Richardson preconditioners for elliptic problems in numerical models of weather and climate

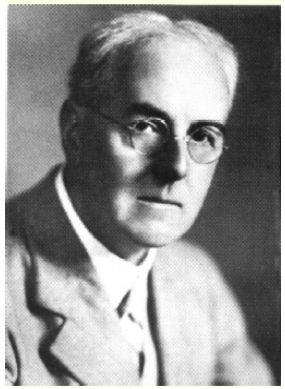
Piotr K. Smolarkiewicz
NCAR, Boulder, Colorado, USA

based on:

Z.P. Piotrowski & P.K. Smolarkiewicz, J. Comput. Phys. 463 (2022) 111296

Zbigniew Piotrowski,
Jan. 22, 2020

Aleksey Krylov,
1863-1945



Lewis Fry Richardson,
1881-1953

Historical remark: Web of Science record from 1965 to August 14, 2022

advection scheme (Topic) & ***MON. WEA. REV.***
Publications: 257; Citing Articles: 10,791;
Times Cited: 14,991; 59 H-Index.

advection scheme (Topic) & ***J. Comput. Phys.***
Publications: 619; Citing Articles: 19,922;
Times Cited: 28,587; 79 H-Index

elliptic solver (Topic) & ***MON. WEA. REV.***
Publications: 18; Citing Articles: 371**;**
Times Cited: 404; 11 H-Index

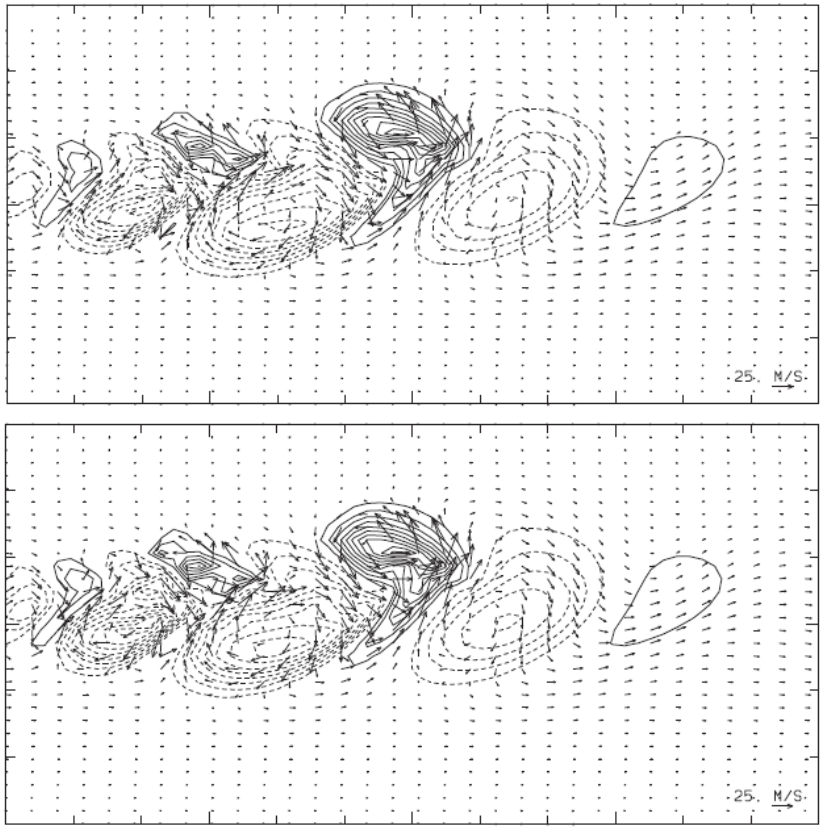
elliptic solver (Topic) & ***J. Comput. Phys***.
Publications: 240; Citing Articles: 5,623;
Times Cited: 6,744; 45 H-Index

preconditioner (Topic) & ***MON. WEA. REV.***
Publications: 6; Citing Articles: 227;
Times Cited: 233; 5 H-Index

preconditioner (Topic) & ***J. Comput. Phys.***
Publications: 320; Citing Articles: 6,976;
Times Cited: 8,291; 44 H-Index

Global baroclinic instability: semi-implicit vs. explicit integrations of all-scale equations; (Smolarkiewicz, Kuehnlein & Wedi, JCP 2014).

8 days, surface θ', 128x64x48 lon-lat grid, 128 PE of Power7 IBM



CPI2, 2880 dt=300 s, wallclock time=2.0 mns

CPEX, 432000 dt=2 s, wallclock time=178.9 mns

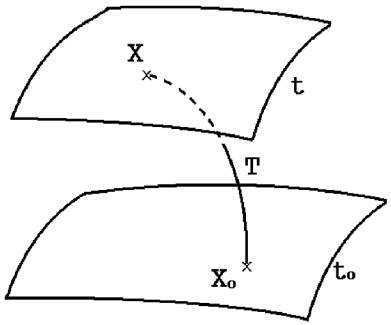
This huge computational-efficiency gain comes at the cost of increased mathematical/numerical complexity, reliant on the solution of an arduous elliptic boundary value problem (BVP)

Semi-implicit integrators (MHD example)

$$\frac{\partial \rho^* \boldsymbol{\Psi}}{\partial \bar{t}} + \overline{\nabla} \cdot (\mathbf{V}^* \boldsymbol{\Psi}) = \rho^* \mathbf{R} \ ,$$

$$\boldsymbol{\Psi} = \{\mathbf{u}, \Theta', \mathbf{B}\}^T$$

$$\mathbf{R} = \{\mathbf{R}_\mathbf{u}, R_{\Theta'}, \mathbf{R}_\mathbf{B}\}^T$$



$$\boldsymbol{\Psi}_\mathrm{i}^n = \mathcal{A}_\mathrm{i}\left(\widetilde{\boldsymbol{\Psi}},\ \widetilde{\mathbf{V}}^*,\ \rho^*\right) + 0.5\delta t \mathbf{R}_\mathrm{i}^n \equiv \widehat{\boldsymbol{\Psi}}_\mathrm{i} + 0.5\delta t \mathbf{R}_\mathrm{i}^n$$

$$\boldsymbol{\Psi}_\mathrm{i}^{n,\nu} = \widehat{\boldsymbol{\Psi}}_\mathrm{i} + 0.5\delta t\ \mathbf{L}\boldsymbol{\Psi}\big|_\mathrm{i}^{n,\nu} + 0.5\delta t\ \mathbf{N}(\boldsymbol{\Psi})\big|_\mathrm{i}^{n,\nu-1} - 0.5\delta t\ \widetilde{\mathbf{G}}\overline{\nabla}\boldsymbol{\Phi}\big|_\mathrm{i}^{n,\nu}$$

$$\boldsymbol{\Phi} \equiv (\pi',\ \pi',\ \pi',\ 0,\ \pi^*,\ \pi^*,\ \pi^*)$$

$$\boldsymbol{\Psi}_\mathrm{i}^{n,\nu} = [\mathbf{I} - 0.5\delta t \mathbf{L}]^{-1} \left(\widehat{\widehat{\boldsymbol{\Psi}}} - 0.5\delta t \widetilde{\mathbf{G}}\overline{\nabla}\boldsymbol{\Phi}^{n,\nu}\right)\Big|_\mathrm{i}$$

$$\widehat{\widehat{\boldsymbol{\Psi}}} \equiv \widehat{\boldsymbol{\Psi}} + 0.5\delta t \mathbf{N}\boldsymbol{\Psi}\big|^{n,\nu-1}$$

➔ elliptic problems for potentials **Φ**

Elliptic problems for generalized atmospheric equations

soundproof:

$$\mathcal{G}^{-1}\nabla\cdot(\mathcal{G}\rho\mathbf{v})=0 \quad \Rightarrow \quad \frac{1}{\mathcal{G}\rho}\nabla\cdot\left(\mathcal{G}\rho(\check{\check{\mathbf{v}}}-\widetilde{\mathbf{G}}^{\mathbf{T}}\mathbf{C}\nabla\varphi)\right)=\mathbf{0}\ ; \quad \text{with } \check{\check{\mathbf{v}}}=\widetilde{\mathbf{G}}^{\mathbf{T}}\,\check{\check{\mathbf{u}}}\equiv\widetilde{\mathbf{G}}^{\mathbf{T}}(\check{\check{\mathbf{u}}}'+\mathbf{u_a})$$

compressible: D/Dt (gas law) ➔

$$-\sum_{\ell=1}^{3}\left(\frac{A^{\star}_{\ell}}{\zeta_{\ell}}\nabla\cdot\zeta_{\ell}(\check{\check{\mathbf{v}}}-\widetilde{\mathbf{G}}^{T}\mathbf{C}\nabla\varphi)\right)-B^{\star}(\varphi-\widehat{\varphi})=0$$

➔ $\boxed{\mathcal{L}(\varphi)-R=0}$

EULAG's "exact-projection" nonsymmetric solver

The approach: left-preconditioned, restarted Generalized Conjugate Residual, "GCR(k)" of Eisenstat et al. SIAM J. Num. Anal. 20 (1983) --- interpreted as an extension of the Richardson iteration (Phil. Trans. Roy. Soc., London A210 (1910)); see also sec. 4.9 Physical Analogies in Birkhoff & Lynch, Numerical Solution of Elliptic Problems, SIAM (1984).

$$\frac{\partial^k \mathcal{P}(\varphi)}{\partial \tau^k} + \frac{1}{T_{k-1}(\tau)} \frac{\partial^{k-1} \mathcal{P}(\varphi)}{\partial \tau^{k-1}} + ... + \frac{1}{T_1(\tau)} \frac{\partial \mathcal{P}(\varphi)}{\partial \tau} = \mathcal{L}(\varphi) - \widehat{\mathcal{R}}$$

The actual algorithm:

In each node $\mathbf{i}$ of the grid, for any initial guess $\varphi^0_{\mathbf{i}}$, set $r^0_{\mathbf{i}} = \mathcal{L}_{\mathbf{i}}(\varphi^0) - \widehat{\mathcal{R}}_{\mathbf{i}}$, $q^0_{\mathbf{i}} = \mathcal{P}^{-1}_{\mathbf{i}}(r^0)$, then iterate:

For $n = 1, 2, ...$until convergence **Do** ; **for** $\nu = 0, .., k-1$ **do** ;

$$\beta = -\frac{\langle r^\nu \mathcal{L}(q^\nu) \rangle}{\langle \mathcal{L}(q^\nu) \mathcal{L}(q^\nu) \rangle} \; ; \quad \varphi^{\nu+1}_{\mathbf{i}} = \varphi^\nu_{\mathbf{i}} + \beta q^\nu_{\mathbf{i}} \; ; \quad r^{\nu+1}_{\mathbf{i}} = r^\nu_{\mathbf{i}} + \beta \mathcal{L}_{\mathbf{i}}(q^\nu) \; ;$$

exit if $\| r^{\nu+1} \| \leq \varepsilon$; $\epsilon_{\mathbf{i}} = \mathcal{P}^{-1}_{\mathbf{i}}(r^{\nu+1})$; $\forall_{\mathbf{i}}$ evaluate $\mathcal{L}_{\mathbf{i}}(\epsilon)$;

$$\forall_{l=0,\nu} \quad \alpha_l = -\frac{\langle \mathcal{L}(\epsilon) \mathcal{L}(q^l) \rangle}{\langle \mathcal{L}(q^l) \mathcal{L}(q^l) \rangle} \; ;$$

$$q^{\nu+1}_{\mathbf{i}} = \epsilon_{\mathbf{i}} + \sum_{l=0}^{\nu} \alpha_l q^l_{\mathbf{i}} \; ; \quad \mathcal{L}_{\mathbf{i}}(q^{\nu+1}) = \mathcal{L}_{\mathbf{i}}(\epsilon) + \sum_{l=0}^{\nu} \alpha_l \mathcal{L}_{\mathbf{i}}(q^l) \; ;$$

end for ; reset $[\varphi,\ r,\ q,\ \mathcal{L}(q)]^k_{\mathbf{i}}$ to $[\varphi,\ r,\ q,\ \mathcal{L}(q)]^0_{\mathbf{i}}$; **End For** .

Why preconditioning? ← problem stiffness; condition numbers & convergence rates; accuracy over a broad range of scales; reduced complexity (Skamarock et al., MWR, 1997)

A simple realization of $\epsilon_{\mathbf{i}} = \mathcal{P}_{\mathbf{i}}^{-1}(r^{\nu+1})$ is the Richardson iteration (RI) $\epsilon_{\mathbf{i}}^{\mu+1} = \epsilon_{\mathbf{i}}^{\mu} + \gamma\left(\mathcal{P}(\epsilon^{\mu}) - r^{\nu+1}\right)_{\mathbf{i}}$, with fixed pseudo-time step γ; $\mathcal{P} \approx \mathcal{L}$ is realized by neglecting off-diagonal entries of $\mathcal{L}$ ➔ polynomial preconditioner of limited value. However, RI can be generalized to semi-implicit time stepping as well as Fourier or multigrid representation. Here we consider the former:

$$\epsilon_{\mathbf{i}}^{\mu+1} = \epsilon_{\mathbf{i}}^{\mu} + \gamma\left(\mathcal{P}^h(\epsilon^{\mu}) + \mathcal{P}^v(\epsilon^{\mu+1}) - r^{\nu+1}\right)_{\mathbf{i}}$$

VIHEX; EULAG's historical standard

$$(D\epsilon^{\mu+1})_{\mathbf{i}} = (D\epsilon^{\mu})_{\mathbf{i}} + \gamma^*\left(\mathcal{P}^h(\epsilon^{\mu}) - D(\epsilon^{\mu+1} - \epsilon^{\mu}) + \mathcal{P}^v(\epsilon^{\mu+1}) - r^{\nu+1}\right)_{\mathbf{i}}$$
$$\epsilon^{\mu+1} = \omega\,\epsilon^{\mu+1}|_{\gamma^* \nearrow \infty} + (1-\omega)\,\epsilon^{\mu}$$

From diagonally preconditioned Dufort-Frankel to weighted line Jacoby; hereafter "reference"

$$\epsilon_{\mathbf{i}}^{\mu+1/2} = \epsilon_{\mathbf{i}}^{\mu} + \hat{\gamma}\left(\mathcal{P}^x(\epsilon^{\mu+1/2}) + \mathcal{P}^v(\epsilon^{\mu}) - \hat{r}^{\mu}\right)_{\mathbf{i}},$$
$$\epsilon_{\mathbf{i}}^{\mu+1} = \epsilon_{\mathbf{i}}^{\mu+1/2} + \hat{\gamma}\left(\mathcal{P}^x(\epsilon^{\mu+1/2}) + \mathcal{P}^v(\epsilon^{\mu+1}) - \hat{r}^{\mu}\right)_{\mathbf{i}},$$

Peaceman & Rachford 2D ADI

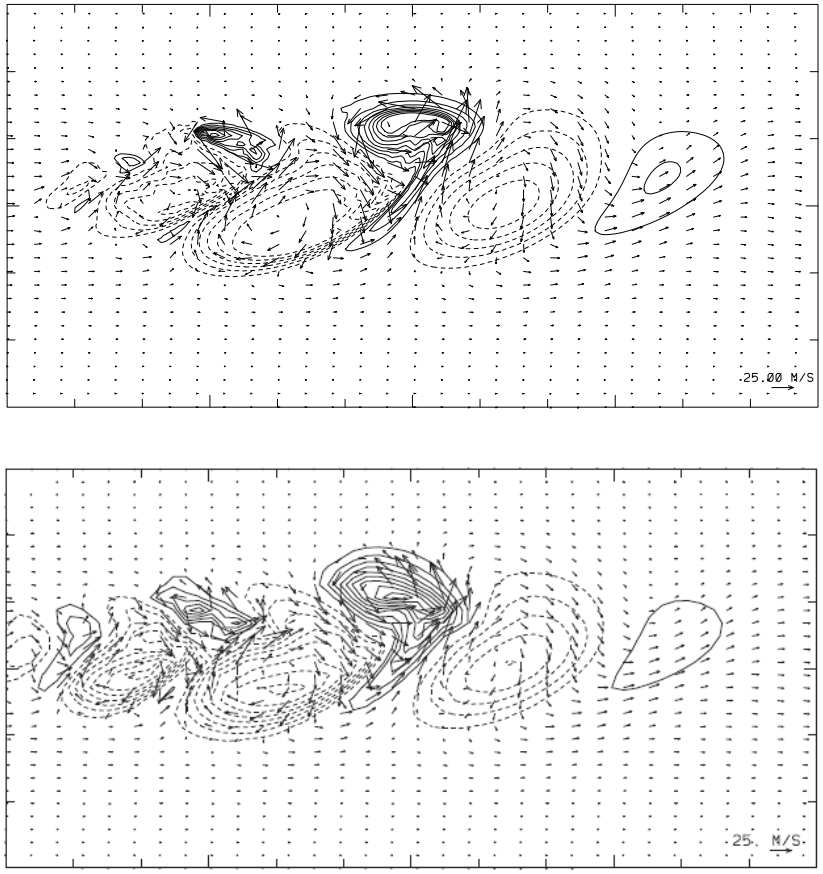
$$\hat{r}^{\mu} \equiv \mathcal{P}^y(\epsilon^{\bar{\mu}}) - r^{\nu+1}$$

$$\epsilon_{\mathbf{i}}^{\mu+1/3} = \epsilon_{\mathbf{i}}^{\mu} + \hat{\gamma}\left[\mathcal{P}^x(\{\epsilon^{\mu}, \epsilon^{\mu+1/3}\}) + \mathcal{P}^y(\epsilon^{\mu}) + \mathcal{P}^v(\epsilon^{\mu}) - r^{\nu}\right]_{\mathbf{i}},$$
$$\epsilon_{\mathbf{i}}^{\mu+2/3} = \epsilon_{\mathbf{i}}^{\mu} + \hat{\gamma}\left[\mathcal{P}^x(\{\epsilon^{\mu}, \epsilon^{\mu+1/3}\}) + \mathcal{P}^y(\{\epsilon^{\mu}, \epsilon^{\mu+2/3}\}) + \mathcal{P}^v(\epsilon^{\mu}) - r^{\nu}\right]_{\mathbf{i}},$$
$$\epsilon_{\mathbf{i}}^{\mu+3/3} = \epsilon_{\mathbf{i}}^{\mu} + \hat{\gamma}\left[\mathcal{P}^x(\{\epsilon^{\mu}, \epsilon^{\mu+1/3}\}) + \mathcal{P}^y(\{\epsilon^{\mu}, \epsilon^{\mu+2/3}\}) + \mathcal{P}^v(\{\epsilon^{\mu}, \epsilon^{\mu+3/3}\}) - r^{\nu}\right]_{\mathbf{i}}$$

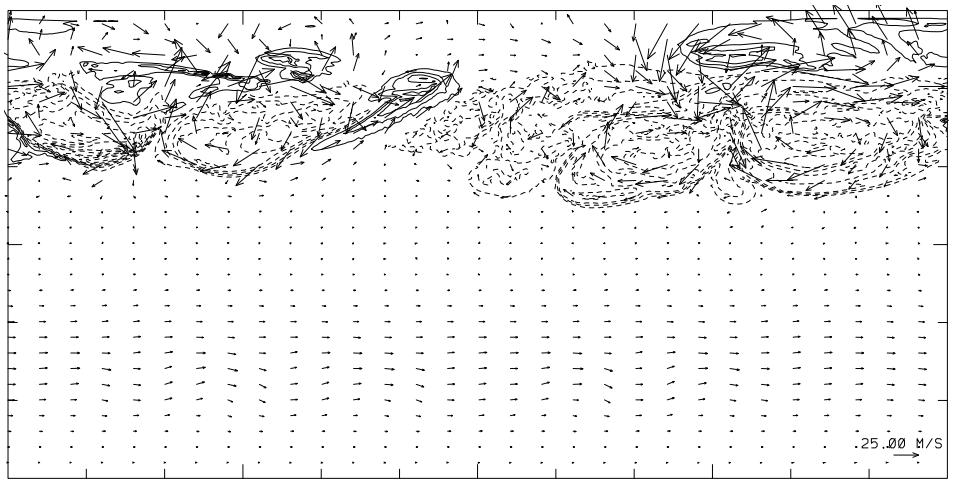
Douglas 3D ADI

Ultimately, all of them rely on the tridiagonal inversion $\epsilon_{\mathbf{i}}^{*} = (\mathcal{I} - \gamma^*\mathcal{P}^*)_{\mathbf{i}}^{-1} r^*$; comments on collocated grids, bcs, and parallelization (Povitsky, J. Parallel Distrib. Comput., 59, 1999).

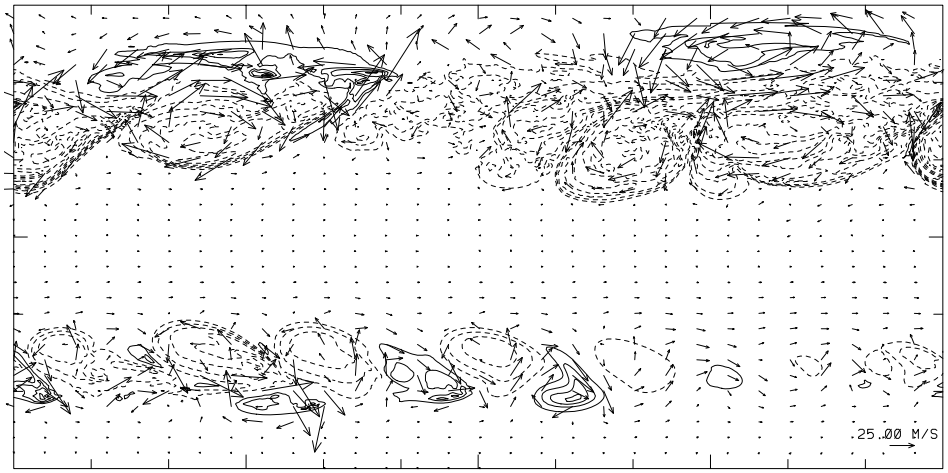
Results: baroclinic instability; domain horizontally global x 23km in radius; 360x180x48 grid, 16 days simulated time with dt = 100 s



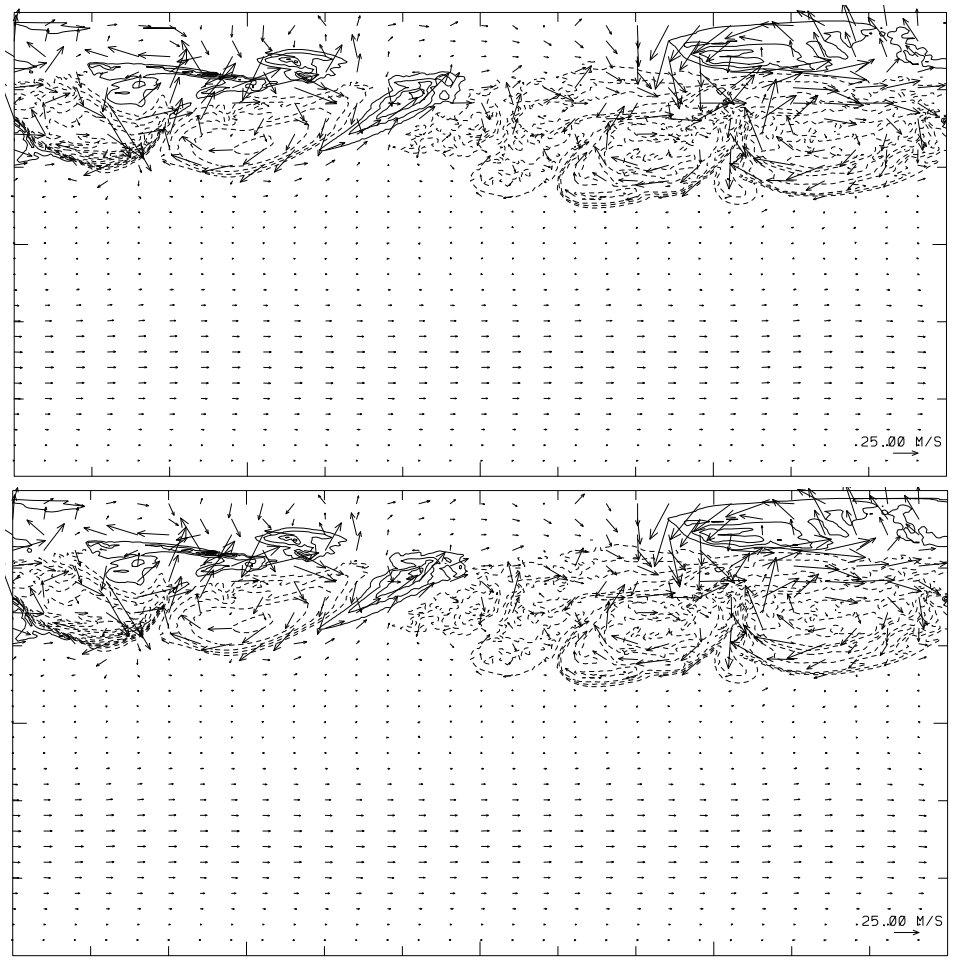
Surface potential temperature, comparison with earlier results on 128x64 grid after 8 days, contour interval 4 K



Weighted line-Jacoby reference preconditioner, 16-day simulation, on the entire horizontal domain [0,360]x[-90,90] deg, contour interval 8 K



As above, but for the vertically-implicit Richardson preconditioner.

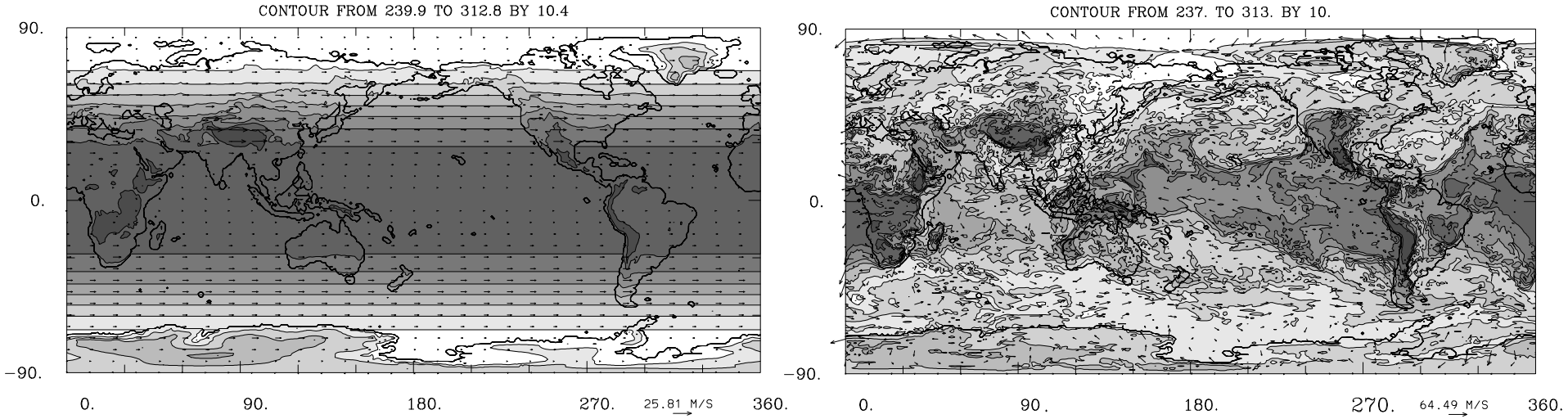


As in the preceding slide, but for xz & xyz ADI preconditioners; cf. the reference preconditioner.

Table 1
Performance of selected preconditioners.

No.	Preconditioner	Equations	Processors	Iterations	WCT [min]
0	none	$\mathcal{P} = \mathcal{I}$	$24 \times 15 \times 1$	89/40	56
1	Richardson impl. in z	(18)	$24 \times 15 \times 1$	28/14	43
2	Jacobi impl. in xyz	(19)-(21)	$24 \times 15 \times 1$	9/8	22
3	Jacobi impl. in xyz	(19)-(21)	$6 \times 15 \times 4$	9/8	19
4	ADI impl. in xz	(23)	$6 \times 15 \times 4$	14/8	20
5	ADI impl. in xyz	(25)	$6 \times 15 \times 4$	15/8	30
3b	Jacobi impl. in xyz	(19)-(21)	$12 \times 30 \times 4$	9/8	8
4b	ADI impl. in xz	(23)	$12 \times 30 \times 4$	15/8	9
3c	Jacobi impl. in xyz	(19)-(21)	$12 \times 30 \times 4$	12/9	54
4c	ADI impl. in xz	(23)	$12 \times 30 \times 4$	8/4	40

0-5 the same 360-180 resolution and the same 360 PE.
3b-4b the same 360-180 resolution but 4xPE of 3-4.
3c-4c 720-360 resolution, dt=50s, and 4PE as in 3b-4b.



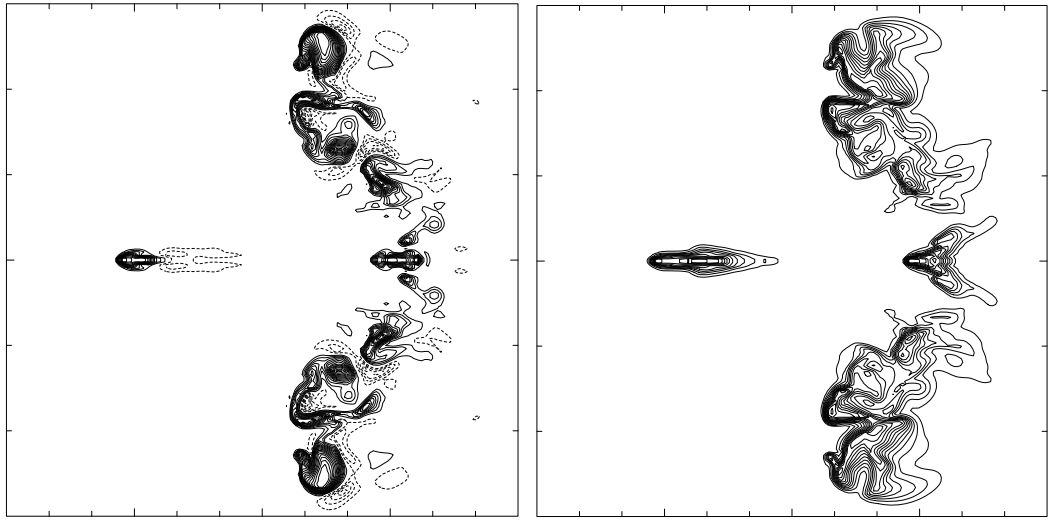
As in bottom panel of the preceding figure (ADI xyz), but along the surface orography, dt= 50 s; left the initial condition, right 16-day result.

Table 2

Supplement of Table 1; orographic forcing.

No.	Preconditioner	Equations	Processors	Iterations	WCT [min]
0t	none	$\mathcal{P}=\mathcal{I}$	6 x 15 x 4	81/62	75
1t	Richardson impl. in z	(18)	6 x 15 x 4	19/15	65
3t	Jacobi impl. in xyz	(19)-(21)	6 x 15 x 4	8/6	35
4t	ADI impl. in xz	(23)	6 x 15 x 4	14/5	30
5t	ADI impl. in xyz	(25)	6 x 15 x 4	14/5	64

Results: tornadic storm splitting; 334x167x20 km^{3,} 720x360x48 grid, 2h simulated time with dt=3 s (Smolarkiewicz, Kuehnlein, Grabowski, JCP 2017)



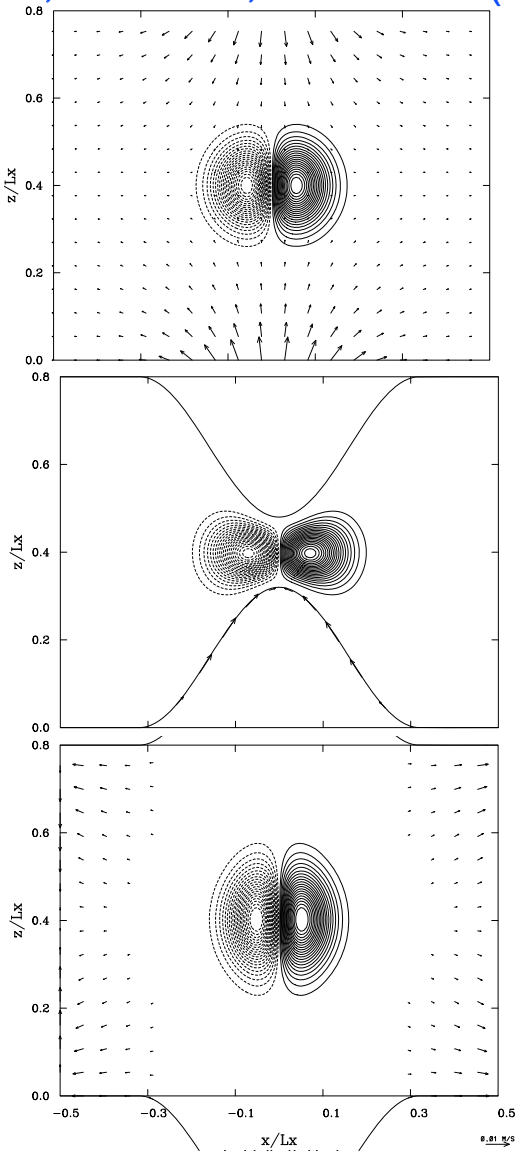
Reference ILES with line Jacoby preconditioner: vertical velocity *w* with contour interval 2 m/s (left), isolines of rainwater mixing ratio with contour interval 1 g/kg. The results are displayed after 2h simulation time at the elevation z ≈ 5km.

Performance of selected preconditioners.

No.	Preconditioner	Equations	Processors	Iterations	WCT [min]
0	none	$\mathcal{P}=\mathcal{I}$	$24\times15\times1$	4/3	14
1	Richardson impl. in z	(18)	$24\times15\times1$	3/3	16
2	Jacobi impl. in xyz	(19)-(21)	$24\times15\times1$	3/3	17
3	ADI impl. in xz	(23)	$24\times15\times1$	3/3	18
4	ADI impl. in xyz	(25)	$24\times15\times1$	3/3	22

Remark: The same grid (re DOF) as in the baroclinic-instability runs 3c and 4c of Table 1, but the solver performs much better. ➔ The significance of the problem conditioning in the spectral sense. For the baroclinic instability the ratio of the longest to the shortest wavelength supported on the grid is $\sim 4\cdot10^4$ ($\approx 4\cdot10^4$ km/1 km), while the same ratio for the tornadic storm is $\approx 4\cdot10^2$. These heuristic estimates can be formally refined, but they already hint large discrepancies in the algebraic condition numbers κ and solver performances for the physics at hand.

Results: weak toroidal magnetic field between flapping membranes, anelastic MHD equations; domain $L_x x L_y x H_0 = [-2.5,2.5]^2 \times [0,4]$, $z_{s0} = 1.6$ and $L_x = L_y = z_{s0}$; 152x152x120 grid, T=17.28, dt=0.036 (after Smolarkiewicz & Charbonneau, JCP 2013)



Reference MHD simulation, B_0=0.5e-2; B_y component and the velocity vectors shown in the central *xz* plane at t=0, 0.25 T, 0.75 T from the top to bottom, with respective contour intervals 0.1012 e-3, 0.1310 e-3 and 0.0893 e-3; the maximal arrow lengths correspond to 0.58, 0.00, and 0.01 velocities.

Table 4
Performance of selected preconditioners for weak field simulations.

No.	Preconditioner	Equations	Iterations **u**	Iterations **B**	WCT [min]
0	none	$\mathcal{P} = \mathcal{I}$	59/3	3/3	6
1	Richardson impl. in *z*	(18)	22/3	3/3	5
2	Jacobi impl. in *xyz*	(19)-(21)	25/3	3/3	5
3	ADI impl. in *xz*	(23)	29/3	3/3	6
4	ADI impl. in *xyz*	(25)	24/3	3/3	8

Conclusions:

✅ GCR(k) solver with a suite of the Richardson preconditioners maintains computational efficacy for all scale problems and becomes a universal tool for a broad range of complex tasks. This universality owes to the increased accuracy in large scales (see ! below). Our approach differs from the pursuit of a sole "super-preconditioner" with capabilities of a universal solver (e.g., spectral or multigrid).

Organizing classical stationary solvers under the idea of Richardson iteration is theoretically justified and practically advantageous for programming. The difficulty in developing the suite, is not in coding its members, but in mastering the parallel tridiagonal inversions with Neumann and periodic boundaries.

Our preconditioners include free parameters: the pseudo-time step, and the number of iterations. The pseudo-time step is determined by the convergence of preconditioner iterations, while the number of iterations is reduced to a necessary minimum. Here, both ADI algorithms used a single pass through the matrix inversion, whereas other schemes used two complete iterations. The task distribution was fixed in each group of intercomparisons. In practice, the preconditioners' setups are optimized for the case at hand.

GCR(k) is tunable via the exit condition. The physics-based criteria are used in practice: e.g., 10^3 and 10^5 for the allowed shortest time scales of the residual errors compared the time scales of physical compressibility and physical displacements, respectively, in compressible and soundproof applications.

! In entries 3t-5t of Table 2, relaxing tenfold the exit criteria had no impact on the results, but considerably reduced the respective wall-clock times. The same relaxation in 1t and 2t entries lead to the solution blowup after 9 days, showing the solution's reliance on the stringency of the stopping criterion diminishing with increasing implicitness of the preconditioners. This reflects the efficacy of the more advanced schemes in preconditioning a broader range of spectral scales. It also illuminates the dependency of the preconditioners' efficacy on the case at hand and attests to the utility of the entire suite.