



ΕΘΝΙΚΟ & ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
Μεταπτυχιακό Πρόγραμμα Σπουδών

Χάρης Γεωργίου (Α.Μ: Μ-177)

ΠΑΡΑΛΛΗΛΗ ΑΡΙΘΜΗΤΙΚΗ ΠΡΟΒΛΕΨΗ ΚΑΙΡΟΥ

Εργασία στο μάθημα: Παράλληλοι Αλγόριθμοι
Διδάσκων: Ν. Μισυρλής, Φ. Τζαφέρης

Αθήνα 1999

1. ΕΙΣΑΓΩΓΗ

Το ζήτημα της πρόβλεψης καιρού αποτελεί αντικείμενο έρευνας σε πολλούς διαφορετικούς τομείς της επιστήμης. Εκτός από το πρόβλημα της μοντελοποίησης των καιρικών φαινομένων, σημαντική προσπάθεια έχει αφιερωθεί στην ανάπτυξη επιστημονικών τεχνικών για την εφαρμογή των θεωρητικών μοντέλων στην πράξη για την προσομοίωση και την έγκαιρη πρόβλεψη της συμπεριφοράς των μοντέλων αυτών. Η δυνατότητα πρόβλεψης των καιρικών φαινομένων αποτελεί μείζον ζήτημα, όχι μόνο σε οικονομικό και κοινωνικό επίπεδο, ιδιαίτερα στις περιπτώσεις έγκαιρης πρόβλεψης ακραίων καταστροφικών φαινομένων (πλημμύρες, τυφώνες), αλλά παράλληλα στη μελέτη κλιματολογικών φαινομένων γενικότερα (μοντέλα κλίματος), καθώς και για την ανάπτυξη αποτελεσματικών μοντέλων ελέγχου της ρύπανσης του περιβάλλοντος (οικολογικές καταστροφές).

Η πολυπλοκότητα των μοντέλων, καθώς και ο μεγάλος όγκος δεδομένων προς επεξεργασία, οδηγεί αναπόφευκτα στη χρησιμοποίηση σύγχρονων υπολογιστικών συστημάτων για την προσομοίωση και μελέτη τους. Πολλές διαφορετικές προσεγγίσεις στο πρόβλημα έχουν προσφέρει ένα πλήθος εφαρμογών, άλλες με μικρότερη και άλλες με περισσότερη επιτυχία.

Μια πρώτη προσέγγιση είναι η αναλυτική διατύπωση του μοντέλου, η εφαρμογή φυσικών νόμων της θερμοδυναμικής και της μηχανικής ρευστών και η επίλυση των αντίστοιχων εξισώσεων. Για την αναπαράσταση και την επίλυση μέσω υπολογιστή το μοντέλο θα πρέπει να "κβαντιστεί", να αναπαρασταθεί δηλαδή συναρτήσει στοιχειωδών μονάδων-περιοχών, και στη συνέχεια να γίνει προσομοίωση του συστήματος με διακριτό τρόπο. Οι αριθμητικές αυτές τεχνικές μοντελοποίησης έχουν το πλεονέκτημα του ορθού θεωρητικού υπόβαθρου, όμως λόγω της εξαιρετικά μεγάλης πολυπλοκότητας σε χρόνο επεξεργασίας και όγκου δεδομένων, η αποτελεσματική υλοποίησή τους σε υπολογιστικά συστήματα έγινε δυνατή σχετικά πρόσφατα. Έχουν διατυπωθεί διάφορα αριθμητικά μοντέλα πρόβλεψης καιρού, τα οποία διαφέρουν όχι τόσο στον τρόπο διατύπωσης του μοντέλου όσο στον τρόπο επίλυσής του. Δύο από τα σημαντικότερα μοντέλα είναι το HIRLAM και το ETA.

Το μοντέλο HIRLAM (High Resolution Limited Area Model) αφορά στην κατάτμηση μιας περιορισμένης γεωγραφικής περιοχής και εφαρμογή ενός συστήματος διακριτών διαφορικών εξισώσεων ρευστών και θερμοδυναμικής (εξισώσεις διαφορών), το οποίο προσομοιώνει με αποδεκτή ακρίβεια το πραγματικό φυσικό μοντέλο. Το σύστημα μπορεί να επιλυθεί με εξισώσεις διαφορών διακριτού χρόνου, με φασματική ανάλυση ή άλλες τεχνικές, ανάλογα με τις εκάστοτε απαιτήσεις ακρίβειας και ταχύτητας.

Το μοντέλο ΕΤΑ αφορά κυρίως γεωγραφικές περιοχές με απότομες υψηλεπτικές διαφορές (ορεινοί όγκοι) και χρησιμοποιεί διαφορετικό τρόπο κατάτμησης της περιοχής από ότι το HIRLAM, όμως σε γενικές γραμμές έχει την ίδια μορφή και οι μέθοδοι επίλυσης είναι παρόμοιες. Το μοντέλο ΕΤΑ αποτελεί τη βάση του συστήματος SKIRON, το σύστημα πρόβλεψης καιρού που αναπτύχθηκε από το Τμήμα Πληροφορικής του Πανεπιστημίου Αθηνών για την Εθνική Μετεωρολογική Υπηρεσία και βρίσκεται ήδη σε επιχειρησιακή χρήση.

Εκτός από τις αριθμητικές μεθόδους, αξίζει να αναφερθούν προσπάθειες που κατά καιρούς έχουν γίνει σε άλλους τομείς, όπως η εφαρμογή Τεχνητών Νευρωνικών Δικτύων για τη μοντελοποίηση καιρικών φαινομένων με βάση ιστορικά "πρότυπα" συμπεριφοράς, η εφαρμογή δυναμικής χαοτικών συστημάτων όσο αφορά στη διατύπωση κανόνων για την πρόβλεψη χρονοσειρών, καθώς και προσπάθειες ανάπτυξης ευφυών Έμπειρων Συστημάτων με στόχο την προσπάθεια ανακάλυψης κανόνων συμπεριφοράς παρόμοιων συστημάτων, ειδικότερα όσο αφορά ακραία καιρικά φαινόμενα.

Σήμερα η προσοχή έχει εστιαστεί στα αριθμητικά μοντέλα πρόβλεψης καιρού, κυρίως λόγω της αξιοπιστίας και των δυνατοτήτων που προσφέρουν τα σύγχρονα υπολογιστικά συστήματα. Παραλλαγές των μοντέλων HIRLAM και ΕΤΑ χρησιμοποιούνται από την πλειοψηφία των Κρατικών Μετεωρολογικών Υπηρεσιών στις χώρες της Ευρώπης και της Αμερικής. Οι συνεχώς αυξανόμενες απαιτήσεις σε ακρίβεια και απόδοση, καθώς και η ανάγκη κλιμάκωσης παρόμοιων μοντέλων σε ευρύτερες περιοχές, οδηγούν στην προσπάθεια υλοποίησής τους σε παράλληλες μηχανές. Στη συνέχεια θα παρουσιαστούν οι ερευνητικές προσπάθειες που έχουν γίνει για την παραλληλοποίηση εφαρμογών του μοντέλου HIRLAM, όπως αναπτύχθηκε αρχικά για σειριακές και ανυσματικές αρχιτεκτονικές.

2. ΠΑΡΑΛΛΗΛΕΣ ΥΛΟΠΟΙΗΣΕΙΣ

Όπως αναφέρθηκε ήδη, τα αριθμητικά μοντέλα αποτελούν αναλυτική αναπαράσταση που προσεγγίζει με συγκεκριμένη ακρίβεια τις πραγματικές φυσικές διαδικασίες. Οι περιορισμοί κάθε υπολογιστικού περιβάλλοντος όσο αφορά την διακριτή αναπαράσταση, καθώς και η εν γένει πολυπλοκότητα του μοντέλου καθορίζει την μέγιστη δυνατή ακρίβεια που μπορεί να επιτευχθεί. Από την άλλη πλευρά, η χρησιμότητα της εφαρμογής εξαρτάται άμεσα από την ακρίβεια και την αξιοπιστία στα αποτελέσματα. Η δυνατότητα εκμετάλλευσης τέτοιων εφαρμογών σε ευρεία κλίμακα προϋποθέτει -θεωρητικά- άπειρη υπολογιστική ισχύ, κάτι φυσικά μη πραγματοποιήσιμο. Η υλοποίηση παρόμοιων μοντέλων σε παράλληλες αρχιτεκτονικές είναι η καλύτερη λύση από θεωρητική και πρακτική άποψη.

Αν και τα παράλληλα συστήματα προσφέρουν ισχύ ανάλογη με την αθροιστική ισχύ των επιμέρους μονάδων τους, η πρακτική -κυρίως- εμπειρία στην ανάπτυξη βέλτιστων εφαρμογών για την πλήρη αξιοποίηση των δυνατοτήτων τους είναι σχετικά μικρή. Η δομή και η συμπεριφορά τέτοιων συστημάτων, αν και έχει μελετηθεί διεξοδικά τα τελευταία χρόνια, παρουσιάζει εν γένει έναν επιπλέον βαθμό δυσκολίας στην ανάπτυξη αποδοτικών εφαρμογών.

Σε σχέση με τα αριθμητικά μοντέλα πρόβλεψης καιρού, η ανάγκη για όλο και μεγαλύτερη υπολογιστική ισχύ μπορεί να επικεντρωθεί σε τρία κύρια σημεία, τα οποία οφείλουν να εξεταστούν σε κάθε παράλληλη υλοποίησή τους:

1. **Ακρίβεια πρόβλεψης:** η ακρίβεια της πρόβλεψης εξαρτάται άμεσα από την ακρίβεια (ανάλυση) του μοντέλου και κατά συνέπεια από τον όγκο των προς επεξεργασία δεδομένων.
2. **Ταχύτητα επεξεργασίας:** κάθε μοντέλο πρέπει να λειτουργεί αποδοτικά και να παράγει αποτελέσματα σε καθορισμένο χρονικό πλαίσιο, ώστε να είναι έγκυρα και αξιοποιήσιμα.
3. **Κλιμάκωση:** κάθε επέκταση του πεδίου εφαρμογής του μοντέλου οδηγεί αναπόφευκτα σε μη γραμμική αύξηση της πολυπλοκότητας στο χρόνο υπολογισμού των αποτελεσμάτων.

Εκτός από τα παραπάνω κριτήρια αξιολόγησης της εφαρμογής, μια παράλληλη υλοποίηση θα πρέπει να παρουσιάζει μια σειρά από χαρακτηριστικά που θα την καθιστούν "καλύτερη" από τον αντίστοιχο ακολουθιακό αλγόριθμο, ειδικότερα όσο αφορά στην κλιμάκωση, την απόδοση και το κόστος. Τα κριτήρια που χαρακτηρίζουν ένα παράλληλο μοντέλο είναι:

- "Ποια η απαιτούμενη προσπάθεια για μια υλοποίηση χρήσιμη και αποδοτική;"
- "Ποια η αποτελεσματικότητα της λύσης ως προς το κόστος/προσπάθεια υλοποίησης;"

- "Ποια η δυνατότητα κλιμάκωσης του μοντέλου;"

Πριν μελετηθούν οι διαθέσιμες παράλληλες υλοποιήσεις, θα πρέπει να εξεταστεί ο τρόπος διατύπωσης και επίλυσης των ίδιων των φυσικών συστημάτων μέσω των μοντέλων κίνησης ρευστών.

3. ΜΟΝΤΕΛΑ ΚΙΝΗΣΗΣ ΡΕΥΣΤΩΝ

Θεωρητικό υπόβαθρο για την κατασκευή αριθμητικών μοντέλων αποτελεί η δυναμική γεωφυσικών ρευστών (geophysical fluid dynamics) και ειδικότερα οι εξισώσεις Navier-Stokes, οι οποίες συχνά προσεγγίζονται με διάφορες μεθόδους. Το σύστημα των εξισώσεων για ένα καθολικό μοντέλο ωκεανών έχει ως εξής:

$$\frac{du}{dt} - \frac{uv \tan \phi}{\alpha} - f v = -\frac{1}{\rho \alpha \cos \phi} \frac{\partial p}{\partial \lambda} + F_\lambda \quad (1) \text{ (momentum equation)}$$

$$\frac{dv}{dt} - \frac{u^2 \tan \phi}{\alpha} - fu = -\frac{1}{\rho \alpha} \frac{\partial p}{\partial \phi} + F_\phi \quad (2) \text{ (momentum equation)}$$

$$\frac{\partial p}{\partial z} = -\rho g \quad (3) \text{ (hydrostatic equation)}$$

$$\frac{1}{\alpha \cos \phi} \left[\frac{\partial u}{\partial \lambda} + \frac{\partial (v \cos \phi)}{\partial \phi} \right] + \frac{\partial w}{\partial z} = 0 \quad (4) \text{ (continuity equation)}$$

$$\rho = \rho(T, p, \dots) \quad (5) \text{ (equation of state)}$$

$$\frac{dT}{dt} = f(T, \dots) \quad (6) \text{ (thermodynamic equation)}$$

$$\frac{d}{dt} \equiv \frac{\partial}{\partial t} + \frac{u}{\alpha \cos \phi} \frac{\partial}{\partial \lambda} + \frac{v}{\alpha} \frac{\partial}{\partial \phi} + w \frac{\partial}{\partial z} \quad (D)$$

Στις παραπάνω εξισώσεις αναπαριστάται η φυσική συμπεριφορά του μοντέλου. (Ω) είναι η γωνιακή ταχύτητα, (λ) και (ϕ) είναι το γεωγραφικό μήκος και πλάτος, $f=2\Omega \sin \phi$ είναι η παράμετρος Coriolis και F_λ , F_ϕ είναι οι δυνάμεις τριβής. Η τελευταία εξίσωση παρουσιάζει τον κανόνα διαφόρισης για όλες τις εξισώσεις. Οι άγνωστες παράμετροι είναι οι συνιστώσες της ταχύτητας u , v , w , η πίεση p , η πυκνότητα ρ και η θερμοκρασία T . Για μοντέλα ατμόσφαιρας

συνήθως προστίθεται άλλη μια εξίσωση υγρασίας (μ), ενώ για μοντέλα ωκεανού χρησιμοποιείται μια εξίσωση περιεκτικότητας σε αλάτι. Πάντως σε γενικές γραμμές το σύστημα εξισώσεων που χρησιμοποιείται σε κάθε περίπτωση έχει αυτή τη μορφή, μια και η δυναμική και στα δύο μοντέλα είναι σχεδόν η ίδια.

Αξίζει να σημειωθεί πως, αν και η θεωρητική διατύπωση των μοντέλων ατμόσφαιρας και ωκεανών είναι το ίδιο, παρόλα αυτά η πρακτική εφαρμογή τους διαφέρει αρκετά. Αυτό οφείλεται κυρίως στο γεγονός ότι τα ατμοσφαιρικά φαινόμενα εξελίσσονται με πολύ γρηγορότερους ρυθμούς και περιλαμβάνουν πιο πολύπλοκες φυσικές διαδικασίες, όπως για παράδειγμα οι νεφώσεις. Σαν αποτέλεσμα, οι κλίμακες χώρου και χρόνου στις δύο περιπτώσεις διαφέρουν σημαντικά. Για τα ατμοσφαιρικά μοντέλα η έκταση του πεδίου εφαρμογής του μοντέλου είναι της τάξης των μερικών δεκάδων μέχρι χιλιάδων χιλιομέτρων, ενώ στα μοντέλα ωκεανού ξεκινούν από μερικές εκατοντάδες μέχρι μερικές χιλιάδες χιλιόμετρα. Αντίστοιχα, η χρονική περίοδος εξέλιξης των ατμοσφαιρικών μοντέλων κυμαίνεται μεταξύ ημερών και ετών, ενώ στα μοντέλα ωκεανού από εβδομάδες ως και αιώνες. Η συνδυασμένη εφαρμογή ατμοσφαιρικών μοντέλων με μοντέλα ωκεανού και η μελέτη της συμπεριφοράς τους σε σχετικά μεγάλες χρονικές κλίμακες δίνουν τα λεγόμενα μοντέλα κλίματος για μια ευρεία γεωγραφική περιοχή ή για ολόκληρο τον πλανήτη.

Οπως έχει ήδη αναφερθεί, η υλοποίηση παρόμοιων μοντέλων σε υπολογιστή απαιτεί την τροποποίησή τους κατά τέτοιο τρόπο ώστε να είναι δυνατή η διακριτή αναπαράσταση των δεδομένων ως προς το χώρο και ως προς τον χρόνο. Έτσι, τα γενικευμένα μοντέλα μετατρέπονται σε διακριτά μοντέλα και οι διαφορικές εξισώσεις σε εξισώσεις διαφορών, για τις οποίες η ακρίβεια και ο χρόνος υπολογισμού της λύσης εξαρτάται άμεσα από την ανάλυση του διακριτού μοντέλου.

Για την επίλυση του συνόλου των εξισώσεων σε αυτά τα μοντέλα χρησιμοποιούνται τρεις κυρίως τεχνικές:

- **Μέθοδος πεπερασμένων διαφορών** (finite difference methods - **FDM**), όπου οι εξισώσεις επιλύονται με κάποια παραλλαγή των μεθόδων Euler.
- **Μέθοδος φασματικής ανάλυσης** (spectral methods - **SM**), όπου οι διακριτές τιμές των δεδομένων μετασχηματίζονται στο πεδίο των συχνοτήτων και στη συνέχεια επιλύονται οι εξισώσεις.
- **Μέθοδος πεπερασμένων στοιχείων** (finite element methods - **FEM**), όπου χρησιμοποιείται η προσέγγιση της τροχιάς στοιχειωδών (διακριτών) σωματιδίων.

Η επιλογή της καταλληλότερης μεθόδου επίλυσης των εξισώσεων του μοντέλου αποτελεί ερευνητικό ζήτημα, όχι μόνο ως προς την ακρίβεια και την απλότητα της μεθόδου, αλλά και ως προς τη δυνατότητα αποδοτικής υλοποίησής της σε παράλληλες αρχιτεκτονικές.

3.1 Το μοντέλο HIRLAM

Το μοντέλο HIRLAM (High Resolution Limited Area Model) είναι αποτέλεσμα ερευνητικού προγράμματος για την κατασκευή ενός συστήματος πρόβλεψης καιρού για περιορισμένο γεωγραφικό πεδίο, με δυνατότητα πρόβλεψης ως και 48 ώρες. Το πρόγραμμα αναπτύχθηκε για χρήση από πολλές μετεωρολογικές υπηρεσίες Ευρωπαϊκών χωρών, όπως Δανία, Φιλανδία, Ισλανδία, Ιρλανδία, Ολλανδία, Νορβηγία και Σουηδία. Η εφαρμογή αναπτύχθηκε κυρίως σε γλώσσα Fortran-77 (εκτός από τμήματα κώδικα I/O σε C) και σήμερα βρίσκεται σε χρήση σε διάφορες αρχιτεκτονικές, κυρίως ακολουθιακών και ανυσματικών υπολογιστικών συστημάτων.

Σαν αριθμητικό μοντέλο, βασίζεται σε ένα σύστημα διαφορικών εξισώσεων που μοντελοποιούν τη συμπεριφορά των φυσικών καιρικών φαινομένων, σύμφωνα με τα γενικευμένα μοντέλα ρευστών που παρουσιάστηκαν παραπάνω. Το σύστημα έχει τροποποιηθεί έτσι ώστε να αναπαριστά καλύτερα το ατμοσφαιρικό μοντέλο:

$$\frac{\partial u}{\partial t} = (f + \xi)v - \eta \frac{\partial u}{\partial \eta} - \frac{R_d T_v}{\alpha \cos \theta} \frac{\partial \ln p}{\partial \lambda} - \frac{1}{\alpha \cos \theta} \frac{\partial}{\partial \lambda} (\phi + E) + P_u + K_u \quad (1)$$

$$\frac{\partial v}{\partial t} = -(f + \xi)u - \eta \frac{\partial v}{\partial \eta} - \frac{R_d T_v}{\alpha} \frac{\partial \ln p}{\partial \theta} - \frac{1}{\alpha} \frac{\partial}{\partial \theta} (\phi + E) + P_v + K_v \quad (2)$$

$$\frac{\partial \phi}{\partial \eta} = -\frac{R_d T_v}{p} \frac{\partial p}{\partial \eta} \quad (3)$$

$$\frac{\partial T}{\partial t} = -\frac{u}{\alpha \cos \theta} \frac{\partial T}{\partial \lambda} - \frac{v}{\alpha} \frac{\partial T}{\partial \theta} - \eta \frac{\partial T}{\partial \eta} + \frac{\kappa T_v \omega}{(1+(\delta-1)q)p} + P_T + K_T \quad (4)$$

$$\frac{\partial q}{\partial t} = -\frac{u}{\alpha \cos \theta} \frac{\partial q}{\partial \lambda} - \frac{v}{\alpha} \frac{\partial q}{\partial \theta} - \eta \frac{\partial q}{\partial \eta} + P_q + K_q \quad (5) \text{ (moisture eq.)}$$

$$\frac{\partial}{\partial \eta} \frac{\partial p}{\partial t} + \nabla \left(\vec{v}_h \frac{\partial p}{\partial \eta} \right) + \frac{\partial}{\partial \eta} \left(\eta \frac{\partial p}{\partial \eta} \right) = 0 \quad (6)$$

$$\xi = \frac{1}{\alpha \cos \theta} \left(\frac{\partial v}{\partial \lambda} - \frac{\partial}{\partial \theta} [\cos \theta u] \right) , \quad E = \frac{1}{2} (u^2 + v^2) \quad (D)$$

Οι παραπάνω εξισώσεις αποτελούν παραλλαγή των βασικών εξισώσεων Navier-Stokes, με την προσθήκη μιας εξίσωσης υγρασίας (5). Για την καλύτερη κατανόηση της δομής και της

συμπεριφοράς του μοντέλου, συχνά χρησιμοποιείται μια απλούστευμένη μορφή του που αναφέρεται ως "Shallow Water Model".

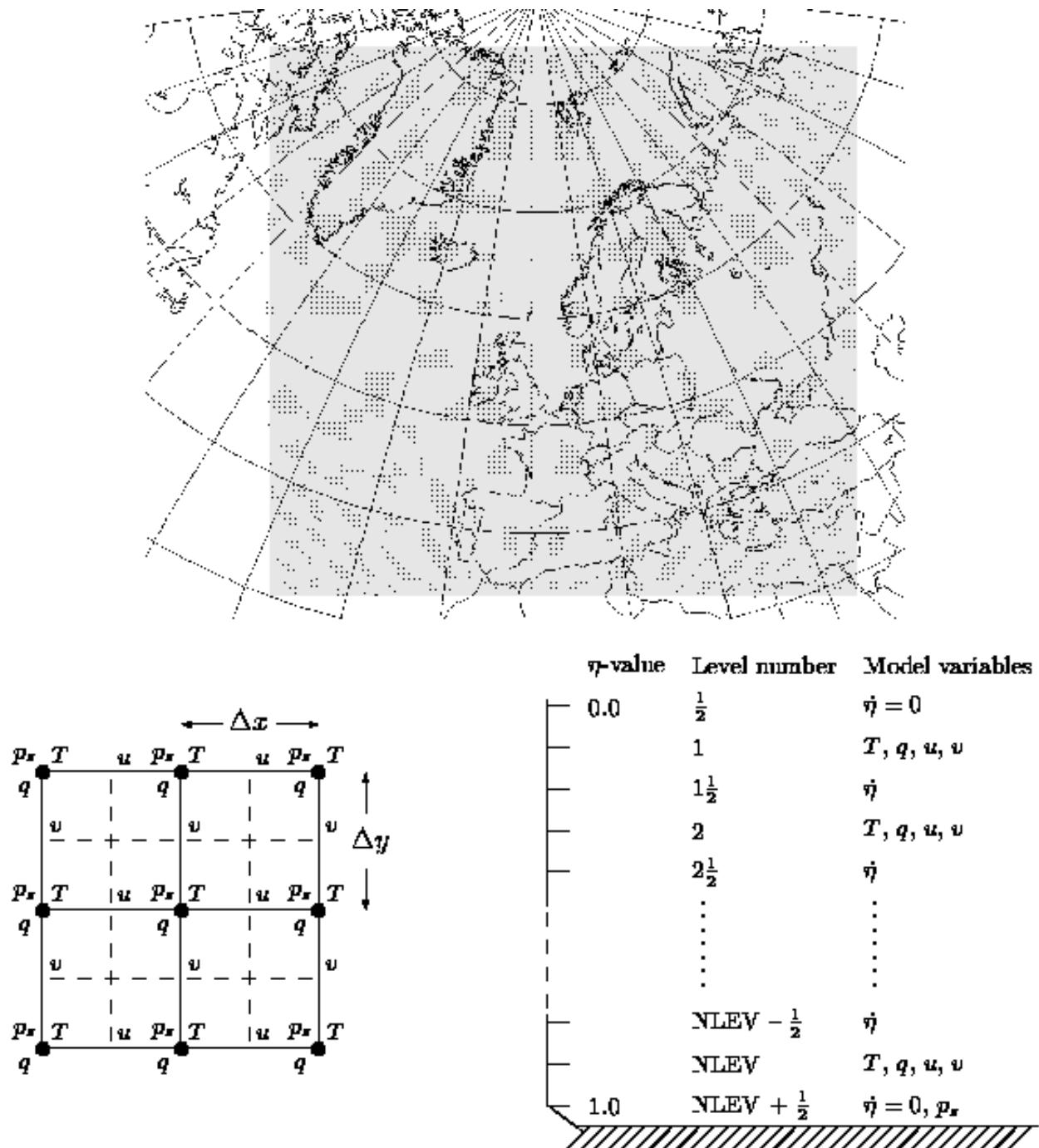
Shallow Water Model

Το μοντέλο αυτό θέτει μια σειρά περιορισμούς στο αρχικό σύστημα με στόχο την απλούστευση, τόσο της κατάτμησης της γεωγραφικής περιοχής, όσο και των υπολογισμών για την επίλυσή του.

Η υπό μελέτη γεωγραφική περιοχή θεωρείται επίπεδη (όχι σφαιρική) και η ατμόσφαιρα θεωρείται ως ρευστό ενός επιπέδου (1-layer), με σταθερή και ομοιόμορφη πυκνότητα. Με αυτό τον τρόπο το αρχικό σύστημα εξισώσεων απλοποιείται σε τρεις εξισώσεις 1^ο βαθμού, οι οποίες λύνονται σχετικά εύκολα.

Σκοπός του μοντέλου είναι κυρίως η μελέτη της συμπεριφοράς και της απόδοσης σε μια συγκεκριμένη αρχιτεκτονική, όπως για παράδειγμα η αρχική αξιολόγηση μιας παράλληλης υλοποίησης αριθμητικού μοντέλου πρόβλεψης καιρού σε κάποια SIMD αρχιτεκτονική.

Στο διακριτό μοντέλο HIRLAM, η γεωγραφική περιοχή χωρίζεται σε στοιχειώδη τμήματα για τα οποία θεωρείται ότι οι παράμετροι κατάστασης στο εσωτερικό τους είναι σταθερές με ομοιόμορφη κατανομή. Η περιοχή μελέτης είναι τώρα σφαιρική και εμπεριέχει όλες τις ιδιότητες της γήινης επιφάνειας (περιστροφή, δυνάμεις τριβής), ενώ η ατμόσφαιρα διαχωρίζεται στο κατακόρυφο επίπεδο σε 16 στρώματα (layers). Το οριζόντιο επίπεδο χωρίζεται σε ένα διπλό πλέγμα στοιχειωδών περιοχών (Arakawa C-grid) το οποίο αποτελεί το βάση για τις εξισώσεις διαφορών.



Η οριζόντια κατάτμηση της περιοχής είναι της τάξης των 60 με 120 gridpoints ανά διάσταση, ενώ ο χρόνος ολοκλήρωσης του κάθε time step εξαρτάται από τη μέθοδο επίλυσης που χρησιμοποιείται, καθώς και τις δυνατότητες του υπολογιστικού συστήματος. Πάντως η κατάτμηση της γεωγραφικής περιοχής, δηλαδή η χρησιμοποίηση διακριτού μοντέλου έχει ως αποτέλεσμα την αναπόφευκτη απόκλιση από την πραγματική συμπεριφορά του φυσικού συστήματος, που οφείλεται στις λεγόμενες "subgrid processes".

Subgrid Processes

Η απόκλιση του διακριτού μοντέλου από το πραγματικό σύστημα οφείλεται σε φυσικές διαδικασίες οι οποίες δεν καλύπτονται επαρκώς από την ανάλυση της κατάτμησης. Αυτό σημαίνει πως είτε εξελίσσονται σε χρονικά διαστήματα μικρότερα του time step, είτε σε περιοχές μικρότερες του gridpoint. Παράδειγμα τέτοιων διαδικασιών αποτελεί η ακτινοβολία που διαπερνά τα στρώματα της ατμόσφαιρας.

Για τη σωστή συμπεριφορά του μοντέλου απαιτείται η εφαρμογή τεχνικών "διόρθωσης" του συστήματος με προσεγγιστικό υπολογισμό των διαδικασιών αυτών. Η ευρύτερα χρησιμοποιούμενη τεχνική αναφέρεται ως "parameterization assumption", όπου οι subgrid processes προσεγγίζονται σε μεγαλύτερη κλίμακα, της τάξης δηλαδή του gridpoint, μέσω της εφαρμογής ενός συνόλου παραμέτρων που ενσωματώνονται στο αρχικό μοντέλο. Η εκτίμηση των subgrid processes και η ενσωμάτωσή τους στο αριθμητικό μοντέλο είναι εξίσου σημαντική με το βασικό σύστημα εξισώσεων και συνήθως είναι περισσότερο χρονοβόρα διαδικασία.

Με την ενσωμάτωση των subgrid processes ολοκληρώνεται η διατύπωση του διακριτού HIRLAM μοντέλου. Η επεξεργασία διαχωρίζεται σε δύο φάσεις, αυτή της επίλυσης του βασικού συστήματος εξισώσεων ή αλλιώς "δυναμικό" τμήμα (DYN) και αυτή του υπολογισμού των subgrid processes ή αλλιώς "φυσικό" τμήμα (PHYS). Αξίζει να σημειωθεί πως το τμήμα DYN αποτελεί το 30% του συνολικού χρόνου επεξεργασίας, σε αντίθεση με το τμήμα PHYS που είναι αρκετά χρονοβόρο (57.7%).

```

Program HIRLAM ;
Read data ;
Initialization ;
For each time step
    DYN ; {dynamical part}
    PHYS ; {physics part}
    6 hourly input of new boundary data
End For ;
Output results.

```

Η κυριότερη δυσκολία στην προσπάθεια διατύπωσης μιας παράλληλης υλοποίησης του HIRLAM είναι το γεγονός ότι, ενώ στο τμήμα DYN η εξάρτηση των δεδομένων (gridpoints) είναι κυρίως οριζόντια, στο τμήμα PHYS η εξάρτηση είναι κατά κύριο λόγο κάθετη. Αυτό σημαίνει

αναγκαστικά πως η κατάτμηση των δεδομένων σε μια (data-parallel) παράλληλη υλοποίηση θα ευνοεί μια από τις δύο φάσεις της επεξεργασίας, όσο αφορά στην ελαχιστοποίηση των επικοινωνιών. Εφ' όσον η φάση PHYS είναι το πιο πολύπλοκο τμήμα επεξεργασίας, αναγκαστικά η κατάτμηση των δεδομένων θα πρέπει να γίνει με κάθετη ομαδοποίηση και οριζόντιο διαχωρισμό του πλέγματος. Το ενδιαφέρον, έτσι, σε μια παράλληλη υλοποίηση επικεντρώνεται στη διαδικασία DYN και τη διατύπωση αποδοτικών τρόπων παράλληλης εκτέλεσής της. Τέλος, θα πρέπει να σημειωθεί πως η φάση ανάλυσης των αποτελεσμάτων του μοντέλου πρόβλεψης καιρού χαρακτηρίζεται από πολυπλοκότητα τουλάχιστον ίδιας τάξης με αυτή του ίδιου του μοντέλου.

3.2 Μέθοδοι Επίλυσης

Το σημαντικότερο στοιχείο που καθορίζει την αποδοτικότητα μιας παράλληλης υλοποίησης του μοντέλου HIRLAM είναι οι απαιτούμενες επικοινωνίες μεταξύ των επεξεργαστικών μονάδων. Πέρα από τον τρόπο κατάτμησης των δεδομένων, οι επικοινωνίες εξαρτώνται άμεσα από τη χρησιμοποιούμενη μέθοδο επίλυσης που χρησιμοποιεί η διαδικασία DYN. Όπως αναφέρθηκε, τρεις είναι οι γενικές μέθοδοι επίλυσης (FDM, SM, FEM) από τις οποίες προκύπτουν παραλλαγές του βασικού μοντέλου που χρησιμοποιούν Eulerian (explicit, semi-implicit), spectral και semi-Lagrangian συστήματα εξισώσεων.

Explicit Eulerian

Η explicit Eulerian μέθοδος επίλυσης χρησιμοποιεί την απλή (άμεση) Euler μέθοδο για εξισώσεις διαφορών με σταθερό time step. Είναι η πιο απλή μέθοδος επίλυσης και για το λόγο αυτό χρησιμοποιείται συχνά, όχι μόνο στο μοντέλο HIRLAM αλλά και σε άλλα παρόμοια αριθμητικά μοντέλα (ETA).

Το σημαντικότερο πλεονέκτημα της μεθόδου είναι η απλότητα υλοποίησης ακόμη και για πολύπλοκα συστήματα εξισώσεων διαφορών. Παρόλα αυτά, το μέγιστο χρησιμοποιούμενο time step είναι σχετικά μικρό, αφού περιορίζεται από τις μέγιστες (απόλυτες) τιμές των παραμέτρων του μοντέλου και ιδιαίτερα των συνιστωσών της ταχύτητας.

Semi-implicit Eulerian

Αποτελεί παραλλαγή της απλής μεθόδου Euler. Εδώ χρησιμοποιείται ημι-έμμεσο σχήμα για την επίλυση των εξισώσεων διαφορών, κάνοντας την μέθοδο πιο ακριβής αλλά ταυτόχρονα πιο πολύπλοκη.

Στα έμμεσα (implicit) σχήματα χρησιμοποιούνται επαναληπτικές μέθοδοι για την επίλυση των διαφορικών εξισώσεων, κάτι που κάνει τη μέθοδο περισσότερο αποδοτική αλλά και ασταθής. Οι semi-implicit τεχνικές συνδυάζουν τα πλεονεκτήματα των άμεσων (explicit) τεχνικών σε επαναληπτικές μεθόδους επίλυσης.

Ως μέθοδος Euler παραμένει απλή στην υλοποίηση, ενώ το έμμεσο σχήμα επίλυσης δίνει τη δυνατότητα για μεγαλύτερο time step, ακόμα και πέντε φορές το αντίστοιχο time step για την explicit μέθοδο. Η κυριότερη δυσκολία υλοποίησης, σε σχέση πάντα με την πολυπλοκότητα της μεθόδου, είναι η ανάγκη χρησιμοποίησης επαναληπτικών (έμμεσων) μεθόδων για την αριθμητική επίλυση της εξίσωσης του Helmholtz.

Spectral

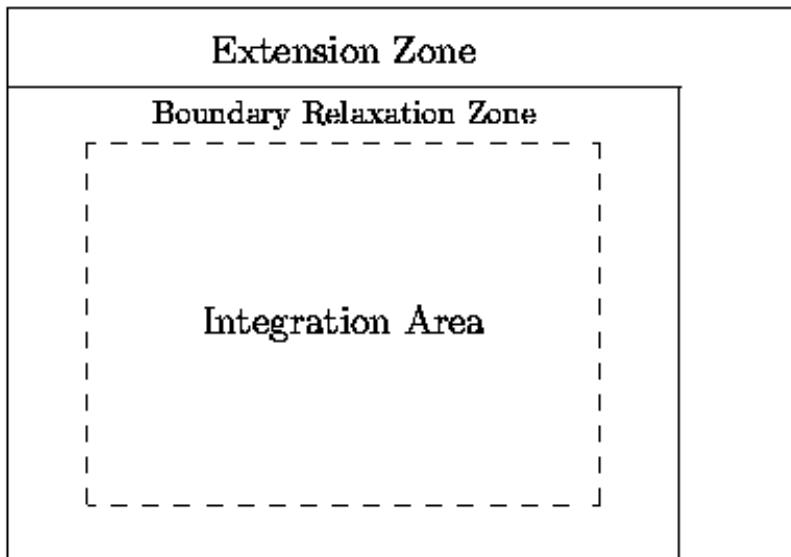
Οι μέθοδοι Euler, αν και χαρακτηρίζονται από την σχετική απλότητα στην υλοποίηση, προσφέρουν περιορισμένη σχετικά απόδοση εν γένει λόγω του τρόπου επίλυσης των διαφορικών εξισώσεων. Ένας εναλλακτικός τρόπος επίλυσης του συστήματος που διαφέρει αρκετά από τις κλασικές μεθόδους για εξισώσεις διαφορών, είναι μέσω φασματικής ανάλυσης.

Τα δεδομένα του πλέγματος που σχετίζονται με τις παραμέτρους του μοντέλου σε κάθε σημείο αρχικά μετασχηματίζονται μέσω Fast Fourier Transformations από το πεδίο του χρόνου στο πεδίο των συχνοτήτων. Το σύστημα επιλύεται στο πεδίο των συχνοτήτων και τελικά με αντίστροφους FFT μετασχηματισμούς λαμβάνονται οι αντίστοιχες τιμές του πλέγματος.

Ο μετασχηματισμός του συνόλου των δεδομένων στο πεδίο των συχνοτήτων προσφέρει τη δυνατότητα επίλυσης των διαφορικών εξισώσεων με μεγαλύτερη ακρίβεια, αφού δεν υπόκειται στους ίδιους περιορισμούς με τις κλασικές μεθόδους διαφορών. Επιπλέον, ο έλεγχος του φάσματος μέσω εφαρμογής κατάλληλων φίλτρων δίνει τη δυνατότητα απόρριψης φαινομένων θορύβου. Η πειραματική απόδοση των spectral μεθόδων είναι ανάλογη με αυτή των seim-implicit Eulerian μεθόδων.

Ένα σημαντικό μειονέκτημα των φασματικών μεθόδων είναι το γεγονός ότι για την εφαρμογή αποδοτικών (παράλληλων) μετασχηματισμών FFT στα δεδομένα απαιτείται η επέκταση του βασικού πλέγματος (integration area + boundary relaxation zone) κατά ένα ποσοστό περίπου 25% (10% ανά διάσταση) που αποτελεί τη ζώνη επέκτασης (extension zone). Αυτό έχει συχνά αποτέλεσμα τη συνολική αύξηση των εξεταζόμενων gridpoints περίπου δύο φορές ($(4/3)^2$) σε σχέση με τις τυπικές μεθόδους διαφορών, για να επιτευχθεί η ίδια ακρίβεια. Αξίζει, επίσης, να σημειωθεί ότι, αν και η πειραματική συμπεριφορά των παράλληλων διαδικασιών FFT παρουσίασαν πολλές φορές super-linear αύξηση στην απόδοση με την αύξηση των επεξεργαστικών μονάδων,

όμως σε γενικές γραμμές οι φασματικές μέθοδοι δεν παρουσιάζουν καλή κλιμάκωση. Η απόδοση των φασματικών μεθόδων εξαρτάται σχεδόν αποκλειστικά από την απόδοση των παράλληλων διαδικασιών FFT, μια και οι επικοινωνιακές απαιτήσεις σε αυτές είναι πολύ μεγάλες σε σχέση με τις απλές διακριτές μεθόδους.

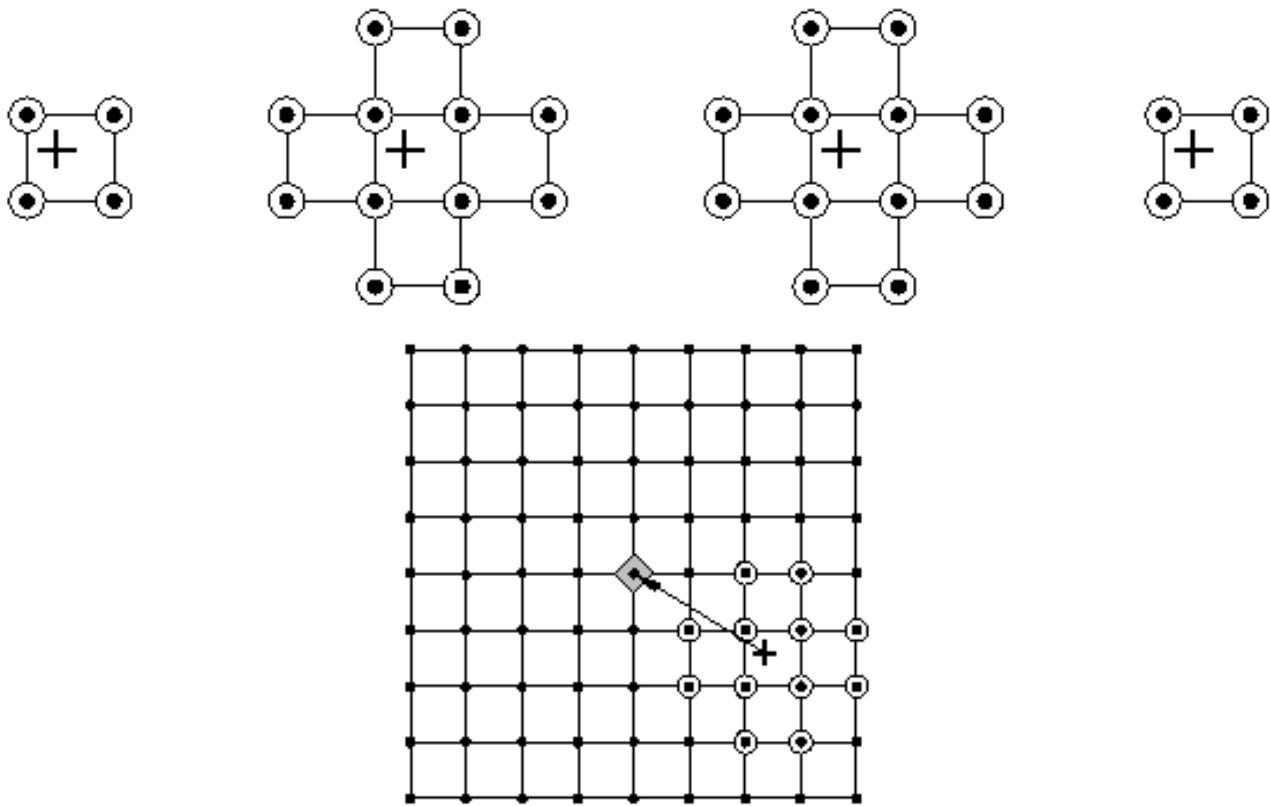


Semi-Lagrangian

Τόσο στις απλές διακριτές, όσο και στις φασματικές μεθόδους, το ενδιαφέρον επικεντρώθηκε στην τροποποίηση των τιμών των παραμέτρων του μοντέλου σε κάθε gridpoint στατικά, χρησιμοποιώντας ένα σύνολο εξισώσεων. Στις μεθόδους διακριτών στοιχείων (finite element methods) το ενδιαφέρον μεταφέρεται στην παρακολούθηση της εξέλιξης του μοντέλου ως ένα σύνολο στοιχειωδών σωματιδίων. Η semi-Lagrangian μέθοδος αποτελεί υλοποίηση της τεχνικής αυτής.

Παρόλο που το στοιχείο επεξεργασίας εδώ δεν είναι άμεσα το gridpoint αλλά το σωματίδιο, η μορφή του μοντέλου θα πρέπει, όπως και πριν, να βασίζεται στην κατάτμηση του πεδίου επεξεργασίας σε gridpoints. Ο κυριότερος λόγος για την απαίτηση αυτή είναι η ανάγκη ομοιομορφίας της κατανομής των τιμών των παραμέτρων του μοντέλου, κάτι που χωρίς την προσαρμογή των καταστάσεων των σωματιδίων στο ομοιόμορφο πλέγμα θα ήταν αδύνατη. Κατά συνέπεια, η τεχνική που χρησιμοποιείται για την ενημέρωση του συστήματος μοιάζει με αυτή των φασματικών μοντέλων. Αφού καθοριστεί η τρέχουσα κατάσταση ενός σωματιδίου, το οποίο γενικά δεν συμπίπτει με κάποιο διακριτό σημείο του πλέγματος, στη συνέχεια υπολογίζεται μια προσέγγιση της τροχιάς του μέσα στο καθορισμένο time step. Η ενημέρωση των παραμέτρων στην

τελική κατάσταση γίνεται σε συνδυασμό με την προσαρμογή στο διακριτό πλέγμα του πεδίου επεξεργασίας.



Η χρησιμοποίηση semi-Lagrangian μεθόδων σε αριθμητικά μοντέλα πρόβλεψης καιρού είναι σχετικά περιορισμένη. Ο κυριότερος λόγος είναι η πολυπλοκότητα της υλοποίησης, ιδιαίτερα σε παράλληλα υπολογιστικά περιβάλλοντα. Η ανομοιόμορφη ενημέρωση των τιμών των παραμέτρων του πλέγματος δημιουργεί γενικά μη προβλέψιμες επικοινωνίες μεταξύ των επεξεργαστικών μονάδων, κάτι που δημιουργεί πολλά προβλήματα στη σχεδίαση και ανάπτυξη αποδοτικών παράλληλων υλοποιήσεων. Πάντως, αξίζει να σημειωθεί πως η ακρίβεια και η απόδοση της λύσης εξαρτάται μόνο εν μέρει από την κατάτμηση της περιοχής έρευνας, ενώ το χρησιμοποιούμενο time step είναι ως και τρεις φορές μεγαλύτερο από τις Eulerian μεθόδους γενικά (μεγαλύτερο ακόμα και από τις implicit παραλλαγές). Αν και οι πειραματικές μελέτες έχουν αναδείξει μονάχα ελάχιστες αποδοτικές υλοποιήσεις, είναι φανερό πως παρόμοιες μέθοδοι μπορούν να επιτύχουν παράλληλη απόδοση συγκρίσιμη με τις καλύτερες semi-implicit και spectral εφαρμογές.

4. ΚΑΤΑΝΟΜΗ ΔΕΔΟΜΕΝΩΝ & ΕΠΙΚΟΙΝΩΝΙΑ

Αν και η μέθοδος επίλυσης του μοντέλου αποτελεί το κύριο αντικείμενο έρευνας στο μοντέλο HIRLAM, οι ίδιες οι υλοποιήσεις του μοντέλου σε διαφορετικά παράλληλα περιβάλλοντα οδηγούν στη μελέτη αποδοτικής κατανομής της εργασίας ή/και των δεδομένων. Ανάλογα με τις δυνατότητες της κάθε παράλληλης αρχιτεκτονικής, εξετάζονται διαφορετικοί τρόποι παραλληλοποίησης του μοντέλου.

4.1 Αναπαράσταση & Αποθήκευση των Δεδομένων

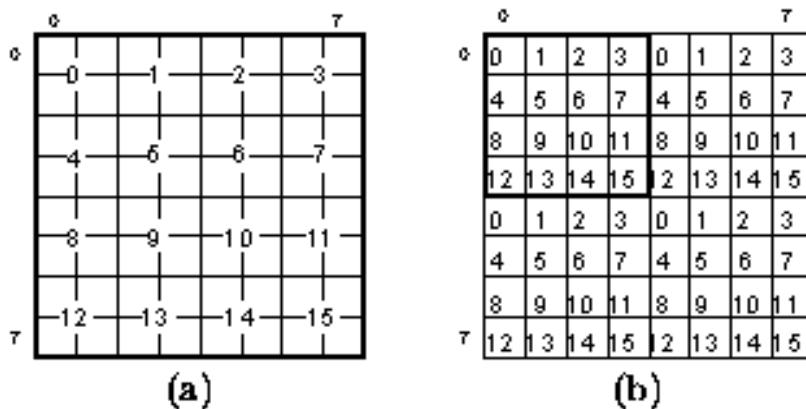
Η αρχική υλοποίηση του μοντέλου HIRLAM έγινε με βάση συμβατικούς (ακολουθιακούς) ή ανυσματικούς επεξεργαστές σε γλώσσα προγραμματισμού Fortran-77. Από τη μια πλευρά οι περιορισμοί του περιβάλλοντος επεξεργασίας και από την άλλη οι ελλιπείς μηχανισμοί ανάπτυξης δομημένων εφαρμογών στη συγκεκριμένη γλώσσα, οδήγησε αναπόφευκτα σε μια υλοποίηση δύσκολα μεταφέρσιμη σε διαφορετικές αρχιτεκτονικές.

Ένα σημαντικό ζήτημα στην παραλληλοποίηση του κώδικα της αρχικής υλοποίησης ήταν η μελέτη της επίπτωσης που έχει ο τρόπος αναπαράστασης του πλέγματος του μοντέλου. Ενώ η φυσική αναπαράσταση είναι σε πίνακες δύο διαστάσεων, η αρχική υλοποίηση χρησιμοποιούσε διανύσματα με γραμμική απεικόνιση. Αν και η διαφοροποίηση αυτή δεν έχει ιδιαίτερη επίπτωση στο ίδιο το μοντέλο, πειραματικές μελέτες έχουν δείξει ότι ο βαθμός επίπτωσης στην τελική απόδοση εξαρτάται από την εκάστοτε αρχιτεκτονική. Πάντως από άποψη σωστής δόμησης του λογισμικού της εφαρμογής και απλοποίησης του παράλληλου κώδικα, οι νεότερες υλοποιήσεις εφαρμόζουν την αναπαράσταση με πίνακες (δύο διαστάσεων).

4.2 Κατανομή Δεδομένων στους Επεξεργαστές

Σε μαζικά παράλληλες μηχανές με πίνακες επεξεργαστικών στοιχείων, η κατανομή των στοιχείων του 2-D πλέγματος στους επεξεργαστές μπορεί να γίνει με πολλούς τρόπους. Για ομοιόμορφη κατανομή, οι δύο σημαντικότερες παραλλαγές είναι η κατανομή κατά τμήματα (blocked) και η διάσπαρτη κατανομή (scattered).

Στην blocked κατανομή σε κάθε επεξεργαστή ανατίθεται ίσο ποσοστό του συνολικού πεδίου πλέγματος με συνεχόμενο τρόπο. Αυτό σημαίνει ότι όλα τα gridpoints που αναλαμβάνει ο κάθε επεξεργαστής είναι διαδοχικά (γειτονικά). Αντίθετα, στην scattered κατανομή, κάθε επεξεργαστής λαμβάνει ένα gridpoint σύμφωνα με τη θέση που έχει στον πίνακα των μονάδων επεξεργασίας. Δηλαδή, σε αυτή την περίπτωση διαδοχικά gridpoints ανατίθενται σε γειτονικούς επεξεργαστές.



Κατανομή δεδομένων στους επεξεργαστές: (a) blocked, (b) scattered.

Η διαφορά στους δύο τρόπους κατανομής έγκειται στον τρόπο επικοινωνίας γειτονικών gridpoints σε σχέση με γειτονικούς επεξεργαστές. Πειραματικές μελέτες έχουν δείξει ότι, αν και η μέση απόδοση των παράλληλων υλοποιήσεων είναι στα ίδια πλαίσια, η scattered κατανομή εμφανίζει σταθερή κλιμάκωση σε σχέση με το πλήθος των επεξεργαστών από ότι η blocked κατανομή, αλλά μικρότερη απόδοση, ιδιαίτερα σε μαζικά παράλληλα συστήματα, λόγω του πρόσθετου επικοινωνιακού φόρτου. Η επιλογή του βέλτιστου τρόπου κατανομής εξαρτάται από τις δυνατότητες του εκάστοτε παράλληλου συστήματος.

4.3 Υλοποίηση Παράλληλου Κώδικα

Η δυνατότητα υλοποίησης του μοντέλου αρχικά σε μηχανές SIMD (MasPar MP-x) και στη συνέχεια και σε μηχανές MIMD (Cray T3D), οδήγησε στη διατύπωση διαφορετικών τρόπων προσέγγισης του προβλήματος της παραλληλοποίησης του κώδικα. Η πιο λογική προσέγγιση για παραλληλοποίηση σε περιβάλλον SIMD ήταν φυσικά η data-parallel προσέγγιση, ενώ για MIMD περιβάλλον η ισοκατανομή της εργασίας (work sharing) είναι πιο φυσική.

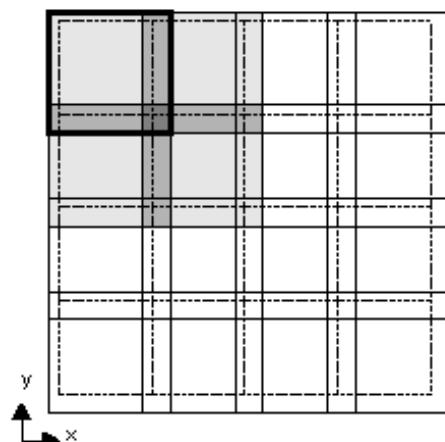
Οι αρχικές προσπάθειες υλοποίησης παράλληλου κώδικα για SIMD μηχανές έγιναν για συστήματα MP-1 και MP-2, τα οποία είναι μαζικά παράλληλα συστήματα με πίνακες επεξεργαστικών στοιχείων. Έτσι η φυσική προσέγγιση ήταν η κατανομή των δεδομένων σε επίπεδο gridpoints. Αντίθετα, η μετέπειτα υλοποίηση παράλληλου κώδικα για MIMD μηχανές έγιναν με βάση το σύστημα Cray T3D, ένα πολύ-υπολογιστικό σύστημα. Η προσέγγιση σε αυτή την περίπτωση είναι η κατανομή της εργασίας, δηλαδή του κώδικα επεξεργασίας, ανάμεσα στις μηχανές.

Και στις δύο περιπτώσεις η υλοποίηση φαίνεται να δεσμεύεται στη συγκεκριμένη αρχιτεκτονική, με σκοπό τη βέλτιστη αξιοποίηση των δυνατοτήτων του συστήματος μέσω ειδικών οδηγιών για τη μεταγλώττιση της εφαρμογής (compiler directives) και εξειδικευμένων εργαλείων παραλληλοποίησης του αρχικού κώδικα. Μια άμεση σύγκριση της απόδοσης των δύο διαφορετικών προσεγγίσεων δεν είναι εφικτή, μια και πρόκειται για εντελώς διαφορετικές παράλληλες αρχιτεκτονικές. Από τεχνολογικής άποψης, η υλοποίηση σε σύστημα Cray T3D εμφανίζει σίγουρα μεγαλύτερη απόδοση (τουλάχιστον μια τάξη μεγέθους), αλλά η χρησιμοποίηση μαζικών συστημάτων εμφανίζει καλύτερο λόγο κόστους προς απόδοση.

4.4 Διαχωρισμός Πεδίου (Sub-Domain Splitting)

Όπως αναφέρθηκε προηγούμενα, η βέλτιστη αξιοποίηση MIMD παράλληλων συστημάτων απαιτεί διαφορετική προσέγγιση στο πρόβλημα παραλληλοποίησης του κώδικα του μοντέλου. Μια προφανής επιλογή για τον τρόπο κατανομής της εργασίας σε ένα τέτοιο σύστημα είναι η μέθοδος του sub-domain splitting.

Η μέθοδος αυτή, όπως φανερώνει και το όνομα, βασίζεται στην υποδιαίρεση ολόκληρου του πεδίου του πλέγματος σε αυτόνομα τμήματα και την ανάθεση της επεξεργασίας κάθε ενός από μια μηχανή του συστήματος. Η ίδια η μέθοδος δεν απαιτεί καμιά ριζική αλλαγή στον κώδικα της εφαρμογής (εύκολη συντήρηση) παρά μόνο στα δεδομένα εισόδου και την περιοχή εξόδου για κάθε αυτόνομη μηχανή. Δεν απαιτείται απόλυτος συγχρονισμός ανάμεσα στις μηχανές, αλλά για τη σωστή λειτουργία και την ελάττωση του επικοινωνιακού κόστους εφαρμόζονται τεχνικές buffering στα άκρα των υποπεριοχών με επέκτασή τους, όπως στην περίπτωση του extension zone στις SM μεθόδους.



Αν και το σχήμα αυτό δημιουργεί αρκετά πλεονάζοντα στάδια επεξεργασίας, η πειραματική μελέτη του μοντέλου έχει δείξει πως είναι πολύ αποδοτικό σε αρχιτεκτονικές MIMD. Εκτός από την ελαχιστοποίηση των επικοινωνιών, σημαντικό πλεονέκτημα εδώ αποτελεί η δυνατότητα ανεξάρτητης σχετικά επεξεργασίας από τις μηχανές του συστήματος.

4.5 Explicit Message Passing

Ένα ακόμα ζήτημα που σχετίζεται με την πραγματική απόδοση της παράλληλης εφαρμογής είναι η πρακτική υλοποίηση των επικοινωνιών. Αν και οι επικοινωνίες εξαρτώνται σε μεγάλο βαθμό από την ίδια την αρχιτεκτονική και την οργάνωση του συστήματος, ένα μέρος του overhead εξαρτάται από τον ίδιο τον κώδικα της εφαρμογής.

Γενικά, έχει διαπιστωθεί πως σε πολλές περιπτώσεις οι αυτόματη μεταγλώττιση παράλληλου κώδικα σε διάφορα περιβάλλοντα προκαλεί την εισαγωγή επιπρόσθετου overhead στην τελική εφαρμογή, με αποτέλεσμα να απαιτείται η επέμβαση του ίδιο του προγραμματιστή για τον καθορισμό του μοντέλου επικοινωνίας. Μάλιστα σε πολλές περιπτώσεις κάτι τέτοιο επιβάλλεται όταν θα πρέπει να εξασφαλιστούν κάποιες προδιαγραφές συμβατότητας για τις εφαρμογές (PVM).

Η προσθήκη ρητών εντολών επικοινωνίας με σύγχρονο (blocked) τρόπο έχει δείξει πως γενικά βελτιώνει το συνολικό χρόνο επικοινωνίας κατά ένα ποσοστό. Προς το παρόν, τα πιο διαδεδομένα μοντέλα παράλληλου προγραμματισμού σε διαφορετικές πλατφόρμες είναι τα πρότυπα PVM και MPI. Στις περισσότερες περιπτώσεις η χρήση αυτών των προτύπων γίνεται για την υλοποίηση επικοινωνιών (message passing) με σύγχρονο τρόπο, αν και κάτι τέτοιο δεν είναι δεσμευτικό. Πάντως, αξίζει να αναφερθεί πως σε κάθε περίπτωση οι επικοινωνίες αυτού του τύπου επιφέρουν επιπλέον επικοινωνιακά κόστη, κυρίως λόγω του επαναλαμβανόμενου handshake.

Ειδικά για τα συστήματα που εφαρμόζουν shared memory management (Cray T3D), η χρησιμοποίηση ασύγχρονων επικοινωνιών αποτελεί ένα σημαντικό πλεονέκτημα. Αν και η ταχύτητα εγγραφής (put) είναι μικρότερη από την ταχύτητα ανάγνωσης (get), λόγω του cache invalidation, οι αυτόματοι μηχανισμοί διαχείρισης διαμοιραζόμενης μνήμης επιφέρουν σημαντικότατες βελτιώσεις στον τομέα των επικοινωνιών. Το σημαντικότερο μειονέκτημα είναι το γεγονός ότι η για τη βέλτιστη αξιοποίηση παρόμοιων μηχανισμών απαιτείται η σχεδόν αποκλειστική δέσμευση του κώδικα της εφαρμογής στη συγκεκριμένη αρχιτεκτονική της μηχανής.

5. ΖΗΤΗΜΑΤΑ ΥΛΟΠΟΙΗΣΗΣ

Αν και η έρευνα όσο αφορά την υλοποίηση μοντέλων όπως το HIRLAM σε παράλληλες μηχανές επικεντρώνεται σε ζητήματα όπως ο τρόπος επίλυσης του μοντέλου και η κατανομή των δεδομένων, αξίζει να αναφερθούν μερικά σημαντικά στοιχεία που σχετίζονται με τις πειραματικές μελέτες που έχουν γίνει για διαφορετικές υλοποιήσεις σε διαφορετικά υπολογιστικά περιβάλλοντα.

Η παράλληλη υλοποίηση του μοντέλου HIRLAM βασίστηκε σχεδόν αποκλειστικά στον αρχικό κώδικα της εφαρμογής που είχε αναπτυχθεί για ανυσματικές αρχιτεκτονικές. Επειδή και τα χαρακτηριστικά μιας εν γένει παράλληλης εφαρμογής, και τα χαρακτηριστικά της ίδιας της γλώσσας προγραμματισμού (Fortran-77) θέτουν πολλούς σχεδιαστικούς περιορισμούς σε μια νέα υλοποίηση, μεγάλη προσπάθεια αναλώθηκε στην τροποποίηση του αρχικού κώδικα σε πιο δομημένη και πιο μεταφέρσιμη μορφή (Fortran-90), πολλές φορές χωρίς επιτυχία. Αξίζει να αναφερθεί πως συγκεκριμένες διεργασίες του μοντέλου δεν υλοποιήθηκαν καθόλου στις πρώτες παράλληλες υλοποιήσεις (MP-1), επειδή η πολυπλοκότητα και η εσφαλμένη ανάπτυξη του κώδικα δεν επέτρεπε την μεταφορά του σε παράλληλο περιβάλλον.

Πέρα από τους περιορισμούς της ίδιας της εφαρμογής, σημαντικές δυσκολίες παρουσιάστηκαν στην μετατροπή του αρχικού κώδικα (περίπου 28.000 γραμμές) ώστε να αξιοποιεί την παράλληλη μηχανή. Στις πρώτες υλοποιήσεις του μοντέλου σε συστήματα MP-1 χρειάστηκε η προσθήκη οδηγιών μεταγλώττισης (compiler directives) στον αρχικό κώδικα σε τέτοια έκταση, ώστε ο κώδικας της τελικής εφαρμογής αυξήθηκε σχεδόν 100% σε μέγεθος, κάτι καθόλου επιθυμητό για μια εφαρμογή που έπρεπε να είναι μεταφέρσιμη σε πολλά διαφορετικά παράλληλα συστήματα. Ανάλογες ήταν και οι προσθήκες στις υλοποιήσεις για μηχανές Cray T3D.

Μερικά σημαντικά προβλήματα παρατηρήθηκαν στην αυτόματη παραλληλοποίηση και μεταγλώττιση του κώδικα σε μερικά περιβάλλοντα, ακόμη και με τη χρήση οδηγιών μεταγλώττισης. Πολλοί μεταγλωττιστές ή εργαλεία απέτυχαν στην μετατροπή του αρχικού κώδικα, είτε λόγω κακής ανάπτυξης του κώδικα, είτε λόγω κακού παραγόμενου κώδικα (περιττές επικοινωνίες δικτύου Xnet σε MP-1). Το γεγονός αυτό φανερώνει την σχετικά μικρή εμπειρία που υπάρχει στην πρακτική υλοποίηση αποδοτικών εφαρμογών, μεταγλωττιστών και αυτόματων εργαλείων για παράλληλα περιβάλλοντα.

Τέλος, οι πειραματικές μελέτες και οι δοκιμές που έγιναν σε πραγματικές συνθήκες λειτουργίας κατέδειξαν το σημαντικό ρόλο των σειριακών τμημάτων επεξεργασίας σε μια παράλληλη εφαρμογή, ιδιαίτερα όσων σχετίζονται με διεργασίες I/O. Σε μια εφαρμογή που περιλαμβάνει διαχείριση μεγάλου όγκου δεδομένων, όπως το μοντέλο HIRLAM, η απόδοση επηρεάζεται δραματικά από τις ακολουθιακές διαδικασίες I/O. Σε πραγματική λειτουργία του μοντέλου HIRLAM σε σύστημα MP-1, η μη ύπαρξη παράλληλων συσκευών I/O για την αποθήκευση και ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ. ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΧΑΡΗΣ ΓΕΩΡΓΙΟΥ, Μ-177

διαχείριση των ενδιάμεσων δεδομένων είχε σαν αποτέλεσμα την κατανάλωση ως και του 70% του συνολικού χρόνου επεξεργασίας σε λειτουργίες I/O, κάτι αρκετά απογοητευτικό. Πάντως, τα προβλήματα αυτά αφορούν κυρίως το τεχνολογικό παρά το θεωρητικό υπόβαθρο της εφαρμογής.

6. ΣΥΜΠΕΡΑΣΜΑΤΑ

Τα συμπεράσματα από τη μελέτη σχετικά με την παράλληλη υλοποίηση του μοντέλου HIRLAM σε πολλές διαφορετικές παράλληλες αρχιτεκτονικές, δεν αφορούν μονάχα στο συγκεκριμένο μοντέλο αλλά γενικότερα σε όλα τα αριθμητικά μοντέλα πρόβλεψης καιρού.

Ως γενικό συμπέρασμα προκύπτει πως η φύση του προβλήματος της αριθμητικής πρόβλεψης καιρού, οι σημερινές απαιτήσεις, καθώς και οι σημερινές τεχνολογικές δυνατότητες, κάνουν το πρόβλημα αυτό ιδανικό για μελέτη με στόχο την υλοποίηση τέτοιων μοντέλων σε παράλληλες μηχανές. Πέρα από την μεγάλη υπολογιστική ισχύ που απαιτείται σε τέτοιες εφαρμογές, το σημαντικότερο, ίσως, πλεονέκτημα είναι η δυνατότητα κλιμάκωσης που μπορεί να επιτευχθεί, τόσο σε SIMD, όσο και σε MIMD αρχιτεκτονικές.

Οι αποδοτικότερες υλοποίησεις για το μοντέλο HIRLAM φαίνεται να είναι οι semi-implicit Eulerian, οι spectral, καθώς και κάποιες semi-Lagrangian μέθοδοι. Όσο για τις τελευταίες μεθόδους, δεν υπάρχει επαρκές ερευνητικό υπόβαθρο σε σχέση με παράλληλες υλοποίησεις. Οι φασματικές μέθοδοι εμφανίζουν καλή απόδοση, αλλά εμφανίζουν μικρό βαθμό κλιμάκωσης και η απόδοσή τους εξαρτάται σχεδόν αποκλειστικά από τις παράλληλες FFT διεργασίες. Σήμερα, τα περισσότερα παράλληλα μοντέλα χρησιμοποιούν Eulerian (FDM) μεθόδους, semi-implicit ή ακόμα και explicit.

Σημαντικό ρόλο στο συνολικό χρόνο επεξεργασίας εμφανίζουν τα σειριακά τμήματα κώδικα και οι σειριακές διεργασίες γενικότερα σε ένα παράλληλο σύστημα. Τα προβλήματα επικεντρώνονται κυρίως στις διεργασίες I/O δεδομένων, είτε ως επικοινωνιακός φόρτος στο δίκτυο (μεταφορά δεδομένων αρχικών/τελικών/ενδιάμεσων), είτε σε σχέση με αποθηκευτικά μέσα (πρόσβαση σε αρχεία δίσκου).

Η υλοποίηση σε πολλά διαφορετικά υπολογιστικά περιβάλλοντα κατέδειξε τη συσχέτιση της απόδοσης με το κόστος της υλοποίησης. Σε μαζικά παράλληλα περιβάλλοντα (MasPar MP-x), οι εν γένει περιορισμοί του συστήματος, με κυριότερο το μεγάλο φόρτο επικοινωνιών, η απόδοση κυμαίνεται σε μέτρια ως ικανοποιητικά επίπεδα, όσο αφορά στο βαθμό αξιοποίησης του συστήματος. Ο λόγος, όμως, του κόστους ως προς την απόδοση είναι σχετικά υψηλός. Αντίθετα, σε πολύ-υπολογιστικά περιβάλλοντα (Cray T3D) η απόδοση είναι σχετικά σταθερή και σε πολύ

υψηλά επίπεδα, τόσο σε ακολουθιακές όσο και σε παράλληλες διεργασίες, αλλά το κόστος ως προς την απόδοση είναι σχετικά υψηλό.

Ως τελικό συμπέρασμα, αξίζει να αναφερθεί πως η μικρή, σχετικά, εμπειρία όσο αφορά στην ανάπτυξη εργαλείων και μεταγλωττιστών για παράλληλα περιβάλλοντα, καθώς και στη σχεδίαση και ανάπτυξη αποδοτικού παράλληλου κώδικα, δημιουργεί προβλήματα στην απόδοση και τη δυνατότητα μεταφοράς της τελικής εφαρμογής. Η εν γένει πολυπλοκότητα του προβλήματος της παραλληλοποίησης, καθώς και η ανάγκη αξιοποίησης στο έπακρο των προσφερόμενων παράλληλων αρχιτεκτονικών, μπορεί να αντιμετωπιστεί σε κάποιο βαθμό από τη σωστή ανάπτυξη του λογισμικού και την εκμετάλλευση σύγχρονης προγραμματιστικής μεθοδολογίας (δομημένος ή/και αντικειμενοστραφείς εφαρμογές).

ΑΝΑΦΟΡΕΣ

Παρακάτω παρουσιάζονται οι σημαντικότερες βιβλιογραφικές αναφορές που χρησιμοποιήθηκαν για το κείμενο της εργασίας και τις διαφάνειες της παρουσίασης. Οι πρώτες οκτώ αναφορές ([1]-[8]) αποτελούν μέρος των δημοσιεύσεων που υπάρχουν στην ηλεκτρονική διεύθυνση του High Performance Computing Division, Computer Science Department, Leiden University (NL) και οι κωδικοί "TRxx-xx" αποτελούν τα αντίστοιχα αναγνωριστικά των αρχείων που υπάρχουν εκεί σε ηλεκτρονική μορφή (PostScript). Οι υπόλοιπες αναφορές, εκτός της τελευταίας ([9]-[13]) αφορούν στα αντίστοιχα άρθρα από το "Proceedings of the Symposium on Regional Weather Prediction on Parallel Computer Environments" (Πανεπιστήμιο Αθηνών, Οκτώβριος 1997). Επίσης, στις ηλεκτρονικές διευθύνσεις περιλαμβάνεται το κεντρικό site (WWW) του συστήματος SKIRON στο Τμήμα Πληροφορικής του Πανεπιστημίου Αθηνών.

- [1] "Atmosphere and Ocean Circulation Simulation on Massively Parallel Computers" - TR92-15 (14p), Lex Wolters
- [2] "A Parallel Implementation of the HIRLAM model" - TR92-29 (15p), Lex Wolters, Gerard Cats
- [3] "Finite Difference and Spectral Models for Numerical Weather Forecasting on a Massively Parallel Computer" - TR93-07 (14p), Gerard Cats, Nils Gustafsson, Lex Wolters
- [4] "A Data Parallel HIRLAM Forecast Model" - TR94-47 (14p), Lex Wolters, Robert van Engelen, Gerard Cats, Nils Gustafsson, Tomas Wilhelmsson
- [5] "Data-Parallel Numerical Methods in a Weather Forecast Model" - TR94-48 (12p), Lex Wolters, Gerard Cats, Nils Gustafsson, Tomas Wilhelmsson
- [6] "Data-Parallel Numerical Weather Forecasting" - TR94-49 (15p), Lex Wolters, Gerard Cats, Nils Gustafsson
- [7] "Dataparallel Semi-Lagrangian Numerical Weather Forecasting" - TR94-50 (7p), Lex Wolters, Gerard Cats, Nils Gustafsson, Tomas Wilhelmsson
- [8] "A Comparison of Parallel Programming Paradigms and Data Distributions for a Limited Area Numerical Weather Forecast Routine" - TR95-10 (8p), Robert van Engelen, Lex

Wolters

- [9] "The Regional Weather Forecasting System SKIRON: An Overview"
G. Kallos, S. Nickovic, A. Papadopoulos, D. Jovic, O. Kakaliagou, N. Misirlis, L. Boukas, N. Mimikou, G. Sakellaridis, J. Papageorgiou, E. Anadranistakis, M. Manousakis.

- [10] "A Domain Decomposition Approach for the ETA Model: Implementation Issues and Results"
L. A. Boukas, N. Th. Mimikou, N. M. Misirlis

- [11] "Evaluation of the Performance of SKIRON Model"
G. Sakellaridis, E. Anadranistakis, T. Charantonis, P. Fragouli, G. Konstantinides, M. Manousakis, J. Papageorgiou, T. Petroliagis, M. Refene, D. Sgouros.

- [12] "Sensitivity Studies of the Surface and Radiation Parameterization Schemes of the SKIRON System"
A. Papadopoulos, G. Kallos, O. Kakaliagou, D. Jovic.

- [13] "Potential Use of ETA/NMC (SKIRON Project) Regional Forecasting on an Extreme Weather Event Over Greece"
T. Petroliagis, M. Anadranitsakis, V. Fragouli, M. Manousakis, I. Papageorgiou, M. Refene, G. Sakellaridis, S. Sgouros.

- [14] "Design and Implementation Issues of Parallelizing Atmospheric Models"
L. A. Boukas, N. Th. Mimikou, N. M. Missirlis.

☒ High Performance Computing Division, Computer Science Dept., Leiden Univ. (NL)

<http://www.wi.leidenuniv.nl/TechRep/>

☒ SKIRON - Weather Prediction System for the Mediterranean Region

<http://skiron.di.uoa.gr/>

Χάρης Γεωργίου (AM: M-177)

mailto: grad0177@di.uoa.gr

<http://www.di.uoa.gr/~grad0177/>

Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής (M.Sc.)

Αθήνα, Ιανουάριος 1999.