WAF, **22**, 1148–1154.

⁵Mi è parso di capire che non tutti condividessero questo punto. Ad esempio Victor Homar (Università delle Baleari) dice che se verificassero campi al suolo invece che il geopotenziale a 500 hPa troverebbero risultati diversi. . .

4 Thomas M. Hamill

tom.hamill@noaa.gov,
NOAA/ESRL, Physical Sciences Division, Boulder
(CO), USA

<http://www.cdc.noaa.gov/people/tom.hamill>



Come avevo scritto nel resoconto della conferenza AMS 2006, Tom Hamill è uno dei meteorologi emergenti in America, perché, pur essendo relativamente giovane, ha già scritto molte pubblicazioni importanti, come ad esempio: Wilks e Hamill 2007: *Comparison of ensemble–MOS methods using GFS reforecasts*, Mon. Wea. Rev., **135**, 2379–2390, Hamill et al. 2004: *Ensemble Reforecasting: Improving Medium-Range Forecast Skill Using Retrospective Forecasts*, Mon. Wea. Rev. **132**, 1434–1447; Hamill 2001: *Interpretation of rank histograms for verifying ensemble forecasts*, Mon. Wea. Rev. **129**, 550–560; Hamill et al. 2000: *A comparison of probabilistic forecasts from bred, singular vector, and perturbed observation ensembles*, Mon. Wea. Rev. **128**, 1835–1851; Quasi tutti i suoi articoli sono disponibili on–line al seguente link: <http://www.cdc.noaa.gov/people/tom.hamill/cv.html>

Ha iniziato la sua lezione con una panoramica dei metodi utilizzati per verificare le previsioni probabilistiche. In particolare ha descritto l'uso del *rank histogram*. Per costruirlo si procede come segue: supponiamo di avere un ensemble con n membri. Bisogna disegnare un istogramma con $n + 1$ colonne e calcolare il *rango* di ogni osservazione rispetto alle previsioni ordinate in modo crescente. Se per esempio il valore osservato è inferiore a tutti quelli previsti dai n membri allora si incrementa di uno la prima barra (più a sinistra) dell'istogramma. Se invece il valore osservato è maggiore dei primi m valori previsti, ma minore dei successivi $n - m$ previsioni, allora si incrementa di uno la colonna $(m + 1)$ –esima. Se infine il valore osservato è superiore a tutti i valori previsti allora si incrementa la colonna $n + 1$.

Dopo aver fatto questa verifica su un numero elevato di casi si ottiene un istogramma che è mediamente piatto nel caso in cui lo spread degli EPS sia sempre ben centrato intorno al valore osservato. Se invece il modello tende ad avere un *bias* (errore sistematico) allora l'istogramma avrà le colonne di un lato più alte di quelle dell'altro. Infine, se lo spread delle previsioni probabilistiche è mediamente troppo basso e l'osservazione casca spesso fuori del range previsto (più alta del massimo o più bassa del minimo) allora il rank histogram avrà la prima e l'ultima colonna molto più alte, assumendo la forma a U.

Quando si verifica l'altezza geopotenziale a 500 hPa si ottiene tipicamente un rank histogram piatto (previsioni ottime), mentre se si verifica la temperatura al suolo si ottiene tipicamente un rank histogram a U! La temperatura a 850 hPa è di solito un parametro con rank histogram intermedio tra queste due situazioni (una specie di V schiacciata).

Il rank histogram è costruito per verificare una variabile scalare alla volta. Se invece si volesse verificarne due ci sono delle tecniche più recenti come il *Minimum Spanning Tree* (vedi Smith e Hansen 2004: *Extending the Limits of Ensemble Forecast. Verification with the Minimum Spanning Tree*, Mon. Wea. Rev., **132**, 1522–1528), o il *Multivariate rank histogram* che calcola il *pre-rank* di due variabili standardizzate usando dei rettangoli che per separare i punti su un piano bidimensionale.

Relativamente alla relazione tra bias (legato al valor medio degli ensemble) e standard deviation (spread degli ensemble) hanno trovato che non sono così indipendenti: ad esempio per la pioggia, si vede che ensemble con pioggia media alta hanno di solito una bassa dispersione, e viceversa (vedi Whitaker e Lough, 1998: *The relationship between ensemble spread and ensemble mean skill*, Mon. Wea. Rev., **126**, 3292–3302).

In Hamill e Juras 2006: *Measuring forecast skill: is it real skill or is it the varying climatology?*, QJRMS, **132**, 2905–2923, si mette in guardia dal verificare campi spaziali che si riferiscono a zone climatologicamente diverse, visto che molti indici di verifica (tutti aggiungerei io) dipendono dalla *prior probability* dell'evento. La soluzione che propongono è quella di dividere in sottozone climaticamente omogenee e poi fare la verifica singolarmente per ogni zona. Ad esempio, per quanto riguarda l'area sottesa dal ROC consigliano di mediare i valori delle aree relative ad ogni zona⁶. Una variante della verifica con la curva di ROC molto utilizzata è la curva del Potential Economic Value, descritta molto bene in Richardson 2003: *Economic Value and Skill* in Jolliffe e Stephenson 2003: *Forecast Verification*, Wiley, 240 pp.

Dal punto di vista della visualizzazione operativa degli ensemble le cose che più mi hanno colpito sono i diagrammi skew-T con tutti i 50 membri di ECMWF sovrapposti, che si aprono quando si clicca su una qualsiasi località della mappa, oppure le mappe di *precipitazione totale* fatte sommando punto per punto tutte le piogge dei 50 membri, per dare un'indicazione della probabilità totale che si verifichi della precipitazione. Ha anche dato una referenza generica su come fare dei grafici di valori numerici, che è Tufte 2001: *The visual display of quantitative information*, Graphics Press, considerato il miglior libro del genere.

Ha poi affrontato il discorso, molto di moda, dei modelli locali ad ensemble, in particolare per le previsioni a breve termine (Short-Range Local Area Ensemble Forecast). In particolare ha sottolineato i problemi relativi alle condizioni al contorno, che devono a loro volta essere perturbate, per non “forzare” le perturbazioni interne a collassare su previsioni molto vicine. Inoltre c'è il problema che le perturbazioni interne non possono espandersi fuori dal dominio (“one way”) e quindi consiglia di prendere un dominio molto più grande dell'area veramente d'interesse.

⁶Personalmente credo che sia più difficile avere alti incrementi del ROC quando la curva è già alta che non quando sia vicina alla diagonale e quindi sarebbe meglio pesare di più i valori alti di quelli bassi, ad esempio con una trasformata Z di Fisher.

Aggiunge che nel caso di nesting successivi, non ha molto senso avere la convezione esplicita solo in quello più interno e parametrizzarla in quello precedente (vedi Nutter, Stensrud e Xue, 2004: *Effects of coarsely resolved and temporally interpolated lateral boundary conditions on the dispersion of limited-area ensemble forecasts*, Mon. Wea. Rev., **132**, 2358–2377). George Bryan dell'NCAR ha recentemente mostrato come la convezione simulata con risoluzioni molto diverse (in particolare passando da 5 km a 500 m o meno) porta a risultati molto diversi (vedi anche Bryan, Wyngaard e Fritsch 2003: *Resolution Requirements for the Simulation of Deep Moist Convection*, Mon. Wea. Rev., **131**, 2394–2416).

In conclusione Hamill suggerisce di usare domini molto estesi, con risoluzioni molto spinte (massimo 1 km) e aggiornamenti frequenti del *lateral boundary layer*. Queste sono condizioni computazionalmente molto pesanti già per un modello ad area limitata deterministico, figuriamoci per uno ad ensemble.

Relativamente agli eventi estremi ha citato gli esempi di previsioni di *severe storms* fatte dall'NSSL usando la previsione contemporanea delle seguenti condizioni: $CAPE > 1000 J/kg$, $Shear_{6km} > 20 m/s$, Storm Relative Helicity $> 100 m^2/s^2$, $LCL < 1000 m$ e pioggia convettiva prevista maggiore a 0.3 mm.

Tribbia e Baumhefner 2004: *Scale Interactions and Atmospheric Predictability: An Updated Perspective*, Mon. Wea. Rev., **132**, 703–713 investigano la propagazione di errori che nascono a scale piccole e poi si propagano nel tempo (sub-sinottiche). Hohenegger, Lüthi e Schär 2006: *Predictability mysteries in cloud-resolving models*, Mon. Wea. Rev., **134**, 2095–2107, invece investigano il caso di eventi convettivi intensi sulle Alpi e la loro predicibilità. Sembra che le *onde di gravità* in questi casi giochino un ruolo molto importante, riuscendo a contrastare il moto del vento in quota se inferiore ad una soglia critica e quindi portando alla genesi *stazionaria* delle celle convettive. L'argomento andrebbe approfondito per quanto riguarda le previsioni nella nostra regione, studiando anche Hohenegger e Schär 2007: *Atmospheric Predictability at Synoptic Versus Cloud-Resolving Scales*, BAMS, **88**, 1783–1793.

Una delle lezioni è stata dedicata anche all'uso delle previsioni probabilistiche nel campo idrologico, in particolare per stimare i tempi del *run-off*. In particolare in queste applicazioni Hamill suggerisce di perturbare anche le condizioni al suolo (es. il *soil moist*, *land cover parameterization*, ecc.) invece che solo l'atmosfera. Questo potrebbe migliorare la variabilità nelle previsioni sui campi al suolo, che è molto più bassa di quella trovata sulle analisi.

L'ultima lezione è stata dedicata al problema della *calibrazione*. Ha citato diverse tecniche, tra cui *linear*, *logistic* e *Non-Homogeneous Gaussian* regression (quest'ultima descritta in Gneiting et al. 2005: *Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation*, Mon. Wea. Rev., **133**, 1098–1118.), filtri di Kalman, *dressing methods*, *Bayesian Model Averaging* (Raftery et al. 2005: *Using Bayesian Model Averaging to Calibrate Forecast Ensembles*, Mon. Wea. Rev., **133**, 1155–1174), ecc. I risultati differiscono a seconda del tipo di variabile. Ad esempio, se la variabile ha errori gaussiani (es. la

temperatura) Wilks e Hamill 2007 trovano che la regressione NHG usando giusto la media e la varianza delle previsioni vada molto bene. In molti casi, anche un semplice *downscaling* per rimuovere il bias (errore sistematico) è un metodo che raggiunge da solo un risultato molto migliore del model output.

Per calibrare i modelli per gli eventi estremi (rari) servono dei database molto lunghi di previsioni omogenee e questo è stato uno dei motivi per lo sviluppo delle *reforecast*, ovvero far girare sul passato il modello NCEP più recente (come nelle *hindcast*, ma nelle reforecast anche le analisi iniziali vengono rifatte con lo stesso tipo di modello).

Mi è parso molto interessante quanto fatto da Hamill e Whitaker 2006: *Probabilistic quantitative precipitation forecasts based on reforecast analogs: theory and application*, Mon. Wea. Rev., **134**, 3209–3229, che applicano la teoria dell’analogo ad un campione di 25 anni di reforecast per fare un downscaling quantitativo della pioggia prevista.

L’uso delle reforecast sta diventando sempre più diffuso, ad esempio sono state usate in India per la previsione dei monsoni tramite *stepwise logistic regression*. Interessante notare che in questo caso il primo predittore scelto per il downscaling è l’acqua precipitabile e non la pioggia prevista dal modello, come trovato anche da altri autori.

Per quanto riguarda il futuro, sembra che NCEP comincerà a produrre ogni giorno le reforecast calibrate del modello deterministico, mentre ECMWF non intende perseguire la strada delle reforecast deterministiche (alta risoluzione), ma da quest’estate dovrebbe rendere operativo un ensemble di 5 membri a risoluzione T399L62.

5 Seminari

5.1 Tiziana Paccagnella

ARPA Emilia–Romagna e Servizio IdroMeteorologico (SIM), Bologna, Italia.

tpaccagnella@arpa.emr.it

<http://www.arpa.emr.it/SIM/>

Il progetto TIGGE significa “THORPEX Interactive Grand Global Ensemble”, dove THORPEX significa “THE Observing System Research and Predictability EXperiment” ed è un progetto del WMO per aumentare il “Weather Watch Warning”. I siti di riferimento sono rispettivamente

<http://tigge.ecmwf.int> e <http://www.wmo.ch/pages/prog/arep/thorpex>.

In poche parole, il progetto TIGGE salva operativamente le previsioni probabilistiche dei 10 seguenti modelli globali: BoM (Australia), CMA (Cina), CMC (Canada), CPTEC (Brasile), ECMWF (Europa), JMA (Giappone), KMA (Corea), Meteo-France (Francia), NCEP (USA) e UKMO (Regno Unito). Ovviamente non vengono salvati tutti i livelli, tutte le variabili alla massima risoluzione possibile, ma un sot-

toinsieme di ciò, che comunque occupa qualcosa come 240 GB al giorno, scritto in 3000 files in formato GRIB2.

L'idea di base è quella di usare questi modelli come un multimodel ensemble, oppure come un super-ensemble (ovvero dando pesi diversi ai singoli modelli), per ottenere previsioni migliori del miglior modello considerato da solo (di solito ECMWF EPS). Ovviamente bisogna capire se il miglioramento che si ottiene giustifica lo sforzo compiuto e le enormi risorse impiegate (si pensi che l'archivio viene replicato in Europa, America e Cina attualmente occupa più di 100 TB).

Nel contesto di questo progetto, il SIM dell'ARPA Emilia-Romagna ospita il sottoprogetto TIGGE-LAM, il cui ambizioso scopo è quello di far girare un ensemble LAM (Local Area Model), partendo dalle condizioni iniziali dei modelli globali archiviati in TIGGE. L'idea sarebbe quella di attivare la previsione "locale" sulle zone dove i modelli globali prevedono il rischio di eventi estremi, in modo da poterne migliorare la risoluzione e il dettaglio. Tutto ciò verrebbe gestito a livello "globale", senza quindi utilizzare gli eventuali ensemble LAM che già girano in diverse parti del mondo, ad es. in Europa (ma in altre non ce ne sono proprio).

5.2 Marino Marrocu

Centro di Ricerca Sviluppo e Studi Superiori in Sardegna, Pula (CA), Italia.

<http://www.crs4.it/>

Questo ricercatore del CRS4 ha fatto diversi lavori insieme coi meteorologi del SAR che riguardano le previsioni ad ensemble, quali ad esempio Ficca et al. 2001: *BOLEPS: A Limited Area Probabilistic Prediction System of small-scale severe rainfall events*, proc. European Geophysical Society, Nice, o Chessa et al. 2005: *Multimodel-multianalysis mesoscale ensemble (MusE)*, proc. First TIGGE Workshop, ECMWF Reading, o anche Marrocu e Chessa 2008: *A multi-model/multi-analysis limited area ensemble: calibration issues*, Met. App. **15**, 171–179.

In particolare nell'ultimo lavoro citato comparano diverse tecniche di calibrazione (Bayesian model averaging, ensemble model output statistics –NHG regression– e dressing kernel) per un multi-model ensemble basato su BOLAM, MM5 e RAMS, inizializzati sia con l'analisi ECMWF che con quella di NCEP, al fine di migliorare la previsione di temperatura al suolo in Sardegna. Tutto sommato i risultati migliori li ottengono con il Bayesian Model Averaging (i BMA sono disponibili anche come pacchetto per **R** al link <http://cran.r-project.org/web/packages/BMA/index.html>).

Marrocu ha anche accennato all'uso di un super-ensemble, idea introdotta da Krishnamurti et al. 1999: *Improved skills for weather and seasonal climate forecasts from multi-model superensemble*, Science, **285**, 1548–1550, e Krishnamurti et al. 2000: *Multi-model superensemble forecasts for weather and seasonal climate*, J. Clim., **13**, 4196–4216. L'idea di base è semplicemente quella di fare una media pesata dei diversi model output utilizzando un periodo di training per trovare i pesi migliori dei singoli membri con una multiregressione lineare.

5.3 José Antonio García–Moya Zapata

Agencia Estatal de Meteorología (AEMet, Madrid, Spagna.

j.garciamoya@inm.es

<http://www.aemet.es/en/portada>

L'Agencia Statale di Meteorologia è il nuovo ente che ha sostituito lo storico Istituto Nacional de Meteorología (INM). Le persone sono sempre le stesse e il cambio riguarda solo il nome.

Al di là del suo seminario, vorrei riportare anche le osservazioni che ha fatto rispondendo al sollecito del dr. Franco Zardini (ARPAV) su quali siano le ricadute più operative delle previsioni probabilistiche. García–Moya ha enfatizzato il fatto che alcuni previsori “vecchio stampo” guardano la media dell'ensemble come se fosse uno dei tanti modelli deterministici a loro disposizione, invece che come una stima dell'incertezza sulla previsione. Questo modo di vedere le cose è secondo lui dovuto al modo “deterministico” di ragionare, che in futuro dovrà lasciare il campo al modo “probabilistico”, ove la previsione sarà solo sulla probabilità che i diversi eventi si verifichino (es. pioggia o vento superiori a delle soglie di riferimento prefissate). Secondo lui è un problema “educativo” e nel suo istituto fanno fare dei corsi specifici ai loro previsori per insegnargli a usare le previsioni probabilistiche come stime sull'incertezza della previsione, piuttosto che come previsioni *tout court*.

Un altro elemento che ha sottolineato è l'importanza delle piccole probabilità per eventi estremi. Spesso un previsore non dà particolare peso a una previsione probabilistica che prevede solo il 10% di probabilità di accadimento, ma se l'evento a cui si riferisce ha una probabilità climatica molto bassa (come le piogge estreme, di cui ricorda il caso di Siviglia 10/09/1996 con 520 mm in un giorno) allora anche una bassa probabilità (ma magari più alta della frequenza di accadimento) diventa importante, soprattutto se la previsione è reliable (modello calibrato). In qualche modo questo ragionamento si lega a quello pubblicato in Manzato 2007, dove il maximum Pierce Skill Score si ottiene per una posterior probability (previsione) pari alla prior probability (climatologia dell'evento) e questo porta ad un alto numero di falsi allarmi nel caso di eventi molto rari.

Nella sua presentazione ha invece mostrato l'approccio multi-model implementato dal servizio meteorologico spagnolo, sulla falsa riga di quanto descritto in Hou, Kalnay e Droegemeier, 2001: *Objective verification of the SAMEX'98 ensemble forecasts*, Mon. Wea. Rev., **129**, 73–91 o Palmer et al. 2004: *Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER)*, BAMS, **85**, 853–872. I modelli che fanno girare sui loro computer sono 5: Hirlam, HRM (del DWD), MM5, UM (di UKMO) e LM del consorzio COSMO, mentre le condizioni iniziali sono prese da 4 modelli globali: ECMWF, UKMO, NCEP e DWD (stanno lavorando per usare anche quelle canadesi). In questo modo riescono a costruire un “multi-model multi-boundaries” Short Range Ensemble Prediction System (SREPS) di $5 \times 4 = 20$ membri diversi.

I risultati sono visualizzati nella loro Intranet, poiché le “policy” per usare i

modelli sviluppati dagli altri non gli permettono di pubblicare in Internet i risultati. Nonostante questo, sono disponibili a consentire l'accesso gratuito via FTP ai loro server per poter scaricare i dati in formato GRIB! Basta chiedergli l'accesso via email e questa potrebbe essere una buona opportunità per avere anche all'OSMER un "very-poor's man ensemble".

La verifica di questo sistema è stata fatta per 3 mesi usando due diversi riferimenti: il primo è l'analisi prodotta da ECMWF e il secondo sono le osservazioni puntuali delle stazioni meteo. Ovviamente nel secondo caso si pone il problema di reperire più dati possibili e per questo motivo ha chiesto la collaborazione di tutti i centri meteo operativi, comprese le diverse ARPA italiane, a fornirgli i dati da loro misurati. Per l'Italia finora ha risposto solo il SIM dell'ARPA Emilia Romagna. Per verificare la calibrazione dell'ensemble usano i rank histograms, la curva di reliability e il diagramma spread-skill (che mostra la varianza della previsione vs. l'errore medio), mentre per verificare le previsioni di eventi binari (soglie di piogge e vento) usano il ROC, il relative Value e il Brier Skill Score. In generale la verifica con le analisi sono migliori di quelle con le osservazioni, in particolare per la reliability delle piogge più deboli. Per quanto riguarda il futuro intendono adottare una calibrazione con tecnica di Bayesian Model Averaging e investigare un "time lagged super-ensemble".

5.4 Jutta Thielen

JRC Joint Research Centre, European Flood Alert System, Ispra, Italia.

<http://efas.jrc.it/index.html>

All'interno del Joint Research Centre di Ispra ci sono alcuni gruppi di lavoro in ambito meteorologico, tra cui quello della Thielen che si occupa di prevedere le alluvioni su tutta Europa, su mandato della Commissione Europea di Bruxelles.

In particolare presenta i progetti *European Flood Alert System* (EFAS, vedi <http://efas.jrc.it/>) e PREVIEW (vedi <http://www.preview-risk.com/>) sviluppato nel contesto del programma GMES (Global Monitoring for Environment and Security). Per approfondimenti vedi i siti Internet relativi o Thielen et al. 2008: *The European Flood Alert System. Part 1: Concept and development*, Hydrol. Earth Syst. Sci. Discuss., **5**, 257–287 e Bartholmes et al. 2008: *Skill assessment of probabilistic operational EFAS forecasts*, Hydrol. Earth Syst. Sci. Discuss., **5**, 289–322.

Anche nel campo idrologico le previsioni probabilistiche stanno prendendo piede, come ad esempio nel progetto Hydrologic Ensemble Prediction EXperiment (HEPEX, vedi <http://hydis8.eng.uci.edu/hepex/>).

6 Conclusioni

Questa scuola è stata molto stimolante, in particolare per l'elevato livello dei docenti, che spesso disputavano fra di loro su quali fossero le tecniche migliori disponibili

al giorno d'oggi. Durante i primi giorni di corso pareva che si delineasse la classica competizione USA–Europa (NCEP vs. ECMWF) con Kalnay che sosteneva a spada tratta i Bred Vectors e gli ETKF, mentre Talagrand avrebbe sostenuto il 4D–var e i Singular Vectors. Hamill era stato assunto al ruolo di giudice imparziale. Invece, la conclusione delle lezioni del prof. Talagrand è stata un vero colpo di scena, con un noto collaboratore di ECMWF che riconosceva la superiorità degli ETKF rispetto ai Singular Vectors e in particolare la superiorità di un approccio “integrato”, ovvero che generi *contestualmente* sia l'analisi che le perturbazioni iniziali.

A voce, poi Talagrand mi ha detto che non è sicuro che gli ETKF saranno l'unica possibilità e che ECMWF sta valutando lo sviluppo di un metodo variazionale probabilistico, una specie di “ensemble 4D-var”, che però potrebbe essere molto oneroso computazionalmente. D'altra parte tutti hanno riconosciuto la superiorità del modello ECMWF rispetto agli altri modelli globali (vedi ad esempio Buizza et al. 2005: *A comparison of the ECMWF, MSC and NCEP global ensemble prediction systems*, Mon. Wea. Rev., **133**, 1076–1097) e pare che una buona dose di questo vantaggio sia dovuta proprio all'analisi iniziale fatta tramite il 4D–var, per cui ECMWF non vuole buttare alle ortiche tutte le tecniche sviluppate finora per questa metodologia. Viceversa, nella sua conclusione, la prof. Kalnay esternava tutta la sua tristezza per la scelta di NCEP di inseguire ECMWF sulla strada del 4D–var, nonostante tutti gli sviluppi più recenti fatti dal suo gruppo indichino che gli ETKF (o loro varianti) possano raggiungere gli stessi risultati del 4D–var con costi computazionali molto minori, perché non necessitano di calcolare l'operatore aggiunto.

Un'altra conclusione che è emersa “tra le righe” è che tutti considerino un ensemble di tipo multimodel, ovvero fatto usando diversi modelli (PEPS: Poor man's Ensemble Prediction System), superiore ad un ensemble fatto con lo stesso modello e solo le condizioni iniziali diverse (EPS), indipendentemente dalla tecnica usata per generare le perturbazioni iniziali. Questo giustifica il notevole sforzo associato al progetto TIGGE, che salva quotidianamente tutte le previsioni fatte dai modelli globali probabilistici.

Per quanto riguarda la verifica dei modelli probabilistici secondo me non è ancora emersa una pratica “standard”, al di là del diagramma di reliability e del valore del Brier Skill Score e varianti (Ranked e Continuous Ranked Probability Score). È anche emersa la necessità, ma anche la relativa facilità d'implementazione, di fare la calibrazione delle previsioni e questo è un ulteriore motivo che mi spinge a focalizzare l'attenzione sull'accuratezza (“sharpness” o “resolution”) piuttosto che sulla reliability, per cercare di ridurre il confronto tra modelli diversi ad un solo numero scalare e svincolarsi dalla “legge di A. Murphy” secondo la quale “forecast verification is a multi–faceted thing”.

Dal punto di vista strettamente operativo non ci sono state particolari indicazioni da seguire, se non quella di non ridurre le previsioni probabilistiche semplicemente ad un'ulteriore previsione deterministica (ad esempio facendone la media) ma a considerarle per quello che vorrebbero essere: ovvero una previsione sull'incertezza

delle previsioni meteo. Se, come mi pare di capire, il futuro ci riserverà sempre più previsioni probabilistiche e sempre meno previsioni deterministiche, allora è il caso che il previsore cominci a “pensare le previsioni” in modo diverso e anche a educare il pubblico a ricevere questo tipo di informazione. Pur non essendo un previsore, io stesso resto comunque più legato alle previsioni deterministiche, ma tutti dicono che quelle probabilistiche hanno più valore economico per l’utenza, che può avere dei parametri di “cost/loss” molto diversi a seconda delle diverse esigenze.

Palmanova, 17 giugno 2008.

I docenti discutono tra di loro: Sigular vectors o Bred Vectors?



Il dr. Buzzi dopo aver saputo l'onorario della prof. Kalnay.



Hamill discute col prof. Romero (Università delle Baleari).





(a) Porto Corte Ricerche



(b) Brainstorming

(a) La bella sede dove si è tenuta la scuola estiva MSMM.

(b) Un momento di *tutoring* pomeridiano.



(c) Fast Singular Vector perturbation



(d) Perturbed system

(a) Tom Hamill subisce le ingiurie del forte vento al largo di Alghero.

(b) La faccia contrariata del ricercatore americano per la mancanza di rispetto del meteo italiano...