**Data Analysis Framework**

**Project:** *Sharks from Space*  
**Objective:** Identify shark movement patterns, foraging areas, and mating zones using NASA satellite data.

**1. Scientific Overview**

Earth’s oceans are dynamic ecosystems where **top predators like sharks** play a key role in maintaining ecological balance.  
Their behavior — migration, feeding, and reproduction — is tightly linked to **physical (temperature, currents)** and **biological (plankton, prey density)** parameters of the ocean.

While direct tracking (tags) provides individual movement data, **NASA’s satellite missions** offer the global, high-resolution environmental context necessary to understand *why* sharks move as they do and *where* they are likely to be found.

By combining NASA ocean datasets, we can:

* **Model habitat suitability** (where sharks are likely to forage or breed)
* **Detect environmental triggers** (eddies, fronts, temperature thresholds)
* **Correlate shark presence with physical and biological variables**

**2. NASA Data Sources and Their Analytical Use**

| **NASA Resource** | **Data Type** | **Key Variables** | **Analytical Use for Shark Ecology** | **Scientific Insight Derived** |
| --- | --- | --- | --- | --- |
| **PACE (Plankton, Aerosols, Clouds, and Ecosystems)** | Ocean color spectrometry (multi-spectral) | Chlorophyll-a, phytoplankton abundance & community composition | Map **primary productivity**; identify feeding grounds where plankton → zooplankton → fish chain initiates | Foraging hotspots; trophic chain foundation |
| **MODIS-Aqua / Terra** | Ocean color, SST | Chlorophyll-a, SST (Sea Surface Temperature), photosynthetic active radiation | Establish **long-term trends** (20+ years) in productivity & temperature | Seasonal & interannual migration patterns; preferred temperature windows |
| **SWOT (Surface Water and Ocean Topography)** | Radar altimetry | Sea Surface Height (SSH), eddy kinetic energy, ocean fronts | Detect **eddies and frontal systems** where prey accumulate | Pathways for foraging and long-distance migrations |
| **GHRSST (Group for High Resolution Sea Surface Temperature)** | Combined multi-sensor SST dataset | Sea Surface Temperature, daily variability | Identify **thermal corridors** and **mating comfort zones** | Reproductive migration triggers; surface vs. deep preference |
| **SMAP (Soil Moisture Active Passive, ocean salinity mode)** | Microwave radiometer | Sea Surface Salinity (SSS) | Track **freshwater inputs**, salinity-driven productivity | Identifies estuarine or nursery zones (juvenile habitat) |
| **GEBCO / NASA Ocean Data Portal** | Bathymetry and seafloor topography | Depth, slope, shelf boundaries | Characterize **continental shelf edges and seamounts** | Common breeding or resting habitats near slopes and reefs |
| **NOAA / NASA Ocean Color Climatologies** | Derived products | Chlorophyll anomalies, temperature anomalies | Detect **ecological anomalies** (blooms, upwellings) | Predict irregular shark aggregations; detect “event years” |
| **NASA Earth Observations (NEO)** | Cross-mission visualization portal | Sea Surface Height, Temperature, Chlorophyll | Integrate and visualize all datasets regionally | Conceptual link between environment and predator movement |

**3. Analytical Methodology**

**Step 1 — Data Integration and Preprocessing**

1. **Data Ingestion:**  
   Download daily or weekly composites from NASA’s *Ocean Color Web (OCWeb)* and *Earthdata* portals.  
   Each raster is geo-referenced and stored as NetCDF or GeoTIFF.
2. **Co-registration:**  
   Align datasets (PACE, MODIS, SWOT, GHRSST, SMAP) to a **common spatial grid (0.1°)** and **temporal scale (daily)**.
3. **Normalization and Masking:**
   * Chlorophyll → log-scaled, normalized per region.
   * SSH → calculate spatial gradient (∂SSH/∂x, ∂SSH/∂y) for eddy detection.
   * SST → normalized anomalies from climatology.
   * Mask land and poor-quality pixels using QA flags.
4. **Feature stacking:**  
   Combine all layers into a single multidimensional data cube:

**Step 2 — Ecological Variable Extraction**

| **Variable Derived** | **From Source(s)** | **Purpose / Interpretation** |
| --- | --- | --- |
| **Primary productivity (PP)** | PACE / MODIS | Indicates prey abundance; base for foraging zones |
| **Eddy kinetic energy (EKE)** | SWOT (SSH variance) | Indicates ocean mixing & aggregation of prey |
| **Thermal comfort index (TCI)** | MODIS / GHRSST | Optimal temperature range for species activity |
| **Front intensity (FI)** | SWOT gradients | Measures transition zones attracting prey |
| **Bathymetric slope (BS)** | GEBCO | Indicates potential nursery or mating sites |
| **Chlorophyll anomaly (ΔChl)** | PACE / MODIS | Signals bloom events (temporary feeding hotspots) |
| **Temperature anomaly (ΔSST)** | MODIS / GHRSST | Identifies environmental stress or attraction zones |
| **Salinity deviation (ΔSSS)** | SMAP | Tracks estuarine or freshwater influence on habitats |

**Step 3 — Correlation and Modeling**

We analyze relationships between shark presence (from tagging databases or literature) and environmental variables:

**Techniques:**

* **Spatio-temporal correlation matrices** between presence density and environmental variables.
* **Principal Component Analysis (PCA)** to reduce redundancy among ocean variables.
* **Cluster Analysis (K-Means / DBSCAN)** to identify recurring **“ecological provinces”** — zones that share environmental signatures preferred by sharks.
* **Dynamic Habitat Modeling (AI-based):**
  + Train a **Foraging Suitability Index (FSI)** model using the derived variables.
  + Incorporate **time-lagged data** (e.g., plankton bloom precedes shark arrival by 3–5 days).

**4. Identifying Ecological Zones**

| **Ecological Behavior** | **Data Indicators** | **NASA Resources Used** | **Interpretation** |
| --- | --- | --- | --- |
| 🦈 **Foraging Zones** | High chlorophyll + high eddy activity + moderate SST | PACE, SWOT, MODIS | Areas where prey density is high; sharks aggregate for feeding |
| 💞 **Mating / Breeding Grounds** | Stable SST + low EKE + specific bathymetric depth (200–1000 m) | GHRSST, GEBCO | Thermal comfort and structural protection ideal for reproduction |
| 🧭 **Migration Corridors** | Persistent SST gradients + recurrent eddies (SSH) | MODIS, SWOT | Predictable physical pathways used seasonally |
| 🐣 **Nursery Areas** | Low salinity + shallow slope (<100 m) | SMAP, GEBCO | Juvenile habitats near river outflows or continental shelves |
| ⚠️ **Stress / Avoidance Zones** | High SST anomaly + low chlorophyll | MODIS, PACE | Low productivity, unsuitable or warming waters cause avoidance |

**5. Example Analytical Workflow (Conceptual)**

1. **Daily NASA Data Layers:**
   * PACE Chlorophyll (OCx algorithm)
   * SWOT SSH & EKE
   * GHRSST SST
   * GEBCO Bathymetry
2. **Derived Composite Map:**
   * Overlay normalized maps
   * Compute FSI = w₁·Chl\_norm + w₂·EKE\_norm + w₃·SST\_band + w₄·Bathy\_edge
3. **Temporal Smoothing:**
   * Apply moving average (±3 days) to detect stable zones.
4. **Output:**
   * Hotspot raster (probability 0–1)
   * Clustered zones (feeding, mating, migration)
5. **Validation:**
   * Cross-check with known tagging data (NOAA, Global Shark Movement Project)
   * Evaluate correlation R² between predicted vs. observed presence density.

**6. Visualization and Interpretation**

* **Heatmaps**: show high FSI areas (foraging).
* **Contour maps**: represent temperature & salinity bands (mating comfort).
* **Time-lapse animations**: show how hotspots shift seasonally or with ENSO events.
* **3D bathymetric overlays**: identify nursery zones along continental slopes.

These visualizations are directly **integrated in the Web + VR App**, allowing users to explore data interactively and intuitively.

**7. Data Accessibility**

| **Dataset Portal** | **Access URL** | **Data Use License** |
| --- | --- | --- |
| NASA Ocean Color Web (PACE, MODIS) | https://oceancolor.gsfc.nasa.gov | Public, scientific use |
| SWOT Data Portal | https://podaac.jpl.nasa.gov/SWOT | Public, with attribution |
| GHRSST (NOAA/NASA SST) | https://www.ghrsst.org | Open, NASA-approved |
| GEBCO Bathymetry | https://www.gebco.net | Open, academic use |
| SMAP Ocean Salinity | https://podaac.jpl.nasa.gov/SMAP | Public, attribution required |

**8. Summary Table**

| **Behavioral Feature** | **NASA Datasets Used** | **Analytical Output** | **Ecological Interpretation** |
| --- | --- | --- | --- |
| Shark Movement Patterns | SWOT, MODIS, PACE | Time-lagged spatial correlation | Dynamic migration routes |
| Foraging Activity | PACE (Chl), SWOT (Eddy), GHRSST (SST) | Foraging Suitability Index (FSI) | Feeding hotspots |
| Reproductive Behavior | GHRSST (SST), GEBCO (Depth), SMAP (Salinity) | Thermal & structural comfort zones | Mating / nursery areas |
| Climate Stress / Avoidance | MODIS (SST anomalies), PACE (ΔChl) | Negative anomaly detection | Habitat degradation tracking |

**9. Key Scientific References**

* Braun, C. D., Gaube, P., Sinclair-Taylor, T. H., Skomal, G. B., & Thorrold, S. R. (2019). *Mesoscale eddies release pelagic sharks from thermal constraints to foraging in the ocean twilight zone.* *PNAS, 116*(35), 17187–17192.
* Gaube, P., et al. (2018). *Mesoscale eddies influence the movements of mature female white sharks in the Gulf Stream and Sargasso Sea.* *Scientific Reports, 8*(7363).
* NASA (2024). *PACE Science Data Overview.* NASA Goddard Space Flight Center.
* NASA JPL (2023). *SWOT Mission Oceanography Data User Handbook.*

**10. Conclusion**

NASA’s ocean-observing satellites provide the environmental context that explains and predicts shark movement patterns.  
By fusing **biological (PACE)**, **physical (SWOT, GHRSST)**, and **geophysical (GEBCO)** data, we can infer:

* **Where sharks feed** (high productivity + eddies),
* **Where they breed** (stable, protected thermal zones),
* **How they migrate** (along persistent frontal corridors).

This multi-source approach transforms NASA’s orbital view into actionable ecological insight — the foundation for both our **AI model** and **immersive educational experience**.