

Stochastic expansions and Hopf algebras

Simon J.A. Malham and Anke Wiese
(*Heriot–Watt University, Edinburgh*)

Collaborators: Kurusch Ebrahimi–Fard (Zaragoza), Hans
Munthe–Kaas (Bergen) and Alexander Lundervold (Bergen)

Utrecht 2009: December 3rd

Brown report

Applied mathematics at the US Department of Energy 2008:

Develop new approaches for efficient modeling of large stochastic systems.

... particularly spatially dependent systems.

Invest in analysis and algorithms for stochastic optimization.

Stochastic differential equations

$$dy_t = \tilde{V}_0(y_t) dt + V_1(y_t) dW_t^1 + \cdots + V_d(y_t) dW_t^d$$

Four approaches to approximation, solve:

- ▶ PDE (transition probability distribution)
- ▶ for a weak approximation (Monte–Carlo)
- ▶ for a strong approximation (Monte–Carlo)
- ▶ pathwise (rough paths)

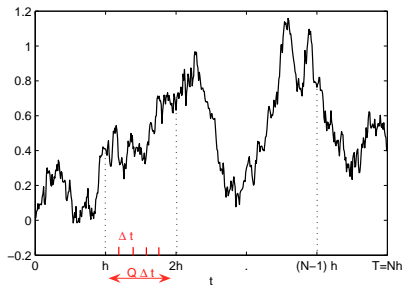
First three: expectation and higher moments of solution sought.

Basic setting

$$y_t = y_0 + \int_0^t \tilde{V}_0(y_\tau) d\tau + \sum_{i=1}^d \int_0^t V_i(y_\tau) dW_\tau^i$$

- ▶ Itô form
- ▶ Non-commuting vector fields: $V_i = \sum_{j=1}^n V_i^j(y) \partial_{y_j}$
- ▶ d -dimensional driving signal: (W^1, \dots, W^d)
- ▶ Convention: $W_t^0 \equiv t$
- ▶ Solution process: $y: \mathbb{R}_+ \rightarrow \mathbb{R}^N$

Wiener process



- ▶ $W_t - W_s \sim N(0, \sqrt{t - s})$
- ▶ Independent increments
- ▶ Continuous, nowhere differentiable

Applications

- ▶ *Finance*: Heston model for pricing options. Stock price is a stochastic process u with stochastic volatility v :

$$\begin{aligned} du_t &= \mu u_t dt + \sqrt{v_t} u_t dW_t \\ dv_t &= \kappa(\theta - v_t) dt + \varepsilon \sqrt{v_t} dZ_t \end{aligned}$$

- ▶ Molecular simulation: Langevin dynamics
- ▶ DNA damage dynamics (Chickarmane *et al.*)
- ▶ Neuronal dynamics (Coombes & Lord)
- ▶ Biochemical reactions (Burrage)
- ▶ Ocean/weather modelling: Bayesian inference (Stuart, Jones)
- ▶ Oil extraction: porous media
- ▶ Subspace tracking: inference on manifolds (Srivastava)

Itô's lemma

$$dy = \tilde{V}_0 \circ y dt + \sum_{i=1}^d V_i \circ y dW^i$$

$$\begin{aligned} \Rightarrow d(f \circ y) &= \tilde{V}_0 \cdot \partial_y(f \circ y) dt + \sum_{i=1}^d V_i \cdot \partial_y(f \circ y) dW^i \\ &\quad + \frac{1}{2} \sum_{i=1}^d (V_i \otimes V_i) : \partial_{yy}(f \circ y) \text{ "dt"} \end{aligned}$$

$$\Rightarrow d(f \circ y) = \mathcal{L}(f \circ y) dt + \sum_{i=1}^d V_i \cdot \partial_y(f \circ y) dW_t^i$$

where

$$\mathcal{L} := \tilde{V}_0 \cdot \partial_y + \frac{1}{2} \sum_{i=1}^d (V_i \otimes V_i) : \partial_{yy}$$

Related PDE

- ▶ Consider Itô SDE for $f \circ y_t$ with initial data $y_0 \equiv y$:

$$\begin{aligned} \Rightarrow f \circ y_t &= f \circ y + \int_0^t \mathcal{L}(f \circ y_\tau) d\tau \\ &\quad + \sum_{i=1}^d \int_0^t V_i \cdot \partial_y(f \circ y_\tau) dW_\tau^i \end{aligned}$$

- ▶ Feynman–Kac \Rightarrow solution of PDE

$$\partial_t u = \mathcal{L} u \quad \text{with} \quad u(0, y) = f(y)$$

$$\text{is } u(t, y) = \mathbb{E}(f(y_t))$$

- ▶ High dimensional diffusion simulation

Weak approximation

- ▶ Replace Gaussian increments $\Delta W^i(t_n, t_{n+1})$ by simpler RVs $\Delta \hat{W}^i(t_n, t_{n+1})$ with appropriate moment properties, eg. by branching process:

$$P(\Delta \hat{W}^i(t_n, t_{n+1}) = \pm\sqrt{h}) = \frac{1}{2}$$

- ▶ Expectation of approximation \hat{y}_T across all paths at the final time T is close to the expectation of the true solution:

$$\|\mathbb{E}(y_T) - \mathbb{E}(\hat{y}_T)\| = \mathcal{O}(h^p)$$

- ▶ No pathwise comparison: paths not “close” to Wiener paths

Strong approximation

- ▶ Generate approximate Wiener process paths
- ▶ Pick independent increments from

$$\Delta W^i(t_n, t_{n+1}) \sim N(0, \sqrt{h})$$

- ▶ Since we have followed Wiener path approximations, we can compare y_T with \hat{y}_T , they're close in the sense:

$$\mathbb{E} \|y_T - \hat{y}_T\| = \mathcal{O}(h^{\frac{p}{2}})$$

Exponentiation of a vector field

Consider the initial value problem

$$\frac{dy}{dt} = V(y), \quad y(0) = y_0.$$

Chain rule \Rightarrow

$$\frac{d}{dt} f \circ y = V \circ f \circ y.$$

$$\Rightarrow f \circ y_t = f \circ y_0 + \int_0^t V \circ f \circ y_\tau d\tau.$$

$$\Rightarrow y_t = y_0 + tV \circ y_0 + \frac{1}{2}t^2 V^2 \circ y_0 + \frac{1}{6}t^3 V^3 \circ y_0 + \dots$$

$$\Rightarrow y_t = (\exp tV) \circ y_0$$

Stochastic Taylor series

$$dy_t = V_0(y_t) dt + V_1(y_t) dW_t^1 + \dots + V_d(y_t) dW_t^d$$

Stratonovich form: $V_0 = \tilde{V}_0 - \frac{1}{2} \sum_{i=1}^d (V_i \cdot \partial_y V_i)$

Stochastic Taylor series

$$y_t = y_0 + \sum_i \underbrace{\int_0^t dW_{\tau_1}^i}_{J_i(t)} V_i \circ y_0 + \sum_{i,j} \underbrace{\int_0^t \int_0^{\tau_1} dW_{\tau_2}^j dW_{\tau_1}^i}_{J_{ji}(t)} V_j \circ V_i \circ y_0 + \dots$$

- ▶ Feynman–Dyson path ordered exponential, Neumann series, Peano–Baker series, Chen–Fleiss series, stochastic B-series,...
- ▶ Euler–Maruyama and Milstein methods
- ▶ Need approximations for iterated integrals

Flow map/exponential Lie series

$$y_t = \varphi_t \circ y_0$$

$$\varphi_t = \text{id} + \sum_i J_i(t) V_i + \sum_{i,j} J_{ji}(t) V_j \circ V_i + \sum_{i,j,k} J_{kji}(t) V_k \circ V_j \circ V_i + \dots$$

Set $\varphi_t = \exp \psi_t$ then \Rightarrow

$$\begin{aligned} \psi_t &= (\varphi_t - \text{id}) - \frac{1}{2}(\varphi_t - \text{id})^2 + \frac{1}{3}(\varphi_t - \text{id})^3 + \dots \\ &= \sum_{i=0}^d J_i V_i + \sum_{i>j} \frac{1}{2}(J_{ij} - J_{ji})[V_i, V_j] + \dots \end{aligned}$$

- Magnus 1954, Chen 1957, Kunita 1980, Strichartz 1987, Ben Arous 1989, Castell 1993, Castell–Gaines 1995, Moan 2004

Castell–Gaines (ODE) method

Truncated exponential Lie series across $[t_n, t_{n+1}]$:

$$\hat{\psi}_{t_n, t_{n+1}} = \hat{J}_1 V_1 + \hat{J}_2 V_2 + \hat{J}_0 V_0 + \frac{1}{2}(\hat{J}_{12} - \hat{J}_{21})[V_1, V_2]$$

Approximate solution:

$$y_{t_{n+1}} \approx \exp(\hat{\psi}_{t_n, t_{n+1}}) \circ y_{t_n}.$$

Castell–Gaines: solve ODE

$$u'(\tau) = \hat{\psi}_{t_n, t_{n+1}} \circ u(\tau)$$

across $\tau \in [0, 1]$ with $u(0) = y_{t_n}$ gives $u(1) \approx y_{t_{n+1}}$.

Signature

Stochastic Taylor expansion:

$$\varphi_t = \text{id} + \sum_i J_i(t) V_i + \sum_{i,j} J_{ji}(t) V_j \circ V_i + \sum_{i,j,k} J_{kji}(t) V_k \circ V_j \circ V_i + \dots$$

Strip away unnecessary labels:

$$\sum_{w \in \mathbb{A}^*} J_w V_w \longrightarrow \sum_{w \in \mathbb{A}^*} w \otimes w$$

\mathbb{A}^* is free monoid of words over $\mathbb{A} = \{0, 1, \dots, d\}$

Hopf algebra of words

Shuffle relations:

$$\begin{aligned}J_{a_1} J_{a_2} &= J_{a_1 a_2} + J_{a_2 a_1} \\J_{a_1} J_{a_2 a_3} &= J_{a_1 a_2 a_3} + J_{a_2 a_1 a_3} + J_{a_2 a_3 a_1} \\J_{a_1 a_2} J_{a_3 a_4} &= J_{a_1 a_2 a_3 a_4} + J_{a_1 a_3 a_2 a_4} + J_{a_1 a_3 a_4 a_2} \\&\quad + J_{a_3 a_1 a_4 a_2} + J_{a_3 a_1 a_2 a_4} + J_{a_3 a_4 a_1 a_2}\end{aligned}$$

Shuffle product: $J_u J_v \longrightarrow u \sqcup v$

Concatenation product: $V_u V_v \longrightarrow uv$

Two Hopf algebra structures, set $\mathcal{H} = \mathbb{R}\langle \mathbb{A} \rangle \otimes \mathbb{R}\langle \mathbb{A} \rangle$ with

$$(u \otimes x)(v \otimes y) = (u \sqcup v) \otimes (xy)$$

Recoding the signature

$$\varphi = \sum_{w \in \mathbb{A}^*} w \otimes w$$

$$\begin{aligned} \Rightarrow \psi &= \sum_{k \geq 1} C_k (\varphi - 1 \otimes 1)^k \\ &= \sum_{k \geq 1} C_k \left(\sum_{w \in \mathbb{A}^+} w \otimes w \right)^k \\ &= \sum_{k \geq 1} C_k \sum_{u_1, \dots, u_k \in \mathbb{A}^+} (u_1 \sqcup \dots \sqcup u_k) \otimes (u_1 \dots u_k) \\ &= \sum_{w \in \mathbb{A}^*} \left(\sum_{k=1}^{|w|} C_k \sum_{w=u_1 \dots u_k} u_1 \sqcup \dots \sqcup u_k \right) \otimes w \end{aligned}$$

Identity endomorphism

$$K(w) = \sum_{k=1}^{|w|} C_k \sum_{w=u_1 \dots u_k} u_1 \sqcup \dots \sqcup u_k$$

Linear map: $\mathbf{i}: w \mapsto w$

$\mu = (\mu_1, \dots, \mu_k)$, $w = u_1 u_2 \dots u_k$, with $|u_i| = \mu_i$:

$$\mathbf{i}^{\sqcup \mu} \circ w = (\mathbf{i}^{\mu_1} \sqcup \mathbf{i}^{\mu_2} \sqcup \dots \sqcup \mathbf{i}^{\mu_k}) \circ w = u_1 \sqcup u_2 \sqcup \dots \sqcup u_k.$$

$$\Rightarrow K = \sum_{\mu(n)} C_{\ell(\mu)} \mathbf{i}^{\sqcup \mu}.$$

Partial integration and antipode formula

$$a_1 \dots a_n - (a_1 \dots a_{n-1}) \sqcup a_n + (a_1 \dots a_{n-2}) \sqcup (a_n a_{n-1}) + \dots + (-1)^n a_n \dots a_1 = 0.$$

$$\Leftrightarrow \mathbf{i}^n + \mathbf{i}^{n-1} \sqcup \alpha + \mathbf{i}^{n-1} \sqcup \alpha^2 + \dots + \mathbf{i} \sqcup \alpha^{n-1} + \alpha^n = 0.$$

$$\Leftrightarrow -\alpha^n = \sum_{\mu(n)} (-1)^{\ell(\mu)} \mathbf{i} \sqcup \mu.$$

Sinh-log series

F sinh-log function \Rightarrow

$$C_k = \begin{cases} 1, & k = 1, \\ \frac{1}{2}(-1)^{k-1}, & k \geq 2. \end{cases}$$

$$\begin{aligned} \Rightarrow K &= \frac{1}{2} \mathbf{i}^n + \frac{1}{2} \sum_{\mu(n)} (-1)^{\ell(\mu)} \mathbf{i}^{\cup \mu} \\ &= \frac{1}{2} (\mathbf{i}^n - \alpha^n) \end{aligned}$$

$$\Rightarrow K(w) = \frac{1}{2} (J_w - J_{\alpha \circ w})$$

Sinh-log series remainder I

Recall strong error measure:

$$\|R\|_{L^2}^2 = \mathbb{E}(R^T R)$$

Set

$$\bar{R} := R^{\text{st}} - R^{\text{sl}}$$

$$\Rightarrow \|R^{\text{st}}\|_{L^2}^2 = \|R^{\text{sl}}\|_{L^2}^2 + E$$

where

$$E = \mathbb{E}(\bar{R}^T R^{\text{sl}}) + \mathbb{E}((R^{\text{sl}})^T \bar{R}) + \mathbb{E}(\bar{R}^T \bar{R})$$

Sinh-log series remainder II

$$R^{\text{st}} = \sum_{|w|=n+1} J_w V_w + \dots \quad \text{and} \quad R^{\text{sl}} = \sum_{|w|=n+1} K_w V_w + \dots$$

$$\text{Set } \bar{J}_w := J_w - K_w = \frac{1}{2}(J_w + J_{\alpha \circ w})$$

$$\Rightarrow E = \sum_{\substack{u, v \in \mathbb{A}^+ \\ |u|=|v|=n+1}} \mathbb{E} (\bar{J}_u K_v + K_u \bar{J}_v + \bar{J}_u \bar{J}_v) V_u^\top V_v$$

$$\begin{aligned} \mathbb{E} (\bar{J}_u K_v + K_u \bar{J}_v + \bar{J}_u \bar{J}_v) &= \mathbb{E} \frac{1}{2} (J_u J_v - J_{\alpha \circ u} J_{\alpha \circ v}) \\ &\quad + \mathbb{E} \frac{1}{4} (J_u + J_{\alpha \circ u})(J_v + J_{\alpha \circ v}) \\ &= \mathbb{E} \frac{1}{4} (J_u + J_{\alpha \circ u})(J_v + J_{\alpha \circ v}) \end{aligned}$$

Take home message

The abstraction we've presented gives us a handle on the shuffle algebra underlying stochastic calculus and expansions. In particular it has given us some insight that has allowed us to derive a new result and perspective. We now intend to apply these ideas to investigate the underlying algebraic and combinatorial structure of SDEs driven by other processes and SPDEs.

More challenges

- ▶ Numerical stability (Buckwar, ...)
- ▶ Symplectic methods (Tretyakov, Bou–Rabee, ...)
- ▶ Driving fractional Brownian motions (Baudoin, ...)
- ▶ Driving processes with jumps, eg. Lévy processes
- ▶ Backward stochastic differential equations (Protter, ...)

Introductory references

- ▶ Theory:

L.C. Evans: *An introduction to stochastic differential equations*

<http://math.berkeley.edu/~evans>

- ▶ Numerics:

D.J. Higham: *An algorithmic introduction to numerical simulation of stochastic differential equations*

SIAM Review 43 (2001), pp. 525–546

Itô or Stratonovich?

- ▶ Itô:

$$\int_0^T W_\tau^i dW_\tau^i = \frac{1}{2}(W_T^i)^2 - \frac{1}{2}T$$

- ▶ Stratonovich:

$$\int_0^T W_\tau^i \circ dW_\tau^i = \frac{1}{2}(W_T^i)^2$$

- ▶ Itô to Stratonovich:

$$V_0 = \tilde{V}_0 - \frac{1}{2} \sum_{i=1}^d V_i^2$$

- ▶ Stratonovich calculus familiar, easier, swapping to/back to Itô form trivial

Quadrature

How do we estimate $J_{12}(t_n, t_{n+1})$?

We can approximate it strongly using:

- ▶ Its conditional expectation $\hat{J}_{12}(t_n, t_{n+1})$
- ▶ Karhunen–Loeve expansion (Fourier expansion)
- ▶ Wiktorsson's method (most promising) — looks at the joint probability distribution function for J_1 , J_2 and J_{12}
- ▶ Stump and Hill's method (analogous)

Quadrature error I

With $\tau_q = t_n + q\Delta t$, $q = 0, \dots, Q - 1$, $Q\Delta t = h$

$$\begin{aligned} J_{12}(t_n, t_{n+1}) &= \int_{t_n}^{t_{n+1}} \int_{t_n}^{\tau} dW_{\tau_1}^1 dW_{\tau}^2 \\ &= \sum_{q=0}^{Q-1} \int_{\tau_q}^{\tau_{q+1}} W_{\tau}^1 - W_{t_n}^1 dW_{\tau}^2 \\ &= \sum_{q=0}^{Q-1} \int_{\tau_q}^{\tau_{q+1}} (W_{\tau}^1 - W_{\tau_q}^1) + (W_{\tau_q}^1 - W_{t_n}^1) dW_{\tau}^2 \\ &= \sum_{q=0}^{Q-1} J_{12}(\tau_q, \tau_{q+1}) + \sum_{q=0}^{Q-1} (W_{\tau_q}^1 - W_{t_n}^1) \Delta W^2(\tau_q) \end{aligned}$$

Quadrature error II

With $\tau_q = t_n + q\Delta t$, $q = 0, \dots, Q - 1$, $Q\Delta t = h$

$$\begin{aligned}\|J_{12}(t_n, t_{n+1}) - \hat{J}_{12}(t_n, t_{n+1})\|_{L_2}^2 &= \sum_{q=0}^{Q-1} \|J_{12}(\tau_q, \tau_{q+1})\|_{L_2}^2 \\ &= \sum_{q=0}^{Q-1} (\Delta t)^2 \\ &= Q(\Delta t)^2 \\ &= h^2/Q\end{aligned}$$

$$\Rightarrow \|J_{12}(t_n, t_{n+1}) - \hat{J}_{12}(t_n, t_{n+1})\|_{L_2} = h/\sqrt{Q}$$

Wiktorsson: simulation of Lévy area I

- ▶ Conditional distribution of $\xi = J_{12}(h)$ given $J_1(h)$, $J_2(h)$:

$$\phi(\xi) = \frac{\frac{1}{2}h\xi}{\sinh(\frac{1}{2}h\xi)} \exp\left(-a^2\left(\frac{1}{2}h\xi \coth(\frac{1}{2}h\xi) - 1\right) + ih\xi b\right)$$

where $a^2 = (J_1^2(h) + J_2^2(h))/(2h)$ and $b = J_1(h)J_2(h)/2h$

- ▶ Need to sample from this

Wiktorsson: simulation of Lévy area II

- ▶ Karhunen–Loève $X_{ij}, Y_{ij} \sim N(0, 1)$:

$$A_{ij}(h) = \frac{h}{2\pi} \sum_{k=1}^{\infty} \frac{1}{k} \left(X_{ik} (Y_{jk} + \sqrt{\frac{2}{h}} J_j(h)) - X_{jk} (Y_{ik} + \sqrt{\frac{2}{h}} J_i(h)) \right)$$

- ▶ Truncate after Q terms, mean-square error $\mathcal{O}(h^2/Q)$
- ▶ Q large, tail sum \sim multivariate Gaussian distribution
- ▶ Approximate with Gaussian RV: mean-square error $\mathcal{O}(h^2/Q^2)$
- ▶ SDELab: Gilling & Shardlow

Stochastic integral properties

Stratonovich-to-Itô relations:

$$w = a_1 a_2: J_w = I_w + \frac{1}{2} I_0 \delta_{a_1=a_2 \neq 0}$$

$$w = a_1 a_2 a_3: J_w = I_w + \frac{1}{2} (I_{0a_3} \delta_{a_1=a_2 \neq 0} + I_{a_1 0} \delta_{a_2=a_3 \neq 0})$$

$$w = a_1 a_2 a_3 a_4: J_w = I_w + \frac{1}{4} I_{00} \delta_{a_1=a_2 \neq 0} \delta_{a_3=a_4 \neq 0} \\ + \frac{1}{2} (I_{0a_3a_4} \delta_{a_1=a_2 \neq 0} + I_{a_1 0a_4} \delta_{a_2=a_3 \neq 0} + I_{a_1 a_2 0} \delta_{a_3=a_4 \neq 0})$$

Expectations of Itô integrals: $\mathbb{E}(I_{\bullet j \bullet}) = 0$ for $j \neq 0 \Rightarrow$

$$\mathbb{E}(J_i) = 0$$

$$\mathbb{E}(J_{ij}) = \frac{1}{2} h \delta_{i=j \neq 0}$$

$$\mathbb{E}(J_{ijk}) = 0$$

$$\mathbb{E}(J_{ijkl}) = \frac{1}{8} h^2 \delta_{i=j \neq 0} \delta_{k=l \neq 0}$$

Basic idea

- ▶ Construct the new series $\psi_t = F(\varphi_t)$
- ▶ Truncate and use \hat{J}_i, \hat{J}_{ij} to produce $\hat{\psi}_t$
- ▶ Approximate flow-map is $\hat{\varphi}_t = F^{-1}(\hat{\psi}_t)$
- ▶ “Flow error” is the flow remainder $R_t = \varphi_t - \hat{\varphi}_t$
- ▶ Approximate solution is $\hat{y}_t = \varphi_t \circ y_0$
- ▶ Error/remainder is $R_t \circ y_0$
- ▶ Mean-square error measure is

$$\|R_t \circ y_0\|_{L^2}^2 \equiv \mathbb{E}(R_t \circ y_0)^T (R_t \circ y_0)$$

Positivity preservation

Recall Heston model for pricing options:

$$\begin{aligned} du_t &= \mu u_t dt + \sqrt{v_t} u_t dW_t \\ dv_t &= \kappa(\theta - v_t) dt + \varepsilon \sqrt{v_t} dZ_t \end{aligned}$$

- ▶ Square-root diffusions prototypical Langevin dynamics
- ▶ Degrees of freedom $\nu := 4\kappa\theta/\varepsilon^2$
- ▶ Transition probability $P(v_t < x: v_0)$ is:

$$F_{\chi^2_\nu}(\lambda)(z) = \frac{e^{-\lambda/2}}{2^{\nu/2}} \sum_{j=0}^{\infty} \frac{(\lambda/2)^j}{j! 2^j \Gamma(\nu/2 + j)} \int_0^z \xi^{\nu/2+j-1} e^{-\xi/2} d\xi$$

where $z = x \hat{\eta}(t)$ and $\lambda = v_0 \eta(t)$ (see e.g. Andersen)

Zero boundary attracting, attainable?

i.e. solution of PDE $u_t = \mathcal{L}u$ is:

$$F_{\chi_\nu^2(\lambda)}(z) = \frac{e^{-\lambda/2}}{2^{\nu/2}} \sum_{j=0}^{\infty} \frac{(\lambda/2)^j}{j! 2^j \Gamma(\nu/2 + j)} \int_0^z \xi^{\nu/2+j-1} e^{-\xi/2} d\xi$$

Zero boundary, for:

- ▶ $\nu > 1$: non-attracting \Rightarrow implicit numerical methods*
- ▶ $\nu \leq 1$: attracting and attainable \Rightarrow transition prob

*with extreme caution depending on vector fields

Extension to SPDEs modelling biochemical reactions?

Gyöngy, Lord, Shardlow, Davie, Gaines, Stuart, Kloeden, . . .

- ▶ Brown report \Rightarrow major important challenge
- ▶ Stuart *et. al.* \Rightarrow path sampling, e.g. of bridge process
- ▶ Lots of physical/ocean/weather and defence applications